LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

COMPARISON OF CARDIAC REHABILITATION OUTCOME MEASURES

David Hevey

Thesis submitted for the Degree of Doctor of Philosophy (Ph.D.)

Department of Psychology,
University of Dublin.

TRINITY COLLEGE
9 NOV 2000
LIBRARY DUBLIN
Thesis
5912

Declarations

This thesis is submitted by the undersigned to Trinity College, University of Dublin, for examination for the degree of Ph.D. I hereby certify that this thesis has not been submitted as an exercise for a degree at this or any other university. This thesis is entirely my own work and the publications of others involved with the study are duly acknowledged within the text wherever included. I agree that the Library of the University of Dublin may lend or copy the thesis upon request.

Date 28/8/00..
Signature of candidate ..David Hevey.....

Summary

The aim of the present study was to investigate the responsiveness of psychosocial instruments to cardiac rehabilitation intervention. Responsiveness refers to the ability of an instrument to detect change over time and it is a key component of evaluative instruments. The study comprised two main components:

1. a systematic review of the literature on psychosocial outcome after cardiac rehabilitation to identify those instruments with empirical evidence documenting their responsiveness to multifactorial cardiac rehabilitation intervention
2. a prospective study to examine the responsiveness of 9 psychosocial instruments to cardiac rehabilitation intervention in a sample of cardiac patients.

Chapter 1 describes coronary heart disease (CHD) and its impact on morbidity, mortality and financial costs to society is outlined. Cardiac rehabilitation plays an important role in the secondary prevention of CHD and the structure of such services is outlined. Psychological aspects of cardiac rehabilitation are discussed, with a particular focus on the evidence that negative psychosocial states are associated with poor prognosis for CHD patients. Explicit goals of cardiac rehabilitation are to promote secondary prevention and to improve quality of life. Therefore, as an aspect of audit and quality assurance, cardiac rehabilitation programmes should document the degree to which the aim of improving quality of life is achieved.

Contributory factors to the growth of quality of life research in health-care are described in Chapter 2 and the implications of outcomes assessment for health services are considered. A distinction is made between quality of life and health–related quality of life (HRQOL) and approaches to assessing HRQOL are described. The applications of such instruments are discussed and the lack of agreement on appropriate instruments selection is highlighted. Appropriate instrument selection depends upon the purpose of the assessment.

Chapter 3 proposes the appropriate criteria for psychosocial instrument selection for outcomes assessment. In addition to the standard psychometric criteria of reliability and validity, responsiveness is discussed in detail. The main statistical approaches to the assessment of responsiveness are described and the literature comparing these different approaches is reviewed.

Chapter 4 highlights the importance of using responsive instruments to assess psychosocial outcome of cardiac rehabilitation service users. Instruments that are not sufficiently responsive to change are unsuitable as outcome tools as they underestimate the benefits of cardiac rehabilitation programme attendance. The failure of studies to take account of responsiveness is highlighted and the need for studies to examine responsiveness of psychosocial instruments in cardiac rehabilitation is outlined.

Chapter 5 describes the results of a systematic review conducted to identify the most responsive instruments in cardiac rehabilitation populations. The systematic review revealed little

consensus on psychosocial evaluation instrument use in the cardiac rehabilitation literature. However, a number of instruments were identified as showing significant potential for routine outcome assessment: Hospital Anxiety and Depression Scale (HADS), Health Complaints Checklist (HCS), Global Mood Scale (GMS), Speilberger State Anxiety Inventory (SSAI), Quality of Life after Myocardial Infarction questionnaire (QLMI), Heart Patient's Psychological Questionnaire (HPPQ), Cardiac Depression Scale (CDS), Profile of Mood States (POMS), Beck Depression Inventory (BDI), Multiple Affect Adjective CheckList (MAACL), General Health Questionnaire (GHQ-30), and the Psychosocial Adjustment to Illness Scale (PAIS). Based on the findings of the systematic review and on considerations regarding recently developed instruments, the responsiveness of nine instruments to multifactorial cardiac rehabilitation was evaluated.

Chapter 6 describes the design of a prospective study to assess the responsiveness of nine psychosocial instruments: HADS, HCS, CDS, SSAI, QLMI, GMS, HPPQ, Quality of Life Index Cardiac Version III (QLI-CV III) and the MOS 36-item short form health survey (SF-36). The nine questionnaires were randomly dived into three colour coded questionnaire batteries, each comprising three questionnaires: HADS, HCS, QLI-CVIII (white), STAI, QLMI, CDS (yellow), and GMS, HPPQ, SF-36 (blue). The questionnaire batteries were administered to 786 participants (411 cardiac rehabilitation and 375 comparison participants) and each participant completed one questionnaire battery at two time points. Cardiac rehabilitation participants completed the questionnaire before and after cardiac rehabilitation and comparison participants completed the questionnaire at equivalent time points. Responsiveness was assessed using five different statistics (Effect size (ES), Standardised Response Mean (SRM), Tuley responsiveness index, Relative Efficiency (RE), Guyatt responsiveness index) and the similarity of these statistics was compared.

The results of the study are provided in Chapter 7. The ranking of instruments provided by the SRM, ES, Tuley index and RE provided approximately the same results; the Guyatt index however was not correlated with the other statistics. Using the overall ranking provided by the statistics, the GMS emerged as the most responsive instrument. A number of other instrument subscales (GMS na, HADs dep, HPPQ d, SF swb, SF ewb, SF rlp, SF rle, SF p, HPPQ f, HCS cog) were also classified as being responsive (SRM >0.20).

Chapter 8 summarises the main findings of the study. The responsiveness of the instruments in the study is discussed and recommendations concerning instrument selection are made. The utility of combining instrument subscales to create questionnaire batteries is considered. Furthermore, the comparability of the different responsiveness statistics is discussed and issues regarding the calculation of the minimum important difference for the Guyatt index are examined. A number of methodological and theoretical limitations of the present study are described. The implications of the results for cardiac rehabilitation practice are considered and possible directions for future research are suggested.

Acknowledgements

I wish to thank the following for their support and assistance

I am very grateful to Professor Hannah McGee for her enthusiasm, unending support and encouragement throughout the course of the thesis.

I wish to sincerely thank Professor Howard Smith for his patient supervision, helpful suggestions and perceptive insights into the study, in particular his invaluable comments regarding the thesis text.

I wish to thank Professor John Horgan for both academic and financial support throughout the course of the project.

To all at the Department of Cardiac Rehabilitation, Beaumont Hospital, my gratitude for their assistance with data collection- Helen, Ali, John, Barbera, Joan, Sinead, Thelma and also for all the emotional support throughout the course of my pestering and annoyance.

A special word of thanks to all at the Department of Psychology, RCSI for offering invaluable help throughout the course of the study.

My deepest thanks to all the patients who took the time to complete an ungainly questionnaire and then took the time to complete yet another monstrous questionnaire later

The study would not have been possible without Dr D Murray, A McGowan, Dr. G Fitzgerald, B O'Brien, Dr J McMahon, M Kiernan, I Byrne, P Kelly, Dr. W. Fennell, Dr. J Fennel, Dr P Sullivan -their kind and generous assistance facilitated data collection at their centres.

To Molly, Nessa, Niamh, Louise, and Michael for assistance with the project administration.

To Erasmus 97, Erasmus 98 and Encore for teaching me that research can actually be enjoyable and that Irish people speak too quickly.

A note to thanks to my friends for their support (intellectually and emotionally) and encouragement over my many years of study.

I would like to especially thank my parents and family who have supported me in every way over the duration of all my college study and in completing the thesis – I suppose I'll have to get a proper job now!

TABLE OF CONTENTS

		Page
	Declaration	(i)
	Summary	(ii)
	Acknowledgements	(iv)
	TABLE OF CONTENTS	(v)
Chapter 1	**Coronary Heart Disease and Cardiac Rehabilitation**	1
1.0.0	Introduction	2
1.1.0	Coronary Heart Disease	3
1.1.1	Impact of CHD	5
1.1.2	CHD and lifestyle	7
1.1.3	CHD risk factors	8
1.1.4	Multiple risk factors	10
1.1.5	Summary	11
1.2.0	Definition of cardiac rehabilitation	11
1.2.1	Structure of cardiac rehabilitation	12
1.2.2	Service provision	13
1.2.3	Summary	16
1.3.0	Effectiveness of cardiac rehabilitation	16
1.3.1	Mortality	17
1.3.2	Morbidity	17
1.3.3	Return to work	18
1.3.4	Cost-effectiveness	18
1.3.5	Summary	19
1.4.0	Psychological aspects of cardiac rehabilitation	19
1.4.1	Depression	20
1.4.2	Anxiety	20
1.4.3	Negative psychosocial states	21
1.4.4	Psychosocial interventions	21
1.4.5	Summary	25
1.5.0	Summary of Key Points	25
Chapter 2	**Quality of life and Health Care**	**26**
2.0.0	Introduction	27
2.1.0	Background	28
2.1.1	Increasing demands on health services	28
2.1.2	Rise in chronically ill populations	29
2.1.3	Consumerism	30
2.2.0	Outcomes assessment	31
2.2.1	Traditional outcome measures	32
2.3.0	Defining Health-Related Quality of Life (HRQOL)	33
2.3.1	History of HRQOL assessment	36
2.3.2	Summary	38

2.4.0	Measuring HRQOL	39
2.4.1	Generic HRQOL instruments	39
2.4.2	Disease-specific instruments	40
2.4.3	Predictive ability of subjective health ratings	41
2.4.4	Criticisms of HRQOL research	43
2.4.5	Summary	44
2.5.0	Applications of HRQOL instruments	45
2.5.1	Research purposes	45
2.5.2	Health Services Research and HRQOL	46
2.5.3	Individual patient care	47
2.5.4	Audit	50
2.5.5	Population studies	50
2.5.6	Resource allocation	51
2.5.7	Summary	53
2.6.0	Summary of Key Points	54
Chapter 3	**Psychosocial Instrument Assessment Criteria**	**55**
3.0.0	Introduction	56
3.1.0	Instrument criteria	56
3.1.1	Appropriateness	57
3.1.2	Conceptual and measurement model	59
3.2.0	Reliability	59
3.2.1	Internal reliability	60
3.2.2	Reproducibility	61
3.3.0	Validity	61
3.3.1	Criterion and predictive validity	62
3.3.2	Face and content validity	62
3.3.3	Construct validity	63
3.4.3	Precision	63
3.4.1	Acceptability	64
3.4.2	Feasibility	65
3.4.3	Screening capabilities	65
3.4.4.	Cross-cultural considerations	66
3.5.0	Responsiveness	66
3.6.0	Assessment of responsiveness	67
3.6.1	Effect size	68
3.6.2	Standardised Response Mean (SRM)	69
3.6.3	Guyatt's Responsiveness Index	70
3.6.4	Minimum Important Difference (MID)	71
3.6.5	Tuley's responsiveness index	72
3.6.6	Relative Efficiency	72
3.6.7	Receiver-operating characteristics	72
3.6.8	Floor and ceiling effects	73

3.7.0	Comparison of responsiveness statistics	74
3.7.1	Summary	76
3.8.0	Limitation of previous research on responsiveness	76
3.9.0	Other contributions of responsiveness to outcomes research	77
3.9.1	Power	77
3.9.2	Statistical significance	78
3.9.3	Clinical significance	79
3.9.4	Summary	81
3.10	Summary of Key Points	82
Chapter 4	**Cardiac Rehabilitation and Outcomes Assessment**	**83**
4.0.0	Cardiac rehabilitation and outcomes assessment	84
4.1.0	Cardiac rehabilitation and audit	85
4.2.0	Benefits of HRQOL assessment in cardiac rehabilitation	86
4.3.0	Cardiac Rehabilitation and HRQOL instrument selection	87
4.3.1	Guidelines for instrument selection	88
4.4.0	Responsiveness of psychosocial measures to cardiac rehabilitation	92
4.5.0	Appropriate psychosocial states for cardiac rehabilitation outcome assessment	96
4.5.1	Positive and negative psychosocial states	97
4.5.2	Cardiac rehabilitation instrument selection	98
4.6.0	Summary of Key Points	100
Chapter 5	**Psychosocial outcomes assessments for use in cardiac rehabilitation service evaluation: A 10-year systematic review**	**101**
5.0.0	Introduction	102
5.1.0	Method	102
5.2.0	Results	103
5.3.0	Discussion	108
5.4.0	Conclusion	111
5.5.0	Aims of the present study	112
Chapter 6	**Methods**	**113**
6.0.0	Design	114
6.0.1	Questionnaire batteries	115
6.2.0	Sample size	116
6.3.0	Participants	117
6.3.1	Inclusion and Exclusion Criteria	117
6.3.2	Cardiac rehabilitation participants	118
6.3.3	Comparison participants	119
6.4.0	Procedure	120
6.5.0	Questionnaires	122
6.5.1	Demographics	122

6.5.2	Duke Activity Status Index	122
6.6.0	Questionnaire Battery A	123
6.6.1	Hospital Anxiety and Depression Scale (HADS)	124
6.6.2	Health Complaints Scale (HCS)	124
6.6.3	Quality of Life Index- Cardiac Version III (QLI-CV III)	125
6.7.0	Questionnaire Battery B	126
6.7.1	State-Trait Anxiety Inventory Form X-1 (SSAI)	126
6.7.2	MacNew Quality of Life after Acute Myocardial Infarction Questionnaire (MacNew QLMI)	127
6.7.3	Cardiac Depression Scale (CDS)	128
6.8.0	Questionnaire Battery C	129
6.8.1	Global Mood Scale (GMS)	129
6.8.2	Heart Patients Psychological Questionnaire (HPPQ)	130
6.8.3	MOS 36-item short-form health survey (MOS SF-36)	130
6.9.0	Global Rating of Change Questionnaire	131
6.10	Responsiveness Statistics	136
6.11	Statistical Analyses	137
Chapter 7	**Results**	**138**
7.0.0	Abbreviations	139
7.1.0	Response rate	140
7.2.0	Comparison of baseline demographic variables for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires	141
7.2.1	Comparison of baseline psychosocial variables for those completing both questionnaires with those completing only time 1	142
7.3.0	Demographic details of the study sample	142
7.4.0	Responsiveness of Instruments	145
7.4.1	Correlation of instrument rankings by responsiveness statistics	148
7.4.2	Overall responsiveness	148
7.4.3	Statistical significance of instrument changes	152
7.5.0	Effect of excluding each centre from the overall analysis	153
7.6.0	Classification of the minimum important difference (MID)	154
7.7.0	Comparison of parametric and non-parametric versions of the ES	156
7.8.0	Floor and ceiling effects	158
7.9.0	Relationship between time since event and SRM of each subscale	161
7.10	Missing data	162
Chapter 8	**Discussion**	**164**
8.0.0	Overview	165
8.1.0	Responsiveness of instruments	165
8.1.1	Generic versus specific instruments	171
8.1.2	Floor and ceiling effects	173

8.1.3	Cardiac rehabilitation psychosocial outcome – inappropriate instruments or small intervention effect?	175
8.1.4	Response shift	178
8.1.5	Non-specific effects	179
8.2.0	Comparison of responsiveness statistics	181
8.2.1	Classification of the MID	183
8.2.2	Responsiveness statistics	185
8.2.3	Parametric versus non-parametric effect size statistics	187
8.3.0	Limitations of the present study	188
8.3.1	Non-random design	188
8.3.2	Self-report questionnaires	190
8.3.3	Postal questionnaires	191
8.3.4	Response bias	192
8.3.5	Questionnaire ordering effects	192
8.3.6	Theoretical limitations	193
8.3.7	Other psychosocial states	194
8.4.0	Implications for routine psychosocial outcome assessment in cardiac rehabilitation	195
8.4.1	Moratorium on psychosocial instrument development?	199
8.5.0	Summary of Key Points	202
8.6.0	Overall summary	203

	List of Tables	x
	List of Figures	xi
	List of Appendices	xii
	Publications arising from this thesis	xiii
	References	178
	Appendices	210

List of Tables

Table 1.0 Established cardiovascular disease risk factors (p.8)

Table 3.0. Criteria for HRQOL instrument selection (p.57)

Table 5.1. Psychosocial assessment instruments identified as responsive (effect size difference between intervention and control ≥0.3) in one or more cardiac rehabilitation study (p.104)

Table 5.2. Summary characteristics of potential cardiac rehabilitation evaluation instruments (p. 105)

Table 6.0. Details of study centres (p.114)

Table 6.1. Format of the questionnaire battery by colour (p.115)

Table 6.2. Number of questionnaires administered at each cardiac rehabilitation centre (p.119)

Table 6.3 Number of questionnaires posted to cardiac patients from each comparison centre (p.119)

Table 6.4. Summary of psychosocial instruments (p.133)

Table 7.1 Response rates of the cardiac rehabilitation and the comparison group at time 1 and time 2 (p.140)

Table 7.2. Demographic profile of study sample (n = 786) (p.143)

Table 7.3. Ranking of responsiveness of each instrument according to each statistic (p.147)

Table 7.4. Spearman rank correlation of instruments by responsiveness index (p.148)

Table 7.5 Overall subscale ranking based on the ES, SRM, Tuley index and RE statistics (p.150)

Table 7.6. Two-way repeated measures ANOVA for each scale (p.152)

Table 7.7. Correlation between each responsiveness statistic based on all 7 centres and the results obtained after excluding each centre individually from the analysis (p.153)

Table 7.8. Mean change in psychosocial instrument scores over time for levels of self-reported change in overall quality of life (p.155)

Table 7.9. Responsiveness statistics using non-parametric versions of effect size and subsequent instrument ranking (p 157)

Table 7.10. Percentage of respondents scoring at the floor and ceiling of each instrument at time 1 (p.159)

Table 7.11. Relationship between responsiveness ranking and the presence of floor, ceiling, 10% floor and 10% ceiling effects in instruments (p.160)

Table 7.12 Percentage of missing data for instruments at each time point and overall rate of missing data.(p.162)

List of Figures

Figure 7.1 SRMs for instrument subscales (p.161)

Figure 7.2. Relationship between time since event and SRMs for each subscale.(p.151)

List of Appendices

APPENDIX 1: Summary details of intervention studies assessing psychosocial outcomes before and after multi-factorial cardiac rehabilitation intervention

APPENDIX 2: Summary details of intervention studies assessing psychosocial outcomes before and after exercise –only cardiac rehabilitation intervention

APPENDIX 3: Letter of invitation to comparison group from physician or consultant cardiologist

APPENDIX 4: Letter included in each questionnaire envelope outlining the study

APPENDIX 5: Demographic and clinical status questionnaire

APPENDIX 6: Duke Activity Status Index (DASI)

APPENDIX 7: Hospital Anxiety and Depression Scale (HADS)

APPENDIX 8: Health Complaints Scale (HCS)

APPENDIX 9: Quality of Life Index-Cardiac Version III (QLI-CVIII)

APPENDIX10: Spielberger State-Trait Anxiety Inventory Form X-1 (SSAI)

APPENDIX 11: MacNew Quality of life after Acute Myocardial infarction Questionnaire (MacNew QLMI)

APPENDIX 12: Cardiac Depression Scale (CDS)

APPENDIX 13: Global Mood Scale (GMS)

APPENDIX 14: Heart Patients Psychological Questionnaire (HPPQ)

APPENDIX 15: MOS 36-item short –form health survey (MOS SF-36)

APPENDIX 16: Global rating of change questionnaire

APPENDIX 17: Comparison of the response rate of each of the study centres

APPENDIX 18: Comparison of the response rate for each questionnaire at each centre over time

APPENDIX 19: Comparison of baseline demographic details for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires

APPENDIX 20: Responsiveness statistics for each subscale

APPENDIX 21: Percentage of responsiveness scoring at the lowest 10% and top 10% range of scores possible for each instrument at Time 1.

Publications arising from this thesis

McGee, H.M., Hevey, D., and Horgan, J.H. (1999). Psychosocial assessment in cardiac rehabilitation: A systematic review 1986-1995. *Social Science and Medicine*, **48**, 1373-1393.

Hevey, D., and McGee, H.M. (1998). The effect size statistic: Useful in health outcomes research? *Journal of Health Psychology*, **3**, 163-170.

Chapter 1
Coronary Heart Disease and Cardiac Rehabilitation

1.0.0 Introduction

The present thesis focuses on the responsiveness of psychosocial instruments to multifactorial cardiac rehabilitation intervention. Responsiveness is a critical psychometric issue that has implications for evaluative research. To date, the issue of responsiveness has received little empirical investigation within health psychology and the comparability of evaluative instruments remains largely unknown. In an era of assessment and accountability in health care, the need to examine responsiveness becomes a critical consideration. Instruments that are not responsive may not adequately assess the impact of health services on psychosocial states such as depression, energy, wellbeing, and anxiety. The importance of such psychosocial states in cardiac rehabilitation has been widely accepted (e.g. Wenger et al., 1995; World Health Organisation, 1993) and increasingly, the impact of such services on psychosocial factors requires documentation (e.g., American Association of Cardiovascular and Pulmonary Rehabilitation, 1995) for audit purposes. Thus, the need to routinely assess psychosocial states forces service providers to focus attention on the necessary psychometric properties of psychosocial instruments. Responsiveness is clearly a key property. Before discussing the role of psychosocial outcome assessment in health services (Chapter 2) and the appropriate instrument properties (Chapter 3), the present chapter outlines the medical background to the main study.

The main clinical manifestations of coronary heart disease (CHD) are briefly described. The impact of CHD in terms of morbidity, mortality and financial costs to society is outlined. Risk factors for CHD are presented and the multiplicative nature of these risks is discussed. Cardiac rehabilitation plays an important role in the secondary prevention of CHD and the structure of such services is described. Current levels of cardiac rehabilitation service provision are described and the effectiveness of cardiac rehabilitation in reducing morbidity, mortality and the financial burden of CHD are considered. Psychological aspects of cardiac rehabilitation are outlined, with a particular focus on the evidence that negative psychosocial states are associated with poor prognosis for CHD patients.

1.1.0 Coronary Heart Disease (CHD)

The cardiovascular system comprises the heart and blood vessels (arteries and veins) which transport oxygenated blood to the body's organs and return deoxygenated blood to the lungs. While a wide variety of diseases can affect the cardiovascular system (e.g., congenital heart diseases, diseases of the heart muscle), the most common type of disease in the cardiovascular system results from atherosclerosis (hardening of the arteries). Atherosclerosis of the arteries that supply the heart with blood and oxygen produces CHD.

The term CHD is slightly misleading as it does not represent a singular disease. Rather, CHD refers to a diverse set of clinical symptoms and there are a variety of clinical manifestations of CHD. Thus for example, clinical syndromes such as angina pectoris, myocardial infarction, cardiac arrest and heart failure are all considered manifestations of CHD. It has recently been suggested that the term ischaemic heart disease (IHD) is preferable to CHD as approximately 10% of CHD cases are attributable to non-atherosclerotic diseases (Stehbens, 1999) and diagnosis of IHD does not depend upon the presence of atherosclerosis. However, the term CHD continues to be more widely used and for the sake of consistency the present thesis uses the more familiar term CHD to refer to the different manifestations of ischaemia and atherosclerosis.

Clinical manifestations of CHD result primarily from the progressive development of atherosclerotic plaques and subsequent thrombus formation. The plaques vary in their composition and characteristics. Some are mainly composed of fatty tissue with a large lipid core and if they are covered by a thin fibrous cap then they may be likely to rupture (Kullo, Edwards & Schwartz, 1998). Rupturing of a vulnerable plaque may lead to subsequent platelet and fibrin deposition and the resulting coronary thrombosis (clot) formed may completely block the affected artery subsequently causing a heart attack. In other cases, the plaques may grow and limit the supply of blood beyond this obstruction and medical or surgical intervention may be required.

The common pathophysiological process underlying the symptoms of CHD is myocardial ischaemia, which refers to the relative lack of oxygen supplied to the myocardium (heart muscle) when demand exceeds supply. Angina pectoris is a discomfort in the chest and adjacent areas due to a transient inadequate blood supply to the heart, which typically

results from disease in one or more coronary arteries (Julian & Cowan, 1992). In addition, many episodes of transient myocardial ischemia are "silent" occurring without angina or other symptoms. Coronary artery bypass grafting (CABG) surgery and percutaneous transluminal coronary angioplasty (PTCA) are both highly effective methods of improving blood supply to the heart. In CABG, a section of artery or vein is used to re-route blood supply around an obstructed artery and thus provides the heart muscle with adequate blood supply. PTCA is a procedure in which a catheter with a balloon at its tip is positioned at an area of narrowing of the artery. The balloon is inflated and the stretching of the artery allows blood to flow more freely. It is quite frequent nowadays for a stent (a metal tube or coil) to be placed into the artery to prevent the artery becoming blocked again. While recovery from PTCA is more rapid compared with CABG, symptoms are more likely to recur and patients are also more likely to need further revascularisation. Thus, there are advantages and disadvantages to each approach and both technical considerations and the clinical status of the patient help determine the choice of intervention.

Angina may be regarded as being unstable if it has developed for the first time recently at relatively low workload or at rest, or if pre-existing angina has worsened for no apparent reason (Julian & Cowan, 1992). Myocardial infarction (MI) is the term applied to the death (necrosis) of a portion of the myocardium as a result of inadequate blood supply (i.e., prolonged and severe ischemia). The diagnosis is based on the history, changes in the electrocardiogram (ECG) and measurement of enzymes that are released into the blood. Cardiac failure is more difficult to define because the clinical and the physiological criteria differ (Julian & Cowan, 1992). However, in general, cardiac failure is characterised by an inadequate circulation for the body's needs. The increase in numbers of patients with heart failure reflects the larger numbers of people surviving into old age and also the improved survival from acute cardiovascular conditions. Sudden cardiac arrest results from a severe deterioration in the heart's rhythm (ventricular fibrillation). If left untreated, the cardiac rhythm deteriorates and death occurs within minutes. If no resuscitation is applied or if the resuscitation is unsuccessful, the cardiac arrest is called sudden death. The phrase "heart attack" covers both myocardial infarction and sudden cardiac death. Thus, from this brief overview, it is clear that there are a variety of different manifestations of CHD.

1.1.1 Impact of CHD

Coronary heart disease (CHD) continues to be the leading cause of death in adulthood in most of the industrialised countries and accounts for one quarter of all deaths in developing countries (Reddy, 1998). While premature death rates from heart disease in the industrialised nations have been declining since the mid-1960s/1970s, compared to other countries such as the US and Canada, the Irish and UK rate of reduction has been less pronounced.

Cardiovascular disease is the main cause of death in Ireland, accounting for approximately 35% of all deaths (Department of Health, 1999). The mortality rate of heart attack per 100,000 of the population (all ages) is much higher than the EU average (238 per 100,000 in Ireland compared with 98 per 100,000 in EU overall). In addition, Ireland also has the highest death rate from heart disease in the under 65 age group in the EU, with almost double the EU average (60.7 per 100,000 in Ireland compared with 32.6 per 100,000 on average in EU).

Cardiovascular disease is the single largest cause of mortality in Ireland. The reasons for the high rates of mortality are not yet fully established but it is likely that both genetic and lifestyle factors contribute. The identification of genes associated with CHD is underway and the Human Genome Project offers hope of determining the relative contribution of specific genetic markers of CHD. The recently commenced multi-annual Survey of Lifestyles, Attitudes and Nutrition (SLÁN) project will provide ongoing information on risk factors for CHD in Ireland. High levels of smoking, blood pressure, dietary fat, blood cholesterol, BMI, and physical inactivity are considered established risk factors for the development of CHD (Department of Health, 1999). Future challenges for research include investigating the interaction between genetic and lifestyle factors.

It has been argued that many of the behaviours that constitute risk factors for CHD emerge early in life and reflect genetic can environmental influences (National Heart, Lung, and Blood Institute (NHLBI), 1998). Several risk factors appear to show significant heritability (e.g., tobacco and alcohol dependence, and obesity). In addition, both psychosocial and environmental factors contribute to the development and maintenance of lifestyle risk factors. For example the role of cultural factors in CHD needs to be considered

given the marked cultural variability in rates of CHD. The similar rates of CHD reported in Ireland and the UK may reflect a common genetic disposition or a common perspective on health which is not conducive to maintaining healthy lifestyles, or a combination of both. Notably, international cardiovascular statistics demonstrate two clusters of countries that dominate league tables for cardiovascular mortality: the five countries that constitute Ireland and the United Kingdom and Eastern European countries. As yet, the balance of genetic and cultural factors that create these clusters is unclear.

Social risk factors such as social deprivation, life stress and lack of social support are increasingly recognised as risk factors for CHD (NHLBI 1998; Rozanski et al., 1999). Therefore in addition to the lifestyle factors, wider social factors contribute to the increased risk for CHD. For example, the well documented socio-economic status "gradient" for health outcomes is not accounted for entirely in terms of lifestyle risk factors (Adler et al, 1994). It has been argued that patterns of health and illness cannot be adequately understood in purely individual terms and thus far, it has proved impossible to reduce cultural and structural effects to lifestyle differences (Abraham & Hampson, 1996). Indeed, there is a danger that focus on the individual may result in neglecting wider socio-economic factors. In addition, significant ethnic differences have been related to CHD risk in a number of multicultural societies (e.g. USA, UK) and recent demographic trends across Europe mean that Ireland will also face the same issues.

In addition to the tremendous human costs of CHD, the financial costs to the state in terms of premature disability and death are substantial. Extrapolation from UK data suggests that CHD costs at least £30 million in Ireland annually with over £100 million of potential national revenue lost as an indirect result of the disease each year (McGee, Graham, Newton & Horgan, 1994). In the UK it has been estimated that CHD costs the NHS £1,600 million in 1996 (British Heart Foundation, 1998). Furthermore, production losses from death and illness in those of working age and from the informal care of people with CHD costs the UK economy £8,500 million. Thus, taken together, CHD is estimated to cost the UK a total of £10,000 million annually (British Heart Foundation, 1998). In the US, latest data from the American Heart Association (AHA) estimates that the cost of cardiovascular disease and stroke is approximately $286 billion (AHA, 1998). In summary, CHD places an enormous

burden on health care costs in many industrialised societies and remains a major public heath issue.

1.1.2 CHD and lifestyle

The national outline of current illness patterns in Ireland show that the main causes of morbidity and mortality are diseases in which lifestyle and environmental factors play a key role. Thus, it is assumed that intense and prolonged health promotion programmes should demonstrate considerable benefit in reducing premature mortality from CHD. Large-scale community health promotion strategies have been developed in a number of countries, including the North Karelia Project in Finland and the Stanford Three Communities Study in the USA. The Kilkenny Health Project (Shelley, Daly, Graham et al, 1991) evaluated a community prevention intervention in an Irish context. During the study's time frame, there were substantial reductions in CHD mortality rates and risk factors in both the intervention and the reference populations. Similar results have been reported from the UK, with the Heartbeat Wales programme documenting that in comparison to the intervention group, the reference group had broadly similar health service policies in place (Tudor-Smith, Nutbeam, Moore & Catford, 1998). Therefore, for some of the large-scale primary prevention programmes, evaluation of the effectiveness of the programme was hindered by the simultaneous improvement in the non-intervention reference areas. It has been suggested that the large-scale interventions have occurred at a time when knowledge and willingness to change was reasonably widespread as a result of media interest in healthy lifestyles (Department of Health, 1999). While the reason for the decline in CHD mortality is not fully understood, changes in lifestyle and more aggressive medical treatment of risk factors are widely accepted as contributory factors. Much emphasis has been placed on primary prevention efforts to modify the established risk factors such as smoking, diet, obesity and lack of exercise. In addition to primary prevention strategies, secondary prevention and rehabilitation are also considered essential components of the WHO strategy for controlling CHD (WHO, 1993). The present thesis focuses on secondary prevention of CHD.

To-date, the causes of coronary artery disease are not fully established. However, a number of risk factors have been identified and before considering the role of cardiac

rehabilitation in secondary prevention, it may prove useful to briefly consider the risk factors for CHD as cardiac rehabilitation explicitly aims to favourably influence these risk behaviours.

1.1.3 CHD risk factors

The cornerstone of optimal care for cardiac patients is risk factor management (27th Bethesda Conference, 1996), typically involving a multidisciplinary team including cardiovascular specialists and other physicians in conjunction allied health professionals such as nurses, exercise physiologists, dietitians, psychologists, and physical and vocational therapists. Cardiovascular disease risk factors can be divided into modifiable and nonmodifiable factors (Table 1.0). The major modifiable risk factors for recurrent cardiac events are the same as those for the initial development of coronary artery disease.

Table 1.0 Established cardiovascular disease risk factors

Non-modifiable factors
Age
Sex
Family history
Modifiable risk factors
Cigarette smoking
Hypertension
Elevated cholesterol
Elevated low-density lipoprotein (LDL) cholesterol
Reduced high-density lipoprotein (HDL) cholesterol
Obesity
Diabetes mellitus
Physical inactivity

In the USA, nearly one in five deaths from cardiovascular diseases are attributable to smoking (AHA, 1998). The relative risk of death from cardiovascular disease in smokers

compared to non-smokers is roughly doubled and tobacco smoking is a critical modifiable risk factor, with smoking cessation significantly reducing the risk for cardiac events (Fowler, 1998). Hypertension is one of the major risk factors for CHD and the risk increases continuously with increasing levels of both systolic and diastolic blood pressure. Research has demonstrated that increasing serum cholesterol is an important cause of CHD and that lowering serum cholesterol reduces this risk (American Heart Association, National Heart Lung and Blood Institute, 1990). A number of recent trials of lipid lowering drugs ("statins") have shown that decreases in LDL-cholesterol are associated with decreases in fatal and non-fatal CHD, and also CHD mortality (e.g., Scandinavian Simvastatin Survival Study group, 1995). In addition the statins increase HDL-cholesterol, which would be expected to decrease CHD mortality by a further 10% beyond the large effect of the reduction in LDL-cholesterol (Law, 1998). Thus, risk factor management is effective as both primary and secondary prevention interventions for CHD.

While obesity, diabetes and physical inactivity are clearly associated with increased risk of cardiovascular disease, the evidence is less clear that primary prevention modification of these factors decreases CHD risk (Manson et al., 1992). However, two-thirds of people with diabetes mellitus die from some form of heart or blood vessel disease (AHA, 1998). The relative risk of CHD associated with physical inactivity ranges from 1.5 to 2.4, an increase comparable with to that observed for high blood cholesterol, high blood pressure or cigarette smoking. A comprehensive review by the Agency for Health Care Policy and Research (Wenger et al., 1995) demonstrated the benefits of exercise training in cardiac patients in relation to morbidity, mortality, exercise tolerance and symptoms.

Recent years have witnessed increased attention on the role of psychological factors in the aetiology of CHD. There is now a growing body of evidence suggesting a possible role for depression, anxiety, social support, vital exhaustion, hostility and chronic life stress (Rozanski, Blumenthal & Kaplan, 1999; National Heart, Lung, and Blood Institute, 1998) in the development of CHD. Furthermore, in CHD populations, prospective studies suggest that depression and social support may predict future outcomes (Hemingway & Marmot, 1998). The present thesis will not comprehensively review this literature, although some studies of relevance to these issues will be discussed later.

1.1.4 Multiple risk factors

Numerous decades of research have demonstrated that CHD is multifactorial in causation. The coexistence of multiple risk factors confers a magnified risk which is multiplicative rather than merely additive (Reddy, 1998). Of note, this finding also holds for psychosocial risk factors (Rozanski et al., 1999). However, it is only relatively recently that trials to change multiple risk factors in cardiac patients have emerged within the literature. One of the first randomised clinical trials of multiple risk factor on post-MI patients was known as the MULTIFIT trial (DeBusk et al., 1994). The MULTIFIT programme incorporated a nurse case management approach to cardiovascular risk management, focusing on exercise training, smoking cessation and dietary and drug management of hyperlipidemia.

Similarly, the Stanford Coronary Risk Intervention Program (SCRIP) evaluated the effects of multifactor risk modification in the progression of angiographic coronary artery disease (CAD) in 300 patients (Haskell et al., 1994). The intervention consisted of exercise, dietary modifications, weight loss, lipid lowering pharmacotherapy and smoking cessation. At 4 years follow-up, patients in the intervention group had on average a 20% increase in exercise capacity, a 4% decrease in weight, and a 22% decrease in LDL-cholesterol compared with the usual care group. In addition, the intervention group showed a significant reduction in coronary disease progression and a decrease in the composite endpoint of death, non-fatal MI, PTCA or CABG. Intensive and comprehensive lifestyle changes have been shown to produce decrease in atherosclerosis in patients with documented CAD (Ornish et al., 1990). Recently published follow-up data from this study demonstrated that these benefits are maintained in those who adhere to the treatment programme (Ornish et al., 1998). However, this intervention was quite extreme (e.g., moderate aerobic exercise, a 10%-fat vegetarian diet, stress management training, smoking cessation and group psychological support) and may not be applicable in typical secondary prevention interventions in cardiac rehabilitation. It should be noted that while significant benefits can be achieved through risk factor management, nearly 25% of patients with premature cardiovascular disease do not have one of the established risk factors (Hoeg, 1997) and additional risk factors are currently being investigated.

Patients with established coronary disease are ten times more likely to develop a (second) myocardial infarction than persons without known coronary disease (Pekkanen et al., 1990). Risk factor intervention in patients with cardiac disease is of great importance because the event rate in patients with established coronary disease is five to seven times the rate of patients with similar risk factors but with no evidence for cardiac disease (Pekkanen et al., 1990). However, while intensive risk factor modification is of utmost importance in CHD patients, evidence shows that risk factor management is far from optimal. Less than 50% of patients with established CHD receive optimal risk factor management in Europe or in the USA (Gohlke & Gohlke-Bärwolf, 1998). Of particular relevance, only 5% of patients in the USA participate in cardiac rehabilitation (Pearson, McBride, Miller & Smith, 1996) and in Europe most countries provide cardiac rehabilitation services to less than 50% of eligible patients (Vanhees, McGee, Dugmore, Vuori & Pentilla, 1999).

1.1.5 Summary

The various manifestations of CHD have substantial consequences for society in terms of mortality, morbidity and financial costs. A number of risk factors for development of CHD have been identified and primary prevention attempts to favourably influence risk factors such as smoking, diet, obesity and lack of exercise. Despite the evidence in favour of their efficacy, secondary prevention strategies for CHD are frequently under-utilised. Cardiac rehabilitation is one such form of secondary prevention and it has been suggested that it is an essential part of care that should be available to all cardiac patients (WHO, 1993).

1.2.0 Definition of cardiac rehabilitation

Cardiac rehabilitation may be defined as

> "the sum of activities required to influence favourably the underlying cause of the disease, as well as to ensure patients the best possible physical, mental and social conditions so that they may, by their own efforts, preserve, or resume when lost, as normal a place as possible in the life of the community. Rehabilitation cannot be regarded as an isolated form of therapy but must be integrated into the whole treatment of which it forms only one facet" (WHO, 1993)

In order to achieve these aims, some patients may require advice and encouragement to achieve a healthy lifestyle, while others may have psychological issues which must be addressed. The US Agency for Health Care and Policy Clinical Guidelines for Cardiac Rehabilitation (Wenger et al., 1995) stated that cardiac rehabilitation is characterised by comprehensive, long-term services involving medical education, prescribed exercise, cardiac risk factor modification, counselling, and behavioural interventions.

These definitions clearly highlight the multifactorial nature of cardiac rehabilitation. In order to provide a service that comprises a number of factors, a multidisciplinary team of health care professionals may be required. As discussed previously, there are a number of well-established risk factors for coronary heart disease and thus, cardiac rehabilitation typically comprises of multifactorial interventions. The evidence on the effectiveness of cardiac rehabilitation will be highlighted before considering the role of psychosocial aspects of cardiac rehabilitation. However, before describing the evidence on cardiac rehabilitation effectiveness, it is necessary to outline the structure of cardiac rehabilitation services currently provided to cardiac patients.

1.2.1 Structure of cardiac rehabilitation

Four phases can typically be distinguished in the structure of cardiac rehabilitation in Europe. Phase I refers to the acute treatment of patients admitted to the intensive care unit and the in-hospital stage in recovery. Phase II typically follows discharge from hospital. It is a period of convalescence for the patient and may involve gradual increases in exercise, such as daily walking. However, there are differences between countries as to what happens in phase II (Gohlke & Gohlke-Bärwolf, 1998). Residential rehabilitation including risk factor modification in specialised clinics is traditionally practised in Germany, Austria, and partially in Switzerland, France and Hungary. Alternative approaches include ambulatory rehabilitation in a specialised department of rehabilitation, associated with a large community hospital or managed by a cardiologist, or non-structured ambulatory care by private physicians. Phase III involves an exercise-based out-patient rehabilitation programme in the hospital. Some programmes commence almost immediately following discharge, but in the

UK most programmes commence between the fourth and eighth week depending upon the amount of cardiac damage (Horgan et al., 1992).

Exercise is frequently linked to education, group discussion, counselling and risk factor management. In most European countries, patients travel three times a week for a period of 8 to 16 weeks (Maes, 1992). At the end of this stage, the patient is considered to have completed rehabilitation and is ready to resume their daily activities. One notable exception to this format is Germany where cardiac rehabilitation patients are treated in rehabilitation clinics for six to eight weeks after discharge form the general hospital (Maes, 1992). Phase IV refers to the long-term maintenance period and community based exercise sessions are typically regarded as Phase IV cardiac rehabilitation. There are important differences both between and within European countries regarding cardiac rehabilitation service delivery (Vanhees et al., 1999).

1.2.2 Service provision

Cardiac rehabilitation services vary greatly between centres in the UK (Horgan et al., 1992) with little consistency in level of funding, selection criteria, programme content, or programme evaluation. Exercise training was the primary component of the early programmes and remains the core of most cardiac rehabilitation programmes to this day. There is now widespread consensus that a more comprehensive approach to cardiac rehabilitation should also include secondary prevention, relaxation, education, stress management and psychosocial aspects (WHO, 1993). In addition, there is also widespread agreement that cardiac rehabilitation should also involve partners and close family members of patients in the rehabilitation process.

Previously, only patients recovered from an uncomplicated MI participated in cardiac rehabilitation. However, patients with complications of the acute episode, those with angina pectoris of varying severity, and those who have undergone CABG and PTCA are now also considered candidates for cardiac rehabilitation participation. In addition, recent years have witnessed calls for heart transplantation (e.g., Kobashigawa et al., 1999) and heart failure patients (e.g., McLeod & Coats, 1995) also to be given access to cardiac rehabilitation services.

The overall level of provision in the UK and Ireland has increased rapidly over the last 10 to 15 years. There is wide variation in the practice and organisation and management of cardiac rehabilitation services. Most programmes are outpatient, hospital-based, concentrating on low-risk patients who have had MI, although many now include CABG or PTCA patients. Although women account for over one-third of CHD patients, they are less likely to receive cardiac rehabilitation then men.

In 1998 the Irish Association of Cardiac Rehabilitation (IACR) surveyed the 40 public hospitals in Ireland which provide services to cardiac patients (McGee, Hevey, & Horgan on behalf of the IACR, manuscript in preparation). Only 12 centres (30%) provided outpatient cardiac rehabilitation services. Furthermore, there was marked variation in the distribution of such services, with certain health board regions having no cardiac rehabilitation centres available. To illustrate the dramatic growth in service provision in Ireland, ten of the twelve cardiac rehabilitation services had been established in the past five years. Among the centres without cardiac rehabilitation services, 12 had plans towards programme development while 16 had no such plans. This highlights the underdeveloped nature of cardiac rehabilitation services in Ireland at present. It is hoped that the recent report of The Cardiovascular Health Strategy Group *"Building Healthier Hearts"* (Department of Health, 1999) will provide much needed impetus towards the development of a national cardiac rehabilitation strategy.

The IACR study also highlighted that the current level of service provision means that only a minority of eligible patients are receiving cardiac rehabilitation. Less than one in ten patients for whom referral to cardiac rehabilitation should be considered are currently participating in cardiac rehabilitation programmes. This figure compares unfavourably with other EU countries such as Finland (50%), the Netherlands (60%) and Austria (95%). Indeed the recent CARINEX study (Vanhees et al., 1999) reported that those countries with the highest mortality from cardiovascular disease have the lowest rates of cardiac rehabilitation services.

Two recent studies have examined the structure and format of cardiac rehabilitation programmes offered to patients in the UK. A survey of 244 programmes in the UK found that the majority of programmes are exercise-based, usually providing group aerobic exercise

sessions once a week for an average of 6 to 10 weeks (Thompson, Bowman, Kitson, de Bono & Hopkins, 1997). Patient education is provided in 70-80% of programmes, either on a one-to-one basis or in group sessions. The majority of programmes also provide relaxation training either as a single session or more frequently. Other forms of psychological treatments are provided in 13% of centres.

Thompson et al. (1997) selected a random sample of 25 centres from a listing of hospitals admitting cardiac patients in England and Wales and each centre was visited by one of the researchers and completed a detailed interview schedule. The aspects of relevance to the present study will be reported. Half of the centres dealt with perceived psychological problems by prescribing medication, 10 offered some form of counselling and four referred patients to a psychologist. In terms of psychosocial assessment, only 12% of this sample attempted to assess the psychological state of the patients. Two centres used the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), one centre as part of The Heart Manual programme, and the other centre in this sample used a stress scale. Unfortunately, the authors did not report whether these assessments were once-off measures or whether there was an attempt to examine the effects of cardiac rehabilitation on these psychological states over time.

According to a similar survey, psychosocial states are poorly assessed in the UK (Lewin, Ingelton, Newens & Thompson, 1998). Validated psychosocial assessment (defined as the use of a published scale or of a standardised procedure with known properties) was reported in only 27% of cardiac rehabilitation centres. In addition, a further 4% used informal assessment (defined as any other method not included in the pervious definition). Thus, just under 70 % of UK cardiac rehabilitation programmes report not collecting standardised psychological data on their participants, despite the fact that the UK national guidelines (Thompson, Bowman, Kitson, deBono & Hopkins, 1996) explicitly call for routine psychological assessment. Unfortunately, Lewin et al.'s (1998) study does not provide a breakdown of the variety of psychological measures used throughout cardiac rehabilitation. The importance of standardised psychosocial data collection in cardiac rehabilitation will be discussed in more detail in section 4.2.0.

In general, it appears that contrary to the UK national guidelines (Thompson et al., 1996), cardiac rehabilitation programmes have not as yet engaged in much psychosocial audit of their service.

1.2.3 Summary

The practice of cardiac rehabilitation in Europe differs greatly in various countries and thus, cardiac rehabilitation as currently practised in Europe is difficult to define and to evaluate. Cardiac rehabilitation in the UK and Ireland has tended to develop as a consequence of local staff enthusiasm. Resources are scarce and programmes have adapted to local circumstances. Thus, cardiac rehabilitation services in the UK and Ireland vary enormously in terms of content and structure. Variability in funding has also shaped cardiac rehabilitation development. In general, cardiac rehabilitation programmes tend to be highly regimented with all patients receiving the same components as part of a fixed programme, regardless of their individual needs; there is still however a great deal of disparity between programmes.

Current evidence suggests that cardiac rehabilitation programmes are not collecting psychosocial data on cardiac rehabilitation participants in a standardised manner. Little research has evaluated the optimal frequency and duration of cardiac rehabilitation programmes and there is a need to evaluate different formats to determine the optimal method of delivering the service. The lack of standard information gathering makes objective assessment of the benefits of cardiac rehabilitation difficult and the literature demonstrating the effectiveness of cardiac rehabilitation will now be discussed.

1.3.0 Effectiveness of cardiac rehabilitation

The goals of cardiac rehabilitation are to promote secondary prevention and to improve quality of life (WHO, 1993) and the effectiveness of cardiac rehabilitation in meeting these goals should be routinely evaluated. Research has typically focused on outcomes such as morbidity, mortality, return to work and cost. Each of these areas will now be discussed.

1.3.1 Mortality

The evidence for the beneficial effects of cardiac rehabilitation on mortality comes primarily from two meta-analyses published in the late 1980s. Oldridge and colleagues included 10 randomised controlled trials of cardiac rehabilitation (Oldridge, Guyatt, Fischer & Rimm, 1988). They reported that cardiac rehabilitation was associated with a significant decrease in cardiac mortality (odds ratio (OR) 0.76) and all cause mortality (OR 0.75). Similar findings were reported in a meta-analysis of 22 randomised controlled trials (O'Connor et al., 1989). At three years follow up, cardiac rehabilitation patients showed a significantly better cardiac mortality (OR 0.78), all cause mortality (OR 0.80), fatal re-infarction rate (OR 0.75) and sudden death rate (0.63). It has been estimated that cardiac rehabilitation results in between 1.9 to 9.2 potential lives saved per 1,000 patients (Meyers, 1996).

However, it is important to consider that these meta-anlayses focused on low-risk patients and subsequently the scope for demonstrating benefits in this population may be small. While it is possible that results illustrating greater reduction in mortality would be obtained using a more high-risk group of patients, caution should be exercised when generalising these results to other populations. As yet, the survival benefits of cardiac rehabilitation attendance among women and elderly patients have not been established.

1.3.2 Morbidity

Neither of the meta-analyses described previously showed a significant decrease in the rate of non-fatal MI. These findings may be interpreted as showing the cardiac rehabilitation does not reduce the risk of subsequent MI. However, it may be that after completing a cardiac rehabilitation programme, patients are more likely to experience a non-fatal MI instead of fatal MI. Such a conclusion would support the benefits of cardiac rehabilitation. It should be noted that research on beneficial effects of cardiac rehabilitation on mortality and morbidity is characterised by small numbers, high treatment crossover, and drop-out rates.

1.3.3 Return to work

Re-employment was previously considered an important measure of the success of cardiac rehabilitation. However, given the current demographic profile of cardiac patient groups (including many elderly patients and also many women), this measure is no longer considered appropriate. Assessing the impact of cardiac rehabilitation on return to work is complicated by the fact that many additional factors influence this outcome. Rates of return to work after MI vary considerably. Of patients in employment prior to suffering a MI, 62-92 % return to work following cardiac rehabilitation (Shanfield, 1990). A recent study examined return to work rates among 1252 employed CHD patients receiving CABG or PTCA (Mark et al., 1992). The authors reported that approximately 20% of patients did not return to work after revascularisation. Of note, in addition to demographic and socioeconomic factors, the patient's perception of activity status was predictive of return to work. Thus, failure to return to work is often attributable to psychological factors rather than to underlying physical constraints. The effects of cardiac rehabilitation on return to work rates have been mixed and this may in part be attributable to the different populations used.

1.3.4 Cost-effectiveness

To date a limited number of studies have examine the cost-effectiveness of cardiac rehabilitation. While it has been reported that cardiac rehabilitation is cost-effective (Oldridge et al., 1993), many of the control group in the study attended some form of rehabilitation and thus these results may have underestimated the true benefit of cardiac rehabilitation. Levin, Perk and Hedback (1991) calculated that the cost of cardiac rehabilitation was offset by a subsequent reduction in hospital re-admissions. In addition, the higher rates of employment among the cardiac rehabilitation patients in comparison with the control group produced a net societal saving of $10,000 per patient. However, this study was performed in Sweden and differences in health care may give the findings a lack of generalisability. Moreover, this study was not a randomised controlled trial (RCT). Indeed, it should be noted that many of the cost-effectiveness studies were not RCTs and there is a need for such studies to be conducted on this crucial topic.

As health care moves into an era increasingly characterised by constraints in spending, it is essential that cardiac rehabilitation provides the best possible scientific evidence of its cost-effectiveness. Chua and Lipkin (1993) concluded that cardiac rehabilitation programmes are cost effective and should be made available to all who would benefit from them. However, the evidence to support such a recommendation is not as scientifically conclusive as Chua and Lipkin (1993) suggested, particularly given the absence of RCT data. In addition, cardiac rehabilitation is clearly not a homogenous service and there are a range of factors that influence the costs and cost-effectiveness of the process. Cardiac rehabilitation may not be cost-effective in all formats for all patients and there is a need to assess the relative costs and outcomes of different formats of cardiac rehabilitation.

1.3.5 Summary

The clinical practice guidelines established for cardiac rehabilitation in the USA (Wenger et al., 1995) reported that cardiac rehabilitation is associated with a range of benefits including increased exercise tolerance, improved symptoms, improved blood lipids, decreased cigarette smoking, improved well-being, and reduced mortality. In addition, there is some evidence to suggest that cardiac rehabilitation is a cost-effective intervention. The key components of cardiac rehabilitation identified in the guidelines are exercise, education and risk factor modification. It should be noted that most of the evidence concerning the efficacy of cardiac rehabilitation is based on studies conducted almost exclusively on a select group of middle-aged males and this limitation raises some concerns regarding the generalisability of the results.

1.4.0 Psychological aspects of cardiac rehabilitation

The needs of people recovering from an acute cardiac event vary. Some have psychological problems or misconceptions about their health that may make it difficult for them to return to a normal life. Some require help in modifying pre-existing risk factors such as smoking, poor diet or lack of exercise. Most are likely to benefit from lifestyle changes such as increasing physical activity and following a low-fat diet. Coronary artery disease is often associated with adverse psychological sequelae. Psychological reactions to

the cardiac event are not consistently related to infarct size or any measure of cardiac status (Ladwig et al., 1992) and psychological adaptation strongly influences recovery from cardiac events.

The present thesis will discuss the psychological states that have been most frequently associated with poor outcome in the literature: depression, anxiety and negative psychosocial states.

1.4.1 Depression

It has been estimated that 16-22% of patients who have suffered a MI are affected by major depression (Carney, Freedland, Sheline, & Weiss, 1997) with minor depression present in a much larger proportion of patients. Depression places individuals at increased risk for developing CHD (e.g., Ford et al., 1998) and this risk exists along a continuum according to the magnitude of depressive symptoms (Rozanski et al., 1999).

In terms of overall recovery from MI, depression has been identified in a number of studies as being independently associated with poor clinical outcomes, including mortality (Carney et al., 1988; Silverstone, 1990; Ladwig et al., 1992; Frasure-Smith, Lespérance & Talajic, 1993). For instance, depression has been associated with a 5-fold increase in risk of mortality at 6 months and an 8-fold increase at 18 months following MI in a large Canadian study (Frasure-Smith et al., 1993; Frasure-Smith, Lespérance & Talajic, 1995a). A more recent study by this group demonstrates that anxiety in hospital and history of major depression make independent contributions, alongside symptoms of depression in hospital, to increased risk of recurrent cardiac events in the first year following MI (Frasure-Smith, Lespérance & Talajic, 1995b).

1.4.2 Anxiety

High levels of anxiety are evident in 13-20% of MI patients in the early post-MI period (Havik & Maeland, 1990; Houston Miller & Barr Taylor, 1995). Research by Kawachi and colleagues provided evidence linking anxiety to the development of cardiac events in general populations (Kawachi, Sparrow, Vokonas, & Weiss, 1994). It should be noted that anxiety is associated with sudden cardiac death. Research on the relationship

between anxiety in cardiac patients is not as common as that for depression. The available evidence suggests that high levels of anxiety may reduce over time after the cardiac event (Wrzesniewski, 1977; Thompson, Webster, Cordle & Sutton, 1987). However, a number of studies have demonstrated a relationship between anxiety and a variety of cardiac outcomes (e.g., Denollet & Brutsaert, 1998), and there is a need of further research on the pathophysiological effects of anxiety in cardiac populations (Rozanski et al., 1999).

1.4.3 Negative psychosocial states

While depression and anxiety are the most widely documented psychosocial phenomena in cardiac populations, there is some evidence to suggest a role for other negative affective states. For example, the Symptom CheckList-90R was used to identify patients in psychological distress in a cardiac rehabilitation programme (Allison et al., 1995). At 6 month follow-up, distressed patients were 15 times more likely to be re-hospitalised and 12 times more likely to suffer from another cardiac event. In addition, the distressed group had additional re-hospitalisation costs that totalled more than $300,000. Recently, Denollet has delineated a distressed personality type (type-D) characterised by a tendency to experience negative emotions and also to inhibit self-expression. Type-D cardiac patients have a high mortality risk (Denollet et al., 1996) and an increased risk of cardiac events (Denollet & Brutsaert, 1998) compared to non-type-D patients.

1.4.4 Psychosocial Interventions

Psychosocial interventions may affect rehabilitation outcomes in two ways. Firstly they may facilitate psychosocial recovery, including return to everyday activities. Secondly, they may play an important role in secondary prevention by improving adherence with medial advice concerning medication and life-style change. Thus, psychologists can contribute to cardiac rehabilitation service delivery. For example, psychologists can be involved in training relevant health care personnel to help improve their communication and psychological intervention skills. Alternatively, psychologists can become more involved in the cardiac rehabilitation team and psychological services may be offered on either an individual or group basis.

There is widespread variation across Europe in the involvement of psychologists in cardiac rehabilitation (Maes, 1992). In Ireland, only 5 (12.5%) centres have psychologists involved in the cardiac rehabilitation programme (McGee et al., manuscript in preparation). In a recent UK cardiac rehabilitation survey, one in five centres reported that a psychologist was involved in the cardiac rehabilitation programme (Lewin et al., 1998). In general, psychologists do not seem to play a substantial role in the provision of psychological services to cardiac patients (Maes, 1992).

Comprehensive multifactorial cardiac rehabilitation has been associated with a decrease in depression in a sample of cardiac patients (Milani, Lavie & Cassidy, 1996). In this study, depressed patients showed marked improvements in depression scores and other behavioural parameter assessed and increased quality of life. In addition, depressed patients also improved exercise capacity, plasma lipids, percent body fat, triglycerides, and high-density lipoprotein cholesterol. Indeed, depressed patients showed similar improvements in most cardiac risk factors, with the exception of certain behavioural and quality of life parameters, where they showed statistically greater improvements than non-depressed patients. The mechanisms by which cardiac rehabilitation favourably impacted depression remains to be resolved, though some possible candidate pathways include the emotional impact benefits of exercise (Byrne & Byrne, 1993), and the impact of educational sessions (Mullen, Mains & Velez, 1992). However, the absence of a control group limits the strength of conclusions drawn from the study.

The effects of exercise-only cardiac rehabilitation on anxiety and depression in cardiac patients were evaluated in a meta-analysis (Kugler, Seelbach & Krüskemper, 1994). Positive effects for the exercise programmes were reported for both anxiety (d mean = 0.31) and depression (d mean = 0.46). However, these effects are classified as being less than medium (defined as $d = 0.50$) according to Cohen's criteria (Cohen, 1977). It must be noted that there are some limitations with this review. For example, the interventions included varied in terms of duration and intensity, and also the quality of these studies varied from randomised controlled trials to pre-post test studies with no control group. These issues raise questions about the validity of the study findings. Furthermore, the mechanism underpinning the observed effects remains to be determined. The authors concluded that

exercise programmes alone should not be considered the only treatment for emotional distress in cardiac patients but that they can be psychologically beneficial as an additional treatment.

It is worth noting that two recent papers have raised concerns about the efficacy of psychosocial interventions for cardiac patients. Supportive and educational home-based psychosocial nursing interventions in comparison to usual care had no significant benefit on survival 1-year post MI in a large sample of Canadian patients (Frasure-Smith et al., 1997). Indeed, the authors noted a poorer overall outcome for female participants in the intervention group compared with in the control group. In the UK, a large multicentre trial of psychological treatment starting two to six weeks post MI failed to show any benefit over usual care (Jones & West, 1996). Patients with confirmed MI were randomly assigned to an intervention or usual care group. The intervention comprised psychological therapy, counselling, relaxation training, and stress management training for seven weeks for both patients and spouses. At 1 year follow up, there were no significant differences in clinical complications, clinical outcomes, or mortality.

A number of concerns have been raised concerning the applicability of these findings to standard cardiac rehabilitation programmes. It is widely accepted that exercise is the cornerstone of cardiac rehabilitation and neither study included a regular exercise component. Furthermore, in contrast to most cardiac rehabilitation services, Frasure-Smith et al. (1997) used individualised interventions and consequently neglected the potentially supportive group nature of cardiac rehabilitation. It should also be noted that this intervention was not performed by a psychologist or by specialists trained in psychosocial intervention. A previous review of the cardiac rehabilitation literature concluded that better designed studies which used more highly skilled or trained service providers demonstrated greater benefit (Wenger et al., 1995). However, taken together these studies do suggest that programmes based only on psychological intervention are not successful and that multifactorial rehabilitation is required. This statement further supports the WHO model of multifactorial cardiac rehabilitation. Further research is clearly required on the efficacy of psychosocial interventions in changing psychosocial factors associated with negative outcomesin cardiac

patients (Williams & Chesney, 1993). One such study currently underway is the Enhancing Recovery in Coronary Heart Disease (ENRICHD) Patients study.

ENRICHD was initiated by and funded by the US National, Heart, Lung, and Blood Institute (NHLBI) and it is the largest-ever controlled clinical trial of psychosocial intervention for depression and social isolation. The study will help determine whether cognitive behavioural therapy reduces the risk of death or subsequent heart attack in a group of 3,000 MI patients who suffer from depression or low social support. The interdisciplinary study involves a number of centres across the US recruiting patients and requires close co-ordination between cardiology and psychology. The findings of this study may have important implications for psychological aspects of cardiac rehabilitation.

In general, psychosocial interventions comprise patient education, counselling, and behaviour interventions or a combination of approaches. A recent meta-anlaysis reported that psychosocial interventions may significantly reduce morbidity and mortality in patients with coronary artery disease (Linden, Stossel & Maurice, 1996). However, a number of issues have been raised concerning the validity of this meta-analysis in relation to cardiac rehabilitation as the studies included in the meta-analysis may not be representative of typical cardiac rehabilitation interventions. In addition, a number of methodological issues have also been raised in relation to the literature search, with little detail regarding validity or quality assessment of studies included, limited study details are provided and duplicate publications are included in the analysis (Effective Health Care, 1998).

In response to some of these shortcomings, a more recent meta-analysis has focused on the effects of psychoeducational programmes (health education and stress management) on variety of outcomes for cardiac patients (Dusseldorp, van Elderen, Maes, Meulman & Kraaij, 1999). Psychoeducational programmes were associated with a 34% decrease in cardiac mortality, a 29% reduction in recurrent MI and significant positive effects on blood pressure, cholesterol, body weight, smoking behaviour, physical exercise and dietary habits. Thus, psychoeducational programmes are associated with a number of beneficial effects for cardiac patients.

1.4.5 Summary

Depression, anxiety and negative psychosocial states have been most frequently associated with poor outcome in cardiac patients. Psychologists can contribute to cardiac rehabilitation by facilitating psychosocial recovery in patients and also through increasing adherence to secondary prevention. Psychoeducational components of cardiac rehabilitation have beneficial effects on mortality, morbidity and in promoting secondary prevention health behaviours.

1.5.0 Summary of Key Points

- Cardiac rehabilitation is considered an important component of the modern comprehensive care plan for many patients with heart disease.
- A combination of exercise, psychological and educational intervention is regarded the most effective form of cardiac rehabilitation
- According to the WHO (1993) the goals of cardiac rehabilitation are to promote secondary prevention and to improve quality of life.

Chapter 2
Quality of Life and Health Care

2.0.0 Introduction

Quality of life research in relation to health care has dramatically increased in recent years. It is worth noting that it was only in the mid-1970s that the subject heading "Quality of Life" entered MEDLINE. Since then, the number of publications using quality of life has risen each year, with approximately 4,500 articles listed as having quality of life as a keyword in 1998.

A number of recent national Department of Health (e.g., 1994, 1995, 1997) policy documents emphasise the need for health services to focus on measurable improvements in health status and quality of life. Furthermore, the recent report of the national Cardiovascular Health Strategy Group (1999) to the Department of Health and Children stated explicitly that

> "A strategic approach to the prevention of cardiovascular disease and to treatment and rehabilitation of patients with CHD requires a full spectrum of interventions. We need to.....ensure the best survival and quality of life outcome for those who recover from an acute event." (p. 2)

From a cardiovascular perspective, it is also worth noting that the mission statement of the European Society of Cardiology is to improve the quality of life of cardiac patients. Thus it is clear that quality of life is now regarded as a key concern in health care at both a basic research level and also in determining health policy strategies. Given such pervasive interest in quality of life, it is reasonable to ask why quality of life has become so important.

2.1.0 Background

The background to the rise of interest in quality of life will first be considered. Increasing costs of health care provision and increased demand on these services, the rise of chronically ill populations, and the consumer-focused approach to health care service have all contributed to the rise in outcomes assessment in health care. Each of these factors will be considered below.

2.1.1 Increasing demands on health services

In recent decades, critical evaluation of the evidence for the effectiveness of medical interventions has become essential. Health care resources are finite and thus, health care supply can never meet the ever-increasing demands placed on health care services. Increases in life expectancy and reductions in birth rate have combined to increase the average age in Western countries thus placing more pressure on health service resources as chronic conditions require more treatment and care (Normand, 1991).

For example, the cost of congestive heart failure, a chronic illness, exceeds $8 billion a year in the USA (Garg, Packer, Pitt & Yusuf, 1993). In the UK, CHF accounts for 1% of the total NHS costs and approximately 70% of these costs are attributable to hospitalisation admission and outpatient follow up (McMurray & Hart, 1993). The rate of re-hospitalisation for CHF patients is relatively high, with research reporting that up to between one-third and 46% of all patients are re-admitted within 12 months of discharge (Burns et al., 1997; Gooding & Jette, 1985). The incidence of new onset of heart failure increases with advancing age and in addition, the increase in survival following MI is

likely to further increase the prevalence of CHF. Thus, the number of patients with compromised cardiac function is expected to increase dramatically over the next few decades and CHF will place an enormous burden on health care costs.

The example of CHF is intended to highlight that with the increasing ability to extend life and the concurrent ageing of the population, health care services will increasingly come under severe pressure to provide an appropriate range of services. In such an environment priority issues come more to the fore. Indeed, as Maynard (1996) stated

> "clinicians and managers in the NHS are seeking to allocate the NHS budget using the benefit principle: patients will be prioritised in relation to their capacity to benefit from care (improve their length and quality of life) per unit cost." (p. 1499)

As this quote makes explicit, quality of life is considered an important component in the allocation of scarce health resources in the UK. This issue will be further considered in section 2.5.6.

2.1.2 Rise in chronically ill populations

This century has witnessed a rise in chronic diseases and a reduction of mortality from infectious diseases, which has resulted in changes in both clinical and public health practices. Previous to these changes, mortality rates provided an adequate description of the burden of disease in a population and these indicators were also acceptable measures of the impact of health policy and therapy at a population and an individual level respectively (Ebrahim, 1995). However, the increase in chronically ill populations in Western

industrialised societies has provoked a reappraisal of outcomes assessment. Within this population, where cure is not the primary treatment option or aim, focus has shifted to consideration of the patient's subjective rating of quality of life. Traditional outcome markers of mortality and morbidity are inappropriate for these chronically ill populations.

In response to the need to find ways of defining and measuring the goals of treatment for chronic illnesses, quality of life concepts have received widespread use in recent years. As the goals of health care change it seems reasonable that the evaluative criteria to assess the impact of health care interventions will also change. Chronic disorders require long-term medical management and this form of management will have long-term implications for quality of life. A number of treatment options are currently available for a variety of illnesses and in such a situation, comparison of potential positive and negative effects of the treatment on quality of life may influence the treatment decision making process.

2.1.3 Consumerism

It has also been suggested that quality of life assessment is essential in the consumer-driven approach to health care that increasingly characterises health care in the USA and other Western industrialised countries (Ware, 1995). Consumers need to be fully informed of the costs and benefits of available treatment choices before making an informed decision in relation to service use. Thus, there is need for patient-based assessments of outcomes. Indeed, as health care moves towards a consumer-provider model there is growing interest in the value of consumer perceptions of service impact. A

number of studies have reported low levels of agreement among doctors when rating the patient's quality of life, and also low levels of agreement between doctors and patients when rating patient's quality of life (Sprangers & Aaronson, 1992). Thus obtaining the views of the patient helps provide a patient-centred baseline which can be used to assess treatment. The shift to a patient-centred outcomes assessment is a reflection of a growing recognition of the role of the patient's quality of life concerns in health care. Clearly, patients and their carers need to be aware of potential post-intervention changes in quality of life related to the treatment.

2.2.0 Outcomes assessment

Just over a decade ago, Relman (1988) stated that healthcare was entering an "era of assessment and accountability" and the rapidly rising costs in health care had focused attention on cost-benefit analyses of health care (Ellwood, 1988). Data testifying that health care services provide value for money is increasingly demanded by health care providers. There is a growing awareness that medical treatments should be provided only when there is information available that they are effective (Jenkinson & McGee, 1998). If the effectiveness of an intervention is to be measured, then the intervention's outcomes must be systematically assessed. The outcomes assessment movement in health care has provided much impetus to quality of life assessment in health care.

Relman (1988) termed the outcomes movement as the "third revolution in medical care" and Ellwood (1988) has called for a major initiative in "outcomes management" in which clinical standards and guidelines are systematically based on patient outcomes.

Indeed, the Agency for Health Care Policy and Research (AHCPR) was established in the USA as part of a programme directed at assessing the effectiveness of medical interventions and developing guidelines for medical practice through the assessment of patient outcomes. It should be noted that some have questioned whether outcomes data is sufficient for the construction of practice guidelines and for making health care more efficient (Epstein, 1990). Nevertheless, it is clear that outcomes assessment is a major theme in healthcare at present. One of the most important effects of the outcomes movement has been a broadening of focus to include a wider range of outcomes. There has been a dramatic increase in the range of outcomes that health care providers and practitioners are willing to consider valid indicators of health. This raises questions concerning which are the most appropriate outcomes to assess?

2.2.1 Traditional outcome measures

Traditionally, health researchers have been concerned with the more "objective" outcomes of mortality, morbidity and the financial costs of the health service. Mortality is the most commonly assessed health outcome variable in clinical studies and there is widespread agreement that effects on mortality are critical in assessing the effectiveness or appropriateness of health care services. However, the emphasis on mortality as an outcome measure from clinical trials would suggest that once participants emerged alive from the study, then it was a success. Similarly, the emphasis on morbidity reflected a concern with purely physiological outcomes. The costs of health services are also an important consideration as scarce resources dictate that not all services can be funded.

Therefore, it is important to assess the financial costs of providing a certain service in order to make informed budget decisions. This has been termed the *"death, disease, and dollars"* approach to outcomes assessment (Ware, 1995). As Jenkinson (1994) observed, however, there is now a general consensus that traditional measures do not provide a comprehensive picture of the impact of health care interventions. What is lacking is a consideration of the patient's perspective on the impact of medical care.

Measures of quality of life are important as these outcomes are significant both to the patient (Jette, 1993) and the health service provider. Thus, there is growing recognition that the patient's perspective is of critical importance if we are to comprehensively assess the outcomes of health care use. Furthermore, as Ware (1995) noted, there are other issues related to quality of life in addition to the absence of disease, how long the disease lasts and how much the disease costs. This basic assumption has underpinned the development and application of questionnaires for assessing subjective health status in a meaningful and reliable way.

2.3.0 Defining Health-Related Quality of Life (HRQOL)

While most people intuitively seem to understand roughly what the phrase "quality of life" means, providing an exact definition has thus far eluded social scientists, health services researchers and others involved in such research. There has been much controversy concerning the conceptualisation, definition and measurement of quality of life. Quality of life is an all encompassing term that can incorporate numerous aspects of an individual's life (e.g., education, housing, health, income, perceptions of immediate

environment) and clearly, it is unreasonable to expect health care services to favourably influence all aspects of a person's quality of life. Therefore, Kaplan and Bush (1982) introduced the term health-related quality of life (HRQOL) to distinguish the broad concept of quality of life from those aspects of life quality specifically relevant to health status.

HRQOL is a concept that attempts to encompass the WHO (1947) definition of health as "a state of complete physical, mental, and social well-being and not merely the absence of disease" by incorporating both personal health status and social well-being in assessing the health of individuals and populations (Guyatt, Feeny & Patrick, 1993). According to a widely cited definition, HRQOL refers to

> "the value assigned to the duration of life as modified by the social opportunities, perceptions, functional states and impairments that are influenced by disease, injuries, treatments or policy." (Patrick & Erickson, 1993)

At a theoretical level, the concept of HRQOL has not been well developed as it evolved from a loosely integrated body of research on health status, functional performance and social well being. This body of work comprised diverse studies of diagnostic category, health status, and functional ability; psychological factors like pain, emotions, locus of control, sense of self-worth and intellectual functioning; social variables such as social role performance and support networks; and composite measures of these elements in assessing HRQOL (Bowling, 1997). Historically, such research was oriented towards solving practical problems and was not formally derived from an

elaborated theory. Thus, a pragmatic approach to HRQOL assessment has dominated and theoretical work has lagged behind instrument construction and validation.

It may be useful to consider quality of life assessment in health care representing a continuum ranging from clinical assessments to physical functioning to self-rated HRQOL to individual quality of life assessments. At the one extreme, clinical assessments rely on physiological data, particularly mortality and morbidity data. Assessments of physical functioning may rely on a combination of physiological and self-rated functioning (e.g., self-reported activities of daily living). HRQOL assessments allow the self-rating of the quality of life aspects related to health status. At the other extreme, individual quality of life assessment measures individual values and perceptions of quality of life not necessarily related to health status. For the purposes of the present chapter, the phrase HRQOL will be used to encompass all the related constructs that focus on self-rated psychosocial and health status.

As Shumaker & Naughton (1995) observed, one source of the confusion surrounding quality of life (and HRQOL) arises from the limited number of definitions and the lack of empirical research testing such definitions. Indeed, quality of life is rarely defined by researchers and quite often only an operational definition is given.

One of the few areas of agreement with quality of life research is that it should be regarded as a multidimensional construct. In general, the definitions available within the literature tend to agree on the components that ought to be included: general health, cognitive function, mental health, emotional state, subjective well-being, life satisfaction and social support (Jenkinson & McGee, 1998). For example, investigation of important

consequences of cardiac surgery revealed that five major dimensions seemed to be important to patients: morale, symptoms of illness, neuropsychological function, interpersonal relationships and economic-employment (Jenkins, Jono, Stanton, & Stroup-Benahm, 1990). While these factors may change across illness populations, it is clear that multidimensional approach permits analysis of the strengths and weaknesses of particular interventions. A profile can be elicited for each individual that highlights specific therapeutic needs.

2.3.1 History of HRQOL assessment

The evolution of HRQOL assessment has taken place in the historical context of research on the broader concept of quality of life health care. While outcomes assessment came to the fore of US healthcare policy in the 1980s, it should be noted that health status assessment has been a topic of concern for much longer.

White (1967) argued that resources should be directed towards systematically and objectively studying the operations of health care services. He stated that

> "Where resources are scarce and the needs are great, it is essential to introduce improvements in the information available on which to base critical decisions."
>
> (p. 848)

White describes five basic levels that have to be considered in an evaluation of the health services system – death, disease, disability, discomfort, and dissatisfaction (the 5 "D"s). It is interesting to note the similarities of the five "D"s to current perspectives on outcome assessment. For example, White defined disability in terms of the ability of a health

service to at least to keep people out of hospitals and to maintain functional capacity, productivity and well-being. This definition has many parallels in contemporary definitions of HRQOL. White also recognised that while other sources of such data (e.g., doctors) had received much attention, there was a need for obtaining the consumers' concerns and interests in relation to health care services provided. In the next decade researchers from a variety of backgrounds commenced work on consumer-driven outcomes assessment.

Health care evaluation research initially evaluated quality of care from a process perspective. However, this approach was challenged by those emphasising the outcome of care and this paved the way for the development of health status indicators in the 1960s and 1970s and the explosive growth in measures of quality of life in the 1980s and 1990s (Rosser, 1997).

The 1970s saw the development of numerous scales which could provide a generic profile of health status. These measures were to be utilised with the general public as well as those with illness or disabilities. Three of the most widely used measures were developed in this period with this specific purpose in mind: Sickness Impact Profile (SIP) (Gibson & Gibson, 1975), Nottingham Health Profile (NHP) (Hunt, McEwen & McKenna, 1980) and the McMaster Health Index Questionnaire (MHIQ) (Chambers, MacDonald, Tugwell, Buchanan & Kraag, 1982). It is noteworthy that all of these measures are behaviourally based and indeed, some have questioned the validity of over-extending their use and calling them quality of life measures (e.g., Hunt, 1999).

The pioneering work of the RAND Corporation in the 1970s and early 1980s resulted in large-scale research on health status scale conceptualisation, development and testing. Indeed, the RAND Medical Outcomes Study in the US was seminal in illustrating the potential benefits of HRQOL assessment (Tarlov et al., 1989). This was the first major project to use subjective assessments to evaluate variations in health care provision. The processes and outcomes of care for patients with chronic disease were assessed according to variations in the style and practice in different systems of health care.

In general, HRQOL assessment has evolved to meet the demands for service evaluation and consequently a rather pragmatic philosophy has tended to dominate to date. Thus, practical utility has dominated over theoretical rigour. Indeed, there has been much criticism that the field tends to lacks any theoretical foundation. Consequently, while numerous domains and dimensions of subjective health assessment have been described and measured, this work has not been situated within an overall coherent theoretical framework. However, such approaches are beginning to emerge in the literature (e.g., Hyland, 1992; Patrick & Erickson, 1993; Wilson & Cleary, 1995) and one of the challenges facing subjective health assessment in the coming decades will be to rigorously test conceptual approaches to HRQOL.

2.3.2 Summary

Recent years have witnessed the rise of outcomes assessment in health care. This rise is attributable to (a) the increasing costs of health care provision and increased demand on these services, (b) the rise of chronically ill populations, and (c) the growth of

consumer-focused approaches to health care. There is a growing awareness that the effectiveness of health care services must be measured and that a variety of outcomes, including HRQOL must be systematically assessed. HRQOL incorporates both personal health status and social well-being in the assessment of service outcomes. HRQOL instrument development has evolved to meet service evaluation needs and the measurement of HRQOL will now be considered.

2.4.0 Measuring HRQOL

A number of different approaches to assessing HRQOL have been developed. While early approaches to health status assessment involved the use of proxy measures and performance ratings, they do not provide a subjective rating of HRQOL by the individual patient. One of the basic premises of HRQOL assessment is that the individual must provide the rating. These ratings typically involve completion of a self-report questionnaire and such questionnaires can be classified as generic or specific instruments. However, as Jenkinson, Bardsley & Lawrence (1994) noted,

"Not all questionnaires are appropriate in all settings for all purposes or for all respondents" (p. 178)

2.4.1 Generic HRQOL instruments

Generic HRQOL instruments assess a range of important aspects of quality of life. Such instruments may include scales assessing a number of physical, social and emotional dimensions. Commonly used examples include the SIP (Gibson & Gibson, 1975), NHP

(Hunt et al, 1980) and Medical Outcomes Study Short Form 36 (MOS SF-36) (Ware & Sherbourne, 1992). The advantages to using such an instrument include the fact that they deal with a number of dimensions and can be used in virtually any population irrespective of underlying condition. Therefore, generic profiles can applied to a wide variety of populations and facilitate broad comparisons of the relative impact of various health care programmes. However, it has been suggested that generic profiles may be less responsive to changes in specific conditions or to response to treatment.

Health utility instruments are derived from economic and decision theory and these instruments reflect the preferences of patients for treatment process and outcome. While utility measures are useful for determining whether patients are better off as the result of therapy, they may fail to reveal the dimension of quality of life on which patients improved versus those on which they deteriorated. Thus such instruments are not appropriate for routine assessment of psychosocial outcome.

2.4.2 Disease-specific instruments

Disease-specific instruments focus on aspects of health status that are specific to particular illness populations. These instruments may show increased responsiveness to changes in patients as these instruments include only important aspects of quality of life that are relevant to the patient population being studied. This contrasts with generic approaches where many of the items may not be appropriate or relevant to the population under investigation. In addition to the likelihood of improved responsiveness, specific

measures have the advantage of relating closely to areas routinely explored by clinicians. However, the fundamental disadvantage to using specific instruments is that it precludes comparison across different illness populations. It has also been suggested that generic instruments allow theoretical frameworks to be developed and subsequently tested across different clinical conditions (Johnston, 1999). Furthermore, the only limit to the number of specific instruments that can be developed is the number of diseases that can be categorised and, a number of different instruments can be developed for each disease.

2.4.3 Predictive ability of subjective health ratings

There is growing evidence that subjective health assessments are independent predictors of clinically important outcomes. The Manitoba Longitudinal Study on Aging revealed that elderly respondent's self-assessments of health were better predictors of survival during the follow-up period than standard objective health data (Mossey & Shapiro, 1982). This finding indicated that self-report survey responses, often considered to be too vague or prone to measurement error/response bias, did possess predictive validity. This association held after controlling for other variables such as objective health status. As Idler (1992) noted, despite substantial differences between studies in terms of their location, population studied, follow-up periods, differences in mortality rates, measurement of subjective and objective health variables, and inclusion of covariates, the research findings are remarkably consistent. Self-assessed health (poor vs. good vs. excellent) predicted mortality throughout follow periods of 7 to 14 years (Idler, 1992).

More recent evidence supports the predictive value of routine health status assessment. In a large scale prospective cohort study, the Physical Component Scale (PCS) of the SF-36 completed before surgery was an independent risk factor for mortality following CABG (Rumsfeld et al., 1999). The physical component summary score comprises the physical functioning, physical role functioning, bodily pain, and general health subscales of the SF-36. The SF-36 was completed within 72 hours prior to surgery by 2480 patients. Adjusting for traditional clinical risk factors, preoperative PCS score was an important independent risk factor for 6-month mortality. An important aspect to this study was the quantification of the magnitude of risk associated with the quality of life variables and thereby placing them in the context of traditional clinical risk factors. In multivariate analysis, a 10-point decrease in a baseline PCS score was associated with a 39% increase in the risk of 6-month mortality.

The important role that patient self-report questionnaire may play in risk stratification before CABG is clearly illustrated in this study. A baseline PCS score helps explain the variance in the outcome above and beyond standard clinical variables. Thus, self-report instruments shows that patients can, in effect, assist in the classification of their functional status.

2.4.4 Criticisms of HRQOL research

A number of criticisms have been raised concerning the use of HRQOL assessment. For example, Jung and Conn (1995) suggested that while the overall medical management of a patient may benefit from the use of a well-defined subjective HRQOL measure, these measures are not scientifically rigorous enough for medicine. Their limitations arise because they are based on subjective situational evaluations and are too variable to be used in comparative efficacy studies. This is contrasted with biological measures which are more precise, consistent and sensitive to interventions than psychological measures, thus making them more appropriate for determining treatment efficacy, especially if predefined beneficial physiological changes are of interest.

In reply, Guyatt & Cook (1995) challenged the claims made by Jung and Conn (1995) in so far as there is evidence that the measurement of HRQOL is more precise, consistent and responsive than some biological measures (e.g., Guyatt, Feeny & Patrick, 1993). In addition, focusing on biological measures may miss the point as patients are often interested in living longer or feeling better, not in changes in the status of their biological markers. However, subjective health assessment can supplement traditional measures of efficacy and should not supplant biological outcomes if these are the outcomes of interest.

Nevertheless, it must be acknowledged that patients sometimes get confused about the meanings of words, may answer in ways they believe will please the interviewer, or may not wish to share their true feelings or symptoms. In addition, patient evaluations of their health may depend as much or more on psychological or emotional factors as on

objective indicators of health. These limitations notwithstanding, patient self-reports provide the most appropriate assessment of HRQOL.

The use of HRQOL questionnaires can provide information on health status as reliable as more technologically sophisticated intermediate measures (McDowell & Newell, 1987). A standard self-administered questionnaire can focus on important domains, can reduce variability in the quality of history taking and from a statistical perspective, the reliability of patient driven data is higher if patients are asked several related questions about a particular dimension of health outcome. Use of questionnaires might, therefore increase the test-retest and inter-rater reliability of patient histories. The test retest reliabilities of patient outcome questionnaires are typically quite high and often exceed reliabilities reported for more objective measures of health status.

2.4.5 Summary

Two main approaches to HRQOL assessment can be distinguished: generic or disease-specific. Each approach has a number of advantages and disadvantages and there is inevitably conflict concerning the desirability of using a measure that allows comparison with other illness and patient groups and using specific measures that are more likely to capture the particular implications of an illness or treatment for some patient group. At present, a generic and a disease specific instrument are commonly used together to assess psychosocial outcomes. However, a number of concerns regarding HRQOL assessment have been raised. Whilst acknowledging these limitations, patient self-reports provide the most appropriate assessment of HRQOL. Self-reported ratings of HRQOL may have

clinical predictive value and thus routine HRQOL assessment may provide valuable information.

2.5.0 Applications of HRQOL instruments

Numerous applications have been suggested for subjective health status measures. These include clinical research, health services research, individual patient care, audit, population studies, and resource allocation. It should be noted that the different contexts in which they are used require different measurement properties and practical features for an instrument to work effectively (Fitzpatrick, 1995).

2.5.1 Research purposes

At present HRQOL instruments are most commonly used for research purposes. While mortality may be the primary outcome of interest for many RCTs, interventions may also impact on HRQOL. Thus, a multidimensional approach to outcomes assessment could incorporate mortality, morbidity, costs and HRQOL. It should be noted that health care intervention may provide a mixture of beneficial and harmful effects. For example, while hypertensive medication may help in controlling blood pressure, there may be negative effects in relation to the patient's mood (Croog et al., 1986). Clearly, such information is essential as it allows the potential benefits and harm of an intervention to be evaluated.

As Jenkinson and McGee (1998) observed, incorporation of HRQOL in such studies is long overdue given that many currently used therapies have no effect on life expectancy but are designed to improve patient health status (e.g., hip replacement). While the use of subjective health status measures in clinical trials is increasing, concern has been expressed over the failure to use standardised and conceptually appropriate measures. This is a fundamental issue and the use of a measure should be psychometrically and theoretically justifiable. The issues of appropriate psychometric properties will be addressed in the next chapter.

2.5.2 Health Services Research and HRQOL

Recent years have witnessed the rise to prominence of Health Services Research (HSR) within the health care system. HSR aims to provide knowledge about the structure, process and or outcome of the interaction between health care service providers and service users. HSR incorporates outcome assessment to monitor the use of resources and policy relevance of structure and process studies (Andersen, Davidson & Ganz, 1994). Research including HSR and HRQOL perspectives provides a more comprehensive picture of the impact of health care services on patients. The incorporation of HRQOL measures into HSR and other health related research might facilitate research findings contributing more to health care policy considerations.

Frankford (1994) raised a number of concerns about the use of outcomes assessment in health care. He stated that although many proponents of this research believe their work need not be linked to the cost containment efforts of health care

providers and governments, this connection seem inevitable because much of this research is funded by agencies who seek financial control. Ware (1995) made a similar point in stating that the principal use of outcomes in this decade will be limited to proving that cost-containment is successful without harming health.

2.5.3 Individual patient care

Three related but distinct clinical applications of HRQOL in the health care of individual patients have been identified. First HRQOL assessment may provide the health professional with useful information regarding the impact of the disease and the treatment on the patient. For example a questionnaire that assess a broad range of health domains may provide a health care professional with valuable information about any unexpected side effects of treatment. Routine HRQOL assessment may help monitor the course of the patient's progress over time and such information may aid in clinical decision making regarding treatment.

The CASPE (Clinical Accountability, Service, Planing, and Evaluation) study examined the feasibility of developing and including routine outcome measures in clinical audit and decision making within the hospital (Bardsley & Coles, 1992). Despite the limited time frame of the study (two years to establish data collection and report back on data to 12 month follow up), a number of changes in clinical practice were found. Favourable changes in clinician's attitudes towards outcome assessment resulted in the acceptance of the importance of patients' perceptions in assessing care. Bardsley and Coles (1992) concluded that such outcomes measures can help not only in the care of

individual patients but may also be of benefit in determining which cases to treat, the choice of treatment for particular patients, the way care is organised for groups of patients, the way a service as a whole is organised, and the choices between services.

Secondly, the process of collecting such data may improve professional-patient communication. Such additional information may be helpful in providing a more complete assessment of the patient and consequently improving communication between doctor and patient. A third use is in screening patients to evaluate their needs for particular services. This type of approach has been used most frequently in screening for psychological distress. In the light of current concerns about health care cost containment, a striking example of the utility of such an approach was illustrated in a study examining the effects of psychological distress in CHD patients in relation to potentially increased morbidity and associated health-care costs (Allison,Williams, Miller, Patten, Bailey, Squires & Gau, 1995).

During the second week in a cardiac rehabilitation programme, 381 patients completed the Symptom Checklist-90–Revised (SCL-90-R) (Derogatis & Lopez, 1983). This 90-item questionnaire consists of multiple choice questions with a 5-point (0 through 4) scale of distress, ranging from "not-at-all" to "extremely" distressed. The primary purpose of the SCL-90-R is to measure general psychological distress but specific symptoms are also classified. Forty-one patients were identified as psychologically distressed according to the SCL-90-R with depression being the most frequently increased individual scale.

Distressed patients were significantly more likely to experience at least one cardiovascular re-hospitalisation in comparison with non-distressed patients. Multiple early re-hospitalisations were also significantly more frequent among the distressed than among the non-distressed patients. Of particular relevance, re-hospitalisation costs were greater (by a mean of $7,358) for the 41 distressed then for the 339 non-distressed patients. The average re-hospitalisation cost for distressed patients was $9,504 compared with $2,146. This study shows that psychological distress, as measured by self-report questionnaire, was associated with increasing morbidity and re-hospitalisation costs. Thus the value of psychological screening can be ascertained from his study.

Ketterer (1993) suggested that simple screening of CHD patients for psychological distress may provide a cost-effective means of identifying those needing further intervention. With computerised scoring and sufficient patient volume, the cost of routine administration of the test should be comparable to a lipid profile. It must be noted that if a subjective health status measure is to be used as a screening instrument, then it must meet the traditional criteria for any screening system. The measure should be accurate, should result in more effective treatment than would have otherwise occurred and must do so with a favourable ratio of costs to benefits. As Fitzpatrick (1995) observed, these parameters are hard to establish in HRQOL assessment compared with conventional screening tests as the entity being screened for (e.g., psychological distress) is often less clearly defined and less precisely measured. In general subjective health status measurement has not been widely used for screening and, thus, there is a lack of empirical evidence concerning the appropriateness of many current subjective health measures to be used in such a way.

2.5.4 Audit

Recent years have witnessed increased recognition of the importance of medical audit and health services must collect data on the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcomes for the patient. Assessment of HRQOL can be applied in programme evaluation and policy analysis to establish priorities, examine the effectiveness of health policy and programmes, and to allocate resources.

Knowledge of outcome and effectiveness of treatment is essential if we are to be able to determine how resources are to be allocated. One method of constructing the appropriate knowledge base to inform decision making is through clinical audit. In measuring outcomes, patient's views are central and the incorporation of this data facilitates consideration of patient generated data into the process of improving care.

2.5.5 Population studies

Meeting the health needs of a particular population is a key issue in health care policy planning. While existing measures of need typically include socio-demographic details and mortality data, it has been suggested that HRQOL measures are a useful source of information about the perceived health of populations (Hunt, McEwen & McKenna, 1985) as these instruments assess the individual's perspective more directly than previously considered. Of note, the routine collection of such data would potentially be relatively inexpensive (probably using a postal questionnaire).

Furthermore, population studies which provide normative data for an instrument serve two important uses. Firstly, such information facilitates monitoring the impact of health care services over time on populations. Thus, a baseline measure could be taken and the impact of a particular service on the HRQOL of a population could be assessed. However, the usefulness of such data in informing health care policy decisions regarding the provision of specific services remains to be proved and it has been questioned if such measures provide sufficiently accurate information for these purposes (Donovan, Frankel & Eyles, 1993). Clearly, empirical evidence is required before the use of such measures is warranted.

Secondly, from a psychometric perspective, population studies can provide extensive normative data regarding the measure. As HRQOL scores cannot be interpreted directly, there is a need to have reference to some standard and the normative data may increase the interpretability of scores on an instrument. However, such norms are interpretable only with respect to some definable reference group and that the reference group depends upon the types of decisions or comparisons under consideration (Nunnally, 1970).

2.5.6 Resource allocation

One of the most controversial uses of quality of life instruments is in the context of the economic appraisal of health care interventions. Increasing demands for health care services coupled with the rising costs of health care place an enormous strain on constrained health services budgets. Given the limited resources available for funding

health care service, not every health service can receive adequate funding. Thus, there is a need for an explicit and rational method of resource allocation in health care.

If the goal is to maximise the amount of health gain that can be obtained for use of health care resources, information about the relevant costs and benefits of interventions in terms of health gain are required. Perhaps the most widely advocated approach to analysing the cost-utility of health care involves the calculation of quality-adjusted life years (QALYs), which integrate both quantity and quality of life to provide an index of the cost-utility of different health care interventions. It should be noted that there are different approaches to calculating QALYs and that each approach is based upon a different set of assumptions.

A patient-based HRQOL assessment provides the opportunity for systematic input from service recipients who often lack access to this decision making process. HRQOL assessments can provide ongoing feedback from these recipients about the outcomes of services and thus influence the further development of service provision. Such information can be used to help decision making in resource allocation in health care based on priorities identified by health care users. However, the use of HRQOL instruments in decision making remains controversial and it is questionable whether the data provided by such measures are appropriate for assessing the health care purchasing requirements of particular populations.

2.5.7 Summary

Numerous applications have been suggested for HRQOL measures. These include clinical research, health services research, individual patient care, audit, population studies, and resource allocation. To date, HRQOL research has mainly occurred in the context of clinical research. However, it has been suggested that HRQOL data can contribute to health policy decision making by providing information about the outcomes of services. It should be noted that the different contexts in which instruments are used requires different measurement properties and practical features for an instrument to work effectively.

2.6.0 Summary of Key Points

- Increasing cost of health care provision and subsequent increased demand on these services, the rise of chronically ill populations, and the consumer-focused approach to health care service have all contributed to the rise in outcomes assessment in health care.
- There is growing recognition that the patient's perspective is of critical importance if we are to comprehensively assess the outcomes of health care use.
- Health-related quality of life (HRQOL) refers to those aspects of life quality specifically relevant to health status.
- HRQOL assessment has developed to meet the needs of service evaluation.
- HRQOL instruments can generally be classified as either generic or disease-specific.
- There are trade-offs in the assessment of HRQOL as no single approach fits the objectives of all studies.
- HRQOL instruments provide complement traditional physiological and biological measures of health status when evaluating the effectiveness of health care.
- A wide range of applications of HRQOL instruments have been proposed and appropriate instrument selection depends on the specific purpose.
- The lack of agreement concerning which standardised measures should be used in which contexts contributes to the slow acceptance of the widespread utilisation of outcomes management and audit.
- Instrument selection involves consideration of the goals of the service.

Chapter 3
Psychosocial Instrument Assessment Criteria

3.0.0 Introduction

The previous chapter highlighted the need for health services to document HRQOL outcomes of health service users. Given the proliferation of HRQOL measures available, how can researchers decide which measure to use? It has been noted that within the outcomes assessment literature that researchers rarely state the criteria they use to select a HRQOL instrument (Gill & Feinstein, 1994). This results in ad hoc instrument selection with little or no theoretical justification and a lack of comparability across studies. Ultimately this practice contributes to a lack of cumulative knowledge in the field of outcomes assessment. Therefore, there is a need for some guidelines for HRQOL instrument selection. If the evaluation of health outcomes is an important concern, then what are the most pertinent criteria?

3.1.0 Instrument criteria

The Scientific Advisory Committee (SAC) of the Medical Outcomes Trust proposed a set of eight attributes that it uses to assess quality of life outcomes instruments (Lohr et al., 1996). They suggest that the integration of strategies for measuring, monitoring, and managing functional outcomes with quality improvement programmes provides the best hope for achieving value from health care funding. The SAC evaluates HRQOL outcome instruments in terms of their intended applications as well as the general significance and contributions of the instruments to the field. The eight criteria required by the SAC are listed in Table 3.0. Of note, Fitzpatrick, Davey, Buxton and Jones (1998) have also recently suggested eight key dimensions in which a HRQOL instrument can be examined (see also Table 3.0).

Table 3.0. Criteria for HRQOL instrument selection

Fitzpatrick et al., 1998	SAC (Lohr et al., 1996)
Appropriateness	Conceptual and measurement model
Reliability	Reliability
Validity	Validity
Responsiveness	Responsiveness
Precision	Alternative forms
Interpretability	Interpretability
Acceptability	Respondent and administration burden
Feasibility	Cultural and language adaptations

There is a high degree of overlap between the two sets of criteria for HRQOL instruments proposed by the two sets of researchers. The following properties are proposed as necessary for selection of a routine evaluative measure of health status: appropriateness, acceptability, feasibility, validity, reliability and responsiveness. In addition, cross cultural considerations will also be discussed. The present thesis will not consider in detail the issue of alternate forms as it is not addressed in the present investigation. Of note, interpretability will be considered in section 3.9.3 as part of the discussion of clinical significance. The key criteria for the present thesis is responsiveness; however, it is useful to briefly consider the other relevant criteria before discussing the role of responsiveness.

3.1.1 Appropriateness

The most fundamental consideration for selecting an instrument is perhaps to identify one that can best answer the research question. As different studies may have different research questions, the content of the instrument must be appropriate to the questions being asked. A critical evaluation of outcome instrument selection in 75 published clinical trials was provided by Guyatt, Veldhuyzen Van Zanten, Feeny and Patrick (1989). They reported that of the 55 studies included in their review that they rated as having important HRQOL implications, 44% of the studies failed to assess HRQOL outcome. Furthermore, of those trials that that did attempt to assess HRQOL, 38% used an untested instrument that was

considered inappropriate. The authors conclude that the use of appropriate HRQOL instruments would have strengthened the recommendations of the relevant clinical trials. Further criticism of the inappropriate use of HRQOL instruments in research articles was provided by Gill and Feinstein (1994). They selected 75 articles that included quality of life assessment randomly from a number of medical journals. They reported that in 85% of articles no definition of quality of life was provided by the authors and furthermore, in over half of the articles the authors provide no justification for the instrument selection. In addition, Gill and Feinstein reported that in only 17% of studies did patients complete a global rating of their quality of life. In a subsequent reanalysis of 15 randomly selected papers from those identified by Gill and Feinstein, Guyatt and Cook (1994) reported that the criteria used by Gill and Feinstein may have been too stringent. While both reviews emphasise the importance of the quality of life instrument asking about issues that are important to patients, there are a number of differences between their conceptualisations of appropriateness (Fitzpatrick et al., 1998).

In general, Gill and Feinstein (1994) argued for a more theoretical and conceptual approach to quality of life assessment. Thus, for example they emphasised the need to base measures on an explicit definition of quality of life, and they suggested that a distinction between global quality of life and HRQOL should be made. Furthermore, they argued that instrument development has been too dominated by psychometric considerations at the expense of providing respondents with sufficient freedom to express their quality of life ratings. The need for quality of life ratings to be based on individual patient values and priorities as opposed to pre-set items was emphasised. In contrast, Guyatt and Cook (1994) adopted a more psychometric approach to quality of life instruments. Thus, they were concerned with the use of previously validated instruments that have proven reliability, validity, and responsiveness. It may be argued that Guyatt and Cooks' approach reflects the pragmatic approach to applied quality of life in health care whereas Gill and Feinstein adopt a more philosophical approach to quality of life issues. The tension between the pragmatic and the philosophical approaches to quality of life assessment has not been fully resolved although, undoubtedly, the pragmatic approach has dominated health care research to–date.

However, more philosophically focused approaches to quality of life assessment have begun to appear in the literature (e.g., Joyce, O'Boyle & McGee, 1999).

It is clear that HRQOL instruments are often uncritically added into health care research protocols with little consideration, if any, of the appropriateness of these instruments in terms of content and relevance to the research aims. The rationale for instrument selection needs to become more explicit if the relevance of findings is to be properly evaluated.

3.1.2 Conceptual and measurement model

The conceptual model refers to the underlying rationale for and description of the concepts that the measure is intended to assess and the relationship between or among those concepts (Lohr et al., 1996). The measurement model is reflected in the instrument scale and subscale structure and the procedures that are used to create the scale scores. Thus, the model specifies the basis for combining items into scales and for providing the descriptive statistics for the scales. The model also provides the rationale supporting the proposed level of measurement (e.g., ordinal, interval or ratio scales) or the scaling assumptions for preference-weighted measures. Furthermore, the model describes and justifies procedures for deriving scale scores from raw scores.

3.2.0 Reliability

Reliability refers to the extent to which the measure gives the same score each time it is administered, assuming that no true change has occurred in the attribute being measured (Hays, Anderson & Revicki, 1995). Or, put in a slightly different way, it refers to the amount of a score that is signal rather than noise. In classical test theory, any observation is held to comprise two components: a true score and an error term. Reliability reflects the extent to which an instrument is free from random error (Nunnally, 1970) and this is important in outcomes research because as the random error increases so too the size of the sample required to obtain precise estimate of the intervention effects in a study will also increase. However, it should always be remembered that reliability of an instrument is not a fixed property but rather it is dependent upon the context and population investigated (Streiner & Norman, 1995).

The reliability requirements of instrument will differ depending on the purposes for which it is used. While a reliability coefficient of 0.70 is acceptable in studies of groups, it has been suggested that a reliability level of 0.90 is required if an instrument is to be used with individuals (Nunnally, 1970). The higher requirement is a result of the wide confidence intervals surrounding an individual's true score if the level is below 0.90. Two different dimensions of reliability can be distinguished: internal reliability and reproducibility.

3.2.1 Internal reliability

Internal reliability or internal consistency relates to the fact that most instruments comprise a number of items to measure a particular dimension or construct. A number of related items will produce a more reliable estimate then one single item. However, for this to be true, the items need to be measuring the same construct rather than several different constructs. Internal consistency involves testing for homogeneity, which assumes that there are correlations between items in a scale which are not the product of random chance but which in fact reflect a real patterning as to how the questions are answered. Individual items should correlate highly with each other and with the total score. Common approaches to assessing internal reliability include split-half reliability, Cronbach's alpha, Kuder-Richardson (KR-20) test, or the correlation of individual items to the overall scale score when that item is omitted.

It should be noted that it has been suggested that excessive attention to internal reliability can result in a overly restricted pool of items which fail to capture the entire range of the construct assessed (Donovan et al., 1993). A high alpha level may simply reflect that effectively the same question is being asked more than once and thus some items may be redundant. In addition, high internal reliability does not mean that a measure is reliable over time.

3.2.2 Reproducibility

Reproducibility refers to whether an instrument produces the same results on repeated applications, assuming that the respondents have not changed on the assessed construct. Consequently, reproducibility is an important consideration in the use of HRQOL instruments in longitudinal research and researchers typically assess test-retest reliability which involves a correlation between the scores at the first assessment with the scores at a later reassessment. There is no agreement as to the length of time that should separate the two assessments. The use of a transition item at the second assessment could identify whether there has been any change on the construct of interest. Such information is necessary because a low correlation between the two assessments could simply reflect the fact that participants have indeed changed on the construct of interest. Thus low reliability may reflect a poor questionnaire, a real change in health status or respondents' familiarity with the questionnaire leading to changes in their response.

Of note, results from two assessments may correlate highly but be systematically different. For example if everybody increases their scores on second administration, then the correlation could be 1.0. In response to this criticism, the use of the intra-class coefficient has been suggested. This approach uses analysis of variance to assess the extent to which the total variability in scores is a result of (a) true differences between individuals and (b) variability in measurement.

3.3.0 Validity

Validity refers to the degree to which the instrument measures what it is supposed to measure and in order to be valid an instrument must be reliable (Nunnally, 1970). The distinction between validity and reliability is important because an instrument may be reliable and give the same score for the same patient at different assessments but it may not be a valid measure as it could be assessing the wrong attribute. Similar to reliability, validity is a matter of degree and should not be regarded as an all-or-none property.

Validation is more accurately regarded as an ongoing process for most instruments and validity is assessed in relation to a specific purpose in a specific population. Therefore, an instrument is typically validated for use in a specific situation and it is meaningless to refer to a valid measure outside of the contexts in which it has been previously been validated. However, concern has been raised that within the field of quality of life research, many instruments have been used for a wide variety of purposes without the evidence for the validity of the instrument for each purpose being established (Revicki & Kaplan, 1995). Three aspects of validity may be distinguished: Criterion and predictive validity, face and content validity, and construct validity (Fitzpatrick et al., 1998).

3.3.1 Criterion and predictive validity

Criterion validity refers to the ability of an instrument to correspond with other measures assumed to be gold standard measures of the attribute being assessed. As Fitzpatrick et al. (1998) observed, few studies can truly claim to have evaluated criterion validity as gold standards are hard to find in HRQOL outcome research. Predictive validity refers to the ability of an instrument to predict the criterion variable. Typically this property is not of concern if we are using an instrument as an outcome variable since we are interested in the outcome in its own right. However, the instrument may also be predictive of other outcomes that are considered important (see section 2.4.3).

3.3.2 Face and content validity

Both face validity and content validity are concerned with the appropriateness of an instrument. It is critical that the content of the instrument in relation to the research question is considered before selecting an outcome measure. Face validity refers to whether the instrument appears to measure what it is intended to measure. Content validity refers to the extent to which the domain of interest is comprehensively sampled by the items or questions in the instrument. Quantitative testing of face and content validity is rarely attempted and typically informal inspection of the instrument is necessary for researchers to make an informed decision regarding the face and content instrument. Assessment of content validity

is not a technical undertaking but rather the application of common sense to the questionnaire under consideration.

3.3.3 Construct validity

A construct is theoretically derived notion of the domain(s) to be measured. HRQOL assessment is concerned with a hypothetical construct rather than a directly observable event. Construct validity involves comparisons between measures and examination of the logical relationships that should exist between a measure and characteristics of patients and patient groups. One single study cannot prove the construct validity of an instrument; rather evidence for construct validity arises from a number of converging studies (Streiner & Norman, 1995).

Testing construct validity often involves examining the convergent (results of the instrument are related to other variables and measures of the construct) and discriminant validity (results are not related to data measuring distinct and unrelated concepts) of the instrument. Basically, the instrument should correlate highly with related constructs and should not significantly correlate with unrelated constructs. A sophisticated approach to assessing convergent and divergent validity involves constructing multitrait-multimethod matrix. However, this approach has received little use in HRQOL research as there may be difficulties obtaining alternative methods of measuring constructs (Fitzpatrick et al., 1998).

3.4.0 Precision

Precision is concerned with the response categories and scoring of an instrument to reflect change. It refers to the ability of the instrument to make distinctions between respondents. The format of the response options will help determine precision. Response options can range from simple dichotomy (yes/no) to seven or more options. While being simple to complete, dichotomous items do not allow respondents to indicate degrees of severity or difficulty that they experience. Therefore, most HRQOL instruments provide respondents with a range of options; typically this involves frequency options or Likert response categories.

The precision of the numbers generated by the instrument is another important concern. Numbers may be derived from weightings of the importance of each item or they may be assigned to Likert scale items on an arbitrary but common sense manner. Clearly there are substantial philosophical issues involved in the assignment of numbers to aspects of subjective experience such as quality of life but these are beyond the scope of the present thesis. Suffice to note that most instruments produce a number that is then statistically analysed. Furthermore, in general this data is treated as if it were at an interval level of measurement

3.4.1 Acceptability

It is essential that instruments are acceptable to patients. While acceptability has not received much formal attention within the research literature, it is concerned with practical issues such as the comprehensibility of the instructions and the items, the length of time taken to complete the instrument, the number of non-respondents and the rates of missing data. High rates of either may indicate that the instrument may be too difficult to understand or that the items are not acceptable to the respondents. However, the use of such proxy data may be misleading as many other reasons can contribute to non-response rates (e.g., respondents not receiving the instrument or respondents not being be able to complete the instrument as a consequence of their health status). The content of the instrument may also contribute to the response rate. It has been reported that elderly respondents were more likely to complete items relating to physical functioning as opposed to emotional items (Guyatt et al., 1993).

Another aspect of acceptability concerns the length of time taken to complete the instrument. Surprisingly few studies have assessed the time required to complete various HRQOL instruments (Fitzpatrick et al., 1998). In one of the few studies to examine this issue in a cardiac population, four instruments were compared in terms of assessment times. The SIP, NHP, HPPQ (Erdman, Duivenvoorden, Verhage, Krazemier & Hugenholtz, 1986) and HADS (Zigmond & Snaith, 1983) were completed by two groups of patients: one with a history of stroke and one with MI (Visser et al., 1995). In the post-MI sample, the mean assessment time of the HADS was 5.8 minutes in comparison with the NHP 8.6 minutes, the

HPPQ 10.4 minutes and the SIP 22.1 minutes. This study showed that HRQOL instruments could be applied within a reasonable time to cardiac patient populations.

3.4.2 Feasibility

Feasibility refers to the burden of instrument administration in collecting and processing the data. Thus, while face-to-face patient interviews may produce higher response rates, they are far more costly both in terms of staff and financial requirements. The complexity of the instrument is also a consideration as more complex instruments may require more staff time to assist and explain the instructions and items.

Once the instrument has been completed and returned, it must be processed. Typically this involves translating the responses to numbers and instruments vary in the complexity of this process. Some instruments involve complex transformations of data (e.g., SF-36) whereas others simply use the response number marked (e.g., HADS). Administration burden includes information abut time needed to administer the instrument, the assumptions about the level of training/expertise needed to administer the instrument and any special resources needed to administer, score or analyse the instruments. A main concern with the routine collection of HRQOL data is the resources required in terms of administering, collecting, processing, cleaning, analysing and storing the patient questionnaire. This issue has received little attention in the literature.

3.4.3 Screening capabilities

One property not considered by the SAC or Fitzpatrick et al. (1998) is that of screening capabilities. It may be desirable that an instrument can identify those individuals in need of particular intervention (e.g., psychological distress or vocational assistance). Allison et al. (1995) used the SCL-90R to identify those patients in psychological distress in a cardiac rehabilitation programme. Distressed patients were 15 times more likely to be rehospitalised and 12 time more likely to suffer from another cardiac event at 6 month follow-up. In addition, the distressed group had additional rehospitalisation costs that totalled more than $300, 000. With the ever increasing pressure for healthcare cost containment, these findings illustrate the potential for health cost savings: by administering a screening measure

to identify those patients with particular needs which can be appropriately targeted, it may be possible to substantially reduce health care costs.

3.4.4 Cross-cultural considerations

Recent years have witnessed a increased concern at both a theoretical and a psychometric level for HRQOL assessment across different cultures. At a theoretical level, the universality of dimensions of HRQOL remains to be clarified and the comparability of these dimensions in culturally different populations is unknown. To date, most instruments have been developed in the USA or in selected European countries. The dimensions of HRQOL identified by researchers in these countries may not prove applicable in different cultural contexts.

In terms of psychometrics, cross-cultural validity of HRQOL assessment involves more than simply translating an instrument into a new language. For example, the content areas and the response options of the instrument may not be appropriate. It has been suggested that the distinction between commonly-used terms like "often" and "frequently" may differ across cultures (Cella et al., 1995). Cultural differences have also been noted in relation to norms for self-disclosure (Bullinger et al, 1995). In addition, items may be literally equivalent in different cultures but the meaning may change. Scale equivalence refers to the extent to which individuals in different cultural groups respond to similar items in similar ways. This has implications for establishing the psychometric properties of instruments across cultures.

3.5.0 Responsiveness

It has been argued that evaluative measures must be able to detect a clinically important change over time, even if this change is small (Guyatt, Deyo, Charlson, Levine & Mitchell, 1989). The ability of a measure to reflect small but important changes in health status is called responsiveness (Deyo, Diehr & Patrick, 1991). Whereas validity and reliability have been extensively investigated for many HRQOL instruments, there is a paucity of information concerning responsiveness. This is quite puzzling since responsiveness to

change over time is considered to be a fundamental property of outcome measures (Guyatt, Walter & Norman, 1987). The terms responsiveness and sensitivity have been used interchangeably with the HRQOL literature. However, the present thesis uses the term responsiveness, as sensitivity has a number of general uses within epidemiology (Deyo et al., 1991.)

Responsiveness has been conceptualised as a separate psychometric property to validity and reliability (e.g., Aaronson, 1989; Ebrahim, 1995; Guyatt et al., 1987). However, it should be noted that others have argued that responsiveness is in fact a component of validity (Hays et al., 1995). An instrument's responsiveness to treatment effects is increasingly considered as evidence for the validity of evaluative instruments (Testa & Nackley, 1994). There is some controversy as to whether a concept of responsiveness is necessary (as distinct from reliability and validity) and how it should be assessed. Confusion arises because the same data may bear simultaneously on the reproducibility, validity and responsiveness of a questionnaire (Deyo et al., 1991). The issue as to whether responsiveness is an aspect of validity or a distinct psychometric property is unresolved but the present thesis considers responsiveness separately from validity and reliability until further psychometric information is available concerning the relation of responsiveness to these properties.

No one single instrument will be appropriate for all research projects as different studies will have different priorities. If the focus of research is to evaluate the impact of a health service on health status, then it is essential that the instrument selected is responsive. A research project may indicate that a given service does not improve patient health status. However, this finding may be attributable to either a real lack of service impact or to the fact that the outcome measure selected was not responsive to the service effects.

3.6.0 Assessment of responsiveness

As yet, there is no single agreed upon method of assessing responsiveness and a variety of approaches have been suggested. However, the literature on responsiveness is not as extensive as that for validity and reliability. The main approaches to responsiveness will be now considered.

3.6.1 Effect size

$$\frac{A - B}{C}$$

where A = mean at time 2

B = mean at time 1

C = standard deviation at time 1

The effect size statistic is the most commonly cited interpretation of change score. The effect size is based on the statistical distribution of the scores in a particular study population. Effect size statistics provide an indicator of the magnitude and the direction of the effect under investigation. Cohen (1977) defined an effect size as the degree to which an effect is present in the population and it offers a quantitative estimate of the difference between two groups. Thus, effect sizes can be used to provide an index of responsiveness for different instruments in a specific study and in a particular context.

This statistic compares the change score on an instrument to the variability of scores on the instrument and it is usually expressed as the pre-treatment mean minus the post-treatment mean, divided by the baseline standard deviation. As this statistic gives some idea of the magnitude of the change rather than the statistical significance, the standard deviation at baseline rather than the standard deviation of the differences between means is used (Kazis, Anderson & Meenan, 1989). It should be noted that a number of different approaches for calculating effect sizes (for a comprehensive description of various estimates of effect sizes see Rosenthal, 1994) have been developed for use in meta-analysis but the present study concerns the use of the effect size statistic to index responsiveness (Kazis et al., 1989).

It has been suggested that effect size statistics can be used to translate changes on health status measures into a standard unit of measurement that facilitates comparison among different measures in a specific study (Kazis et al., 1989). Thus, the effect size statistic provides a standardised means of comparing different instruments (Lydick & Epstein, 1993) and studies can identify which one of a number of potential instruments is most responsive to the changes evident for particular populations or health interventions in a specific context.

An effect size of 1.00 is equivalent to a change of one standard deviation in the sample. As a benchmark for assessing change, it has been suggested than an effect size of 0.2 is small, 0.5 is moderate, and 0.8 is large (Cohen, 1977). A small effect size represents a change of at least one-fifth of a standard deviation of the baseline measure. It must be noted

that Cohen's criteria for interpreting effect sizes are not based on any theoretically or statistically derived principles. Rather, they represent an intuitive estimate. Nevertheless, these values have become the accepted benchmarks against which to judge effect sizes and there is some evidence to support the validity of these values (Sedlmeier & Gigerenzer, 1989). However, whilst these cut-off points are useful as a general rules, some caution is needed for interpretation. For example, it is possible for a very sensitive instrument to indicate large effect sizes for what are subjectively relatively modest changes. It should also be noted that the responsiveness indicies are valid only for specific uses in specific population and similar to validity and reliability, responsiveness is not an inherent property of the instrument (Fitzpatrick et al., 1998).

Kazis et al. (1989) recognised that often quality of life data from an instrument is not normally distributed and suggested that instead of using means and standard deviations, the medians and interquartile ranges can be used to calculate effect sizes. However, as yet, this suggestion has not been applied within the literature.

3.6.2 Standardised Response Mean (SRM)

$$\frac{A - B}{D}$$

where D = standard deviation of score changes

The standardised response mean (SRM) statistic was suggested as a modification to the calculation of the effect size (Liang, Fossel & Larson, 1990). Liang et al. (1990) argue that the baseline standard deviation does not incorporate response variance (variation between respondents as to responses at baseline and then at follow up) and therefore it does not contain information on the instruments ability to detect change over time. Consequently they suggest the use of standard deviation of the mean change score as the denominator. The SRM has been advocated as superior to the effect size as calculated above in the assessment of the sensitivity to change when comparing various instruments because it takes account of response variance that is likely to vary between instruments (Katz, Larson, Phillips, Fossel & Liang, 1992).

3.6.3 Guyatt's Responsiveness Index

$\frac{\text{MID}}{\text{E (for stable subjects)}}$ where MID = minimum important difference

E = the square root of twice the mean square error

Guyatt and colleagues propose that responsiveness is a function of the variability in score changes for stable subjects (Guyatt et al., 1987). Stable subjects are those subjects reporting little or no change in health status. The responsiveness of a test is directly related to the magnitude of the change in subject score which constitutes a clinically important difference. Unlike the other two formulae, this approach accounts for non-specific changes in scores. If an instrument is responsive it must be able to detect the "signal" of intervention changes above and beyond the 'noise' of non-specific change. As the baseline variability increases, one will need larger and larger treatment effects to demonstrate efficacy; therefore responsiveness is inversely proportional to the between subject variability of the individual changes in score over time. Thus, the most appropriate measure of responsiveness would relate the variability in test score in stable subjects to the clinically important difference. A single index of responsiveness can be derived by taking the ratio of the clinically important difference (usually denoted by Δ) to the variability in stable subjects. Using a repeated measures design, the stable subject variability is estimated by the square root of twice the mean square error (see Guyatt et al., 1987) and the responsiveness index is given by the mean change in scores of patients who report improvement divided by the variability of patients reporting no change.

This has the advantage of accounting for (in the denominator) the non-specific score changes often observed in patients who are apparently unchanged. Responsiveness is a function not of the baseline standard deviation of scores, but of the variability in score changes for stable patients. This approach acknowledges that there is some non-specific variability in scores and that to be responsive, an instrument must detect change above and beyond this non-specific degree of change. However, the practicality of this method is limited because it requires quantifying the minimum clinically meaningful change (Liang et al., 1990).

3.6.4 Minimum Important Difference (MID)

In Guyatt et al.'s (1987) formula for estimating responsiveness, the numerator is the minimum important difference (MID). However, as a number of commentators have observed, the estimation of such a clinically important difference is not standardised. Two approaches within the literature have used patient ratings to determine MID. One approach to assessing MID relates change in quality of life scores to a patient's estimate of their global rating of change (e.g., Juniper, Guyatt, Willan & Griffith, 1994). This method compares changes in the same patients at different times and is a "within-patient" method. An alternative approach was developed by Redelmeier which used patient interpersonal comparisons of other individuals with the same condition and thus uses a "between-patient method" (Wells et al., 1993). The comparability of these methods was examined in a sample of chronic obstructive pulmonary disease patients (Redelmeier, Guyatt & Godstein, 1996) and the authors claim that both approaches provide similar results.

However, as Wright (1996) observed, while the average threshold for the MID was similar across the subscales examined, closer analysis of the data shows a high degree of variability across the subscales individually. Thus for example, the MID for the dyspnea subscale is five times higher when assessed using the within patients method compared with the results using the between patient method. Therefore, clearly further research is required to examine the comparability of results. It should also be noted that the estimate of MID depends on the scaling technique used to classify global change. Indeed, the arbitrary selection of such cut-off points can not be theoretically justified and clearly further research is needed to examine the results of using different cut-off points for MID determination. Furthermore, the MID for different interventions in different illness populations may vary according to the invasiveness and magnitude of the intervention and also according to the disease stage. It may be that individuals at the end stage of terminal illness may have lower MIDs than those in the middle stage. Empirical evidence is needed to test this hypothesis.

3.6.5 Tuley's responsiveness index

$$\frac{A\text{-}B}{E \text{ (for stable subjects)}}$$

Tuley, Mulrow and McMahan (1991) developed a method that is related to the Guyatt index. Instead of using the minimal clinical difference as the numerator, the difference of the intervention group is used. The changes in an instrument's score are thus related to the natural variability of scores on that instrument. Despite this approach having potential wider applicability than the Guyatt index, as yet, this method has not been widely used in the literature.

3.6.6 Relative Efficiency

$$\left[\frac{\text{T statistic } a}{\text{T statistic b}}\right]^2$$

where a = t-statistic for instrument a

b = t-statistic for most responsive instrument

Another approach to the assessment of responsiveness involves comparison of HRQOL instruments with a standard instrument in a patient population assumed to have experienced change (Liang, Larson, Cullen & Schwartz, 1985). The instrument that produces the largest paired t-test score for pre and post intervention is considered most responsive. Liang et al. (1985) developed a standardised version of the use of t-statistics (termed "relative efficiency") as the square of the ratio of t-statistics of two instruments being compared. One limitation of the relative efficiency (RE) approach is that it does not provide a responsiveness statistic but, rather, it gives a relative rating of each instrument in comparison with the most responsive instrument.

3.6.7 Receiver-operating characteristics

Receiver-operating characteristic (ROC) curves require an external gold-standard assessment of whether change has actually occurred. This method plots the true positive rate (i.e, a true change has occurred) for an instrument against the false positive rate for all possible cut-off points. The most responsive instrument would have a plot where the true

positive rate sharply increases whilst the false positive rate remains low. The greater the total area under the curve from all cut-off points the greater the responsiveness of the instrument. As yet, the use of receiver operator characteristics has not been widespread within the quality of life literature.

3.6.8 Floor and ceiling Effects

Some instruments will not be able to detect change over time if the scale is limited by "floor" and "ceiling" effects. A floor effect is present when the initial baseline score is so low that change cannot be detected when there is a worsening of health. The MOS-20, a scale designed to detect health status was found to have this floor phenomenon and thus is unsuitable for assessing change in health status in HRQOL studies (Bindman, Keane & Lurie, 1990). This finding provided impetus to the development of a longer version of the scale, the widely-used SF-36 questionnaire. A ceiling effect exists when patients score at the high end of a scale and therefore cannot score any higher to register any improvement in health status. Floor and ceiling effects need to be constantly monitored for in HRQOL instruments as this reduces the ability to detect change over time.

The actual form of the questionnaire may reduce the likelihood of further improvement or deterioration being recorded beyond a certain point. Instruments that do not sample items from the extremes of the domains of interest may not be capable of detecting improvement or deterioration in that domain. The distribution of difficulty levels in the instrument is an important consideration (Fitzpatrick et al., 1998). The impact of excess reliance on inter-item agreement was previously discussed. However, clearly these issues may also be of relevance to responsiveness: if the items on a scale only reflect a narrow range of issues then only minor improvements may be recorded as high levels. There is some evidence from hip arthroplasty (Stucki, Liang, Fossel & Katz, 1995) to suggest that the items on the physical functioning scale of the SF-36 do not reflect extremes of functioning. This is an important issue that merits further investigation within the literature.

3.7.0 Comparison of responsiveness statistics

The first four methods discussed above share a similar design: change in scores is divided by an estimate of the precision of the measurement to provide a responsiveness statistic. The relative efficiency approach simply rates each instrument in comparison to the most responsive instrument. Liang et al. (1990) use the same numerator as Kazis et al. (1989) for calculating the SRM but use the standard deviation of the change scores as the denominator instead of the baseline standard deviation. The baseline standard deviation does not incorporate response variance and therefore, it does not contain information on the accuracy of the instrument in detecting change over time. Indeed, some concern has been expressed regarding the appropriateness of the effect size statistic to compare responsiveness across instruments (Katz et al., 1992). However, to-date, the effect size is more commonly used than the SRM because baseline standard deviations are usually more available (Liang, 1995). With differing approaches assessing responsiveness, which one should be used?

Recent studies have compared the different approaches to responsiveness. Unfortunately, this literature is characterised by a marked lack of consensus. For example, Stucki. Liang, Fossel and Katz (1995) compared the relative responsiveness of condition-specific and generic health status measures to surgical intervention for lumbar spinal stenosis. They assessed the responsiveness of the Sickness Impact Profile (SIP), the Roland Scale and a spinal stenosis measure using three estimates of effect size: SRM, effect size, and the Guyatt responsiveness index. Each responsiveness statistic revealed the same ordering of scales, with the condition-specific measure more responsive than the generic measures. However, this study examined only the overall measure scores and did not investigate the various scales of the measures in detail.

Gliklich and Hilinski (1995) compared the responsiveness of the SF-36 and the Chronic Sinusitis Scale (CSS) in patients undergoing sinus surgery. They used the SRM and effect size statistic to estimate responsiveness of each of the sub scales of each measure. While both statistical approaches generally revealed the same ordering of scales, there were a number of differences. For example, the total CSS was the most responsive scale according to the effect size statistic whereas Sinus Pain was the most responsive according to the SRM value. Similarly, in a sample of hip arthroplasty patients, Wright and Young (1997)

investigated whether different indices of responsiveness provided the same rank ordering of scales in terms of responsiveness. They examined if disease-specific questionnaires were more responsive than generic and patient-specific scales using five estimates of responsiveness: SRM, effect size, responsiveness index, RE and the correlation of each scale with patients' global rating of change in hip function. While the disease-specific scales were generally more responsive than the generic scale, the different responsiveness indices provided different orderings of scales.

One of the few papers to show that different methods to assess responsiveness produce similar results was reported in a sample of older adults (Wagner, LaCroix, Grothaus & Hecht, 1993). In a longitudinal study of self-reported health measures (restricted activity days, bed disability days, the MOS physical function scale, self-evaluated health and positive affect), the SRM and ROC curves produced similar results in terms of responsiveness to intervening illness or injury. In general, research has shown inconsistencies in the ordering of scale responsiveness according to different methodologies (e.g., Beaton, Hogg-Johnson & Bombardier, 1997)

The reproducibility and responsiveness of the Chronic Heart Failure questionnaire (CHQ) was recently examined using a variety of measures of responsiveness: observed change, effect size and Guyatt's responsiveness coefficient (O'Keefe, Lye, Donnellan & Carmichael, 1998). Effect size and responsiveness coefficients were broadly similar but there were some discrepancies. Interestingly, for all measures used in the study, effect sizes for detecting deterioration were greater than effect sizes for detecting improvement. It is possible that ceiling effects may have limited ability to detect improvements.

The comparative responsiveness of two disease specific instruments (Epilepsy Surgery Inventory (ESI-55), Washington Psychosocial Inventory (WPSI) and one generic instrument (SCL-90-R) was recently examined in a sample of temporal lobe epilepsy patients (Wiebe, Rose, Derry & McLachlan, 1997). They used variants of the SRM and the RE to assess responsiveness to change over a year. They reported that the ESI-55, which comprises the SF-36 and 19 additional items, had the most responsive (SF-36) and the least responsive (epilepsy-specific questions) scales. Furthermore, the SCL-90-R also demonstrated higher levels of responsiveness than the specific WPSI. This finding would suggest that in this

population generic instruments may be more responsive than specific scales. Indeed, there is evidence from a number of studies to question the implicit assumption that disease–specific instruments are more responsive (e.g., Guyatt, Feeny & Patrick, 1993)

3.7.1 Summary

The effect size, SRM, RE, Guyatt index and Tuley index are concerned with slightly different aspects of responsiveness and it is not surprising that each approach produces slightly different results. As has been noted previously, there are no agreed external "gold standards" of the extent of real change against which to compare the different measures of responsiveness. Unfortunately, comparisons of the relative merits of a number of HRQOL instruments in a single study are lacking, and indeed, there are even fewer studies examining the responsiveness of instruments. At present the empirical evidence in relation to the assumed superior responsiveness of specific instruments over generic instruments is equivocal.

Ray and Shadish (1996) have recently demonstrated that different methods for calculating effect sizes for meta-analyses are not interchangeable and that more appropriate methods need to be developed. Clearly, such recommendations also apply to the assessment of responsiveness. Given the paucity of available evidence, the use of the SRM is recommended as it takes account of response variance and thus at present offers the best estimate of responsiveness.

3.8.0 Limitation of previous research on responsiveness

One potential problem with responsiveness statistics is that the score changes in the numerator may overestimate treatment effects. This is because some change is often observed even in stable patients and many studies have not included control groups. Rather, they have simply assessed the responsiveness of different instruments to the intervention. However, this approach neglects the natural changes in variables of interest over time. Thus, it is suggested that a more appropriate approach to responsiveness assessment would involve estimating the responsiveness of an instrument to an intervention and also to the change over time in a control group. Subtracting the responsiveness statistic for the control group from

the statistic of the intervention group provides an index of the responsiveness of an instrument to detect intervention effects over and above the change reported in the control group. Such an approach has not previously been used within the literature.

3.9.0 Other contributions of responsiveness to outcomes research

Consideration of responsiveness can contribute to a number of outcomes research issues. The relation of responsiveness to issues of power and of statistical and clinical significance are considered here.

3.9.1 Power

It is essential in health outcomes research that a study has sufficient power to detect treatment effects. The power of a study is the probability of not making a type II error, (i.e. falsely concluding that the treatment has no effect). The traditional approach to increasing power has been to increase the sample size. However, this also means increasing the duration and cost of the study. Others have suggested changing the alpha level from the "sacred" level of 0.05 (Judd, McClelland & Culhane, 1995) but it is rare to find any researcher adopting an alpha level greater than 0.05. As yet, few have explicitly considered the possibility of increasing power by selecting a more responsive measure. A responsive measure is more likely to detect the effects of an intervention, even if these effects are small.

Knowledge of responsiveness is especially important when designing "equivalence trials" or attempting to show no significant differences in treatment effect. One could easily be unaware of a large type II error because the error could be due to low responsiveness rather than small sample size. Low responsiveness is often not taken into account when designing clinical trials and when such an omission occurs, conclusions based upon acceptance of the null hypothesis can be extremely misleading (Testa & Nackley, 1994).

The estimation of effect size has been commonly ignored within the literature. Within the rehabilitation literature, Ottenbacher and Barrett (1990) reviewed a hundred data-based studies and performed a post-hoc power analysis. They reported that the median power to detect a small effect size was 0.08, to detect a medium effect was 0.26 and to detect a large effect was 0.56. These results suggest the possibility of a high rate of type II experimental

errors in the rehabilitation research. In this era of evidence-based medicine, such errors may result in services being cancelled or not financially supported. The psychological literature shows similarly low levels of power. Indeed, following the seminal work of Cohen (1977) identifying the low power of psychology studies, it is somewhat disheartening that a decade ago, Sedlmeier and Gigerenzer (1989) reported that the power of studies had actually decreased in the literature. A responsive instrument increases the power of the study without increasing the cost.

3.9.2 Statistical significance

Health researchers are interested in the effects of health care on health outcomes. Oakes (1986) argued that if the purpose of research is to estimate the effect of an intervention, then traditional statistical testing does not address the issue. Emphasis only on statistical significance is quite inappropriate as a sufficiently large sample size and reliable assessment procedure will result in a statistically significant result for even trivial effects (Cohen, 1990). The fact that a treatment effect is statistically significant reveals little about the magnitude of the effect. Paquin (1983) cautioned against the over dependence on p-values and levels of significance to the detriment of effect size indices when evaluating clinical programmes and research. As Rosnow and Rosenthal (1988) noted, in counselling psychology many practitioners lamented that research findings reported in terms of p values were of little or no clinical use and thus, the use of effect size statistics was recommended. Similarly, recent calls have been made for the reporting of effect size estimates in clinical trials in medicine (Borenstein, 1997). Routine reporting of effect sizes can contribute to a cumulative knowledge base in both health psychology and general medicine.

Effect size refers to the magnitude and direction of the effect under investigation. Effect size estimates can thus supplement traditional statistics to provide a richer understanding of the treatment effects. Indeed, it is worth noting that recently the American Psychological Association (APA) has been asked to ban significance tests from its journals (see Shrout, 1997) and the debate about the appropriateness of statistical significance testing continues (see Harlow & Mulaik, 1997). In recent guidelines issued by the APA Task Force

on Statistical Inference, Wilkinson (1999) recommended that some effect size estimate should always be presented when reporting a p value.

3.9.3 Clinical significance

Change in a HRQOL instrument may be statistically significant but determining the clinical significance of the finding remains problematic. Translating HRQOL change scores into clinically meaningful terms has become increasingly important in health outcomes research. In a review of the literature on health status measures and clinical significance, Lydick and Epstein (1993) identified two main approaches to the assessment of clinical meaningfulness: distribution-based or anchor-based interpretations. Distribution-based interpretations are based on the statistical distribution of scores and, as Deyo and Patrick (1995) observed, the most popular approach is the effect size statistic.

Cohen's criteria for small, medium and large effects may be used to assess the relative magnitude of HRQOL change. However, the effect size value must be interpreted with caution. It may well be that an effect size of 0.5 is medium in one area of research while being a small effect in another (Ottenbacher & Barrett, 1989). Indeed, Cohen (1990) warned against simply substituting sanctified effect size values for the sanctified alpha level of 0.05. Therefore, the effect size statistic by itself cannot tell us all we want to know about clinical significance. Effect size advocates accept the need to compare the statistic with some external measure to increase understanding of the clinical significance of effect size values. Lydick and Epstein (1993) termed this approach anchor-based interpretation.

Anchor based interpretations involve relating ("anchoring") the change in HRQOL to other external changes or results. The most common approach compares the health status score with a single global rating question, which can be rated by either the patient or the clinician. Juniper, Guyatt, Willan and Griffith (1994) reported that for a seven-point Likert scale, a change of one-half point represents the smallest clinically important change to be reported by the patient and clinician. A change of one point represented a moderate change and a change of one and a half points represents a large change. These findings support similar conclusions of Jaeschke, Singer and Guyatt (1989) regarding the interpretation of HRQOL changes. However, this technique requires some arbitrary definition of when

change is clinically important and it seems likely that the significance of a specific (e.g. one point) change may vary by clinical condition, at different points in the natural history of a disease and also at the different ends on a scale (Deyo & Patrick, 1995). Furthermore, the actual scale content in terms of response options is also a critical importance. Indeed, the cautionary words of Cohen (1990) against null hypothesis significance testing may be equally applicable in this situation:

> "Mesmerized by a single all-purpose, mechanized, objective ritual in which we convert numbers into other numbers.....we have come to neglect close scrutiny of where the numbers come from.....surely we must know where the numbers come from and should be far more concerned with why and what and how well we are measuring...." (p. 1310).

There is a risk that by simply using such half-point changes as indicators of MID, the psychological meaning of these numbers may be lost. A more fundamental problem relates to the fact that as effect sizes are dependent upon the variability in a particular population, then if the variance in one population is different from that in another then the effect size from the same amount of clinical significance will be different. Thus researchers need to wary of unwarranted generalisations across populations with differing variabilities. At present this issue has not been addressed in the literature and clearly warrants future investigation. Nevertheless, anchoring the effect size in terms of a patient's global rating of health status change might help increase the interpretation of the effect size statistics and also help determine the minimum important difference.

Alternative approaches to anchoring such changes include comparing scores with external events that may be more clinically familiar. Such approaches would be extremely beneficial by providing insights into the relationship between these changes and established clinical parameters. Scale changes may also be made more interpretable by relating certain changes with prognosis, response to treatment and predicting future events. Thus clinical significance may be associated with predictive validity of instruments. Indeed, as described earlier (see section 2.4.3), there is some evidence to suggest that self-rated health is a good predictor of subsequent mortality (Idler, 1992; Rumsfeld et al., 1999). Thus, the prognostic significance of changes in psychosocial instruments warrants investigation.

Some authors have interpreted changes in instruments in the context of experiencing a particular illness state or to well-established different classes of illness. Thus for example, changes in psychosocial status of an individual may be associated with changes in the NYHA (clinical) classification of the individual. Through widespread use and empirical investigation it may be possible to identify threshold scores that mark changes in clinical status. A variety of external anchors can be used to put the change in psychosocial instruments into perspective. If psychosocial outcome assessment is to become more widely accepted in clinical practice, then ultimately the scores on these instruments must be made widely interpretable. Through increasing use and familiarity with the instruments, it is hoped that the clinical importance of changes in scores can be anchored to more familiar outcomes. Indeed, some commentators have suggested that only a small number of instruments should be widely used in clinical trials and outcomes research (Deyo & Patrick, 1995).

Naylor and Llewellyn-Thomas (1994) discussed the role of determining clinically important differences in sample size estimations for RCTs of treatments. They state that clinically important differences have typically been selected on the basis of biological plausibility, prior clinical data, methodological convention, clinician perception and the feasibility of achieving the corresponding sample size. Patient perspectives about potential benefits and side effects of treatments have been neglected in such decision making. The use of probability trade-off techniques to determine patients' opinions about clinically important differences reveals strengths of preference for different treatment options. While such an approach may be useful in helping patients select between treatment options in terms of mortality and morbidity, its utility in determining clinically important differences in psychosocial states remains to be tested.

3.9.4 Summary

Data on responsiveness can help in the planning of studies, informing instrument selection and sample size calculation to maximise statistical power. A small HRQOL change may be statistically significant, especially with large sample sizes, but the actual clinical meaning of these changes remains uncertain. By anchoring effect size statistics to other outcomes it is intended to clarify the clinical significance of effect size values.

3.10 **Summary of Key Points**

- A multitude of HRQOL instruments is currently available and instrument selection involves careful consideration.
- A number of criteria for routine evaluative instrument selection were outlined in the present chapter: appropriateness, acceptability, precision, feasibility, validity, reliability, and responsiveness.
- Responsiveness refers to the ability of an instrument to detect small but important changes in HRQOL.
- If the primary focus of research is to evaluate health outcomes, then the instrument selected must be responsive to change.
- A number of ways of estimating the responsiveness have been suggested: effect size statistics, SRM, RE, Guyatt responsiveness index, Tuley responsiveness index
- The SRM is the most appropriate index of responsiveness as it incorporates the correlation between the two instruments on different assessment times.
- Subtracting the responsiveness statistic for the control group from the statistic of the intervention group provides an index of the responsiveness of an instrument to detect intervention effects over and above the change reported in the control group.

Chapter 4
Cardiac Rehabilitation and Outcomes Assessment

4.0.0 Cardiac rehabilitation and outcomes assessment

As mentioned previously in chapter 2, the goals of the health care service will dictate the appropriate outcome measures to be used. According to the WHO (1993), the explicit goals of cardiac rehabilitation are to promote secondary prevention and to improve quality of life. These goals should determine the outcome measures selected. In the current health care environment, cardiac rehabilitation services must demonstrate that they represent good use of health care funds by providing empirical data to show the benefits that result from allocating resources to them. Thus cardiac rehabilitation must demonstrate its effectiveness in influencing outcomes of importance to both health care users and providers. As chapter 2 discussed, HRQOL is clearly one such outcome. Cardiac rehabilitation must focus on HRQOL as it adds not just "years to life, but life to years." As has been widely noted, there must be a balance in health care between extending life and improving/restoring HRQOL. HRQOL measures have a particular relevance to CHD as the goal of therapy for most patients with chronic disease is improvement in function not a cure (Wenger, Mattson, Furberg & Elinson 1984).

Increasing health care costs have contributed to a climate of accountability in health care delivery. The demonstration of positive treatment outcomes is the basis upon which economic resources are awarded. Audit is increasingly used in health services research to provide such evidence of service effectiveness.

4.1.0 Cardiac rehabilitation and audit

The recent report of the national Cardiovascular Health Strategy (Department of Health, 1999) stated that audit must become an integral part of cardiac rehabilitation services to facilitate assessment of service effectiveness and efficiency. Such an audit is essential if the quality of service provision is to be assessed. There is a need to develop national audit guidelines for cardiac rehabilitation and to include standardised formal assessment of patient outcomes.

Thompson, Bowman, deBono and Hopkins (1997) provided a draft set of standards for audit of cardiac rehabilitation. The report recommends that services should document the assessment of coronary risk factors for each patient during cardiac rehabilitation. Of note, the report recommends that

> "Patients admitted to hospital with acute cardiac events should have a simple formal assessment for depression/anxiety before discharge" (p. 217)

However, this does not refer to the routine assessment of psychological data before and after cardiac rehabilitation. Furthermore, as the guidelines begin with an explicit recognition that there has been a move towards HRQOL outcome assessment (Todd & Cay, 1997), it is disappointing to find no section on such outcomes assessment of service participants. It is regrettable that the opportunity to recommend routine psychological data collection was missed in this influential report. It is to be hoped that future guidelines will incorporate such recommendations.

According to Lewin (1997), current HRQOL assessment techniques tend to be more suitable for research than for clinical practice and there does not appear to be a suitable British measure for assessing quality of life, for comparing outcomes between interventions and also for conducting clinical audit. The challenge facing quality of life researchers is to evaluate the suitability of available instruments for these purposes. The desirable criteria of an outcome instrument were described in chapter 3. Thus there is widespread agreement within cardiac rehabilitation circles that HRQOL is an important outcome and that it should be routinely assessed. Routine outcomes assessment in cardiac rehabilitation has a number of benefits.

4.2.0 Benefits of HRQOL assessment in cardiac rehabilitation

According to Mayou and Bryant (1993), HRQOL assessment can impact at policy, research and clinical levels. At a policy level, when decisions are made regarding funding of health services, cardiac rehabilitation services must be able to document the relative need for cardiac rehabilitation services and the benefits obtained by patients, in comparison to similar spending on other services. The role of research is to evaluate the impact of cardiac rehabilitation services and to provide information on how services can be delivered in the most effective and efficient manner. For example, research can compare the benefits of home-based versus hospital-based rehabilitation (DeBusk et al., 1985), or can focus on patient subgroups such as elderly versus non-elderly (Lavie & Milani, 1995). Where services are equally effective, considerations such as cost will need to be borne in mind. At a clinical

level, HRQOL evaluation can provide a structured system around which to set the goals of cardiac rehabilitation for individual patients from the outset. Thus it can provide a method of evaluating change in quality of life over time in individual patients and groups of patients. However, the choice of instrument remains critical and the issue now arises as to which instruments should be selected for routine outcomes assessment.

4.3.0 Cardiac Rehabilitation and HRQOL instrument selection

While a proliferation of instruments exist for HRQOL assessment, there are a number of concerns regarding the appropriateness of these instruments for use in routine outcomes assessment. For example, psychometric information concerning these instruments is currently unavailable for cardiac groups. Furthermore, many of the available instruments are either unstandardised or in the process of being validated for cardiac populations. Although informal and unstandardised questionnaires can be clinically useful, it is recommended that an instrument with documented psychometric properties be considered. This should provide greater reliability of results and accuracy in identifying abnormal quality of life levels and assessing change over time. Indeed, there is some evidence that many cardiac rehabilitation services currently use unstandardised measures to assess quality of life. A recent survey by the American Association of Cardiovascular and Pulmonary Rehabilitation Outcomes Committee (AACVPR) (1995) reported that while 62% of US programmes assess quality of life, only 22% of these use standardised measures. Similarly, in the UK, the use of validated

psychosocial assessment (defined as the use of a published scale or of a standardised procedure with known properties) was reported in only 27% of cardiac rehabilitation centres surveyed (Lewin et al., 1998). A final issue relates to the fact that many of these measures are not tailored for clinical use.

4.3.1 Guidelines for instrument selection

The AACVPR (1995) outcome tools resource guide provided a brief synopsis of a number of instruments available for assessing outcomes in cardiac rehabilitation. It recommended that cardiac rehabilitation should routinely assess outcome of cardiac rehabilitation in three general areas: Health (which includes morbidity, mortality, and HRQOL), Clinical (which includes physical, psychological, and social functions), and Behavioural (which includes behaviours that patients attempt to change in order to reduce cardiac risk).

While such an approach by the AACVPR is to be welcomed, there are a number of concerns with the guide. Firstly, there is no definition of HRQOL provided to distinguish it from physical, psychological and social functioning. Consequently the listing of HRQOL instruments includes measures that are quite varied in focus. Thus, for example, despite their differing focus, the Multidimensional Health Locus of Control (Wallston, Wallston & DeVillis, 1978) and the Nottingham Health Profile (Hunt et al., 1980) are listed together as measures of quality of life.

The section on psychological outcome measures contains an equally heterogeneous mixture of questionnaires ranging from the COPE (Carver, Scheier & Weintraub 1989) and the Minnesota Multiphasic Personality Inventory (notably, the feasibility of using a three hour personality test as a routine outcome measure is not addressed). Furthermore, the inclusion of such personality measures as the 16PF and MMPI suggests that cardiac rehabilitation should aim to produce changes in basic personality structure. One may legitimately ask as to whether this is an appropriate goal of cardiac rehabilitation? In addition, the list of measures is incomplete and does not include a number of measures commonly used in Europe such as the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983), the Heart Patients Psychological Questionnaire (HPPQ) (Erdman et al., 1986) or the General Health Questionnaire (GHQ) (Goldberg, 1978). Thus, AACVPR guidelines provide an incomplete listing of measures and no clear theoretical guidelines to instrument selection.

As yet there are no guidelines on HRQOL assessment from the European Association of Cardiovascular Rehabilitation (EACVR). However, the European Cardiac Society working group on Cardiac Rehabilitation and Exercise Physiology is convening a working group to develop consensus guidelines on HRQOL assessment in cardiac rehabilitation. Such initiatives are to be welcomed as the current project to develop a core audit system for cardiac rehabilitation centres in Europe includes a section on HRQOL assessment pre and post cardiac rehabilitation (Sala, personal communication). Such a database would prove invaluable in documenting the effects of different formats of cardiac rehabilitation in different settings.

Thus if consensus on HRQOL assessment could be achieved it would ensure that the HRQOL of cardiac rehabilitation participants across a wide range of settings could be evaluated and the impact of different forms of cardiac rehabilitation service delivery could be assessed. Furthermore, such data could be used to help efforts at quality control in cardiac rehabilitation services. Indeed, in the USA this has been explicitly recognised:

> "Cardiac rehabilitation professionals must continue to develop innovative means to deliver their services and to document what they are doing by using outcome assessment and cost control." (Froelicher, Herbert, Myers & Ribisl, 1996, p. 154)

Such research could examine the outcomes of different formats of cardiac rehabilitation delivery or different programmes. Indeed, one key issue facing cardiac rehabilitation is how different formats of delivery compare with each other in terms of effectiveness. Such evaluations should include HRQOL assessment. Given the literature documenting the adverse influence of psychosocial variables, it is clear that cardiac rehabilitation programmes should routinely include psychosocial outcome when evaluating their effectiveness.

There is widespread agreement that HRQOL is multifaceted. However, the number of components of HRQOL differs from author to author. For example, Thompson, Meadows and Lewin (1998) suggested that the assessment of health-related quality of life with CHD patients should include recreation, work, social, personal and sexual relationships, emotional state, as well as the acute and chronic physical consequences of the disease including fatigue.

However, no rationale is provided for selecting these variables as opposed to other variables. Furthermore, the important scientific considerations for instrument selection highlighted by Thompson et al. (1998) are reliability and validity. Responsiveness is mentioned only briefly and there is no discussion of the methods to assess responsiveness, nor are any references to the literature on responsiveness provided.

The selection of quality of life instrument should be based on careful consideration of its psychometric properties, the comprehensiveness of relevant health domains sampled, and the specific clinical or research purpose for which it is to be used. Thus, decisions regarding the choice of quality of life measures should always be made in terms of the research question to be addressed and the relevant instrument criteria. The criteria of an evaluation instrument outlined in chapter 3 can aid instrument selection: user-friendliness; validity, reliability, provision of clinically significant information, screening capabilities and responsiveness.

Of note, no study has addressed the responsiveness of a number of psychosocial instruments to cardiac rehabilitation interventions using the responsiveness statistics outlined in the previous chapter, although a number of papers have examined different psychometric aspects of psychosocial instruments in cardiac populations.

4.4.0 Responsiveness of psychosocial measures to cardiac rehabilitation

The mixed findings of psychosocial benefits of cardiac rehabilitation participation may reflect the use of inappropriate outcome instruments. Denollet (1993a) noted that the psychosocial instruments used in cardiac rehabilitation outcomes research may not be sufficiently responsive to detect a beneficial psychological effect. He suggested that the use of measures of psychopathology (e.g., Beck Depression Inventory (Beck, Ward, Mendelsoe, Mock & Erbauch, 1961)) may be inappropriate for cardiac populations as these instruments may not address issues that concern the majority of coronary patients who do not experience psychopathological problems. Furthermore, these instruments may be too general to detect significant effects and there may be a poor match between the instrument content and the psychosocial states that actually change over the course of cardiac rehabilitation. Of note, Denollet has developed two specific instruments, the Global Mood Scale (GMS) and the Health Complaints Scale (HCS) for use with cardiac populations. In a series of studies, Denollet has provided evidence that some instruments (e.g., HPPQ, GMS, HCS) may be more appropriate outcome instruments than some of the other commonly used instruments (Denollet, 1993a; 1993b; 1994; Denollet & Brutsaert, 1995). However, while Denollet's research suggested that some instruments are more responsive to change, the responsiveness of the instruments used in these studies has not been examined using the responsiveness statistics outlined in the previous chapter. Rather, Denollet's recommendations are based on

analysis using statistical significance levels and such an approach does not provide an index of responsiveness.

Dempster, Bradley, Wallace and McCoy (1997) compared the results of the SF-36 with the NHP in a sample of 28 participants at the commencement of a cardiac rehabilitation programme. No rationale was provided by the authors for selecting these two particular generic measures other than the fact they are widely used. The authors developed criteria for assessing ceiling (at least 25% of the study population scoring at the maximum) and floor effects (at least 25% of the study population scoring at the minimum). Ceiling effects were reported for five of the six NHP scales (physical mobility, isolation, pain, emotion, energy) and only one subscale of the SF-36 (Role limitations due to emotional health problems). Two of the subscales of the SF 36 (role limitations due to physical health problems, role limitations due to emotional health problems) had floor effects.

The SF-36 scales generally showed higher reliability coefficients (Cronbach's alpha) than those of the NHP. Seven of the eight SF-36 subscales had an alpha greater than 0.70 with the exception being the two-item social functioning scale. In contrast only three (physical mobility, sleep, emotion) of the six NHP scales had an alpha greater than 0.70. The validity of the measures was assessed by examining the intercorrelation matrix of each scale. No strong correlations between any of the subscales on either the SF-36 or NHP were reported suggesting the each subscale is assessing different aspects of quality of life in each measure.

The authors concluded that based on the reliability and validity data reported, the SF-36 appears to be the superior instrument. The NHP is limited by the forced yes/no response format and suffers from floor and ceiling effects. Based on the findings of this study, the authors recommended that

> "in the absence of any contradictory studies, the SF-36 should be the instrument of choice when measuring QOL in cardiac rehabilitation." (p. 216)

One can question whether this recommendation is valid given the small sample size and the fact that the SF-36 was compared only with the NHP. In addition, such a cross-sectional study cannot provide information about the responsiveness of the instruments. However, limited responsiveness data is available from a recent study by Taylor, Kirby, Burdon and Caves (1998) for the NHP, SIP and MHIQ.

Eighty-eight consecutive first myocardial infarction patients completed the SIP, NHP and MHIQ at 6 weeks and 6 months post-discharge. Similar to the study by Dempster et al. (1997), the authors provided no explicit rationale for the selection of measures other than the fact that these measures are widely used. Responsiveness was assessed using the effect size statistic and effect sizes of 0.30 or less were obtained for all subscales. Indeed, for both the NHP and MHIQ, no values of 0.20 or greater were reported. These findings could be accounted for in terms of ceiling effects in the measures at time one. The frequency of ceiling scores at time 1 for the various subscores of the SIP and NHP varied from 33% to 85%. While

the MHIQ did not have such a high proportion of high scores, the authors reported that a similarly high proportion of scores tended towards the ceiling value.

These findings lead the authors to conclude that the three instruments used (SIP, NHP and MHIQ) showed only modest responsiveness and thus are unlikely to detect changes in HRQOL that may occur as a result of recovery or of a particular medical intervention. Furthermore, the authors cautioned against the use of such generic instruments in evaluating changes in patients over the course of cardiac rehabilitation. However, the assertion that these instruments may not be able to detect change as a result of an intervention may be challenged; as the study participants received no interventions and showed little change over time, it does not necessarily follow that a sample of patients receiving a cardiac rehabilitation intervention would fail to show any benefit on the instruments used in the study. In essence, this study provides information on how a control group would perform on the measures selected. Such information is useful in facilitating comparison with the magnitude of change in an intervention group.

In addition, Taylor et al. (1998) stated that their findings highlight the need to improve existing instruments and to develop new ones. Indeed the authors recommended the development of "cardiospecific quality of life measures is required". However, is this recommendation warranted given the paucity of data we have concerning the responsiveness of existing psychosocial instruments?

An alternative approach would be to examine the responsiveness of available instruments to establish which measures are most responsive to cardiac rehabilitation interventions. These measures could then act as benchmarks against which the performance of other scales can be assessed and can help determine the need and possible content of a new cardiac specific measure. Such an approach has the advantage of utilising existing measures and thus avoiding the cost of validating new measures. Indeed, with new scales being developed so often, there is a concern that the actual validation of these scales may not be as extensive as that of the existing measures.

4.5.0 Appropriate psychosocial states for cardiac rehabilitation outcome assessment

As noted in Chapter 2, the term "quality of life" has been used variously as a term to describe the use of diverse measures of health status, psychological well-being and social functioning in the cardiac rehabilitation literature. Since the term quality of life is used ambiguously and since it can also reflect aspects of life not usually considered to be alterable by cardiac rehabilitation (e.g., living conditions), the present thesis will consider more focused terms reflecting those psychosocial variables deemed conceptually appropriate in evaluating the effectiveness of cardiac rehabilitation.

The aim of cardiac rehabilitation from a psychosocial perspective has typically been to reduce those difficulties most associated with cardiac events and those with the most negative consequences for patients. Three types of psychosocial outcomes associated with adverse

consequences for cardiac patients were highlighted previously: depression (section 1.4.1), anxiety (section 1.4.2) and negative psychosocial states (1.4.3).

4.5.1 Positive and negative psychosocial states

A key issue facing cardiac rehabilitation programmes is that to date most research has focused almost exclusively on the amelioration of negative psychosocial states. Indeed, it has been noted that the discipline of psychology in general has concentrated on repairing damage within a disease model of human functioning (Seligman & Csikszentmihalyi, 2000) and this almost exclusive attention to pathology results in a biased understanding of functioning. As Salovey et al. (2000) note, this pathological bias of most health psychology research results in an imbalance between our well-established knowledge of how negative emotions promote illness and a relative paucity of knowledge of how positive emotions may promote health. Indeed, while quality of life models emphasise both the positive and negative aspects of functioning, it must be acknowledged that research has tended to be concerned with the consequences of negative psychosocial states. Positive states such as optimism, a sense of personal control and a sense of meaning have been identified as powerful resources which may be protective of both mental and physical health (Taylor, Kemeny, Reed, Bower & Gruenewald, 2000).

There is evidence that finding meaning in a heart attack has beneficial impact on recovery from such an event (Affleck, Tennen, Croog, & Levine, 1987). Optimism has been associated with better psychological and physical recovery from coronary artery bypass surgery (Scheier et al 1989). A small number of studies have investigated the effects of cardiac rehabilitation on patients' self-efficacy ratings for performing lifestyle behaviours (e.g., Oldridge & Rogowski 1990; Gulanick, 1991; Jeng & Braun 1997). Given the important relationship between self-efficacy and health behaviours, cardiac rehabilitation services can play a critical role in enhancing levels of efficacy. Cardiac rehabilitation typically incorporates the mechanisms by which efficacy can be increased, e.g., verbal persuasion, mastery experiences, physiological feedback and vicarious learning. The standard definition of cardiac rehabilitation explicitly states that it aims to

> "to ensure patients the best possible physical, mental and social conditions so that they may, *by their own efforts*, preserve, or resume when lost, as normal a place as possible in the life of the community." (WHO, 1993) (emphasis added).

Cardiac rehabilitation aims to increase an individuals' sense of control and it is suggested that this increase may reduce emotional distress (Denollet, 1993). This hypothesis requires empirical testing.

4.5.2 Cardiac rehabilitation instrument selection

Some general texts provide psychometric information on quality of life instruments (e.g., Bowling, 1997) and the Outcomes Committee of the AACVPR have published a listing of "selected valid and reliable tools" for cardiac rehabilitation (AACVPR, 1995). However, specific data are not available to compare across measures in the cardiac rehabilitation context and the AACVPR recommendation is based on general psychometric characteristics of measures rather than their performance in a cardiac setting. Given the paucity of data to inform instrument selection in evaluation, and the widespread use of instruments without any psychometric justification in practice, there is a need to evaluate the appropriateness of available psychosocial instruments for routine cardiac rehabilitation service evaluation. Therefore, a review of the responsiveness of psychosocial instruments to cardiac rehabilitation interventions would provide empirical evidence to inform instrument selection.

4.6.0 Summary of Key Points

- The literature on psychosocial outcomes in cardiac rehabilitation is characterised by a lack of consensus on the most appropriate instrument to use for service evaluation.
- The aim of cardiac rehabilitation from a psychosocial perspective must be to reduce those difficulties most associated with cardiac events and those with the most negative consequences for patients.
- Depression, anxiety and general negative psychosocial states have been associated with adverse outcomes for cardiac patients.
- Despite the emphasis on positive states within quality of life research, the impact of positive psychosocial states which enhance capabilities and recovery from cardiac events has not received much empirical investigation
- The responsiveness of psychosocial measures of these states to cardiac rehabilitation has not been empirically examined.
- Responsiveness has not been explicitly considered in recommendations for psychosocial instrument selection in cardiac rehabilitation outcomes assessment.

Chapter 5
Psychosocial Outcomes Assessments for use in Cardiac Rehabilitation Service Evaluation: A 10-year Systematic Review

5.0.0 Introduction

The present chapter is a systematic literature review of psychosocial outcomes after cardiac rehabilitation. A systematic review may be defined as the application of scientific strategies that limit bias to the systematic assembly, critical appraisal and synthesis of all relevant studies on a specific topic (Cook, Sackett & Spitzer, 1995). The aim of this review is to examine the responsiveness of psychosocial instruments most appropriate for cardiac rehabilitation outcomes assessment.

5.1.0 Method

The following systematic search strategy was adopted to identify the relevant research on psychosocial outcomes after cardiac rehabilitation for the period 1986-1995. Computer databases from relevant disciplines Medline (Medicine), Psychlit (Psychology), Cinahl (Nursing and allied health professions) and Sociofile (Social Sciences) were searched using the following terms: angioplasty; coronary artery surgery; coronary artery bypass; myocardial infarction; heart transplant; PTCA; psychological; psychosocial and quality of life. Relevant review articles were examined to obtain further references. The bibliographies of all articles identified by the above methods were also searched. In addition, published conference abstracts from relevant professional associations were examined: American Association of Cardiovascular and Pulmonary Rehabilitation; American College of Cardiology; American Heart Association; American Psychosomatic Association; European Congress of Cardiology; European Health Psychology Society; International Health-Related Quality of Life Society; and National Teaching Institute for the American Association of Critical Care Nurses.

Cardiac rehabilitation was defined broadly, as any programme that consisted of an exercise component alone or an exercise component with psychosocial intervention. Evaluation of exercise alone or multi-factorial programmes was considered separately. The responsiveness of psychosocial instruments was calculated using the formula of Kazis et al. (1989):

$$\text{effect size} = \frac{\text{mean change over time}}{\text{baseline standard deviation}}$$

Traditionally an effect size of 0.2 is considered to represent minimal clinically significant change, 0.5 moderate change and 0.8 or greater major change. The effect size statistic is based on the mean and standard deviation, data that are commonly given in published papers, and it can be used to estimate responsiveness of instruments to cardiac rehabilitation interventions. As it is not possible *a priori* to specify a cut-off point, an effect size that best separated the instruments used into two groups, a "more" and a "less" responsive group, was to be selected based on the results obtained.

5.2.0 Results

This systematic review identified 32 intervention studies which assessed psychosocial outcomes, including depression and anxiety, before and after cardiac rehabilitation programmes and which included a control group (See references 1-32 in References). In 16 of 32 intervention studies the data necessary to calculate effect sizes (means and standard deviations) were not published in the original report. In these cases the corresponding author was contacted and asked to forward the relevant information. The final number of studies that could not be included because of missing information was 11 (See references 22-32 in References). Study details of multi-factorial cardiac rehabilitation are presented in Appendix 1 and details of exercise only cardiac rehabilitation programmes are presented in Appendix 2.

Based on Appendices 1 and 2 it was decided that an effect size difference of ≥0.3 best represented those less and more responsive. An instrument was categorised as responsive if the effect size for the rehabilitation group minus the effect size for the control group was ≥0.3 in more than one study. The review identified only five instruments (BDI, STAI, GMS, HCS, and HPPQ) meeting the modest responsiveness criterion established here. Because of the large number of potentially useful psychosocial instruments and the lack of a clear 'leader' in this context, those other instruments which have shown similar effect size difference of ≥0.3 in one study to date are also listed. Instruments identified as responsive and promising using these criteria are described in Table 5.1.

Table 5.1. Psychosocial assessment instruments identified as responsive (effect size difference between intervention and control ≥0.3) in one or more cardiac rehabilitation study.

INSTRUMENT	*STUDY REFERENCE NUMBER*	EFFECT SIZE DIFFERENCE
DEPRESSION		
Beck Depression Inventory	*5*	0.45
	18	0.34
Cardiac Depression Scale	*6*	1.24
note: intervention group vs. two different control groups		1.10
Hospital Anxiety and Depression Scale - Depression	*14*	0.58
Multiple Affect Adjective Checklist - Depression	*15*	0.57
ANXIETY		
Spielberger State Anxiety Inventory	*3*	0.45
	2	0.43
	18	0.86
Hospital Anxiety and Depression Scale - Anxiety	*14*	0.49
Multiple Affect Adjective Checklist - Anxiety	*15*	0.50
NEGATIVE PSYCHOSOCIAL OUTCOMES		
Global Mood Scale		
Negative affect	*9*	0.57
(Positive affect)		(0.49)
Negative Affect	*11*	0.34
(Positive Affect)		(0.47)
Health Complaints Checklist		
Somatic Complaints	*9*	0.37
Cognitive Complaints		0.40
Somatic Complaints	*10*	0.47
Cognitive Complaints		0.30
Heart Patients Psychological Questionnaire		
Disability	*9*	0.62
(Well Being)		(0.59)
Disability	*11*	0.72
(Well Being)		(0.52)
(Well Being)	*13*	(0.48)
General Health Questionnaire30–Psychological distress	*7*	0.62
Profile of Mood States		
Tension	*5*	0.63
Total Mood Disturbance		0.30
Psychosocial Adjustment to Illness Scale		
Total psychosocial adjustment to illness	*15*	0.49
Quality of Life after Myocardial Infarction Questionnaire		
Restriction	*2*	0.35
Emotion		0.30
(Confidence)		(0.37)
Total		0.37

Note: Study reference number refers to the number assigned to each study included in the review –see References for full listing of studies included in the review and the corresponding number.
Note: Positive outcomes are given in brackets.

Twelve instruments were identified as potential routine cardiac rehabilitation psychosocial evaluation instruments (see Table 5.2 for summary details of instrument characteristics).

Table 5.2. Summary characteristics of potential cardiac rehabilitation evaluation instruments†.

MEASURE	FOCUS OF SCALE & SCREENING CAPABILITIES	RELIABILITY	VALIDITY (variance explained)
DEPRESSION			
Beck Depression Inventory (21 items)	Depression Score 10-15 mild depression, 16-19 mild/moderate, 20-29 moderate-severe, 29+ severe.	see Beck *et al.*, (1961), also Beck *et al.*, (1988)	
Cardiac Depression Scale (26 items)	Depression in cardiac patients	see ref. 6 Cronbach's alpha (0.90) Test-retest (0.86)	First-order factor analysis- 8 factors Second-order factor analysis- 2 factors
Hospital Anxiety and Depression Scale Depression (7 items)	Depression Score above 11 index clinical depression	see Herrmann (1997)	
Multiple Affect Adjective Checklist Depression (12 items)	Depression	see Zuckerman *et al.*, (1983) Cronbach's alpha= 0.82	Factor analysis 5-11%
Profile of Mood States Depression-dejection (12 items)	Depression-dejection	see McNair *et al.*, (1981)	
ANXIETY			
Spielberger State Anxiety Inventory (20 items)	State levels of anxiety	see Spielberger *et al.*, (1970)	
Hospital Anxiety & Depression Scale Anxiety (7 items)	Anxiety Score above 11 index clinical anxiety	See Herrmann (1997)	
Multiple Affect Adjective Checklist Anxiety (12 items)	Anxiety	see Zuckerman *et al.*, (1983) Cronbach's Alpha = 0.80	Factor analysis 5-11%
Profile of Mood States Tension-anxiety (9 items)	Tension-anxiety	see McNair *et al.*, (1981)	

NEGATIVE PSYCHOSOCIAL OUTCOMES			
General Health Questionnaire (30 item version)	General levels of psychiatric distress Score above 4/5 index psychiatric caseness	Cronbach alpha (0.88)- cited in ref. 7 see also Goldberg (1972; 1978)	
Global Mood Scale		Cronbach's alpha- see ref. 8	Factor analysis-2 factors
Positive mood (10 items)	Positive mood states	0.91	18%
Negative mood (10 items)	Negative mood states	0.94	43%
Health Complaints Checklist		Cronbach's alpha- see ref. 10	Exploratory Factor Analysis- 2 factors
Somatic complaints (12 items)	Somatic complaints	0.89	Confirmatory Factor Analysis- 2 factors
Cognitive complaints (12 items)	Cognitive complaints	0.95	Total variance explained- 56%
Heart Patients Psychological Questionnaire		see ref. 12	Factor analysis- 4 factors
Well Being (12 items)		Guttman's Test-retest	Discriminative validity
Feelings of Disability (12 items)	Psychological and physical well-being	Lambda (Pearson r)	Construct validity
Despondency (10 items)	Feelings of disability as result of illness	0.93 0.73	
Social Inhibition (6 items)	Negative mood, displeasure & dysphoria	0.87 0.85	
	Social influence and confidence in	0.80 0.80	
Dummy items (12 items)	social settings	0.64 0.77	
Multiple Affect Adjective Checklist (70 items)		see Zuckerman *et al.*, (1983) Cronbach's alpha	Factor analysis- 5 factors Total variance explained 47 to 51%
Positive affect (21 items)	Positive affect	0.93	17-20%
Sensation seeking (12 items)	Sensation seeking	0.74	4-8%
Hostility (15 items)	Hostility	0.82	7-11%
Profile of Mood States (65 items)			
Fatigue-inertia (15 items)	Fatigue-inertia	see McNair *et al.*, (1981)	
Anger-hostility (8 items)	Anger-hostility		
Vigour-activity (7 items)	Vigour-activity		
Confusion-bewilderment (7 items)	Confusion-bewilderment		

Psychosocial Adjustment to Illness scale (46 items)			
Health care orientation (8 items)	Health care orientation	see Derogatis (1976)	
Vocational environment (6 items)	Vocational environment	also Derogatis & Lopez (1983)	
Domestic environment (8 items)	Domestic environment		
Sexual relations (6 items)	Sexual relations		
Extended family relations (5 items)	Extended family relations		
Social environment (6 items)	Social environment		
Psychological distress (7 items)	Psychological distress		
Quality of Life after Myocardial Infarction Questionnaire (27 items) *		see Valenti *et al.*, (1996) Cronbach's alpha	Factor analysis- 3 factors
Emotional (14 items)	Emotional quality of life (QoL)	0.95	Emotional (28.1%)
Physical (14 items)	Physical QoL	0.93	Physical (17.2%)
Social (13 items)	Social QoL	0.95	Social (21.4%)
* some items load onto more one factor			

† For some instruments examples of data illustrating psychometric properties are provided. For others, references to extensive summaries of data demonstrating the psychometric adequacy of the instruments are provided – see references for listing of studies and the corresponding number.

5.3.0 Discussion

Evaluation of the identified studies demonstrates that there is little consensus on instrument use - most instruments were used in only a small number of studies. The most responsive instruments demonstrated moderately significant differences between intervention and control groups. Two of the five most responsive instruments were long-established depression and anxiety scales (BDI, STAI) with two of the three more recent instruments coming from one (Belgian) research group (GMS, HCC). While the latter instruments were each developed with cardiac populations, only one is a disease-specific scale - the HPPQ. Of the instruments with the highest responsiveness in one study only, two are cardiac-specific measures developed in this decade (CDS, QLMI), two others are long-established generalised mood scales (MAACL, POMS). The remaining scales are a general measure of psychological distress (GHQ-30), an anxiety and depression scale for use with hospital populations (HADS), and one is a long-established measure of general illness impact (PAIS).

An important consideration is the time between the initial assessment and subsequent assessment. The studies reviewed in Appendices 1 and 2 show great variability in measurement times and it is difficult to offer conclusions concerning responsiveness over time for the reviewed instruments. However, it appears as if the HADS shows a relatively high level of responsiveness up to one year, whereas the QLMI levels remain high up to 8 months but decrease at 12 months. The STAI appears responsive in the short term (i.e., 2 months) but not in the long term. Unfortunately, only a small number of studies have long-term follow up of patients and there is a marked lack of comparability across studies in terms of assessment time. These factors limit any conclusions that can be drawn from the literature to date.

A further issue relates to the wide variation in terms of interventions given to cardiac patients in the studies reviewed. Interventions ranged from exercise only to comprehensive multi-factorial rehabilitation. Given the small number of exercise-only studies covered in this review it is difficult to draw conclusions concerning the differential impact of multi-factorial versus exercise-only on responsiveness. A further complication is the variable time of assessment. Indeed, comparable data are only available for a small number of studies using the BDI and STAI. In the short-term, the BDI was responsive in one of five multi-factorial

programmes and also in the one exercise-only study included. The STAI was most responsive in the short-term to the exercise-only study but the multi-factorial studies also showed evidence of responsiveness in the short-term.

The present review identified a number of disease-specific and generic instruments as being potentially appropriate for cardiac rehabilitation outcomes assessment. In general, the decision to use a disease-specific or a generic questionnaire depends primarily upon the purpose of the study. No one single measure will be appropriate for all research projects as different studies will have different priorities. If the focus of research is to evaluate the impact of cardiac rehabilitation on psychosocial status then it is essential that the measure selected is responsive. If the purpose of the study is to identify those individuals with psychological distress, then the instrument must have screening capabilities and well established norms. Disease specific questionnaires may focus on aspects that are more relevant to specific illness groups while generic measures allow comparability across illness groups. While further evidence is necessary on each of the identified instruments, we suggest that if both a disease specific and generic measure are shown to have equal responsiveness, then it may be preferential to use a generic measure as this facilitates comparison with other patient and general populations. This then provides further opportunity to understand the impact of a cardiac event and rehabilitation on patients. However, the selection of instruments is determined primarily by the research question being addressed.

The instruments listed in Table 5.2 are a mixture of generic and disease-specific scales. A number of the generic instruments have been extensively used in research and have undergone many psychometric analyses. For these instruments, references are provided to summaries of the psychometric adequacy of these instruments. In terms of the disease-specific instruments, the HPPQ is the oldest and evidence is available of its validity (see Erdman et al., 1986). A new factor structure and scoring system for the QLMI have been proposed (Valenti et al., 1996) and as yet there have been few published applications of the revised scale. The GMS, HCC and CDS have only recently been published and they have not been widely used. Further psychometric analysis of these instruments is recommended. Some measures (e.g., BDI, HADS, GHQ) posses screening capabilities and can thus identify

those individuals with particular psychosocial needs. Consequently, appropriate intervention can be provided to best address those needs. Indeed, the benefits of screening instruments to cardiac rehabilitation services have been suggested (Alison et al., 1995).

Of particular note is the fact that only three (BDI, STAI, and POMS) of the twelve measures identified in this focused evidence-based review as demonstrating responsiveness in cardiac rehabilitation studies were included in the list of 18 cardiac quality of life /psychological / sociological instruments identified as being valid in the listing of the Outcomes Committee of the AACVPR (1995). While there is some evidence for the validity of the 18 scales listed, only three of these measures have proven themselves to be responsive in the studies reviewed. Thus, fifteen out of eighteen (83%) of the measures listed by the AACVPR have not yet had their responsiveness evaluated in cardiac rehabilitation populations. This highlights the fact that, in addition to establishing the usual psychometric properties of validity and reliability in the population in question, it is also important to consider evidence of responsiveness. This point is crucial if an instrument is to be used as an outcome measure.

The cut-off point of an effect size difference ≥ 0.3 used in the present review was selected arbitrarily and *post hoc*. This decision was based on analysis of the data presented in Appendices 1 and 2 as there are currently no guidelines available concerning appropriate values. While this figure did prove useful in discriminating between more and less responsive instruments, it is recommended that further research is needed to examine the utility of such a definition of responsiveness.

The main focus of this review was on the responsiveness of available psychosocial measures. Combining this evidence with that on validity, reliability, user-friendliness, and ability to provide clinically significant and screening information, the present review suggests that no one single measure best fits these requirements. However, ongoing research in the area will help better inform instrument selection in cardiac rehabilitation. To this end, the list generated here is not meant to be proscriptive. There are two different ways in which it can be used to further efforts to maximise the quality of psychosocial evaluation in cardiac rehabilitation. Firstly, the present listing provides a benchmark of the best available levels of responsiveness to date. Further studies that document improved levels of responsiveness

with other instruments are to be welcomed. Secondly, the listing can be used by those researchers and service providers wishing to evaluate services in future.

Further use of the identified instruments will increase knowledge of their overall usefulness as even for the most responsive instruments, the numbers of studies available is still very small. As the evidence available to date has been accumulated from many different countries and formats of cardiac rehabilitation service delivery, and as only one or a small number of instruments would typically have been completed in each study, it is likely that some of the variation in instrument performance is influenced by these contextual factors. Hence, a more useful approach for the future would be to evaluate a range of the most responsive instruments conjointly in a large study where a number of measures are completed by each patient. In this way, the relative merits of the various best-performing measures can be determined.

5.4.0 Conclusion

There is little consensus on psychosocial instrument use for programme evaluation in the cardiac rehabilitation literature. The quality of the available evaluative instruments (i.e., valid, reliable and responsive) will determine the ability to document the impact of programmes and to evaluate programme refinements within the cardiac rehabilitation context. The quality of cardiac rehabilitation evaluations will also influence the ability to justify services in increasingly constrained and accountable health financing systems.

The available published data are inadequate to guide instrument selection for routine service evaluation. Many researchers do not publish the basic psychometric data necessary for the estimation of responsiveness. However, some instruments do demonstrate significant potential for psychosocial outcome assessment. Comparative assessment of these instruments is recommended to inform future recommendations about instrument selection for evaluation purposes.

5.5.0 Aims of the present study

The following nine psychosocial instruments were selected for further investigation of responsiveness to cardiac rehabilitation intervention in a prospective study: Cardiac Depression Scale (CDS), Global Mood Scale (GMS), Hospital Anxiety and Depression Scale (HADS), Health Complaints Scale (HCS), Heart Patients Psychological Questionnaire (HPPQ), Quality of Life after Myocardial Infarction Questionnaire (QLMI), Quality of Life Index-Cardiac Version III (QLI CVIII), Spielberger State Anxiety Inventory (SSAI), and Medical Outcomes Trust SF-36 (SF-36)). Seven of the above instruments were identified as responsive in the systematic review. The other two instruments (SF-36, QLI-CVIII) were selected because they may also be potentially applicable in a cardiac population. The SF-36 has become a popular psychosocial outcome instrument and has been widely used across a large number of illness populations. However, as no intervention studies using the SF-36 in a cardiac rehabilitation setting were performed over the review period, the responsiveness of this instrument could not be assessed. Similarly, the QLI–CV III, a disease-specific instrument developed specifically to be used in cardiac populations, had also not been included in an cardiac rehabilitation evaluation study and thus its responsiveness remained to be determined.

In order to inform recommendations about instrument selection for outcomes assessment, the aims of the present study were:

(i) to assess the responsiveness of 9 psychosocial instruments: HCS, HADS, GMS, QLMI, QLI CV III, HPPQ, CDS, STAI, and SF-36 to multifactorial cardiac rehabilitation

(ii) to assess responsiveness using 5 different statistical approaches: Effect size, SRM, Guyatt's responsiveness index, Tuley's responsiveness index, and relative efficiency.

Chapter 6
Methods

6.0.0 Design

A quasi-experimental design was used in the study. While the disadvantages of such quasi-experimental designs have long been documented, a properly randomised trial was not feasible as the efficacy of cardiac rehabilitation in reducing mortality has been established in two meta-analyses and it is considered unethical to withhold a treatment that reduces mortality by 20-25% (Oldridge et al., 1988; O'Connor et al., 1989). Four hospitals nationally that routinely provide cardiac rehabilitation were selected and compared with three hospitals admitting cardiac patients but without cardiac rehabilitation services (Details of the hospitals are provided in Table 6.0). As the study was not a randomised controlled trial, the use of the term "control group" is inappropriate and thus the term comparison group is used in the following sections (Wilkinson, 1999).

Table 6.0. Details of study centres

Hospital	**Location**	**Teaching hospital**	**Number of cardiac beds in CCU (Beds on Cardiology ward)***
CR hospitals			
Centre 1	City	Teaching	12 (28)
Centre 2	Suburban	Teaching	6
Centre 3	Rural	Non-teaching	5 (6)
Centre 4	Rural	Non-teaching	4
Comparison hospitals			
Centre 5	City	Teaching	5 (7)
Centre 6	Rural	Non-teaching	4
Centre 7	Suburban	Teaching	5

*Note: not all hospitals have a Cardiology ward

As can be seen from Table 6.0, both the cardiac rehabilitation and the comparison centres included a mixture of academic/teaching hospitals, a main city hospital, a suburban hospital and a regional hospital.

6.0.1 Questionnaire batteries

In order to reduce respondent burden and maximise co-operation of the centres, it was decided to randomly divide the nine measures of interest into three colour coded questionnaire batteries, each comprising three questionnaires. Each participant completed one of the three colour batteries. The questionnaires were coloured white, yellow and blue. The colour coding was used to facilitate easier tracking of cardiac rehabilitation centres' questionnaire administration. As the use of a random order of presentation to participants was only feasible in one of the cardiac rehabilitation centres (centre 1) (where the researcher was located), it was decided to present the questionnaires in a fixed order to all participants. It was felt that the use of random order would place too much burden on the cardiac rehabilitation co-ordinators in the other centres for tracking questionnaires and the risk of loss of data would be too great. The final questionnaire formats are outlined in Table 6.1.

Table 6.1. Format of the questionnaire battery by colour.

Questionnaire Colour	**Instruments**		
White Questionnaire (A)	HADS	HCS	QLI CV III
Yellow Questionnaire (B)	STAI	QLMI	CDS
Blue Questionnaire (C)	GMS	HPPQ	SF-36

In order to ensure patient confidentiality and ethical integrity, and as regular ethics committee meetings were not scheduled in all hospitals, the proposed study design was sent to the national Data Protection Commissioner for evaluation. The commissioner reported that the design did not violate the Data Protection Act and foresaw no problems with the study design.

6.2.0 Sample size

Based on the findings of the systematic review presented in chapter 4, a power calculation for the study sample was performed with:

a) alpha set at 0.05

b) beta set at 0.80

c) effect size set at 0.35

The results of the power calculations revealed that 130 patients were required in each group to detect a statistically significant ($p<0.05$) difference between instruments. With three questionnaire batteries, an overall sample size of 800 was selected, with 400 (i.e., approximately 130 by three questionnaire batteries) in the cardiac rehabilitation intervention and 400 (i.e., approximately 130 by three questionnaire batteries) in the comparison group receiving standard care. Based on previous research with cardiac rehabilitation patients at Centre 1, a response rate of 60% in the cardiac rehabilitation group was assumed and thus an initial sample of 675 was required to obtain 400 participants completing the questionnaire at both time points. Based on previous experience with national cardiac surveys, a lower

response rate (50%) was assumed in the comparison group, and an initial sample of 800 was required to obtain 400 participants completing the questionnaire at both time points.

6.3.0 Participants

Consecutive cardiac rehabilitation patients at 4 national hospitals were invited to participate from March 1997 until May 1999 in a study assessing quality of life in cardiac rehabilitation patients. Comparison group patients were recruited from 3 hospitals that did not provide a cardiac rehabilitation service. Comparison patients who met the entry criteria (section 6.3.1) were assumed to have been eligible for cardiac rehabilitation if such a service was available in the hospital. The coronary admissions record of these centres were examined by the investigator to identify those cardiac patients discharged alive. The patient's physician or consultant cardiologist then wrote to the patients with an invitation to participate in the study (see Appendix 3 for copy of the invitation letter).

6.3.1 Inclusion and Exclusion Criteria

The main criterion for inclusion in the study was that the participants were considered eligible for standard cardiac rehabilitation. The comparison group were selected to match the characteristics of those patients offered cardiac rehabilitation. Based on consultation with a number of cardiologists and the previous literature, the following criteria were used to recruit participants.

The following categories of cardiac patients are considered appropriate referrals for attending a typical multifactorial outpatient cardiac rehabilitation programme: myocardial infarction, CABG, PTCA, valve surgery, and stable angina. There were no age or gender restrictions on the sample. The other main criterion was that participants also had to possess the literacy skills required for completing a self-report questionnaire in English. Cardiac patients typically not referred for cardiac rehabilitation in the study centres include those diagnosed with unstable angina and heart failure and thus cardiac patients with these conditions were not included in the present study.

6.3.2 Cardiac rehabilitation participants

Six hundred and seventy five questionnaires were administered to cardiac rehabilitation participants at four cardiac rehabilitation centres. In order to achieve a sample size of 400 cardiac rehabilitation participants, a 60% response rate was assumed. Due to differences in the number of patients completing cardiac rehabilitation in each of the study centres, a different number of questionnaires was administered in each centre (Table 6.2). The main investigator administered the instruments at the main study centre and the cardiac rehabilitation co-ordinators in the other centres administered the questionnaire batteries to their respective participants.

Table 6.2. Number of questionnaires administered at each cardiac rehabilitation centre.

	Questionnaire A	Questionnaire B	Questionnaire C	Total
Centre 1	150	150	150	450
Centre 2	25	25	25	75
Centre 3	25	25	25	75
Centre 4	25	25	25	75
Total	225	225	225	675

6.3.3 Comparison participants

For the comparison group, in order to achieve a sample size of 400, 810 patients were contacted with a response rate of 50% was assumed. The number of patients from each centre contacted reflected the number of cardiac patients seen in each of the control centres (Table 6.3)

Table 6.3. Number of questionnaires posted to cardiac patients from each comparison centre.

	Questionnaire A	Questionnaire B	Questionnaire C	Total
Centre 5	210	210	210	630
Centre 6	45	45	45	135
Centre 7	15	15	15	45
Total	270	270	270	810

6.4 Procedure

The study was based at a main suburban teaching hospital that routinely provides multifactorial cardiac rehabilitation to its patients. The other cardiac rehabilitation centres were managed by cardiac rehabilitation co-ordinators trained on a Diploma-level co-ordinator training programme at the main study hospital. All four cardiac rehabilitation programmes were based on the comprehensive model of cardiac rehabilitation provided at the main hospital centre. This model emphasises the multifactorial nature of cardiac rehabilitation and includes group-based exercise training, psychological sessions, educational sessions, dietary sessions, pharmacological sessions and also vocational rehabilitation where necessary. Thus, despite the geographical differences between the cardiac rehabilitation centres, the interventions received by participants in each centre were comparable as each centre followed the same basic multifactorial rehabilitation programme and had similarly trained coordinators.

Cardiac rehabilitation participants at four centres were approached by the investigator at one centre and by the cardiac rehabilitation co-ordinator at the other centres. The purpose of the study was explained to consecutive cardiac rehabilitation participants and they were invited to participate by taking a questionnaire envelope. A standard letter outlining the study in more detail was included in each questionnaire envelope (see Appendix 4). Those who agreed to participate were given the questionnaire to take home and return it at the next cardiac rehabilitation session. Reminders were given by the investigator and cardiac rehabilitation co-ordinators to those patients who forgot to return the questionnaire.

Participants were approached on their tenth week of cardiac rehabilitation and asked to complete the second questionnaire. All participants of the cardiac rehabilitation programmes at the study centres during the study period were invited to participate and those willing to participate were included in the study. All participants were thanked for their assistance in the study.

In two comparison centres, participants were contacted by post directly by the researcher and in one centre (Centre 7), the consultant cardiologist wrote to the potential participants and invited them to contact the researcher if they were interested in participating in the study. All comparison patients were contacted by post inviting them to participate in a national study looking at quality of life of cardiac patients. Interested patients returned the stamp addressed envelope to the main study centre. Following receipt of the letter of consent, patients were randomly assigned to receive one of three colour coded questionnaire batteries. Each questionnaire included a standard letter outlining the study. Those agreeing to participate were posted a copy of the time one questionnaire. To check that none of the comparison group participated in a cardiac rehabilitation programme, all the comparison group questionnaires included an item asking about cardiac rehabilitation attendance. None of the comparison group reported participating in cardiac rehabilitation. The time 2 questionnaire was posted, with a stamped addressed envelope inside, to those who completed the time 1 questionnaire ten weeks after receiving the first questionnaire. Those participants who had not returned either questionnaire after two weeks of the first posting were sent a reminder letter

with another copy of the appropriate (time 1 or 2) questionnaire included.

6.5.0 Questionnaires

The following section outlines the various questionnaire measures used in the present study. Where relevant, empirical investigations demonstrating validity, reliability and alpha-coefficients will be quoted for standardised psychometric measures. All questionnaires were laid out on separate sheets and all had clear instruction labels. A summary table describing the nine psychosocial instruments is provided (Table 6.4) at the end of the chapter.

6.5.1 Demographics

Demographic details of each participant were requested. Some items had a number of response options (e.g., sex) while others were open (e.g., name, occupation). For identification purposes and to allow comparison of pre and post data, the name and address of each participant was recorded. Participants were then assigned a unique identification code which was then used in all subsequent analyses. A copy of the demographic questionnaire is given in Appendix 5.

6.5.2 Duke Activity Status Index

To assess changes in general physical and functional capacity, the present study included the 12-item Duke Activity Status Index (DASI) (Hlatky et al., 1989). The DASI is a

short, self-administered questionnaire that was developed from physical testing of patients on a standard exercise capacity test. The DASI is primarily designed to "accurately measure functional status while providing insight into selected aspects of quality of life" (p. 651). The DASI was included to assess whether cardiac rehabilitation improved functional capacity and to compare the magnitude of these changes to those detected by the psychosocial instruments.

Items on the instrument are based on common physical activities that are carried out in a normal physical routine (e.g., walking, housework, sports). In the initial validation study, the DASI had a significant correlation ($r = 0.58$, $p < 0.0001$) with peak oxygen uptake and other measures of functional capacity (e.g., Canadian Cardiovascular Society Classification and Specific Activity Scale). Items are dichotomously scored ("yes" or "no") and responses are multiplied by a set of weights, the weightings are based on the metabolic costs of each activity in MET units, and items are summed to produce an overall score. Both cardiac rehabilitation and comparison patients completed the DASI (Appendix 6).

6.6.0 Questionnaire Battery A

Questionnaire battery A comprised the HADS, HCS and QLI-CVIII. Each of these instruments is described below.

6.6.1 Hospital Anxiety and Depression Scale (HADS)

The HADS is a widely used 14-item instrument developed to detect states of anxiety and depression in medical outpatients. It consists of both depression (7 items) and anxiety (7 items) subscales (see Appendix 7). Replies relate to affective feelings during the past week and ratings are made on 4-point scales, with different items having different response options. Items are summed on each of the two sub-scales. High scores on each scale indicate the presence of problems (non-cases ≤ 7; doubtful case = 8-10; definite cases ≥ 11). In the initial report of the HADS, Zigmond and Snaith (1983) reported good results for initial tests of reliability and validity, and that the scale scores were not affected by the presence of physical illness. This feature of the scale makes it extremely attractive in clinical research of physical illness. The scale also aims to detect the presence of relatively mild degrees of mood disorders likely to be present in nonpsychiatric patients. The scale is brief, readily comprehensible and easily completed. Furthermore the scale is available in translation into all major European languages and a recent comprehensive review supports the psychometric properties and also the clinical utility of the HADS (Hermmann, 1997).

6.6.2 Health Complaints Scale (HCS)

The HCS assesses both somatic and cognitive health complaints. It comprises 12 health complaints and 12 cognitive complaints (see Appendix 8). In order to make the instrument relevant to cardiac populations, the instrument was developed on the basis of

commonly rated somatic and cognitive complaints in a sample of cardiac patients. The respondent rates how bothered they have been "lately" by each health complaint on a 5-point scale of distress ranging from 0 (not at all) to 4 (extremely). Of note, the time referent used can be modified depending upon the research question. High levels of internal consistency (Cronbach's alpha > 0.80) and test-retest reliability (r >0.69) have been reported in a cardiac rehabilitation population (Denollet, 1994). In addition, the HCS shows good correlations with standard psychosocial instruments such as the SCL-90 and the STAI (Denollet, 1994), and is also brief and easy to administer.

6.6.3 Quality of Life Index- Cardiac Version III (QLI-CV III)

This instrument is a variation of a general quality of life index developed by Ferrans and Powers (1985). The questionnaire consists of 72 items, composed of two parts (see Appendix 9). Part 1 comprises 36 items measuring satisfaction with various domains of life on a six point scale (very dissatisfied to very satisfied). Part 2 comprises 36 items assessing the importance of the same domains to the subject on a six point scale (very unimportant and very important). Domains include Health and Functioning (16 items), Socio-economic (9 items), Psychological/Spiritual (7 items) and Family (4 items). To eliminate negative scores, 15 is added to each score to produce the final item scores. Each domain score is simply the sum of the relevant items in that subscale divided by the number of items in that subscale and the overall QOL is the composite of these domain scores. The domain scores and the total

quality of life score are based only on the completed items and missing data do not render the instrument unusable. A computer programme that calculates scores is available from the instrument's authors.

Internal consistency (Cronbach's alpha) for overall quality of life (0.92), health and functioning subscale (0.87), socioeconomic subscale (0.73), psychological/spiritual (0.88) and family (0.63) have been reported in cardiac rehabilitation sample (Deshotels. Planchock, Dech & Prevost, 1995). Factor analysis of the original 64-item version of the QLI in a sample of hemodialysis patients supported the authors proposed four factor solution (Ferrans & Powers, 1992). However, the QLI-CV III has not received much investigation in a cardiac sample as yet. In comparison to some other instruments, the QLI-CV III is relatively long to administer and complete.

6.7.0 Questionnaire Battery B

Questionnaire battery B comprised the SSAI, QLMI, and the CDS. Each of these instruments is described below.

6.7.1 State-Trait Anxiety Inventory Form X-1 (SSAI)

The SSAI inventory is a commonly used instrument that reflects the degree of current anxiety experienced (Spielberger et al., 1970). It comprises 20 items, each of which have to be answered on a four-point scale (score range 20-80) (see Appendix 10). Higher scores

indicate the higher levels of anxiety. Ten of the items are reverse scored. The total score is the sum of the 20 responses and ranges from 20 to 80. Reliability of the SSAI was tested on undergraduates. Measures of internal consistency coefficients range from 0.83-0.92. It has been adapted into a number of different cultures. The scale is brief and easy to both complete and administer.

6.7.2 MacNew Quality of Life after Acute Myocardial Infarction Questionnaire (MacNew QLMI)

The MacNew QLMI is a self-report modification of the original interviewer-administered questionnaire (Valenti et al., 1996). The 27 items address three quality of life domains: emotional (14-items), physical (14-items) and social (13-items) (see Appendix 11). Some items contribute to the total score of more than one domain and one item contributes to all three domains. Response options are rated on a 7-point scale with higher scores indicating better quality of life. The domain score are the average of the summed domain items and thus, missing data does not contribute to the final score. A global quality of life score can be estimated as the average of all the items. In the initial report using this version of the instrument, high Cronbach's alphas were reported for the emotional (0.95), physical (0.93) and social (0.95) subscales (Valenti et al., 1996). Factor analysis revealed a three factor structure with some items loading onto more than one factor. The QLMI is relatively easy to administer and to complete.

6.7.3 Cardiac Depression Scale (CDS)

The CDS was developed by Hare and Davis (1996) in response to a number of limitations in the then existing measures of depression. Available instruments are not always suitable for cardiac patients because they have been validated in other populations, primarily in more severely depressed psychiatric patients. The CDS was derived from cardiac patients and is self-rating of the level of depression in the range commonly found in cardiac populations. The original questionnaire comprised 35 items with positive and negative items on a scale from 1 to 7 (see Appendix 12). Item order involved mixing the 24 depression items with the 11 positive items to avoid a response set.

Of note, the scale includes somatic symptoms of depression as the authors believed that the inclusion of such items makes the instrument more sensitive for assessing moderate or severe depression in cardiac patients. The CDS comprises 26 items, with 6 items being reverse scored. The total depression score is given by the sum of the individual items. The validation studies of the instrument revealed that it correlates well with the BDI (r=0.73) and with clinical assessment (r=0.67) of depression (Hare & Davis, 1996). Cronbach's alpha reliability for the final 26 items is 0.90. To date the scale has not been widely used within the cardiac literature. The scale is relatively easy to administer but the negatively phrased items may be confusing to some respondents.

6.8.0 Questionnaire Battery C

Questionnaire battery C comprised the GMS, HPPQ, and SF-36. Each of these instruments is described below.

6.8.1 Global Mood Scale (GMS)

The GMS comprises 10 negative and 10 positive mood terms (see Appendix 13). This scale is based on the two-factor model of mood and was developed with a medical population in mind. The respondent is asked to rate on a 5-point scale the extent each mood state has been experienced lately, with response options ranging from 0 (not at all) to 4 (extremely). The scale was developed for use in a cardiac population. High levels of internal consistency (Cronbach's alpha > 0.90) and test-retest reliability (r >0.57) have been found in a cardiac rehabilitation population (Denollet, 1993a). Furthermore, it shows good correlations with a range of subscales from the POMS, the HPPQ and the STAI (Denollet, 1993a). The ability to reliably assess positive affect may be a promising feature of the GMS as many instruments do not assess positive mood but rather focus on negative affect. Although other scales that measure negative and positive affect already exist, the advantage of the GMS is that it was designed for use in cardiac patients. Furthermore, in comparison to many commonly available instruments, the GMS is short and easy to complete and administer.

6.8.2 Heart Patients Psychological Questionnaire (HPPQ)

The HPPQ assesses four psychological variables: Well-being (12 items), Feelings of disability (12 items), despondency (10 items), and social inhibition (6 items) (see Appendix 14). In addition, the instrument also contains 12 dummy items that are not used in subscale calculations. The response option for each of the 52 items is correct/unsure/incorrect. Six items are reverse scored and missing items are scored as 2. The test was validated in the Netherlands on a large (1,649) reference group of cardiac patients and factor analysis revealed the presence of four factors (Erdman et al., 1986). Internal consistency for the four subscales were reported as well-being (0.93), feelings of disability (0.87), despondency (0.80) and social inhibition (0.64). The four subscales reported the following test retest reliabilities over 1 to 2 weeks: well being (0.73), feelings of disability (0.85), despondency (0.80) and social inhibition (0.77). The scale takes approximately 10 minutes to complete (Visser et al., 1995) and administration is relatively straightforward.

6.8.3 MOS 36-item short-form health survey (MOS SF-36)

The MOS SF-36 is one of the most widely used self-reported health instruments in outcomes research (Ware & Sherbourne, 1992). It comprises 36 items which assesses 8 health dimensions: Physical functioning (10 items), bodily pain (2), role limitations due to physical problems (4), role limitations due to personal or emotional problems (3), social functioning (2), energy/fatigue (4), and general health perceptions (5). An additional item, addressing

health change, is recorded but not scored. For each dimension, item scores are coded, summed and transformed to produce a score ranging from 0 (worst quality of life) to 100 (best quality of life). In the original MOS, the Cronbach's alpha for the instrument were reported as physical functioning (0.93), role limitations/physical (0.84), role limitations/emotional, energy/fatigue (0.86), emotional well-being (0.90), social functioning (0.85), pain (0.78) and general health (0.78). The psychometric properties of the SF-36 have been extensively tested in a variety of healthy and illness populations in a number of cultures. In comparison to a number of other scales, the SF-36 is more complicated in terms of coding and scoring.

6.9.0 Global Rating of Change Questionnaire

This instrument was developed to assess the minimal important difference (MID) in change on a quality of life instrument (Juniper *et al.*, 1994). The global rating questionnaire asks respondents to indicate on a 15-point scale (see appendix 16) if there was any change in their overall quality of life since the last assessment. Following the authors' instructions respondents who scored –1, 0, 1 were categorised as unchanged, those scoring –3, -2, 2, 3 were classified as experiencing minimal important difference, scores of –5, -4, 4, 5 indicated moderate change and scores of –7, -6, 6, 7 were considered to have experienced large change. In order to assess the magnitude of change in patients classified as stable, the sign of the scores for those who deteriorated was changed, the scores were summed and the average of these scores was calculated. Similarly, in order to estimate the degree of change reported by

those classified as experiencing MID, the change in scores was combined by changing the sign of the scores for those who deteriorated.

Table 6.4. Summary of psychosocial instruments

Instrument *subscales*	**Number of items**	**Response format**	**Scoring**	**Time frame**	**Minimum score (meaning)**	**Max score (meaning)**
Health Complaints Checklist *Somatic complaints* *Cognitive complaints*	24 *12* *12*	5 point scale 0 (not at all) – 4 (extremely)	Sum items	Lately (but can be modified)	0 (no complaints)	48 (extreme complaints)
Hospital Anxiety & Depression Scale *Anxiety* *Depression*	14 *7* *7*	4 point scale 0 –3	Sum items	Past week	0 (no anxiety) 0 (no depression)	21 (clinically severe anxiety) 21 (clinically severe depression)
Quality of Life Index Cardiac III version *Health & functioning* *Socio-economic* *Psychological/Spiritual* *Family*	72 *32* *18* *14* *8*	6 point scale 1 (very unsatisfied) – 6 (very satisfied) 6 point scale 1 (very unimportant) – 6 (very important)	Transform satisfaction score, multiply by importance score and add 15, Sum items	Unspecified (presumably at present)	0 (low QoL) 0 (low QoL) 0 (low QoL) 0 (low QoL)	30 (high QoL) 30 (high QoL) 30 (high QoL) 30 (high QoL)
Cardiac Depression Scale	26	7 point scale 1 (strongly disagree) 7 (Strongly agree)	7 items reverse scored Sum items	At present	26 (no depression)	182 (high depression)

Spielberger State Anxiety Inventory	20	4 point scale 1 (not at all) 4 (very much)	10 items reverse scored Sum items	At present	20 (no anxiety)	80 (high anxiety)
Mac New Quality of Life after Myocardial Infarction	27	7 point scale	Sum items	During past two weeks		
Emotional	*14*	1 -7			1 (low QoL)	7 (high QoL)
Physical	*14*				1 (low QoL)	7 (high QoL)
Social	*13*				1 (low QoL)	7 (high QoL)
Global Mood Scale	20	5 point scale				
Negative affect	*10*	0 (not at all) 4 (extremely)	Sum items	Lately	0 (low negative mood)	40 (high negative mood)
Positive Affect	*10*				0 (low positive mood)	40 (high positive mood)
Heart Patients Psychological Questionnaire	52	3 point scale	4 items reversed Sum items	Nowadays		
Well being (W)	*12*	3 (correct)			12 (low well being)	36 (high well being)
Feelings of Disability (F)	*12*	2 (?)			12 (low disability)	36 (high disability)
Despondency (D)	*10*	1 (incorrect)			10 (low despondency)	30 (high despondency)
Social inhibition (S)	*6*				6 (low social inhibition)	18 (high social inhibition)
	12 dummy items					

MOS SF-36	36	Varies from	Recode all	Varies:		
Physical functioning	*10*	2 point	items	now,	0 (poor physical	100 (good physical
		Yes/No	Sum items	compared with	functioning)	functioning)
Role functioning	*4*	to 6 point		one year ago,	0 (role limitations due to	100 (no role limitations due
/physical		All of the		in general,	physical problems)	to physical problems)
Role functioning		time to		past 4 weeks	0 (role limitations due to	100 (no role limitations due
/emotional	*3*	None of the			emotional problems)	to emotional problems)
Energy / fatigue	*5*	time			0 (low energy)	100 (high energy)
Emotional well-being	*2*				0 (low emotional well-being)	100 (high emotional well-being)
Social functioning	*2*				0 (low social functioning)	100 (high social functioning)
Pain	*5*				0 (high pain)	100 (low pain)
General health	*1*				0 (poor general health)	100 (good general health)

6.10 Responsiveness Statistics

The present study examines the responsiveness of the selected instruments using 5 different ways of estimating responsiveness. These statistical approaches to responsiveness have been previously used within the literature and each formula is calculated according to the author's specification.

Effect size statistic (Kazis et al., 1989)

$$\frac{\text{Mean time 2} - \text{Mean time 1}}{\text{Standard deviation time1}}$$

Standardised Response Mean (SRM) (Liang et al., 1990)

$$\frac{\text{Mean time 2} - \text{Mean time 1}}{\text{Standard deviation of change scores}}$$

Guyatt Responsiveness Index (Guyatt et al., 1987)

$$\frac{\text{Minimum important difference}}{\sqrt{2 \times \text{Mean Square Error of stable patients}}}$$

The degree of change in the instrument that constitutes the MID is estimated from those classified as MID according to Juniper's method. The mean square error is calculated from a repeated measures ANOVA of those participants reported as stable.

Tuley Responsiveness Index (Tuley et al., 1994)

$$\frac{\text{Mean time 2} - \text{Mean time 1}}{\sqrt{2 \times \text{Mean Square Error of stable patients}}}$$

Relative efficiency (Liang *et al.*, 1985)

$$\left[\frac{\text{T statistic for instrument}}{\text{T-statistic for instrument with largest t-statistic}}\right]^2$$

All of the methods described above provide an estimate of the change recorded by an instrument.

6.11 Statistical Analyses

All statistical analyses were performed using DATA DESK 6 (Data description, Inc, Ithaca, NY) and InStat 2.01 (GraphPad software). Baseline variable descriptive statistics were analysed using t-tests, ANOVAS and MANOVAS where appropriate for interval-ratio data and chi-square tests were used for nominal-scale variables. Responsiveness was estimated using the formulas described above.

Chapter 7
Results

7.0.0 Abbreviations

In the present chapter the following abbreviations are used for each scale:

HCS cognitive complaints	= HCS cog
HCS somatic complaints	= HCS som
HADS depression	= HADS dep
HADS anxiety	= HADS anx
QLI health & functioning	= QLI health
QLI socio-economic	= QLI socio
QLI psychological	= QLI psy
QLI family	= QLI fam
QLI total	= QLI total
Cardiac Depression Scale	= CDS
Spielberger State Anxiety Inventory	= SSAI
QLMI emotional	= QLMI e
QLMI physical	= QLMI p
QLMI social	= QLMI s
QLMI total	= QLMI t
GMS negative affect	= GMS na
GMS positive affect	= GMS pa
HPPQ well-being	= HPPQ w
HPPQ feelings of disability	= HPPQ f
HPPQ despondency	= HPPQ d
HPPQ social inhibition	= HPPQ s
SF-36 physical functioning	= SF pf
SF-36 role limitations due to physical problems	= SF rlp
SF-36 role limitations due to emotional problems	= SF rle
SF-36 energy/fatigue	= SF e
SF-36 emotional well-being	= SF ewb
SF-36 social well-being	= SF swb
SF-36 pain	= SF p
SF36 general health	= SF gh

7.1.0 Response rate

The study sample comprised participants from seven different centres and the response rate of each of the centres was examined using chi-square analysis (Appendix 17). Analysis of the data revealed no significant association between response rate over time and study centre (χ^2 = 1.85, 6 df, p = 0.93). The overall response rates of the cardiac rehabilitation and the comparison groups are given in Table 7.1.

Table 7.1. Response rates of the cardiac rehabilitation and the comparison group at time 1 and time 2.

Group	Total administered at Time 1	Time 1 N (% completion)	Time 2 N (% completion)
Cardiac rehabilitation	675	465 (69%)	411 (88%)
Comparison group	810	446 (55%)	375 (84%)

Chi square analyses revealed no significant difference in the overall response rates between groups (χ^2 = 0.26, 1 df, p = 0.61). The cardiac rehabilitation group response rate was slightly higher than anticipated rate of 60% at time 2, whereas the control group was slightly below the expected rate of 50% at time 2. The response rate of each questionnaire at each study centre is provided in Appendix 18. Chi-square analysis showed no significant overall association between the response rates of the questionnaires and the different study centres (χ^2 = 3.66, 20 df, p = 0.99). To compare the association between

each centre and each questionnaire individually, a further series of chi-square analyses were conducted. No significant associations were reported for the white ($\chi^2 = 0.62$, 6 df, p = 0.99), yellow ($\chi^2 = 2.05$, 6 df, p = 0.91) or blue ($\chi^2 = 0.91$, 6 df, p = 0.99) questionnaires across centres.

7.2.0 Comparison of baseline demographic variables for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires

To investigate if there were any differences on the baseline demographic variables between those who completed both questionnaires compared with those who completed only time 1 questionnaires, a series of chi-squares analyses were performed for categorical data and t-tests were performed for continuous data. The results of this analysis are provided in Appendix 19.

Chi square analysis revealed a significant association between education levels and number of questionnaires completed ($\chi^2 = 7.85$, 3 df, p = 0.049). To determine which categories contributed to this significant relationship, the standardised residuals of each of the categories was estimated (Hinkle, Wiersma & Jurs, 1998). Examination of the residuals showed that the main contributor to the significant relationship was the missing data category (standardised residual = 2.35). There were no other significant differences on any of baseline demographic variables. The main analyses focus only on those participants who completed both questionnaires.

7.2.1 Comparison of baseline psychosocial variables for those completing both questionnaires with those completing only time 1.

In order to examine if there were significant differences between the participants that completed time 1 only in comparison with those who completed both time 1 and time 2 questionnaires, a series of MANOVAS were performed. For each MANOVA, the dependent variables were the psychosocial instruments and the number of questionnaires completed was the between group factor. No significant differences were reported for the white (WILKS $\lambda = 0.93$, $F = 1.72$, $p = 0.09$), yellow (WILKS $\lambda = 0.95$, $F = 0.26$, $p = 0.95$) or blue (WILKS $\lambda = 0.92$, $F = 1.38$, $p = 0.17$) questionnaire.

7.3.0 Demographic details of the study sample

Demographic characteristics of all 786 patients included in the main analyses are presented in Table 7.2.

Table 7.2. Demographic and clinical profile of study sample (n = 786)

Demographic and Clinical characteristics	**Cardiac rehabilitation (N= 411)**	**Comparison group (N= 375)**	**P-value**
Gender	N (%)	N (%)	$\chi^2 = 0.01$, 1df,
Male	310 (75)	284 (76)	p = 0.92
Female	101 (25)	91 (24)	
Age (yrs)	Mean (SD)	Mean (SD)	t = -1.34, 781df,
	62.1 (9.8)	62.7 (10.8)	p=0.18
	Median (range)	Median (range)	
	63 (35-83)	64 (33-89)	
Marital Status	%	%	
Married	83	75	
Single	7	10	$\chi^2 = 7.59$, 3df,
Widowed	9	12	p = 0.06
Other	1	3	
Occupational status	%	%	
Full-time	37	29	
Part-time	6	11	
Retired	44	41	$\chi^2 = 21.58$, 5df,
Unemployed	2	5	p <0.01
Homemaker	6	12	
Other/missing	4	3	
Education	%	%	
Primary	36	47	
2nd level	43	33	$\chi^2 = 12.32$, 3df,
Third level	17	14	p <0.01
Missing	4	5	

Residential location			
City	58	27	
Town	15	30	$\chi^2 = 101.17$, 3 df,
Village	14	9	p = 0.00
Rural	14	34	
Living arrangement			
Alone	8	15	$\chi^2 = 98.73$, 2df,
With others	91	84	p = 0.01
Missing	1	1	
Time (weeks)	Mean (SD)	Mean (SD)	z = -1.65, 757df,
Weeks since the most	18.1 (10.54)	19.13 (6.6)	p = 0.09
recent cardiac event:	Median (range)	Median (range)	
	14 (5-260)	17 (8-52)	
Cardiac event			
MI	19	16	
CABG	51	47	$\chi^2 = 28.63$, 3df,
PTCA	15	30	p <0.01
Other/missing	15	7	
Cardiac History			
None	51	58	
MI	18	18	
CABG	9	7	$\chi^2 = 5.67$, 4df,
PTCA	4	2	p = 0.23
Other/missing	17	14	
Co-morbidity			
No	29	32	χ^2 =0.86, 1df,
Yes	71	68	p = 0.35
Physical function			t-Statistic = -1.17, 745 df
DASI	35.3(12.4)	36.4 (13.7)	p = 0.2422

Significant associations were reported between study group and occupational status, education, residential location, living situation, cardiac event. Post-hoc analyses of these associations assessed which categories contributed to these significant relationships. For occupational status and educational level, none of the categories had a standardised residual greater than 2 and thus the contributing categories cannot be determined. For residential location, only the village category in the comparison group did not have a standardised residual greater than 2. Thus, nearly all of the categories contributed to the significant difference between the cardiac rehabilitation and comparison groups. The significant contributor to the living situation was the living alone category for the comparison group. With respect to the cardiac event, the levels of PTCA and other/missing categories for both groups contributed to the significant association.

In summary, the comparison group were less likely to live in a city and more likely to live in a town or rural area, and were also less likely to live with others and more likely to have had PTCA as the most recent cardiac event. Given these differences, it was decided that the effect of each centre on the overall results would be examined (see section 7.5.0).

7.4.0 Responsiveness of Instruments

The responsiveness statistics for each instrument subscale were calculated according to the formulae presented in section 6.10 for the Effect size (ES) statistics,

Standardised Response Mean (SRM), Relative Efficiency (RE), Tulkey Responsiveness index (Tuley index), and Guyatt Responsiveness index (Guyatt index). The difference between the responsiveness statistics for cardiac rehabilitation and comparison group for each instrument was calculated (Appendix 20). The rank order of each instrument in terms of responsiveness was determined for each responsiveness statistic (Table 7.3) and instruments were ranked from highest (1) to lowest (29).

Table 7.3. Ranking of responsiveness of each instrument according to each statistic

Subscale	ES	SRM	RE	Tuley index	Guyatt index
HCS cog	10	10	11	8	15.5
HCS som	13	12	13	12	1
HAD dep	11	9	9	7	7
HAD anx	19	19	20	20.5	2
QLI health	16.5	18	19	16	11
QLI socio	14	16	17	14	8
QLI psy	18	25	27	17	4
QLI fam	27.5	26.5	27	28	5
QLI total	20	28	25	18.5	9
CDS	21	17	18	20.5	28
SSAI	24.5	24	24	29	20.5
QLMI e	29	29	15	27	15.5
QLMI p	22	20.5	21	24	24
QLMI s	24.5	22	22	24	23
QLMI t	24.5	23	27	24	19
GMS pa	2	1	1	1	12
GMS na	5	3	3	5	25
HPPQ s	27.5	20.5	23	22.5	18
HPPQ d	5	8	6	2	22
HPPQ w	24.5	26.5	29	22.5	17
HPPQ f	7	7	7	6	10
SF pf	16.5	14	14	18.5	27
SF rlp	1	2	2	4	14
SF rle	3	4	12	3	26
SF e	12	13	8	13	3
SF ewb	8	11	10	11	29
SF swb	5	5	4	9	6
SF p	9	6	5	10	20.5
SF gh	15	15	16	15	13

7.4.1 Correlation of instrument rankings by responsiveness statistics

The ranking of each instrument by the different responsiveness indices was given in Table 7.3. In order to examine the similarity of these rankings, a Spearman rank correlation matrix was examined (Table 7.4).

Table 7.4. Spearman rank correlation of instruments by responsiveness index

	SRM	ES	Tuely index	RE	Guyatt index
SRM	1.00				
ES	0.94	1.00			
Tuley index	0.92	0.97	1.00		
RE	0.92	0.87	0.85	1.00	
Guyatt index	-0.10	0.01	0.03	-0.06	1.00

The results indicate that the rankings of responsiveness provided by the SRM, ES, Tuley index and RE are highly correlated and provide approximately the same results. However, the Guyatt Responsiveness index has very low correlations with the rankings produced by the other methods and the ranking of instruments it provides is not related to the ranking provided by the others. Thus, it was decided to base the overall ranking of responsiveness of instruments on the ES, SRM, Tuley index and the RE statistics.

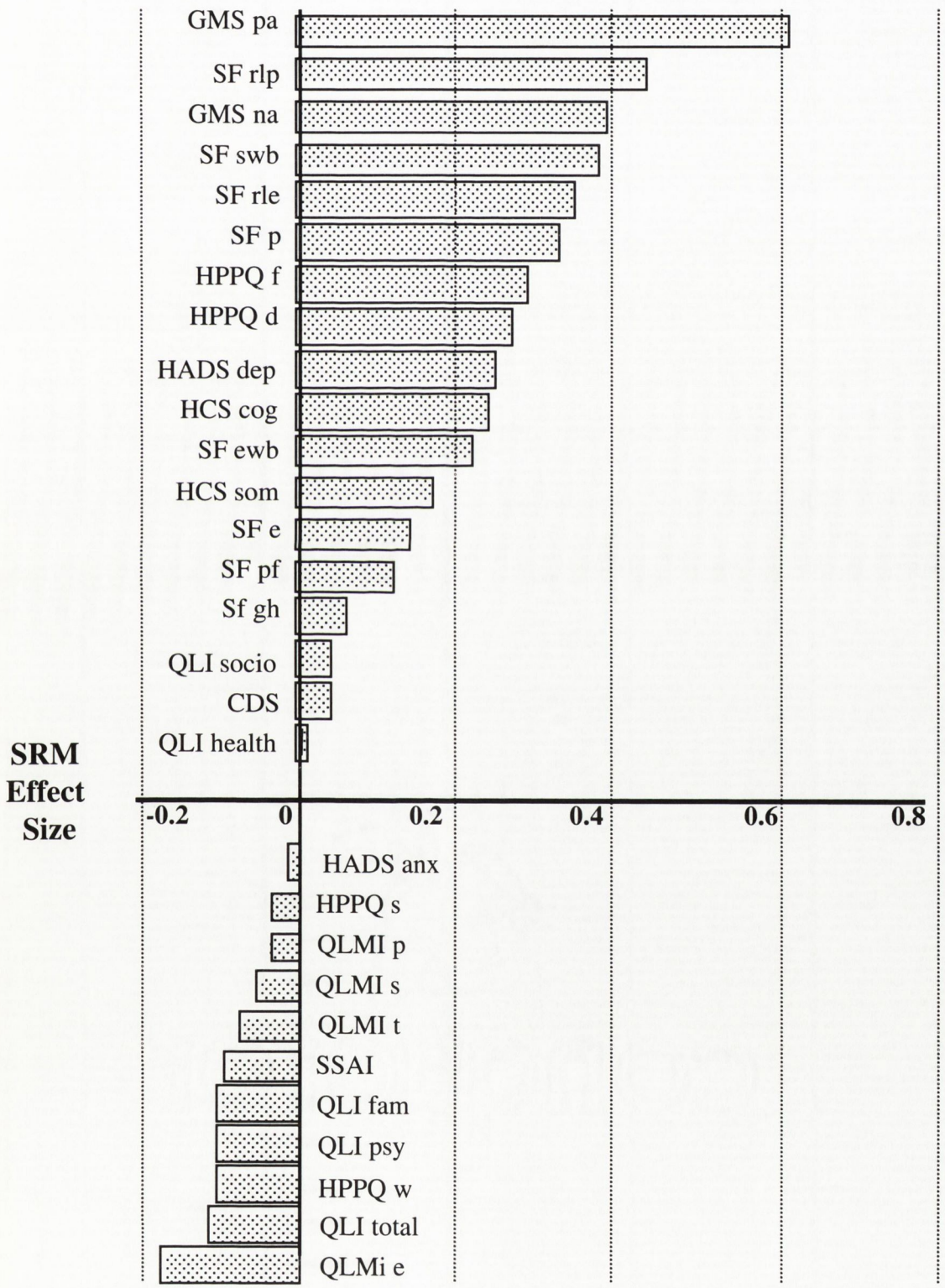
7.4.2 Overall responsiveness

In order to determine the most responsive instrument, the ranking of each instrument as given by the ES, SRM, Tuley index, and RE was summed. Instruments were ranked according to the total sum, with lower totals indicating higher responsiveness (Table 7.5). Of note, the GMS pa was the most responsive subscale. A graph depicting the SRMs for each instrument clearly illustrates the variability in responsiveness of instruments (Figure 7.1)

Table 7.5. Overall subscale ranking based on the ES, SRM, Tuley index and RE statistics

Rank	Instrument subscale
1	GMS pa
2	SF rlp
3	GMS na
4	HPPQ d
5	SF rle
6	SF swb
7	HPPQ f
8	SF p
9	HADS dep
10	HCS cog
11	SF ewb
12	SF e
13	HCS som
14	QLI socio
15	SF gh
16	SF pf
17	QLI health
18	CDS
19	HADS anx
20	QLI psy
21	QLMI p
22	QLI total
23	QLMI s
24	HPPQ s
25	QLMI t
26	QLMI e
27	SSAI
28	HPPQ w
29	QLI fam

Figure 7.1 SRMs for instrument subscales



7.4.3 **Statistical significance of instrument changes**

To assess the statistical significance of the changes reported on the instruments by the cardiac rehabilitation and the comparison groups, a series of 2-way repeated measures ANOVAs were performed for each scale (Table 7.6).

Table 7.6. Two-way repeated measures ANOVA for each scale

Subscale	F-value for Interaction	P-value
HCS cog	5.91	0.016
HCS som	2.56	0.110
HAD dep	4.91	0.027
HAD anx	0.00	0.979
QLI health	0.17	0.676
QLI socio	2.47	0.117
QLI psy	0.02	0.886
QLI fam	1.14	0.286
QLI total	3.70	0.055
CDS	2.26	0.134
SSAI	0.60	0.439
QLMI e	0.84	0.359
QLMI p	1.96	0.163
QLMI s	0.75	0.385
QLMI t	0.97	0.326
GMS pa	43.26	≤0.0001
GMS na	16.25	≤0.0001
HPPQ s	5.89	0.016
HPPQ d	7.19	0.008
HPPQ w	0.00	0.988
HPPQ f	8.98	0.002
SF pf	0.01	0.945
SF rlp	22.52	≤0.0001
SF rle	30.26	≤0.0001
SF e	1.64	0.21
SF ewb	6.33	0.012
SF swb	8.31	0.004
SF p	7.56	0.006
SF gh	0.59	0.443

The data presented in Table 7.6 illustrate the variability in outcome as assessed by each subscale. Some instruments showed highly significant differences between the cardiac

rehabilitation and comparison groups (e.g., GMS pa) whereas other instruments (e.g., HPPQw) showed no differences between the groups.

7.5.0 Effect of excluding each centre from the overall analysis

As a number of demographic differences were reported between the different centres, the impact of each centre on the overall results was assessed. Each centre was individually removed from the analysis and the subsequent responsiveness rankings of the instruments using the remaining 6 centres was compared using Spearman rank order correlations with the results when all 7 centres were included. This process was performed for each of the five responsiveness statistics (Table 7.7). The Guyatt index was included to examine if its lack of association with the other responsiveness statistics may be attributable to the influence of a particular centre.

Table 7.7. Correlation between each responsiveness statistic based on all 7 centres and the results obtained after excluding each centre individually from the analysis

Responsiveness statistic	Correlation with overall ranking when each centre excluded						
	Centre 1 N=276	Centre 2 n=53	Centre 3 N=43	Centre 4 N=39	Centre 5 N=293	Centre 6 N=66	Centre 7 N=16
ES	0.82	0.72	0.75	0.61	0.66	0.87	0.86
SRM	0.83	0.88	0.91	0.86	0.66	0.88	0.87
Tuley index	0.84	0.88	0.94	0.79	0.79	0.92	0.88
RE	0.88	0.79	0.91	0.76	0.84	0.89	0.88
Guyatt index	0.74	0.72	0.95	0.95	0.39	0.82	0.92

Analysis of Table 7.7 showed that all the reported correlations, other the correlation between the overall ranking and the Guyatt index when centre 5 was excluded, were statistically significant ($p<0.001$). The lack of association between the Guyatt index and the overall rankings when centre 5 is excluded can be accounted for by the fact that following the removal of this centre, the MID is consequently based on the responses of only two respondents. Thus, in general these results suggest that the one centre does not seem to be overly influencing the overall results.

7.6.0 Classification of the minimum important difference (MID)

Instrument ratings provided by Guyatt index were unrelated to the ratings provided by the other 4 responsiveness statistics. As this index requires estimation of the (MID) and the MSE for stable participants, the present study used the global transition scale to classify participants as stable, minimum important difference, moderate change and large change. It may be that classification of participants according to this method was not appropriate. Based on the transition scale, a linear relationship between classification of change (e.g., stable, minimum, moderate and large change) and the degree of change reported on each instrument was predicted. To investigate this hypothesis, a mean change for each of the categories was estimated (Table 7.8).

Visual inspection of Table 7.8 reveals that the predicted linear trend only occurs in the 3 scales (the HCS som, SF gh, GMSna). Thus approximately 90% of scales failed to follow the expected response pattern. This suggests that the use of the Juniper scale in the

present study failed to accurately identify those subjects to be classified as reporting a minimum important difference.

Table 7.8. Mean change in psychosocial instrument scores over time for levels of self-reported change in overall quality of life

Instrument Subscale	**No change** Mean (SD)	**MID** Mean (SD)	**Moderate change** Mean (SD)	**Large change** Mean (SD)
HCS cog	2.04 (1.9)	2.74 (2.87)	1.68 (1.83)	2.20 (2.28)
HCS som	1.67 (2.21)	3.48 (3.30)	3.57 (5.11)	3.63 (5.34)
HAD dep	2.78 (2.72)	4.51 (5.75)	2.97 (3.11)	2.67 (2.42)
HAD anx	2.05 (1.96)	3.06 (2.35)	3.24 (4.02)	3.01 (2.99)
QLI health	2.77 (3.08)	3.92 (3.40)	2.29 (2.42)	3.27 (3.56)
QLI socio	1.22 (1.28)	2.23 (1.88)	1.77 (1.81)	2.29 (2.32)
QLI psy	1.53 (1.58)	3.18 (1.37)	2.25 (2.31)	2.16 (2.11)
QLI fam	4.90 (5.12)	8.21 (7.75)	8.05 (9.07)	7.02 (6.56)
QLI total	3.78 (3.65)	6.56 (3.65)	4.16 (3.87)	5.00 (5.41)
CDS	15.31 (17.33)	8.04 (6.46)	11.98 (11.71)	10.30 (9.96)
SSAI	5.52 (6.35)	5.45 (4.93)	6.92 (5.68)	5.43 (5.64)
QLMI e	0.55 (0.59)	0.75 (0.58)	0.70 (0.71)	0.47 (0.49)
QLMI p	0.49 (0.55)	0.45 (0.39)	0.52 (0.47)	0.44 (0.43)
QLMI s	0.63 (0.74)	0.64 (0.49)	0.67 (0.61)	0.42 (0.35)
QLMI t	0.45 (0.59)	0.53 (0.46)	0.51 (0.54)	0.35 (0.30)
GMS pa	3.75 (4.09)	3.17 (3.33)	4.27 (4.12)	4.30 (5.25)
GMS na	4.48 (3.42)	5.3 (2.49)	5.72 (5.51)	7.42 (5.31)
HPPQ s	2.98 (2.00)	3.54 (2.90)	3.07 (2.07)	3.80 (3.82)
HPPQ d	3.79 (3.22)	3.88 (3.05)	3.89 (3.76)	4.23 (4.66)
HPPQ w	2.30 (1.71)	1.88 (1.63)	2.74 (2.04)	3.98 (4.25)
HPPQ f	2.16 (1.88)	2.23 (1.77)	1.18 (1.18)	1.49 (1.45)
SF pf	7.70 (8.42)	6.92 (6.01)	11.91 (7.84)	11.90 (15.95)
SF rlp	22.28 (31.52)	35.58 (33.30)	21.77 (21.53)	38.62 (35.09)
SF rle	14.32 (26.46)	17.25 (32.05)	33.41 (34.56)	26.49 (37.16)
SF e	13.73 (11.11)	23.42 (16.56)	13.38 (12.83)	14.29 (15.29)
SF ewb	11.67 (14.92)	15.58 (14.16)	14.44 (13.79)	24.14 (78.70)

SF swb	14.02 (17.10)	26.63 (20.64)	12.38 (12.50)	12.50 (22.06)
SF p	13.64 (16.79)	14.61 (16.91)	12.24 (12.08)	13.26(18.29)
SF gh	11.55 (9.25)	13.46 (12.39)	13.74 (13.91)	16.60 (16.34)

7.7.0 Comparison of parametric and non-parametric versions of the ES

One of the advantages of the ES statistic is that it can be modified for use if the scale's distribution is not approximately normal. Kazis et al., (1989) recommend using the difference between the medians as the numerator and the interquartile range of the baseline scores as the denominator. As yet, non-parametric versions of the other responsiveness statistics have not been proposed in the literature. In order to examine the comparability of the parametric and non-parametric versions of the ES, the medians and interquartile ranges for all instruments at each time point were calculated. The non-parametric ES and parametric ES statistics and instrument rankings based on these statistics are given below (Table 7.9).

Table 7.9. Responsiveness statistics using non-parametric versions of effect size and subsequent instrument ranking

Instrument Subscale	Non-parametric ES statistic	Parametric ES statistic	Non-parametric ranking	Parametric ranking
HCS cog	0.15	0.18	9	10
HCS som	0.03	0.14	14	13
HAD dep	0.20	0.17	8	11
HAD anx	0.00	0.00	17	19
QLI health	-0.06	0.03	21	16.5
QLI socio	-0.1	0.07	24	14
QLI psy	-0.17	0.01	26	18
QLI fam	0.00	-0.07	17	27.5
QLI total	-0.17	-0.01	26	20
CDS	-0.02	-0.02	20	21
SSAI	-0.01	-0.06	29	24.5
QLMI e	-0.18	-0.11	28	29
QLMI p	-0.09	-0.03	23	22
QLMI s	0.02	-0.06	15	24.5
QLMI t	0.06	-0.06	11	24.5
GMS pa	0.58	0.59	3	2
GMS na	0.33	0.30	6	5
HPPQ s	0.00	-0.07	17	27.5
HPPQ d	0.36	0.30	5	5
HPPQ w	-0.07	-0.06	22	24.5
HPPQ f	0.07	0.26	10	7
SF pf	-0.01	0.03	19	16.5
SF rlp	1.7	0.64	1	1
SF rle	0.67	0.36	2	3
SF e	0.24	0.15	7	12
SF ewb	0.04	0.24	13	8
SF swb	-0.17	0.30	26	5
SF p	0.39	0.21	4	9
SF gh	0.04	0.06	12	15

The correlation between the parametric and non-parametric values of the ES statistic was examined using Pearson product-moment correlation. A statistically significant correlation co-efficient of 0.60 (27df, $p \leq 0.01$) was found between the two versions of the ES statistics. The similarity of instrument rankings according to the parametric and non-parametric responsiveness statistics was analysed using the Spearman rank order correlation. A statistically significant correlation co-efficient of 0.68 (27df, $p \leq 0.01$) was found between the two versions of the ES statistics. Thus there is some association between the parametric and non-parametric rankings of instruments.

7.8.0 Floor & Ceiling Effects

The presence of floor and ceiling effects on each scale at time 1 was assessed. Floor effects were defined as scoring at the lowest possible level on the scale and ceiling effects were defined as scoring at the highest possible score. The percentage of participants scoring at the floor and ceiling of each scale at time 1 was calculated (Table 7.10). Of note, the SF –36 scales are affected by the presence of both floor and ceiling effects. The responsiveness of an instrument may also be constrained by the fact that if a large number of respondents score near the extreme ends of an instrument, then the possible change score is lower than for those who score near the mid-point in the possible scale score. Therefore, in addition, the percentage of respondents scoring in the lowest 10% and highest 10% range of scores possible for each instrument was also estimated (Appendix 21).

Table 7.10. Percentage of respondents scoring at the floor and ceiling of each instrument at time 1

	Cardiac rehabilitation		Comparison group	
Instrument	**% at floor**	**% at ceiling**	**% at floor**	**% at ceiling**
HCS som	5.5	0	6.5	0
HCS cog	3.9	2.3	8.9	0
HADS anx	6.3	0	11.4	0
HADS dep	8.6	0	8.1	0
QLI health	0	3.9	0	1.6
QLI socio	0	10.2	0	14.6
QLI psy	0	12.5	0	16.3
QLI fam	0	39.8	0	35.8
QLI total	0	0	0	0
CDS	0	0	0.8	0
SSAI	12.1	0	10	0
QLMI e	0	0.7	0	0
QLMI p	0	3.4	0	0
QLMI s	0	3.4	0	0.8
QLMI t	0	0	0	0
GMS na	5.2	0	11.4	0
GMS pa	0.7	0	0	2.3
HPPQ s	3.7	2.2	0	4.5
HPPQ d	0	0	13.6	0
HPPQ w	1.4	21.7	3	18.2
HPPQ f	2.2	0	1.5	0
SF pf	0	1.5	0	7.6
SF rlp	59	7.5	18.2	39.4
SF rle	33.6	31.3	22	57.6
SF e	0	2.2	0	4.5
SF ewb	0	2.2	0	10.6
SF swb	1.5	27.6	0	46.2
SF p	1.5	26.9	0	31.8

SF gh	0	6	0	4.5

In order to examine the relationship between the presence of floor and ceiling effects on responsiveness, Spearman rank correlations were calculated (Table 7.11)

Table 7.11. Relationship between responsiveness ranking and the presence of floor, ceiling, 10% floor and 10% ceiling effects in instruments

	Floor	**Ceiling**	**10% Floor**	**10% Ceiling**	**Responsiveness**
Floor	1.00				
Ceiling	0.40	1.00			
10% Floor	0.66	0.37	1.00		
10% Ceiling	0.15	0.67	-0.10	1.00	
Responsiveness	**-0.38**	-0.24	**-0.44**	0.09	1.00

Significant ($p<0.05$) correlations in bold.

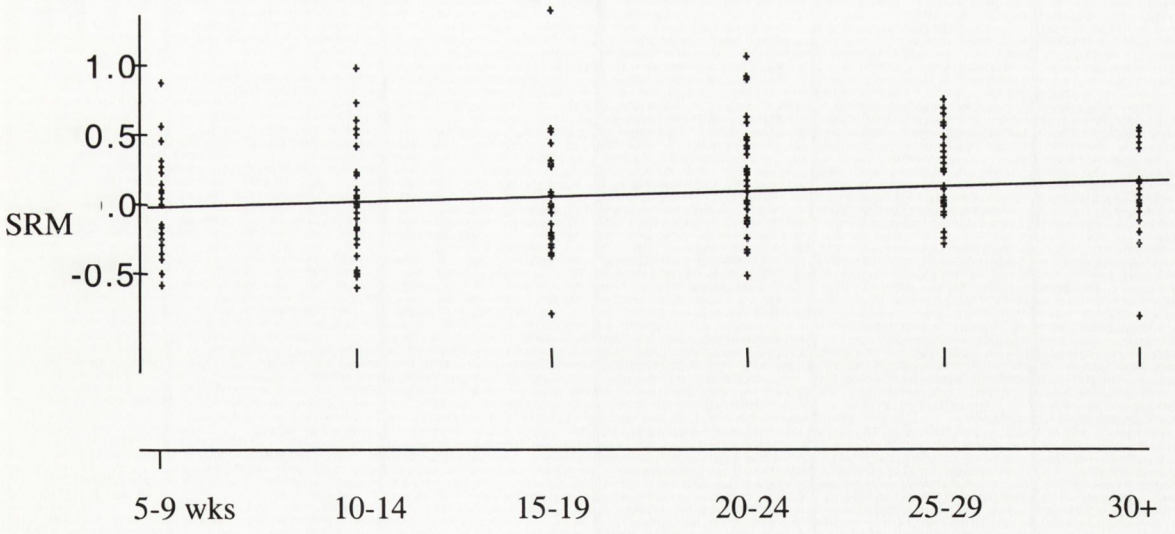
Examination of the data shows that the presence of floor effects on an instrument is related to higher levels of responsiveness (the negative correlation reflects the fact that the most responsive instrument has the lowest rank, i.e., 1). Of note the relationship is stronger for those instruments where respondents score in the lowest 10% scores at time 1. The relationship between responsiveness and both the ceiling effect for 10% ceiling is weaker than that for floor effects and is not statistically significant.

7.9.0 Relationship between time since event and SRM of each subscale

As the time since event had high variability, the relationship between responsiveness and time since event was examined. The SRM of each subscale at each time point was calculated and the time since event was categorised as 5-9 weeks, 10-14 weeks, 15-19 weeks, 20-24 weeks, 25-29 weeks and greater than 30 (30+) weeks.

Figure 7.2 illustrates the variability in responsiveness over time. The Spearman rank correlation (r=0.19) was not-significant and there is a weak relationship between time since event and responsiveness as indexed by the SRM.

Figure 7.2. Relationship between time since event and SRMs for each subscale.



7.10 Missing data

A number of instruments require respondents to complete all items in order to provide a sub-scale or an overall score. The percentage of missing data for these instruments at each time-point and an overall average rate of missing data is given in Table 7.12.

Table 7.12. Percentage of missing data for instruments at each time point and overall rate of missing data.

Instrument subscale	**Time 1 Cardiac rehabilitation** %	**Time 2 Cardiac rehabilitation** %	**Time 1 Comparison group** %	**Time 2 Comparison group** %	**Overall** %
HCS som	21	21.1	23.6	22.8	22.1
HCS cog	21.9	21.9	18.7	18.7	20.3
HADS anx	7.8	7.8	5.7	5.7	6.8
HADS dep	6.2	6.3	4.1	4.1	5.2
CDS	7.9	7.3	6.1	6.7	7
STAI	6.1	6.7	6.8	7.5	6.8
GMS pa	15.7	15.7	12.8	13	13.1
GMS na	16.4	16.4	10.7	10.7	13.6

A number of instruments (e.g., HCS som and HCS cog) have high levels of missing data and this may limit their applicability in health outcomes research. To investigate if the rates of missing data for the subscales were related to their position in the questionnaire battery, the Spearman rank correlation was estimated. A moderate correlation of 0.50 was reported and this finding suggests that there is some relationship between questionnaire order and rates of missing data, with questionnaires that came last in the battery having higher rates of missing data.

Chapter 8
Discussion

8.0.0 Overview

The aim of the present study was to investigate the responsiveness of psychosocial instruments to cardiac rehabilitation intervention. The study comprised two main components:

(i) a systematic review of the literature on psychosocial outcome after cardiac rehabilitation to identify those instruments with empirical evidence documenting their responsiveness to multifactorial cardiac rehabilitation intervention

(ii) a prospective study to examine the responsiveness of 9 psychosocial instruments to cardiac rehabilitation intervention in a sample of cardiac patients.

The responsiveness of these instruments was estimated using 5 statistics (ES, SRM, Tuley index, Guyatt index and RE) and instruments were subsequently ranked in terms of responsiveness. The similarity of rankings of instruments provided by each statistic was compared.

8.1.0 Responsiveness of instruments

The systematic review (Chapter 5) used the ES statistic to identify those instruments with empirical evidence of responsiveness. Based on the findings of the systematic review and on considerations regarding recently developed instruments, 9 instruments were evaluated in the main prospective study. In order to identify the most responsive instrument, the rankings of the instruments according to the ES, SRM, Tuley

index and RE statistics were combined to produce an overall ranking of instruments. The post hoc decision not to include the values obtained from the Guyatt Responsiveness index was made after analysis of the data (Table 7.4) showed that this index was not associated with the other responsiveness statistics. This issue will be discussed in more detail in section 8.2.0. Using the overall ranking provided by the 4 statistics, the GMS emerged as the most responsive instrument to multifactorial cardiac rehabilitation in the present study.

The GMS is a brief, easily administered, easily completed and easily scored instrument (see section 6.8.1). Thus, the GMS has many practical advantages over a number of other commonly used psychosocial instruments. Hlatky and Spertus (1998) noted that quality of life researchers often use lengthy instruments that are returned by only 20% of the sample; they suggest that it would be preferable to use a "really short and snappy" questionnaire that gets to the heart of what is being assessed and that is returned by 90% of the sample. Based on the findings of the present study, the GMS may fit such a description (with a response rate of 87%).

While the majority of psychosocial instruments assessed in the present study focus on negative psychosocial states, the GMS assesses both positive and negative mood. Indeed, the GMS positive mood subscale was the most responsive subscale identified in the present study. Typically, positive psychosocial states are not routinely assessed in cardiac rehabilitation settings. However, the present findings suggest that cardiac rehabilitation may benefit from considering both positive and negative psychosocial states to provide a more comprehensive assessment of psychosocial functioning of service users.

Furthermore, although other scales to assess negative and positive affect already exist, an advantage of the GMS is that it was designed specifically for use in cardiac patients. Thus, the items may be perceived by respondents as being particularly relevant to their cardiac condition. As Denollet (1993a) observed, many currently used outcome assessment instruments may be more accurately considered measures of psychopathological affective states (e.g., BDI) and thus many of the items on such instruments may not be relevant to the majority of cardiac patients. As the GMS was developed specifically for use with cardiac patients, the item content of the instrument may be more appropriate than that of other instruments.

One possible limitation of the instrument is the reliance on all items being answered in order to provide subscale scores. Approximately 13% of respondents failed to complete all items on the GMS and this loss of data may compromise the instrument's appropriateness. Development of methodologies that facilitate calculation of an overall score based on the completed items may prove useful. While the GMS was developed for use in cardiac populations, the psychometric properties of the instrument in other populations may warrant investigation. The item content may generalise to other conditions and the correlations reported previously (e.g., Denollet, 1993b) between the GMS and other generic instruments (e.g. POMS, STAI) suggest that such uses are worth investigating.

Given the positive orientation in the WHO (1993) definition of cardiac rehabilitation, there is a need for cardiac rehabilitation services to consider the role of

positive states. Positive and negative affect are widely accepted as core dimensions of emotional functioning (Watson & Tellegen, 1985). Negative affect comprises health complaints and mental stress whereas positive affect is associated with energy, activity and social engagement. According to Denollet (1993), negative affect tend to remain at fairly low levels. Since most cardiac patients do not suffer from psychopathology, a measure that assesses both positive and negative affect may be more appropriate. The failure of previous studies to report long-term HRQOL benefits of cardiac rehabilitation may be attributed to the use of inappropriate instruments that assessed inappropriate domains. In addition, the focus on positive mood may be more congruent with the model of cardiac rehabilitation. Indeed, given the finding that the positive affect subscale of the GMS emerged as the most responsive subscale in the present study, it may well be that cardiac rehabilitation outcomes research has adopted the wrong model of outcomes assessment as it attempts to show that cardiac rehabilitation decreases negative states but neglects the benefits it may have on positive states. As Diener (2000) observes, positive well-being is not simply the absence of ill-being and these states have different antecedents and consequences. Both positive and negative affect are major components of subjective well-being. Cardiac rehabilitation has the potential to enhance positive psychosocial states and both the psychological and physiological effects of increased levels of positive states warrants further investigation.

Analysis of the overall ranking of instruments in Table 7.5 shows that the most responsive instruments comprise both generic and disease-specific instruments. Of note,

some subscales of particular instruments were more responsive than other subscales of that instrument. For example, SF rlp was the most responsive of the eight SF-36 subscales, with the SF pf scale being least responsive. While two of the HPPQ subscales (d and f) were among the most responsive instruments, the other two scales (s and w) were among the least responsive. The low level of responsiveness of the HPPQ w scale was surprising, given that it demonstrated evidence of responsiveness in the systematic review (see Table 5.1). Similarly, the QLMI also performed poorly in relation to the findings of the systematic review (see Table 5.1). Both the HPPQ and QLMI were selected on the basis of evidence of responsiveness to cardiac rehabilitation.

Of interest, neither the SSAI, a widely used measure of anxiety, nor the HADS anxiety scale, were responsive in the present study. Furthermore, the HADS depression subscale was more responsive than the disease-specific Cardiac Depression Scale. In addition to the higher levels of responsiveness, the relative brevity of the HADS and its screening capabilities provide a strong case for its use in preference to the CDS and the SSAI as a measure of mood. Both of the HCS subscales, somatic and cognitive complaints, were ranked in the top half of instrument subscales. The poor performance of two disease-specific instruments, the QLMI and the QLI-CVIII, is noteworthy. Of the 9 subscales comprising the QLMI and the QLI CVIII, only the socio-economic subscale of the QLI CVIII was ranked in the top half of the overall responsiveness rankings. In general, caution should be exercised in the use of the QLMI, QLI-CVIII and SSAI as outcome measures.

While the GMS pa emerged as the most responsive subscale, a number of other subscales were also shown to be responsive. Indeed, as some instruments comprise a number of subscales, it may be useful for researchers to consider the utility of constructing a battery of instrument subscales rather than the entire instrument to assess HRQOL outcomes. Examination of the data presented in Figure 7.1 shows that eleven subscales (GMS pa, SF rlp, GMS na, SF swb, SF rle, SF p, HPPQ f, HPPQ d, HADS dep, HCS cog, and SF ewb) had an SRM greater than 0.20 (small effect). Thus, an assessment battery comprising a number of these subscales may potentially be responsive to cardiac rehabilitation intervention. The composition of such a battery would depend upon the particular research question under investigation. These subscales may be classified as assessing the following psychosocial states:

(i) Depression (HADS dep)

(ii) Negative psychosocial state (GMS na, HPPQ d)

(iii) Positive psychosocial state (GMS pa, SF swb, SF ewb)

(iv) Subjective health status (SF rlp, SF rle, SF p, HPPQ f, HCS cog)

Given the evidence concerning the importance of depression and general negative psychosocial state, it would be recommended that measures of these constructs be routinely assessed. The high responsiveness of the measures of positive psychosocial state provides a rationale for the inclusion of such a scale. With respect to subjective health status, it may be that some scales are more appropriate for outcomes assessment in cardiac populations (e.g., role limitations, feelings of disability, perceived impact of health) than

others (e.g., pain). For example, a 34–item assessment instrument that assesses depression (HADS dep), negative and positive psychosocial state (GMS) and subjective health status (SF rle, SF rlp) could be routinely used to assess HRQOL outcomes in cardiac rehabilitation. Clearly, other instruments could be constructed using different combinations of subscales. Such instruments may then undergo psychometric investigation to determine their reliability, validity and responsiveness.

8.1.1 Generic versus specific instruments

It is assumed that disease-specific instruments will be more responsive than generic instruments. An instrument developed with cardiac populations (although not technically a disease specific instrument), the GMS, was found to be the most responsive instrument to multifactorial cardiac rehabilitation in the present study. In addition, two subscales (d and f) of the HPPQ, a disease-specific instrument, were also found to be responsive. However, a number of generic instruments were also responsive, in particular some of the subscales of the SF-36 and the depression subscale of the HADS. Thus, use of the SF-36 and the HADS may facilitate general comparisons across illness conditions. Furthermore, Johnston (1999) has recently suggested that the use of generic instruments facilitates model development in health psychology and facilitates the increased use of common psychological constructs.

From the perspective of routine outcomes assessment, if generic instruments are shown to be as responsive as disease-specific instruments, then it will be preferable to use

such generic instruments. This would facilitate comparison across different clinical conditions and the effects of different interventions. At present, ever more specific instruments are emerging to join the multitude already available and indeed, the only limitation to the number of instruments is the number of disease categories that can be identified. It is worth remembering that already numerous cardiac specific instruments are currently available (e.g., QLMI, HPPQ, QLI CV III, CDS). Therefore, it may be more beneficial for the future if research examined how these different instruments are related to each other. Such an approach would seem preferable to the potentially endless prospect of additional disease-specific instruments appearing in the literature. Comparison of the psychometric profiles of current instruments would at least provide a set of benchmarks against which to judge the properties of new instruments. The present study is the first to offer such comparative framework for psychosocial outcomes assessment in cardiac rehabilitation.

The present findings reveal a striking variability (see Figure 7.1) in responsiveness of the different psychosocial instruments. Given this high variability between instruments to detect change, it is very clear that outcome instrument selection must involve an explicit consideration of the empirical evidence documenting responsiveness of an instrument in situations likely to match those to be investigated. For example using only one instrument to assess psychosocial outcome may provide misleading conclusions about health service effectiveness. Given the current health care climate of evidence-based medicine, the use of inappropriate instruments to document outcomes may prove costly in terms of service

funding. This statement must however be tempered by the caveat that researchers need to be wary of the possibility that using a number of instruments may result in incomprehensible results if a battery of questionnaires provide inconsistent or contradictory evidence. The routine reporting of effect sizes (or some other index of responsiveness) would provide researchers with empirical evidence on responsiveness that can be used to inform instrument selection. Similar calls for increased reporting of effect sizes in the general psychological literature have been made as much as a decade ago (e.g., Cohen, 1990).

8.1.2 Floor and ceiling effects

Floor and ceiling effects were found on a number of instruments, in particular the SF-36 (see Appendix 21). The presence of floor effects at time 1 was significantly associated ($r=0.38$, $p<0.05$) with responsiveness. Thus, if respondents score at the lowest possible score at time one, then there is more potential to detect improvement over time than if respondents score elsewhere on the instrument. While the presence of floor effects was associated with increased responsiveness in the present study, previous research has documented that floor effects can reduce the responsiveness of an evaluative instrument (e.g., Bindman et al., 1990). A weaker association ($r=0.24$, $p <0.05$) was reported between ceiling effects and responsiveness. In addition, the present study examined the relationship between scoring in the lowest and highest possible scale scores and responsiveness. This was investigated as responsiveness may be constrained by respondents scoring near the

extreme end point of an instrument as the possible change score is lower than those scoring in the mid-point of scale scores. Indeed, a strong relationship ($r=0.44$, $p<0.05$) was reported between those scoring in the lowest 10% range of an instrument and responsiveness. This further supports the fact that if individuals score at the lower end of instruments (i.e., poorer functioning) then the instrument has greater potential to detect improvement. However, floor effects compromise the ability of the instrument to detect deterioration in function and this raises questions concerning the range of functioning assessed by the instrument. If a scale does not adequately assess the extremes of functioning, then the clinical relevance of the scale may be questioned.

The relationship of items on an instrument to the distribution of the full spectrum of the construct may influence responsiveness (Fitzpatrick et al., 1998). For example, it has been suggested that the SF-36 pf subscale items reflect intermediate levels of physical difficulty rather than the extremes of difficulty for patients undergoing total hip arthroplasty (Stucki et al., 1995). Consequently, patients who improve at an intermediate level will have large change scores on the scale. Thus, the scale may be limited in terms of the range of functioning assessed.

The presence of these effects thus limits the ability of instruments to detect changes in the variables assessed. Furthermore, McHorney and Tarlov (1995) suggested that a floor or ceiling effect higher than 15% may render a questionnaire inadequate for assessing individuals for clinical purposes and the instrument will not be responsive to changes in these patients. In the present study, a number of instruments had floor and ceiling effects

higher than recommended (e.g., QLI fam, SF rlp, SF rle, SF swb, SF p, HPPQ w). If psychosocial instruments are to be used routinely in clinical practice, then issues such as floor and ceiling effects, and the spectrum of items assessed, need to be carefully considered.

8.1.3 Cardiac rehabilitation psychosocial outcome – inappropriate instruments or small intervention effect?

In comparison with Cohen's criteria for small (0.20), medium (0.05) and large (0.80) effects, the magnitude of effect for cardiac rehabilitation reported in the present study is modest. Indeed, it is notable that only 9 (GMS pa, SF rlp, GMS na, SF swb, SF rle, SF p, HPPQ f, HPPQ d, and SF ewb) out of 29 subscales had effect sizes of greater than 0.20. Furthermore, only two scales (GMS pa, SF rlp) had effect sizes greater than 0.50 (moderate change) and no scale exceeded the 0.80 criterion for large change. It should be noted that effect size for the change in the self-reported physical functioning using the DASI in the cardiac rehabilitation participants (ES=0.26) was greater than that reported by the comparison group (ES = 0.06). Overall, the magnitude of effect reported for psychosocial states is similar to that reported for physical functioning.

Examination of the data presented in Table 7.6 reveals the marked variability in psychosocial outcome as assessed by each subscale. For example, less responsive instruments (e.g., STAI, CDS) failed to show a significant difference ($p \geq 0.10$) between the cardiac rehabilitation and comparison group over time. In contrast, responsive instrument

(e.g., GMS, SF rlp, SF rle) showed a highly significant difference ($p \leq 0.001$) between the cardiac rehabilitation and comparison groups. These data illustrate the importance of selecting responsive instruments in service evaluation as the use of inappropriate instruments may result in service benefits being underestimated. The present study assessed the impact of cardiac rehabilitation on psychosocial states using a variety of instruments. The responsiveness of these instruments was assessed to determine those instruments that were most capable of detecting change. However, it must be acknowledged that responsiveness is a marker for validity but cannot illustrate whether a different evaluative approach would be more comprehensive and thus have higher validity. The majority of instruments in the present study focused almost exclusively on negative states and a more appropriate approach to outcomes assessment in cardiac rehabilitation should incorporate positive and negative mood (see section 4.5.1).

It should be noted that as the present study was specifically focused on investigating the responsiveness of a number of psychosocial instruments to cardiac rehabilitation interventions, it was not designed to assess the efficacy of cardiac rehabilitation. Nevertheless, based on the findings of the systematic review (chapter 5), it was expected that higher effect sizes would be detected by the instruments used in the present study. Indeed, seven of the nine instruments assessed had previously shown sufficient effect sizes to suggest that the instruments were responsive. It is however possible that the systematic review overestimated the effects of cardiac rehabilitation as a result of publication bias. Thus, the effect sizes noted in chapter 5 may reflect the results

of only those studies which report a significant difference. However, given that a number of papers included in the review failed to report a statistically significant difference between cardiac rehabilitation and control groups, one can question the extent to which publication bias influenced the review.

Nevertheless, publication bias has serious implications for unbiased data collection in a systematic review and must be considered a potentially confounding factor. The selection of instruments demonstrating the highest effect size difference in the review may have provided an inflated estimate of responsiveness and a process akin to regression to the mean may have occurred in the present study. The sample obtained using only published papers may be biased in favour of those studies showing the largest effects sizes. The development of statistics to estimate the number of negative findings required to overturn the conclusions of a meta-analyses is promising and can be used to caution against the over-enthusiastic adoption of meta-analytic findings (Rosenthal, 1991).

A systematic review of the effectiveness of cardiac rehabilitation in comparison to usual care is currently underway with the assistance of Cochrane Collaboration in the UK. This review explicitly aims to include data from both published and unpublished randomised controlled trials of cardiac rehabilitation (Jolliffe, 1998). The forthcoming Cochrane review may provide insight into the possibility that the present systematic review (Chapter 5) provided a biased estimate of the responsiveness of instruments.

Given the variability in time since most recent cardiac event among the study participants, it may be that stronger effects would be detected among participants who started cardiac rehabilitation sooner post-event in comparison to those who start cardiac rehabilitation later. Examination of the data presented in Figure 7.2, shows that the relationship between time since event and responsiveness was weak (Spearman rank order r=0.19). However, as the study did not explicitly match each cardiac rehabilitation participant with a comparison participant in terms of time since most recent cardiac event, it is possible that this factor contributed to the present findings and consequently limits their generalisability. It should be noted that this factor has not been empirically investigated in the HRQOL assessment literature and further investigation of this issue is clearly warranted.

8.1.4 Response shift

The failure to find larger effects of cardiac rehabilitation may be attributable to the psychological process of response shift. Response shift refers to the fact that when individuals undergo a change in their health state, they may also change their internal standards, their values or their conceptualisations of psychosocial status (Schwartz & Sprangers, 1999). The presence of a response shift may have contributed to the results in a number of possible ways.

Cardiac rehabilitation participants may have changed their psychosocial status standards as a function of programme involvement. Thus, their internal standards may

change as their expectations and beliefs about the cardiac rehabilitation outcome change. The traditional group-based format of cardiac rehabilitation involves a number of potentially powerful psychological effects, including the presence of social comparisons. Indeed, there is some evidence to suggest that cardiac rehabilitation participants engage in social comparisons (e.g., Helgeson & Taylor, 1993) and it has been proposed that social comparison processes may be a mechanism that contributes to response shift (Gibbons, 1999). A substantial body of research documents the use of social comparison strategies by individuals dealing with health problems (Buunk & Gibbons, 1997). However, further empirical testing of the possible contributions of social comparisons to response shift are needed

In the present study, the responsiveness values reported by the majority of instruments are moderate to non-existent. However, if response shift is contributing to the results, then one must question whether some instruments are more severely affected by such a process. As yet there is no empirical evidence to assess whether different response formats and options are differentially influenced by response shift. Furthermore, the relationship between response shift and responsiveness has not been examined within the literature and both theoretical and empirical work is required on this topic. For example, Sprangers and Schwartz (1999) noted that response shift may limit the ability of instruments to detect change over time. In addition to the possibility that response shift influenced the results, the impact of non-specific effects must also be considered.

8.1.5 Non-specific effects

With respect to the comparison group, as individuals follow the natural course of recovery, internal standards, values and conceptualisations of quality of life may change. In addition, the changes in psychosocial status reported by the comparison group over time may reflect non-specific effects on psychosocial ratings in longitudinal studies. Bouchet, Guillemin and Briançon (1996) examined the impact of non-specific effects on quality of life assessments at one time and two months later using the SF-36 in a study on the effects of vitamin supplementation in healthy adults. Participants were randomly allocated to one of four groups: an active treatment group, a placebo group, a two measurement control group to assess the Hawthorne effect (tendency of individuals to change their behaviour because they are participating in a study), and a group measured at the second time point only to assess the effects of natural history. Of note, after correcting for natural history effects, the Hawthorne group reported significant changes over time in the emotional and physical role limitations scales of the SF-36. Thus, simply being asked to fill in two questionnaires over a two-month interval also had a significant effect on the responses of the participants.

In a recent examination of a reduced version of the DASI in a sample of CHD patients before and after coronary angioplasty (Alonso et al., 1997), an effect size of 0.75 was reported for improvement in the reduced DASI. Of note, a stable control group of CHD patients showed an effect size of 0.22 on the same instrument over a two-week re-test period. Thus, while the effect size was more than three times higher for angioplasty

patients than for stable patients, the presence of the Hawthorne effect may account for the reported improvement in the control group.

In the present study, it is possible that the Hawthorne effect contaminated the responses of the comparison group who were aware that they were involved in a study and as such may have altered their responses accordingly. However, as a perfect control is practically impossible (i.e., a group who could fill in a questionnaire without knowing that they were completing a questionnaire), the role of non-specific effects cannot be fully eliminated.

In addition, the role of participant expectancies needs to be considered as changes in psychosocial status can occur as a consequence of being included in a study without any intervention. Thus, research needs to take account of the influence of non-specific effects in longitudinal studies. In the present study it may be that some instruments are better at detecting such non-specific noise and this reduces the responsiveness of the instrument. It would be interesting to replicate the present study with the addition of the natural history group methodology described above.

8.2.0 Comparison of responsiveness statistics

The present study used 5 approaches to assessing responsiveness. A potential weakness in previous research examining responsiveness is that the majority of these studies do not include a control or comparison group. Typically, the responsiveness of an instrument to a health care intervention is estimated. However, this data reveals nothing

about the amount of change that may occur in a non-intervention group. In response to this deficit, the present study adopted a different approach. Inclusion of a comparison group provided a means to estimate the degree of change seen over time in a non-intervention group. The magnitude of this change can then be compared to the magnitude of change seen in a cardiac rehabilitation intervention group. The difference between the two values can be regarded as a measure of the instruments responsiveness to the effects of cardiac rehabilitation. Thus the difference scores between a cardiac rehabilitation and a comparison group were calculated. This methodology was used to assess the responsiveness of instruments using the ES, SRM, Tuley index and RE approaches.

The Guyatt index explicitly assumes that responsiveness is a function of the variability in score changes for stable subjects (Guyatt et al., 1987) and thus provides a single index that accounts for non-specific change. However, it is critical that the stable group is accurately classified. As the present study has highlighted, psychosocial states in cardiac patients change over the course of time and the assumption that a non-intervention group remain stable needs to be empirically evaluated. A comparison group is not necessarily a stable group and conversely, an intervention group does not necessarily show change. The responsiveness of the instrument reflects the ability of the instrument to detect these changes.

In order to identify the most responsive instrument, the rankings of the instruments according to the ES, SRM, Tuley index and RE statistics were combined to produce an overall ranking of instruments. The correlations between the different rankings showed

that the ES, SRM, Tuley Responsiveness index and the relative efficiency approaches to responsiveness were significantly related (Spearman rank correlations ranging from 0.85 to 0.97). Thus in general, these approaches to estimating responsiveness provide similar orderings of instruments. However, the Guyatt responsiveness index was not associated with any of the other statistics (Spearman Rank correlations ranging from -0.07 to 0.07). The reasons for the lack of agreement between the Guyatt index and the other responsiveness statistics remains to be determined. One possibility concerns the use of the Juniper transition instrument to classify the MID.

8.2.1 Classification of the MID

The use of the Juniper index to classify individuals as stable does not appear to be accurate. For some scales (see Table 7.8), those individuals who reported being stable reported higher change scores than those classified as minimum change (e.g SSAI, HPPQ d), moderate change (e.g, HCS som) or even large change (HAD anx, SF rle).

Indeed, the stable group reported the lowest degree of change in approximately half of the subscales (e.g., HCS cog, HAD dep, GMS pa). The use of the transition scale to classify a stable group has implications for estimating the MSE for both Tuley's index and Guyatt's responsiveness index. If the transition item does not accurately identify those participants who remain stable, then the denominator for both Tuley's and Guyatt's index may not be accurate. Therefore, if transition items are used to classify participants as stable, the use of both Tuley and Guyatt indices may be questionable.

It would have been predicted that a linear trend would be observed on all scales according to the self-rated change in quality of life over the previous ten weeks. Thus, those participants coded as 1 (i.e., reporting no change or hardly any change) would be expected to show the least degree of change, those coded as 2 (reporting a little changed or somewhat changed) should on average show a higher mean change than group 1, those coded as 3 (reporting a moderate change or a good deal of change) should change more than group 2 and finally those coded as 4 (reporting a great deal of change or a very great deal of change) should report the highest levels of average change. Examination of Table 7.8 shows that for the 29 scales assessed in the present study, only 3 scales followed this pattern (the HCS cog, SF gh, GMS pa). Thus approximately 90% of scales failed to follow the expected response pattern. While the expected pattern is never reversed (although CDS comes close), there is considerable deviation from the predicted pattern.

Of note, previous research using transition items has also reported a marked inconsistency in the degree of change detected on different instruments and the self-rated degree of change. Considerable between-questionnaire variations in ratings of change were reported in a group of 100 patients at a rheumatology clinic (Fizpatrick, Ziebland, Jenkinson, Mowat & Mowat, 1993). At follow-up the patients described their progress by completing a five-point transition concerning their improvement, stability or deterioration since the last visit. The results showed a striking variance in patterns of responses across questionnaires. For example, among the one third of patients who characterised themselves as having improved (either greatly or slightly) over the three month period, the

change score on one questionnaire indicated substantial changes, while another showed no change and a third showed only slight improvements. Therefore, there is a marked inconsistency between transition items and change scores on psychosocial instruments.

It is difficult to provide a clear explanation for the present findings of a lack of consistency in the degree of change detected on different psychosocial instruments and the self-rated degree of change. Improvements in emotional, social and functional status were not reflected in ratings of perceived changes in quality of life over the study period on the transition scale. The transition item may be assessing something else other than a perceived change in quality of life. Given the high scores on the measure, it may simply be reflecting social desirability or some other expectation bias that resulted in participants recording a high degree of change. In the present study, some participants who rated themselves as stable actually reported quite a high degree of change on the psychosocial instruments. It is hard to explain the apparent inconsistencies in respondent answering tendencies. It may be that the participants felt obliged to report a high quality of life in response to some perceived expectation that they should be reporting such changes. It is impossible to evaluate the influence of such a bias on the results of the present study.

Future research may address such concerns through the use of social desirability measures to identify the extent to which such tendencies are present within a cardiac population. It may be necessary to investigate the use of additional methodologies to classify participants. For example Multi-Trait Multi-Method (MTMM) approaches may provide insight into the comparability of different approaches to classification of change.

In addition, the use of methodologies that encourage respondents to verbalise their thought processes involved in making decisions about answering items may offer a fruitful avenue to pursue such issues.

8.2.2 Responsiveness statistics

The present study used 5 statistics to assess the responsiveness of instruments. As noted previously, the different statistics provided different rankings of instruments and thus classification of instrument responsiveness is dependent upon the statistical approach used. Given such variability in responsiveness according to the statistics, the question arises as to which statistic should be used.

The ES statistic is the most commonly cited measure of responsiveness and is based on the mean change and the standard deviation of scores at baseline. It has been suggested that the ES provides a standardised approach to comparing the responsiveness of different instruments (Lydick & Epstein, 1993). However, the baseline standard deviation does not incorporate response variance, and it has been suggested that this raises concern over the validity of the effect size. In response to this weakness of the effect size, the SRM was suggested as a more appropriate measure of responsiveness. The SRM uses the standard deviation of the change scores to take account of the response variance. In Guyatt and colleagues (1987) formula for estimating responsiveness, the numerator is the "minimum important difference". The denominator is an index of variability in stable subjects. Thus, the responsiveness index is given by the mean change in scores of patients

who report improvement divided by the variability of patients reporting no change. However, estimating the MID is required for the calculation of responsiveness and this limits the practicality of this method (Liang et al., 1990).

As noted previously, estimation of the MID and also classifying the stable group was problematic in the present study. Comparison of the self-rated change on the Juniper scale and the change on the instruments showed marked inconsistencies between these measures. Until the issue of classifying individuals as stable and the issue of minimum important difference receives further theoretical and empirical investigation, the use of the Tuley and Guyatt responsiveness indices remains questionable. The RE statistic is based on t-statistics and consequently, these values may not be appropriate in samples where the distribution is not Gaussian (Fitzpatrick et al., 1998). With the above considerations in mind, the SRM remains the recommended responsiveness statistic.

The ES, SRM, Tuley index and RE approaches provide approximately the same ordering of instruments. However, it is worth noting that the ordering between the different approaches is not exactly the same. This finding highlights the fact that research needs to consider the possibility that reliance on a single approach could provide misleading conclusions about instrument responsiveness. On the other hand, it is possible that the use of several responsiveness statistics may result in incomprehensible results if they provide inconsistent or even contradictory findings. The reported variability in responsiveness values as given by the different statistics clearly warrants further investigation.

8.2.3 Parametric versus non-parametric effect size statistics

The present study compared the responsiveness statistics of both the parametric and the non-parametric versions of the ES statistic. The use of non-parametric responsiveness statistics has not received much empirical attention in the outcomes literature (Fitzpatrick et al., 1998) and it was decided to examine the comparability of these approaches. While the correlation between the two approaches was significant ($r = 0.60$), there are clearly discrepancies between the two approaches. This finding suggests that outcomes research needs to consider the distribution of data before selecting measures of responsiveness. At present only the effect size statistic has a non-parametric equivalent. However, it may be possible to develop non-parametric versions of the SRM (using the median change as the denominator and the interquartile range as the numerator), the RE (using Wilcoxon statistics), the Tuley index (using median change and non-parametric repeated measures ANOVA) and the Guyatt index (using non-parametric repeated measures ANOVA). The utility of such approaches requires empirical validation.

8.3.0 Limitations of the present study

A number of limitations of the present study need to be addressed. Both methodological and theoretical limitations will be considered.

8.3.1 Non-random design

The decision to use a quasi-random design was based on practical considerations. To deny standard treatment with proven efficacy above and beyond usual care to a group of patients in a study that is not comparing the relative efficacy of different treatments raises a number of ethical issues. Thus a quasi-experimental design was employed as it ensured that no patient was denied usual treatment in any of the study centres. Without random assignment the processes that select patients into one group or another may produce systematic differences between the groups. For example, self-selection biases may influence the results.

While the selection of participants from the cardiac rehabilitation and control groups was as inclusive as possible, selection bias may have contributed to the results in a number of ways. For example, only those in the best psychological health may have accepted the invitation to participate in cardiac rehabilitation. Furthermore, among those in cardiac rehabilitation, perhaps only those in the best psychological health returned the questionnaires at both time points. In the comparison group, participants who were not in psychological distress may have replied to the invitation to participate and returned the questionnaire.

In addition, there may have been an initial bias in terms of referral for cardiac rehabilitation and then once the place on cardiac rehabilitation had been offered, the cardiac rehabilitation participants become a self-selected group and may not be

representative of the wider cardiac population. Therefore, the cardiac rehabilitation and comparison groups may have differed systematically on variables that influenced the outcome.

An attempt was made to match the hospitals offering cardiac rehabilitation with similar hospitals not offering such a service. Analysis revealed some patient differences across centres in terms of demographic details. However, comparison of the results when each centre was removed from calculations revealed no significant difference in the overall results. Nevertheless, it must be acknowledged that the success of such a strategy may legitimately be questioned as variables not considered in the matching may contribute to the outcomes of the patients in various centres.

There was no significant difference in participant attrition between the cardiac rehabilitation and comparison group, with similar drop-out rates in both group. However, if the loss of participants was attributable to a systematic factor related to the outcomes of interest then the internal validity of the study is compromised. No significant differences were reported on variables between those completing time 1 and time 2 questionnaires. Nevertheless, it is not known whether the final sample differs in some respect from the original sample on some critical variable that was not assessed.

8.3.2 Self-report questionnaires

The present study used self-completion questionnaires to assess the psychosocial functioning of a large sample if cardiac patients. This method was selected because it offered a number of potential advantages. The main focus of research on health status assessment has been on developing and testing self-report questionnaires that can be used to assess the health outcomes of large numbers of service users. Questionnaires are generally accepted to be more feasible than interviewing techniques for this purpose.

While self-report questionnaires can be a fallible source of data (Schwarz, 1999), psychosocial status assessment is predicated on the assumption that in the majority of cases, only the individual can provide meaningful data about their psychosocial state. Individuals with episodes of psychosis and dementia provide exceptions to this statement but, in general, there is widespread agreement that such ratings are best made by the individual. Thus, self-report questionnaires were used to assess the psychosocial status of the study participants.

8.3.3 Postal questionnaires

Posting questionnaires to the comparison group allowed a large sample of respondents who live in a wide dispersion of locations to be contacted. Such an approach also has the advantage of having low costs of data collection and processing and avoids interviewer bias. However, the disadvantages to this approach must also be acknowledged. While examination of the sample profile shows a similar breakdown to other cardiac

samples in the literature, the participants were a self-selected group and thus a number of biases may limit the generalisations of the findings to a wider cardiac population. Interviews in which participants completed the instrument in the presence of the tester would have offered higher response rate, an opportunity to correct any misunderstandings, facilitated participation of those with poor literacy skills, and provided control over missing data. However, it was decided that direct interviewing of 800 patients on two occasions in such a large study was not feasible. In addition, interviewer bias (since the interviewer would not be blind to the participant's group status), may have influenced the outcome.

Nevertheless, as Oppenheim (1992) noted, postal questionnaires may also be influenced by a bias as a respondent may create an image of the organisation that sent the questionnaires, in this case a Department of Cardiology, and this referent may bias the results. The extent to which this occurred in the present study cannot be estimated but clearly is an issues to be considered in the use of postal surveys.

8.3.4 Response bias

It seems reasonable to assume that there are differences in important characteristics between those that respond to questionnaires and those who do not. Thus, the actual sample obtained in a mail survey may not be representative of the population under investigation. The present study attempted to increase response rates through the use of reminder letters and repeat postings of questionnaires. However, the key issue remains as

to whether this sample is biased and if the reason for non-response is related to the research question. The present study cannot provide any information on differences between those agreeing to participate in the study and those who did not participate.

8.3.5 Questionnaire ordering effects

In the present study the questionnaires were completed in a fixed order. While this reduces a source of variance across participants, such a design may result in one way asymmetrical transfer effects. Thus it is possible that questionnaire order may have influenced the participants' responses, i.e., completing one questionnaire may influence the response given to a subsequent questionnaire in the battery. Indeed, given the usual recommendations that a generic and a disease-specific questionnaire be administered there is a need to compare the response patterns following completion of other questionnaires. Experimental manipulation of questionnaire order may provide insights into the impact of such transfer effects.

It has been well documented that respondents derive information from their answers to comparative scale items (e.g., frequency ratings, satisfaction ratings) and that this information is subsequently used in making comparative judgements (Schwartz, 1999). Therefore the order in which a respondent completes a questionnaire battery has implications for the outcome as answers to a previous item can influence how the subsequent items are responded to. The cognitive and social demands involved in

completing questionnaires have received much empirical investigation (See Schwartz 2000 for a review); however, this literature appears to have not yet impacted on health psychology research.

8.3.6 Theoretical limitations

One of the key weaknesses of the present research is the lack of theoretical model to guide instrument selection. The present study was driven by a pragmatic approach to outcomes assessment insofar as the focus was on those psychosocial states that have empirical evidence to suggest that they are appropriate targets for cardiac rehabilitation. Thus, an evidence-based approach to outcomes assessment was adopted. According to the WHO (1993), the goals of cardiac rehabilitation are to improve quality of life and to promote secondary prevention. By targeting those psychosocial factors (e.g., depression, anxiety, general negative psychosocial states) associated with poor outcome, cardiac rehabilitation services may favourably influence these factors and consequently may contribute to an increased quality of life and also improved secondary prevention.

To-date, pragmatism, rather than theory, has characterised applied HRQOL research. Perhaps the key issue facing quality of life research is that at present the field has a number of competing models of quality of life but little empirical evaluation of the relative merits of each model. Future research could empirically test different models with quality of life outcomes using structural equation modelling and confirmatory factor analysis.

8.3.7 Other psychosocial states

While a number of psychosocial factors were assessed, other important psychosocial states were not considered. Hostility, for example, has been associated with negative outcomes within cardiac populations (Sotile, 1999). However, whilst acknowledging this relationship, the psychosocial states most frequently targeted by cardiac rehabilitation programmes seem to be anxiety and depression, and general negative affective states. Hostility may require specific interventions that are beyond the remit of standard multi-factorial cardiac rehabilitation. Nevertheless, it may be worthwhile exploring the effects of cardiac rehabilitation on hostility and the responsiveness of different hostility instruments to cardiac rehabilitation interventions. In addition, there is a growing body of research documenting the pathophysiological consequences of social isolation (Pitula, Burg & Froelicher, 1999). As yet the impact of cardiac rehabilitation on social isolation has not been systematically assessed (however the ENRICHD study aims to examine the impact of psychosocial intervention on social isolation -see section 1.4.4). In addition to the negative states discussed above, a number of positive states that warrant further investigation include optimism, perceived control and resilience. The impact of cardiac rehabilitation on these states requires investigation.

8.4.0 Implications for routine psychosocial outcome assessment in cardiac rehabilitation

The practice of evidence-based medicine (EBM) means integrating individual clinical expertise with the best available external clinical evidence from systematic research. There are fears that in this new era of accountability in health care, EBM will be used to cut costs. However, as has been observed:

> "Doctors practising evidence based medicine will identify and apply the most efficacious interventions to maximise the quality and quantity of life for individual patients; this may raise rather than lower the cost of their care."
>
> (Sackett, Rosenberg, Gray & Richardson, 1996, p. 72)

Thus, evidence documenting the benefits of cardiac rehabilitation can help promote service provision. The use of appropriate instruments is essential as the use of unresponsive instruments may significantly underestimate the benefits of cardiac rehabilitation attendance. In the current health care environment, even relatively inexpensive services such as cardiac rehabilitation require documentation of accountability and cost (AACVPR, 1995). Thus it is essential that cardiac rehabilitation services use responsive instruments to evaluate the psychosocial benefits of cardiac rehabilitation. Failure to do so could result in studies showing no benefit that consequently could compromise the effectiveness of arguments in favour of increased cardiac rehabilitation service provision.

To-date, the issue of responsiveness in instrument selection for cardiac rehabilitation outcomes assessment has received little consideration. For example, the need for consensus regarding quality of life assessment in vascular disease has been proposed (Beattie, Golledge, Greenhalgh & Davies, 1997). The criteria suggested by these authors

for an ideal quality of life instrument include that it is equally applicable to any disease process or outcome, that it can be applied across all levels of illness and degrees of validity, and it should have proven validity and be capable of being used in different cultural contexts. According to the authors, no single instrument fulfils all these criteria. Indeed, only the NHP and the MOS SF-36 are considered as possible candidates for sole use in CHD. Based on a rather limited review of the literature, the authors conclude that

"the MOS SF-36 health related questionnaire can be recommended for use in vascular disease-related HRQL assessments and outcome analyses." (p. 12)

This reinforces the previous argument that in many cases instruments are being recommended without explicit data on their appropriateness, and in particular their responsiveness, being cited.

The large-scale psychosocial intervention study by Jones and West (1996) reported no significant differences between the experimental and the control groups on psychosocial outcomes. The authors recognised that the results may be attributable to the use of unresponsive instruments and they stated that

> "As patients reported subjective benefit, it is possible that measures sensitive enough to identify improved well-being after myocardial infarction have yet to be developed." (p. 1520).

While this statement suggests that there are no responsive instruments currently available, the authors fail to provide any evidence to support this view. Indeed, justification for the outcomes instruments selected by the authors is not provided in either

of the reports of the study (Jones & West, 1995; 1996). It may also be that the instruments they used in their study (e.g., DSSI/sAD (Bedford, Foulds & Sheffield, 1976) were not responsive and that this may explain the lack of impact on psychosocial well-being. The use of more responsive instruments may have revealed stronger treatment effects.

The practice appears to have been that some instruments become "fashionable" and are used almost indiscriminately by researchers because they have been used by many others (Donovan et al., 1993). Thus, a consensus emerges that these measures are valid and reliable. It may be argued that the NHP and SIP were the "fashionable" instruments of the 1980s. The SF-36 appears to be the current favourite. Many instruments appear to have been selected in previous cardiac rehabilitation studies because they were easily available or because similar studies used the measure, without any consideration of the responsiveness of the instrument. This state of affairs is detrimental to the development of a scientifically rigorous HRQoL discipline as it is difficult to make progress in the field if there is no shared definition of the concept or phenomenon under study. Therefore, it is difficult to assess the appropriateness of an instrument if there is uncertainty concerning what it is supposed to be measuring in the first place. The absence of a clear conceptual basis for quality of life measures (see also Leplege & Hunt, 1997) has resulted in much confusion in the outcomes assessment literature.

The recent development of a Working group for Consensus on Quality of Life Assessment in cardiac rehabilitation by the European Cardiology Society Working Group on Cardiac rehabilitation and Exercise Physiology is testament to the high levels of interest

in identifying a core number of instruments for routine QOL assessment in cardiac rehabilitation. Initiatives such as a European Society of Cardiology's Database on Cardiac Rehabilitation will provide impetus and opportunities to gather common information concerning psychosocial outcomes for cardiac rehabilitation across Europe. Consensus on instrument selection would facilitate greater comparability across studies examining the outcomes of cardiac rehabilitation. Such studies could potentially provide insight into the effects of different aspects of health care (e.g., structure and process) on outcomes and this data could inform both health services research and quality of life research. While such agreement may be impossible at present, an extended research project on this topic could prove useful.

The present study has provided benchmark values of responsiveness for a number of psychosocial instruments. Clearly, other instruments need to be evaluated in a similar manner. Over the course of a number of such studies it may be possible to identify a number of instruments that are most appropriate for routine outcomes assessment. However, such recommendations must be regarded as tentative, as the iterative nature of research will constantly produce evidence on the topic. A psychosocial outcomes panel could provide regular expert reviews of the evidence and produce updated guidelines on instrument selection. Such an approach is currently being undertaken in rheumatology with the development of the OMERACT (Outcome Measures in Rheumatology) conferences and its various committees (e.g., Boers, Brooks, Vibeke Strand & Tugwell, 1998).

OMERACT has provided much impetus to produce an active outcomes movement within rheumatology. OMERACT serves to improve outcome measurement in rheumatology through empirical research on the appropriateness of instruments. It has made the selection of outcome measures more explicit and has placed issues of responsiveness on the agenda of clinical researchers. Such an approach can be adopted within cardiac rehabilitation.

8.4.1 Moratorium on psychosocial instrument development?

The present study included nine psychosocial instruments selected on the basis of the systematic review (Chapter 5). In general, it must be noted that while superior instruments may be available to measure changes in psychosocial status over time, the aim of this thesis was to set a benchmark against which these instruments can be compared. Furthermore, there is a need to replicate and extend these findings.

Since the study began, a number of other cardiac-specific instruments have emerged in the literature. The Cardiac Quality of Life Inventory (CQLI) is a recently developed quality of life scale specific to a cardiac rehabilitation population (Rukholm, McGirr & Potts, 1998). This 20-item scale assesses psychosocial well being, physical well being, symptoms, nutrition, and worry. The CQLI is currently undergoing further psychometric evaluation. In addition, the Multidimensional Index of Life Quality (MILQ) is another disease specific measure that has recently emerged for use with cardiac patients (Avis et al., 1996). This 11-item instrument assesses 11 domains that contribute to quality

of life: physical health, physical functioning, mental health, cognitive functioning, work and productivity, financial status and security, relationship with spouse or partner, family, friends, leisure activities, and relationships with health-care providers. Initial evidence suggests that the MILQ is a psychometrically reliable and valid instrument for measuring quality of life in patients with cardiovascular disease. Clearly, these instruments need to be examined in terms of responsiveness.

At present, numerous psychosocial instruments are available for outcomes research. If instrument development proceeds at the current rate, then there is a risk that this multitude of instruments will simply pile up to form a psychometric Tower of Babel. However, such a tower may lack solid theoretical and empirical foundations and thus risks collapsing. If such a situation develops then it will be to the detriment to all those concerned with providing assessment of the subjective experiences of health care users as it becomes increasingly difficult to provide empirically-informed recommendations about the relative merits of different instruments.

An analogy perhaps can be drawn between the current status of quality of life outcomes research and personality theory until the emergence of the "big 5" model. In personality research, a multitude of different personality inventories were developed and instrument use followed the implicit models of the researchers. It is only in the last decade that widespread consensus has been achieved through rigorous psychometric analyses (though there are still a healthy number of critics of this model). It may be that quality of life outcomes research is at the stage where the number of instruments has reached

saturation point and the time has come to work on developing an integrated yet parsimonious model of quality of life. Indeed, the challenge for the next few decades in outcomes research may be to match the progress of the field in applied contexts with similar progress at a basic theoretical level.

8.5.0 Summary of Key Points

(i) A high degree of variability in instrument responsiveness was reported in the present study

(ii) A number of instruments (e.g. GMS, SF-36, HPPQ, HCS, HADS) showed higher levels of responsiveness than others

(iii) The GMS is recommended if the researcher is interested in assessing the effects of cardiac rehabilitation on negative and positive mood

(iv) A high degree of similarity was found in the ranking of instruments according to level of responsiveness for the ES, SRM, Tuley index and relative efficiency

(v) However the Guyatt responsiveness index was not related to the other indexes

(vi) Further research is needed to examine the estimation of the minimum important difference (MID) for psychosocial instruments

(vii) The present study provides a set of benchmarks against which the performance of other psychosocial outcome instruments for use in cardiac rehabilitation can be compared

(viii) The use of responsive instruments is critical if cardiac rehabilitation is to succeed in demonstrating its psychosocial benefits to participants

(ix) Documentation of psychosocial benefits is important in evaluating the care of individual patients, the effectiveness of specific cardiac rehabilitation programmes and also for cardiac rehabilitation as a scientific health care service.

8.6.0 Overall Summary

"Reliable, valid and *responsive* measures of outcome are essential in the evaluation of the effectiveness of cardiac rehabilitation. Objectively measured information of this type not only assesses the progress of the individual but helps evaluate the effectiveness of the rehabilitation program and the discipline as a whole." (emphasis added)

(*AACVPR, 1995*)

In conclusion, this study has attempted to identify those instruments that are most appropriate for routine psychosocial outcome assessment in cardiac rehabilitation service evaluation. The psychosocial states of depression, anxiety and negative psychosocial status were assessed as these states had previously been associated with negative outcomes for cardiac patients in the literature. The first part of this thesis reviewed the background factors that have contributed to the need for this topic to be investigated. A systematic review was conducted which identified a number of potentially promising instruments. Subsequently, nine psychosocial instruments were assessed in a large-scale study.

Psychometric analysis focused on the responsiveness of these instruments and compared different indices of responsiveness in these measures. The ranking of instruments provided by the SRM, ES, Tuley index and RE provided approximately the dsame results; the Guyatt index however was not correlated with the other statistics. Using the overall ranking provided by the statistics, the GMS emerged as the most responsive instrument. However, a number of other subscales (e.g., HADs dep, HPPQ d, SF swb, SF ewb, SF rlp, SF rle, SF p, HPPQ f, HCS cog) were identified as being responsive and these instruments can be combined to form an evaluative questionnaire battery.

Recommendations concerning instrument appropriateness should be interpreted as preliminary as the research process is iterative, in that such recommendations are open to challenge as new evidence is produced concerning the responsiveness of selected instruments. The present findings can help inform instrument selection at a local level and can aid in setting responsiveness benchmarks for future studies to compare their results. Thus, the thesis also provides a frame of reference against which other instruments can be

compared. In addition, it is also hoped that this research may encourage wider use of appropriate statistics and more consistent reporting of the psychometric properties of health status instruments. Through the development of a cumulative research knowledge base, it is hoped that we can better advise on instrument selection in psychosocial outcome assessment in cardiac rehabilitation.

References

27th Bethesda Conference. (1996). Matching the intensity of risk factor management with the hazard for coronary disease events. *Journal of the American College of Cardiology*, **27**, 957-1047.

AACVPR Outcomes Committee. (1995). Outcome measurement in cardiac and pulmonary rehabilitation. *Journal of Cardiopulmonary Rehabilitation* , **15**, 394-405.

Aaronson, N.K. (1989). Quality of life assessment in clinical trials: Methodologic issues. *Controlled Clinical Trials,* **10**, 195S-208S.

Abraham, C. & Hampson, S.E. (1996) A social cognition approach to health psychology: philosophical and methodological issues. *Psychology and Health*, **11**, 223-241.

Adler, N.E., Boyce, T., Chesney, M.A., Cohen, S., Folkman, S., Kahn, R.L., & Syme, S.L. (1994) Socioeconomic status and health. *American Psychologist*, **49**, 15-24.

Affleck, G., Tennen, H., Croog, S., & Levine, S. (1987). Causal attribution, perceived benefits, and morbidity after s health attack: An eight-year study. *Journal of Consulting and Clinical Psychology*, **55**, 29-35.

Allison, T.G., Williams, D.E, Miller, T.D., Patten, C.A., Bailey, K.R., Squires, R.W., & Gau, G.T. (1995). Medical and economic costs of psychologic distress in patients with coronary artery disease. *Mayo Clinic Proceedings*, **70**, 734-742.

Alonso, J., Permanyer-Miralda, G., Cascant, P., Brotons, C., Prieto, L., & Soler-Sole, J. (1997). Measuring functional status of chronic coronary patients. *European Heart Journal*, **18**, 414-419.

American Association of Cardiovascular and Pulmonary Rehabilitation Outcomes Committee (1995) *AACVPR outcomes tools resource guide*. Wisconsin: AACVPR.

American Heart Association, National Heart Lung and Blood Institute. (1990). The cholesterol facts: a summary of the evidence relating dietary fats, serum cholesterol and coronary heart disease. *Circulation*, **81**, 1721-1733.

American Heart Association. (1998). *Heart and stroke statistical update*. Dallas, Texas: American Heart Association.

Andersen R.M., Davidson, P.L., & Ganz, P.A. (1994). Symbiotic relationships of quality of life, health services research and other health research. *Quality of Life Research*, **3**, 365-371.

Avis, N.E., Smith, K.W., Hambleton, R.K., Feldman, H.A., Selwyn, A., & Jacobs, A. (1996). Development of the Multidimensional Index of Life Quality. A quality of life measure for cardiovascular disease. *Medical Care*, **34**, 1102-1120.

Bardsley, M.J., & Coles, J.M. (1992). Practical experiences in auditing patient outcomes. *Quality in Health Care*, **1**, 124-130.

Beaton, D.E., Hogg-Johnson, S., & Bombardier, C. (1997). Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with muscoskeletal disorders. *Journal of Clinical Epidemiology*, **50**, 79-93.

Beattie, D.K., Golledge, J., Greenhalgh, R.M., & Davies, A.H. (1997). Quality of life assessment in vascular disease: towards a consensus. *European Journal of Endovascular Surgery*, **13**, 9-13.

Beck, A., Ward, C., Mendelsoe, M., Mock, J., & Erbauch, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry,* **4**, 53-63.

Beck, A.T., Steer, R.A., & Garbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, **8**, 77-100.

Bedford, A., Foulds, G.A., & Sheffield, B.F. (1976). A new personal disturbance scale. DSSI/sAD. *British Journal of Social and Clinical Psychology*, **15**, 387-391.

Bindman, A.B., Keane, D., & Lurie, N. (1990). Measuring health changes among severely ill patients: the floor phenomenon. *Medical Care*, **28**, 1142-1152.

Boers, M., Brooks, P., Strand, C.V., & Tugwell, P. (1998). The OMERACT Filter for outcome measures in rheumatology. *Journal of Rheumatology*, **25**, 198-199.

Bornstein, M. (1997). Hypothesis testing and effect size estimation in clinical trials. *Annals of Allergy, Asthma, & Immunology*, **78**, 5-16.

Bouchet, C., Guillemin, F., & Briançon, S. (1996). Nonspecific effects in longitudinal studies: Impact on quality of life measures. *Journal of Clinical Epidemiology*, **49**, 15-20.

Bowling, A. (1997). *Measuring health. A review of quality of life measurement scales* (2nd ed.). Buckingham: Oxford University Press.

British Heart Foundation. (1998). *Coronary heart disease statistics. British Heart Foundation statistics database 1998*. London: British Heart Foundation Education Department.

Bullinger, M., Anderson, R., Cella, D. & Aaronson, N.K. (1995) Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. In S. A.

Shumaker & R. A. Berzon (Eds.), *The international assessment of health-related quality of life: Theory, translation, measurement and analysis* (pp83-92). Oxford: Rapid Communications.

Burns, R. B., McCarthy, E. P., Moskowitz, M. A., Ash, A., Kane, R. L., & Finch, M. (1997). Outcomes for older men and women with congestive heart failure. *Journal of the American Geriatric Society*, **45**, 276-280.

Buunk, B.P. & Gibbons, F.X. (1997). *Health, coping and well-being. Perspectives from social comparison theory*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Byrne, A., & Byrne, D.G. (1993). The effect of exercise on depression, anxiety and other mood states: a review. *Journal of Psychosomatic Research,* **37**, 565-574.

Carney, R.M., Freedland, K.E., Sheline, Y.I., & Weiss, E.S. (1997). Depression and coronary heart disease: a review for cardiologists. *Clinical Cardiology*, **20**, 196-200.

Carney, R.M., Rich, M.W., Freedland, K.E., Saini, J., teVelde, A., Simeone, C., & Clark, K. (1988). Major depressive disorder predicts cardiac events in patients with coronary artery disease. *Psychosomatic Medicine*, **10**, 627-633.

Carver, C.S., Scheier, M.F., & Weintraub, J. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, **56**, 267-283.

Cella, D.F., Wiklund, I., Shumaker, S.A. & Aaronson, N.K. (1995) Integrating health-related quality of life into cross-national clinical trials. In S. A. Shumaker & R. A. Berzon (Eds.), *The international assessment of health-related quality of life: Theory, translation, measurement and analysis* (pp 75-82). Oxford: Rapid Communications.

Chambers, L.W., MacDonald. L.A., Tugwell, P., Buchanan, W.W., & Kraag, G. (1982). The McMaster Health Index Questionnaire as a measure of quality of life for patients with rheumatoid disease. *Journal of Rheumatology*, **9**, 780-784.

Chua, T.P., & Lipkin, D.P. (1993). Cardiac rehabilitation should be available to all who would benefit. *British Medical Journal*, **306**, 731-732.

Cohen, J. (1977). *Statistical power analysis for the behavioural sciences.* New York: Wiley.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist,* **45**, 1304-1312.

Cook, D., Sackett, D., & Spitzer, W. (1995). Methodological guidelines for systematic reviews of randomised controlled trials in health care. *Journal of Clinical Epidemiology* , **48**, 167-71.

Croog, S., Levine, S., Testa, M., Brownm B., Bulpitt, C.J., Jenkins, C.D., Klerman, G.L., & Williams, G.H. (1986). The effects of antihypertensive therapy on the quality of life. *New England Journal of Medicine*, **314**, 1657-1664.

DeBusk, F., Haskell, W.L., Miller, N.H., Berra, T., Taylor, C.B., Berger, W.E. 3rd, & Lew, H. (1985). Medically directed at-home rehabilitation soon after clinically uncomplicated acute myocardial infarction: a new model for patient care. *American Journal of Cardiology*, **55**, 251-257.

DeBusk, R.F., Miller, N.H., Superko, H.R., Dennis, C.A., Thomas, R.J., Lew, H.T., Berger, W.E.3rd, Heller, R.S., Rompf, J., Gee, D., Kraemer, H.C., Bandura, A., Ghandour, G., Clark, M., Fisher, L., & Taylor, C.B. (1994). A case-management system for coronary risk factor modification after acute myocardial infarction. *Annals of Internal Medicine,* **120**, 721-729.

Dempster, M., Bradley, J., Wallace, E., & McCoy, P. (1997). Measuring quality of life in cardiac rehabilitation: comparing the Short Form 36 and the Nottingham Health Profile. *Coronary Health Care*, **1**, 211-217.

Denollet, J. & Brutsaert, D. (1998). Personality, disease severity, and the risk of long-term cardiac events in patients with a decreased ejection fraction after myocardial infarction. *Circulation,* **97**, 167-173.

Denollet, J. (1993a). Sensitivity of outcome assessment in cardiac rehabilitation. *Journal of Consulting and Clinical Psychology*, **61**, 686-695.

Denollet, J. (1993b). Emotional distress and fatigue in coronary heart disease: the Global Mood Scale (GMS) *Psychological Medicine*, **23**, 111-121.

Denollet, J., & Brutsaert, D. (1995). Enhancing emotional well-being by comprehensive rehabilitation in patients with coronary heart disease. *European Heart Journal*, **16**, 1070-1078.

Denollet, J., Sys, S., Stroobant, N., Rombouts, H., Gillebert, T. & Brutsaert, D. (1996). Personality as independent predictor of long-term mortality in patients with coronary heart disease. *Lancet*, **347**, 417-421.

Denollet. J. (1994). Health complaints and outcome assessment on coronary heart disease *Psychosomatic Medicine*, **56**, 463-474.

Department of Health. (1994). *Shaping a healthier future. A strategy for effective health care in the 1990s*. Dublin: Stationery office.

Department of Health. (1995). *A health promotion strategy. Making the healthier choice the easier choice*. Dublin: Stationery office.

Department of Health. (1997). *Statement of strategy*. Dublin: Stationery office.

Department of Health. (1999). *Building healthier hearts. The report of the cardiovascular health strategy group*. Dublin: Stationery office.

Derogatis, L. (1976). *Scoring and procedures manual for PAIS*. Baltimore, Maryland: Clinical Psychometric Research.

Derogatis, L.R., & Lopez, M.C. (1983). *Psychosocial Adjustment to Illness Scale (PAIS & PAIS-SR) Administration, scoring and procedures manual-1*. Baltimore, Maryland: Clinical Psychometric Research.

Deshotels, A., Planchock, N., Dech, Z., & Prevost, S. (1995). Gender differences in perceptions of quality of life in cardiac rehabilitation patients. *Journal of Cardiopulmonary Rehabilitation*, **15**, 143-148.

Deyo, R.A., & Patrick, D.L. (1995). The significance of treatment effects: The clinical perspective. *Medical Care*, **33**, AS286-AS291.

Deyo, R.A., Diehr, P., & Patrick, D.L. (1991). Reproducibility and responsiveness of health status measures. *Controlled Clinical Trials*, **12**, 142S-158S.

Diener, E. (2000) Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, **55**, 34-43.

Donovan, J., Frankel, S., & Eyles, J. (1993). Assessing the need for health status measures. *Journal of Epidemiology and Community Health*, **47**, 158-162.

Dusseldorp, E., van Elderen, T., Maes, S., Meulman, J., & Kraaij, V. (1999). A meta-analysis of psychoeducational programs for coronary heart disease patients. *Health Psychology*, **18**, 506-519.

Ebrahim, S. (1995). Clinical and public health perspectives and applications of health-related quality of life measurement. *Social Science and Medicine*, **41**, 1383-1394.

Effective Health Care Bulletin. (1998). *Cardiac rehabilitation.* **4**, 4.

Ellwood, P.M. (1988). Outcomes management: a technology of patient experience. *New England Journal of Medicine*, **318**, 1549-1556.

Epstein, A.M. (1990). The outcomes movement – will it get us where we want to go? *New England Journal of Medicine*, **323**, 266-270.

Erdman, R., Duivenvoorden, H., Verhage, F., Krazemier, M. & Hugenholtz, P. (1986) Predictability of beneficial effects in cardiac rehabilitation: a randomised clinical trial of psychosocial variables. *Journal of Cardiopulmonary Rehabilitation* , **6**, 206-213.

Ferrans, C.E. & Powers, M.J. (1985). Quality of life index: development and psychometric properties. *Advances in Nursing Science* , **8**, 15-24.

Ferrans, C.E., & Powers, M.J. (1992). Psychometric assessment of the quality of life index. *Research in Nursing and Health*, **15**, 29-38.

Fitzpatrick, R. (1995). Applications of health status measures. In C. Jenkinson (Ed.), *Measuring health and medical outcomes* (pp. 27-41). London: UCL Press

Fitzpatrick, R., Davey, C., Buxton, M.J., & Jones, D.R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment*, **2**, (14).

Fitzpatrick, R., Ziebland, S., Jenkinson, C., Mowat, A., & Mowat, A. (1993). A comparison of the sensitivity to change of several health status instruments in rheumatoid arthritis. *Journal of Rheumatology*, **20**, 429-436.

Ford, D.E., Mead, L.A., Chang, P.P., Cooper-Patrick, L., Wang, N-Y., & Klag, M.J. (1998). Depression is a risk factor for coronary artery disease in men. *Archives of Internal Medicine,* **158,** 1422-1426.

Fowler, G. (1998). Tobacco and cardiovascular disease: Achieving smoking cessation. In S. Yusuf, J.A. Cairns, A. John Camm, E.L. Fallen, & B.J. Gersh (Eds.), *Evidence based cardiology* (pp. 179-190). London: BMJ Books.

Frankford, D.M. (1994). Scientism and economism in the regulation of health care. *Journal of Health Politics, Policy and Law,* **19**, 773-799.

Frasure-Smith, N., Lespérance, F., & Talajic, M. (1993). Depression following myocardial infarction. Impact on 6-month survival. *Journal of the American Medical Association,* **270**, 1819-25.

Frasure-Smith, N., Lespérance, F., & Talajic, M. (1995a). Depression and 18-month prognosis after myocardial infarction. *Circulation*, **91**, 999-1005.

Frasure-Smith, N., Lespérance, F., & Talajic, M. (1995b). The impact of negative emotions on prognosis following myocardial infarction: Is it more than depression? *Health Psychology*, **14**, 388-398.

Frasure-Smith, N., Lesperance, F., Prince, R.H., Verrier, P., Garber, R.A., Juneau, M., Wolfson, C., & Bourassa, M.G. (1997). Randomised trial of home-based psychosocial nursing intervention for patients recovering from myocardial infarction. *Lancet*, **350**, 473-479.

Froelicher ,V.F., Herbert, W., Myers, J., & Ribisl, P. (1996). How cardiac rehabilitation is being influenced by changes in health-care delivery. *Journal of Cardiopulmonary Rehabilitation*, **16**, 151-159.

Garg, R., Packer, M., Pitt, B., & Yusuf, S. (1993). Heart failure in the 1990s: Evolution of a major public health problem in cardiovascular medicine. *Journal of the American College of Cardiology*, **22**, Suppl A, 3A-5A.

Gibbons, F.X. (1999). Social comparison as a mediator of response shift. *Social Science and Medicine*, **48**, 1517-1530.

Gibson, B.S., & Gibson, J.S. (1975). The Sickness Impact Profile – development of an outcome. *Annals of Internal Medicine*, **65**, 1304-1310.

Gill, T.M., & Feinstein, A.R. (1994). A critical appraisal of the quality of quality-of-life measurements. *Journal of the American Medical Association*, **272**, 619-626.

Gliklich, R.E., & Hilinski, J.M. (1995). Longitudinal sensitivity of generic and specific health measures in chronic sinusitis. *Quality of Life Research*, **4**, 27-32.

Gohlke, H. & Gohlke-Bärwolf, C. (1998). Cardiac rehabilitation: where are we going? *European Heart Journal*, **19** (suppl O), O5-O12.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire*. London: Oxford University Press.

Goldberg, D. (1978) *Manual of the General Health Questionnaire*. Windsor: NFER-NELSON Publishing Company.

Gooding, J., & Jette, A. M. (1985). Hospital readmissions among the elderly. *Journal of the American Geriatric Society*, **33**, 595-601.

Gulanick, M. (1991). Is phase 2 cardiac rehabilitation necessary for early recovery of patients with cardiac disease? A randomized, controlled trial. *Heart and Lung*, **20**, 9-15.

Guyatt, G., Deyo, R., Charlson, M., Levine, M., & Mitchell, A. (1989). Responsiveness and validity in health status measurement: a clarification. *Journal of Clinical Epidemiology*, **42**, 403-408.

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, **40**, 171-178.

Guyatt, G.H., & Cook, D.J. (1994). Health status, quality of life and the individual. *Journal of the American Medical Association*, **272**, 630-631.

Guyatt, G.H., & Cook, D.J. (1995). Letter to the Editor. *Journal of the American Medical Association,* **273**, 844-845.

Guyatt, G.H., Feeny, D.H., & Patrick, D.L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine*, **118**, 627-629.

Guyatt, G.H., Joan Eagle, D., Sackett, B., Willan, A., Griffith, L., McIlroy, W., Patterson, C.J., & Turpie, I. (1993). Measuring quality of life in the frail elderly. *Journal of Clinical Epidemiology*, **46**, 1433-1444.

Guyatt, G.H., Veldhuyzen Van Zanten, S.J.O., Feeny, D.H. & Patrick, D.L. (1989) Measuring quality of life in clinical trials: a taxonomy and review. *Journal of the Canadian Medical Association,* **140**, 1441-1448.

Hare, D., & Davis, C. (1996). Cardiac Depression Scale: Validation of a new depression scale for cardiac patients. *Journal of Psychosomatic Research*, **40**, 379-386.

Harlow, L.L., & Mulaik, S.A. (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.

Haskell, W.L., Alderman, E.L., Fair, J.M., Maron, D.J., Mackey, S.F., Superko, H.R., Williams, P.T., Johnstone, I.M., Champagne, M.A., Krauss, R.M., & Farquhar, J.W. (1994). Effects of intensive multidisciplinary risk factor modification on coronary atherosclerosis and clinical cardiac events in men and women with coronary artery disease: the Stanford Secondary Risk Factor Intervention Project (SCRIP). *Circulation*, **89**, 975-990.

Havik, O.E., & Maeland, J.G. (1990). Patterns of emotional reactions after a myocardial infarction. *Journal of Psychosomatic Research* , **34**, 271-285.

Hays, R.D., Anderson, R., & Reviciki, D.A. (1995). Psychometric evaluation and interpretation of health-related quality of life data. In S. Shumaker & R. Berzon (Eds.), *The international assessment of health-related quality of life: Theory, translation, measurement and analysis*. (pp.103-114). Oxford, England: Rapid Communications.

Helgeson, V.S., & Taylor, S.E. (1993). Evaluative and affiliative comparisons and coping among cardiac patients. *Journal of Applied Psychology*, **23**, 1171-1195.

Hemingway, H., & Marmot, M. (1998). Psychosocial factors in the primary and secondary prevention of coronary heart disease: a systematic review. In S. Yusuf, J.A. Cairns, A. John Camm, E.L. Fallen & B.J. Gersh (Eds.), *Evidence based cardiology* (pp. 269-285). London: BMJ Books.

Hermann, C. (1997). International experiences with the Hospital Anxiety and Depression scale - a review of validation data and clinical results. *Journal of Psychosomatic Research*, **42**, 17-41.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). *Applied statistics for the behavioural sciences*. Boston: Houghton Mifflin Company.

Hlatky, M., & Spertus, J.A. (1998). Practical use of quality of life testing. *ACC Current Journal Review*, 90-93.

Hlatky, M.A., Boineau, R.E., Higginbotham, M.B., Lee, K.L., Mark, D.B., Califf, R.M., Cobb, F.R., & Pryor, D.B. (1989). A brief self-administered questionnaire to determine functional capacity (the Duke Activity Status Index). *American Journal of Cardiology*, **64**, 651-654.

Hoeg, J.M. (1997). Evaluating coronary heart disease risk. Tiles in the mosaic. *Journal of the American Medical Association*, **277**, 1387-1390.

Horgan, J.H., Bethell, H., Carson, P., Davidson, C., Julian, D., Mayou, R.A., & Nagle, R. (1992). British Cardiac Society Working Party report on cardiac rehabilitation. *British Heart Journal*, **67**, 41-418.

Houston Miller, N., & Barr Taylor, C. (1995). *Lifestyle management for patients with coronary heart disease*. Champaign, IL: Human Kinetics.

Hunt, S. (1999). The researcher's tale: a story of virtue lost and regained. In C.R. Joyce, C.A. O'Boyle & H.M. McGee (Eds.), *Individual quality of life approaches to conceptualisation and assessment*. (pp. 225-232). Amsterdam: Harwood Academic.

Hunt, S., McEwen, J., & McKenna, S. (1980). A quantitative approach to perceived health. *Journal of Epidemiology and Community Health*, **34**, 281-295.

Hunt, S., McEwen, J., & McKenna, S. (1985). Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners*, **35**, 185-188.

Hyland, M.E. (1992). A reformulation of quality of life for medical science. *Quality of Life Research*; **1**, 267-272.

Idler, E.I. (1992). Self-assessed health and mortality: A review of studies. In S. Maes, H. Leventhal, & M. Johnston (Eds.), *International review of health psychology* (pp 33-54). Chichester: John Wiley & Sons Ltd.

Jaesche, R., Singer, J., & Guyatt, G.H. (1989). Measurements of health status: ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, **10,** 407-415.

Jeng, C., & Braun, L. (1997). The influence of self-efficacy on exercise intensity, compliance rate and cardiac rehabilitation outcomes among coronary artery disease patients. *Progress in Cardiovascular Nursing* , **12**, 13-24.

Jenkins, C.D., Jono, R.T., Stanton, B.A., & Stroup-Benahm, C.A. (1990). The measurement of health-related quality of life: major determinants identified by factor analysis. *Social Science and Medicine*, **31**, 925-931.

Jenkinson, C. (1994). Measuring health and medical outcomes: an overview. In C. Jenkinson (Ed.), *Measuring health and medical outcomes* (pp.1-6). London: UCL Press.

Jenkinson, C., & McGee, H. (1998). *Health status measurement. A brief but critical introduction*. Oxford: Radcliffe Medical Press.

Jenkinson, C., Bardsley, M., & Lawrence, K. (1994). Measurement in subjective health assessment: themes and prospects. In C. Jenkinson (Ed.), *Measuring health and medical Outcomes* (pp. 176-1). London: UCL Press.

Jette, A.M. (1993). Using health-related quality of life measures in physical therapy outcomes research. *Physical Therapy*, **73**, 528-537.

Johnston. M. (1999, October). *Perceived control, mood, and activity limitations: Disease specific versus generic issues*. Paper presented at the 13th Conference of the European Health Psychology Society, Florence.

Jolliffe, J. (1998). Review requests information about unpublished research in the effectiveness of cardiac rehabilitation. *BACR Newsletter*, **2**, 3.

Jones, D. & West, R. (1995). A multi-centre randomised controlled trial of psychological rehabilitation. In D. Jones & R. West (Eds.), *Cardiac rehabilitation* (pp.207-226). London: BMJ Publishing Group.

Jones, D.A., & West, R.R. (1996). Psychological rehabilitation after myocardial infarction. Multicentre randomised controlled trial. *British Medical Journal*, **313**, 1517-21.

Joyce, C,R, O'Boyle C.A & McGee, H. (Eds.) (1999). *Individual quality of life approaches to conceptualisation and assessment*. Amsterdam: Harwood Academic.

Judd, C.M., McClelland, G.H., & Culhane, S.E. (1995). Data analysis: continuing issues in everyday analysis of psychological data. *Annual Review of Psychology*, **46**, 433-465.

Julian, D.G. & Cowan, J.C. (1992). *Cardiology* (6th Edition). London: Bailliere Tindall.

Jung, B.C. & Conn, G. (1995). Letter to the editor. *Journal of the American Medical Association,* **273**, 844.

Juniper, E.F., Guyatt, G.H., Willan, A., & Griffith, L.E. (1994). Determining a minimal important change in a disease-specific quality of life questionnaire. *Journal of Clinical Epidemiology,* **47**, 81-87.

Kaplan, R., & Bush, J.W. (1982). Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychology* , **1**, 61-80.

Katz, J.N., Larson, M.G., Phillips, C.B., Fossel, A.H., & Liang, M.H. (1992). Comparative measurement sensitivity of short and longer health status instruments. *Medical Care*, **30**, 917-925.

Kawachi, I., Sparrow, D., Vokonaas, P.S., & Weiss, S.T. (1994). Symptoms of anxiety and risk of coronary heart disease in the Normative Aging Study. *Circulation*, **90**, 2225-2229.

Kazis, L., Anderson, J., & Meenan, R. (1989). Effect sizes for interpreting changes in health status. *Medical Care* , **27**, S178-S189.

Ketterer, M.W. (1993). Secondary prevention of ischemic heart disease. The case for aggressive behavioural monitoring and intervention. *Psychosomatics*, **34**, 478-484.

Kobashigawa, J.A., Leaf, D.A., Lee, N., Gleeson, M.P., Liu, H., Hamilton, M.A., Moriguchi, J.D., Kawata, N., Einhorn, K., Herlihy, E., & Laks, H. (1999). A controlled trail of exercise rehabilitation after heart transplantation. *New England Journal of Medicine,* **340**, 272-277.

Kugler, J. Seelbach, H., & Krüskemper, G.M. (1994). Effects of rehabilitation exercise programme on anxiety and depression in coronary patients. A meta analysis. *British Journal of Clinical Psychology*, **33**, 401-410.

Kullo, I.J., Edwards, W.D., & Schwartz, R.S. (1998) Vulnerable plaque: Pathology and clinical implication. *Annals of Internal Medicine*, **129**, 1050-1060.

Ladwig, K.H., Lehmacher, W., Roth, R., Breithardt, G., Budde, T., & Borggrefe, M. (1992). Factors which provoke post-infarction depression: results from the post-infarction late potential study (PILP). *Journal of Psychosomatic Research*, 36, 723-729.

Lavie, C.J., & Milani, R.V. (1995). Effects of cardiac rehabilitation programs on exercise capacity, coronary risk factors, behavioural characteristics, and quality of life in a large elderly cohort. *American Journal of Cardiology*, **76**, 177-179.

Law, M. (1998). Lipids and cardiovascular disease. In S. Yusuf, J.A. Cairns, A. John Camm, E.L. Fallen & B.J. Gersh (Eds.), *Evidence based cardiology* (pp. 191-205). London: BMJ Books.

Leplege, A., & Hunt, S. (1997). The problem of quality of life in medicine. *Journal of the American Medical Association*, **278**, 47-50.

Levin, L.A., Perk, J., & Hedback, B. (1991) Cardiac rehabilitation – cost analysis. *Journal of Internal Medicine*, **230**, 427-434.

Lewin, R. (1997). The psychological component. In D.R. Thompson, G.S. Bowman, D. P. deBono & A. Hopkins (Eds.), *Cardiac rehabilitation. Guidelines and audit standards* (pp.21-46). London: Royal College of Physicians of London.

Lewin, R.J., Ingelton, R., Newens, A.J., & Thompson, D.R. (1998). Adherence to cardiac rehabilitation guidelines: a survey of cardiac rehabilitation programmes in the United Kingdom. *British Medical Journal* , **316**, 1354-1355.

Liang, M.H. (1995). Evaluating measurement responsiveness. *Journal of Rheumatology*, **22**, 1191-1192.

Liang, M.H., Fossel, A.H., & Larson, M.G. (1990). Comparison of five health status instruments for orthopedic evaluation. *Medical Care*, **28**, 632-642.

Liang, M.H., Larson, M.G., Cullen, K.E., & Schwartz, J.A. (1985). Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis and Rheumatism*, **28**, 542-547.

Linden, W., Stossel, C., & Maurice, J. (1996). Psychosocial interventions for patients with coronary artery disease: a meta-analysis. *Archives of Internal Medicine*, **156**, 745-752.

Lohr, K.N., Aaronson, N.K., Alonso, J., Burnam, M.A., Patrick, D.L., Perrin, E.B., & Roberts, J.S. (1996). Evaluating quality-of-life and health status instruments: Development of scientific review criteria. *Clinical Therapeutics*, **18**, 979-992.

Lydick, E., & Epstein, R.S. (1993). Interpretation of quality of life changes. *Quality of Life Research*, **2**, 221-226.

Maes, S. (1992). Psychological aspects of cardiac rehabilitation in Europe. *British Journal of Clinical Psychology,* **31**, 473-83.

Manson, J.E., Tosteson, H., Ridker, P.M., Satterfield, S., Herbert, P., O'Connor, G.T., Buring, J.E., & Hennekens, C.H. (1992). The primary prevention of myocardial infarction. *New England Journal of Medicine*, **326**, 1406-1411.

Mark, D.B., Lam, L.C., Lee, K.L., Clapp-Channing, N.E., Williams, R.B., Pryor, D.B., Califf, R.M., & Hlatky, M.A. (1992). Identification of patients with coronary artery disease at high risk for loss of employment: a prospective validation study. *Circulation*, **86**, 1485-94.

Maynard A. (1996). Rationing health care. *British Medical Journal*, **313**, 1499-1501.

Mayou, R., & Byrant, B. (1993). Quality of life in cardiovascular disease. *British Heart Journal*, **69**, 460-466.

McDowell, I. & Newell, C. (1987). *Measuring health: A guide to rating scales and questionnaires*. New York: Oxford University Press.

McGee, H., Graham, T., Newton, H., & Horgan, J.H. (1994). Introduction. *The Irish Journal of Psychology*, **15**, 5-7.

McGee, H., Hevey, D, & Horgan, J.H. on behalf of the IACR. *Cardiac rehabilitation service provision in Ireland*. Manuscript in preparation.

McHorney, C.A., & Tarlov, A.R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research*, **4,** 293-307.

McLeod, A., & Coats, A.J.S. (1995). Medical aspects of cardiac rehabilitation. In A. Coats, H. McGee, H. Stokes, & D. Thompson (Eds.), *BACR guidelines for cardiac rehabilitation* (pp. 40-55). Oxford: Blackwell Science.

McMurray, J., & Hart, W. (1993). The economics of heart failure on the UK National Health Service. *European Heart Journal*, **14**, 133.

McNair, D., Lorr, M., & Droppleman, L. (1971). *Manual for the Profile of Mood States*. San Diego, CA: Educational Testing Service.

Meyers, D.G. (1996). Relative survival benefits of risk factor modifications. *American Journal of Cardiology*, **77**, 298-299.

Milani, R.V., Lavie, C.J., & Cassidy, M.M. (1996). Effects of cardiac rehabilitation and exercise training programs on depression in patients following major coronary events. *American Heart Journal*, **132**, 726-732.

Mossey, J.M., & Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health*, **72**, 800-808.

Mullen, P.D., Mains, D.A., & Velez, R. (1992). A meta-analysis of controlled trials of cardiac patient education. *Patient Education and Counselling*, **19**, 143-162.

National Heart, Lung, and Blood Institute Report of the task force on *Behavioral Research in Cardiovascular Lung, and Blood Health and Disease*. (1998). US Department of Health and Human services, Public Health Service, National Institutes of Health.

Naylor, C.D., & Llewellyn-Thomas, H.A. (1994). Can there be a more patient-centred approach to determining clinically important effect sizes for randomised treatment trials? *Journal of Clinical Epidemiology*, **47**, 787-795.

Normand, C. (1991). Economics, health, and the economics of health. *British Medical Journal*, **303**, 1572-1577.

Nunnally Jr, J.C. (1970). *Introduction to psychological measurement*. Tokyo: McGraw-Hill, Kogakusha.

O'Connor, G.T., Buring, J.E., Yusuf, S., Goldhaber, S.Z., Olmstead, E.M., Paffenbarger, R.S.Jr., & Hennekens, C.H. (1989). An overview of randomised trials of rehabilitation with exercise after myocardial infarction. *Circulation*, **80**, 234-244.

O'Keeffe, S.T., Lye, M., Donnellan, C., & Carmichael, D.N. (1998). Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. *Heart*, **80**, 377-382.

Oakes, M. (1986). *Statistical Inference: A commentary for the social and behavioural Sciences*. New York: Wiley.

Oldridge, N., & Rogowski, B. (1990). Self-efficacy and in-patient cardiac rehabilitation. *American Journal of Cardiology*, **66**, 362-362.

Oldridge, N., Furlong, W., Feeny, D, Torrance, G., Guyatt, G., Crowe, J., & Jones, N. (1993). Economic evaluation of cardiac rehabilitation after acute myocardial infarction. *American Journal of Cardiology*, **72**, 154-161.

Oldridge, N.B., Guyatt, G.H., Fischer, M.E., & Rimm, A.A. (1988). Cardiac rehabilitation after myocardial infarction. Combined experience of randomised clinical trials. *Journal of the American Medical Association*, **260**, 945-950.

Oppenheim, A.N. (1992). *Questionnaire design: interviewing and attitude measurement*. Pinter: London.

Ornish, D., Brown, S.E, Scherwitz, L.W., Billings, J.H., Armstrong, W.T., Ports, T.A., McLanahan, S.M, Kirkeeide, R.L., Brand, R.J., & Gould, K.L. (1990). Can lifestyle

changes reverse coronary heart disease? The Lifestyle Heart Trial. *Lancet*, **336**, 129-133.

Ornish, D., Scherwitz, L.W., Billings, J.H., Brown, S.E., Gould, K.L., Meritt, T.A., Sparler, S., Armstrong, W.T., Ports, T.A., Kirkeeide, R.L., Hogeboom, C., & Brand, R.J. (1998). Intensive lifestyle changes for reversal of coronary heart disease. *Journal of the American Medical Association*, **280**, 2001-2007.

Ottenbacher, K.J., & Barrett, K.A. (1989). Measures of effect size in the reporting of rehabilitation research. *American Journal of Physical Medicine & Rehabilitation*, **68**, 52-58.

Ottenbacher, K.J., & Barrett, K.A. (1990). Statistical conclusion validity of rehabilitation research. *American Journal of Physical Medicine & Rehabilitation*, **69**, 102-107.

Paquin, M.J.R. (1983). Beyond significant yet meaningless results in psychotherapy research. *Psychotherapy: Theory, Research and Practice*, **20**, 38-40.

Patrick, D.L., & Erikson, P. (1993). *Health status and health policy: Allocating resources to health care*. Oxford University Press: New York.

Pearson, T.A., McBride, P.E., Miller, N.H., & Smith, S.C. (1996). Organisation of preventive cardiology service (Task Force 8 of the 27th Bethesda Conference) *Journal of the American College of Cardiology*, **27**, 1039-1047.

Pekkanen, J., Linn, S., Heiss, G., Suchindran, C.M., Leon, A., Rifkind, B.M., & Tyroler, H.A. (1990). Ten year mortality from cardiovascular disease in relation to cholesterol lowering among men with and without preexisting cardiovascular disease. *New England Journal of Medicine*, **322**, 1700-1007.

Pitula, C.R., Burg, M.M., & Froelicher, E.S. (1999). Psychosocial risk factors: Assessment and intervention for social isolation. In N.K. Wenger, L. K. Smith, E.S. Froelicher, & P. McCall Comoss (Eds.), *Cardiac rehabilitation. A guide to practice in the 21st century* (pp. 279-286). New York: Marcel Dekker, Inc.

Ray, J.A.W., & Shadish, W.R. (1996). How interchangeable are different estimators of effect size? *Journal of Consulting and Clinical Psychology*, **64**, 1316-1325.

Reddy, K.S. (1998). Global perspective on cardiovascular disease. In S. Yusuf, J.A. Cairns, A. John Camm, E.L. Fallen, & B.J. Gersh (Eds.) *Evidence based cardiology* (pp. 147-164). London: BMJ Books.

Redelmeier D.A., Guyatt, G.H., & Goldstein, R.S. (1996). Assessing the minimal important difference in symptoms: A comparison of two techniques. *Journal of Clinical Epidemiology*, **49**, 1215-1219.

Relman, A. (1988). Assessment and accountability: the third revolution in medical care. *New England Journal of Medicine*, **319**, 1220-1222.

Revicki, D.A., & Kaplan, R.M. (1995). Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. In S. A. Shumaker & R. A. Berzon (Eds.), *The international assessment of health-related quality of life: Theory, translation, measurement and analysis* (pp. 125-135). Oxford: Rapid Communications.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev.ed). Newbury Park, CA: Sage.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis*. (pp 232-245). New York: Russell Sage Foundation.

Rosnow, R.L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counselling psychology. *Journal of Counselling Psychology*, **35**, 203-208.

Rosser, R. (1997). Quality of life assessment. In A. Baum, S. Newman, J. Weinman, R West, & C. McManus (Eds.), *Cambridge handbook of psychology, health and medicine* (pp. 310-314). Cambridge, England: Cambridge University Press.

Rozanski, A., Blumenthal, J.A., & Kaplan, J. (1999). Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation*, **99**, 2192-2217.

Rukholm, E., McGirr, M., & Potts, J. (1998). Measuring quality of life in cardiac rehabilitation clients. *International Journal of Nursing Studies*, **35**, 210-216.

Rumsfeld, J.S., MaWhinney, S., Carthy Jr, M., Shoyer, A.L.W., VillaNueva, C.B., O'Brien, M., Moritz, T.E., Henderon, W.G., Grover, F.L., Sethi, G.K., & Hammermeister, K.E. (1999). Health-related quality of life as a predictor of mortality following coronary artery bypass graft surgery. *Journal of the American Medical Association*, **281**, 1298-1303.

Sackett, D.L. Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, **312**, 71-72.

Salovey, P., Rothman, A.J., Detweiler, J.B. & Steward, W.T. (2000). Emotional states and physical health. *American Psychologist*, **55**, 110-121.

Scandinavian Simvastatin Survival Study group. (1995). Baseline serum cholesterol and treatment effect in the Scandinavian Simvastatin Survival Study (4S) *Lancet*, **345**, 1274-1275.

Scheier, M.F., Matthews, K.A., Owens, J.F., Magovern, G.J., Sr., Lefebvre, R.C., Abbott, R.A., & Carver, C.S. (1989). Dispositional optimism and recovery from coronary artery bypass surgery: The beneficial effects on physical and psychological well-being. *Journal of Clinical and Consulting Psychology*, **57**, 1024-1040.

Schwartz, C.E., & Sprangers, M.A.G. (1999). Introduction to symposium on the challenge of response shift in social science and medicine. *Social Science and Medicine*, **48**, 1505-1506.

Schwarz, N. (1999). Self-reports. How the questions shape the answers. *American Psychologist*, **54**, 93-105.

Sedlmeier, P. & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, **105**, 309-316.

Seligman, M.P. & Csikszentmihalyi, M. (2000) Positive psychology: An introduction. *American Psychologist*, **55**, 5-14.

Shanfield, S.B. (1990). Return to work after acute myocardial infarction: a review. *Heart and Lung*, **19**, 109-116.

Shelly, E., Daly, L., Graham, I., et al. The Kilkenny Health Project: a community research and demonstration cardiovascular health programme. *Irish Journal of Medical Science*, Suppl.10-16.

Shrout, P.E. (1997). Should significance tests be banned? Introduction to a special section exploring the pros and cons. *Psychological Science*, **8**, 1-2.

Shumaker, S.A., & Naughton, M.J. (1995). The international assessment of health-related quality of life: a theoretical perspective. In S. A. Shumaker & R. A. Berzon (Eds.), *The international assessment of health-related quality of life: Theory, translation, measurement and analysis* (pp 3-10). Oxford: Rapid Communications.

Silverstone, P.H. (1990). Depression increases mortality and morbidity in acute-life-threatening medical illness. *Journal of Psychosomatic Research*, **34**, 651-657.

Sotile, W.M. (1999). Psychosocial risk factors: overview, assessment, and intervention for anger and hostility. In N.K. Wenger, L.K. Smith, E.S. Froelicher, P. McCall Comoss (Eds.), *Cardiac rehabilitation. A guide to practice in the 21st century* (pp. 257-262). New York: Marcel Dekker, Inc.

Spangers, M.A.G., & Aaronson, N.K. (1992). The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *Clinical Epidemiology* , **45**, 743-760.

Spielberger, C., Gorsuch, R., & Lushene, R. (1970). *The State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Sprangers, M.A.G., & Schwartz, C.E. (1999). Integrating response shift into health-related quality of life research: a theoretical model. *Social Science and Medicine*, **48**, 1507-1515.

Stehbens, W.E. (1999). Misuse of "coronary heart disease". *Heart*, **82**, 1-2.

Streiner, D.L., & Norman, G.R. (1995). *Health measurement scales: a practical guide to their development and use* (Second Edition). Oxford: Oxford University Press.

Stucki, G., Liang, M.H., Fossel, A.H., & Katz, J.N. (1995). Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *Journal of Clinical Epidemiology*, **48**, 1369-1378.

Tarlov, A.R., Ware, J.E., Greenfield, S., Nelson, E.C., Perrin, E., & Zubruff, M. (1989). The Medical Outcomes Study: an application of methods for monitoring the results of medical care. *Journal of the American Medical Association*, **262**, 925-930.

Taylor, R., Kirby, B., Burdon, D., & Caves, R. (1998). The assessment of recovery in patients after myocardial infarction using three generic quality-of-life measures. *Journal of Cardiopulmonary Rehabilitation*, **18**. 139-156.

Taylor, S.E., Kemeny, M.E., Reed, G.M., Bower, J.E. Gruenwald, T.L. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, **55**, 99-109.

Testa, M.A., & Nackley, J.F. (1994). Methods for quality-of-life studies. *Annual Review of Public Health* , **15**, 535-559.

Thompson D.R., Meadows, K.A., & Lewin, R.J.P. (1998). Measuring quality of life inpatients with coronary heart disease. *European Heart Journal*, **19**, 693-695.

Thompson, D.R., Bowman, G.S., de Bono, D.P., & Hopkins, A. (1997). *Cardiac rehabilitation: Guidelines and national standards*. London: Royal College of Physicians of London.

Thompson, D.R., Bowman, G.S., Kitson, A.L., de Bono, D.P., & Hopkins, A. (1997). Cardiac rehabilitation services in England and Wales: a national survey. *International Journal of Cardiology*, **59**, 299-304.

Thompson, D.R., Bowman, G.S., Kitson, A.L., de Bono, D.P., & Hopkins, A. (1996). Cardiac rehabilitation in the United Kingdom: guidelines and national standards. *Heart,* **75**, 89-93.

Thompson, D.R., Webster, R.A., Cordle, C.J., & Sutton, T.W. (1987). Specific sources and patterns of anxiety in male patients with first myocardial infarction. *British Journal of Medical Psychology*, **60**, 343-348.

Todd, I & Cay, L. (1997) Overview. In D.R. Thompson, G.S. Bowman, D.P de Bono, & A. Hopkins (Eds.), *Cardiac rehabilitation: Guidelines and national standards*. (pp. 1-7). London: Royal College of Physicians of London.

Tudor-Smith, C., Nutbeam, D., Moore, L.& Catford, J. (1998) Effects of Heartbeat Wales programme over five years on behavioural risks for cardiovascular disease: quasi experimental comparison of results from Wales and a matched reference area. *British Medical Journal*, **316**, 818-22.

Tuley, M.R., Mulrow, C.D., & McMahan, C.A. (1991). Estimating and testing an index of responsiveness and the relationship of the index to power. *Journal of Clinical Epidemiology*, **44**, 417-421.

Valenti, L., Lim, L., Heller, R.F., & Knapp, J. (1996). An improved questionnaire for assessing quality of life after myocardial infarction. *Quality of Life Research*, **5**, 151-161.

Vanhees, L., McGee, H.M., Dugmore, L.D., Vuori, I., & Pentilla, U.R. (1999). *The carinex Survey. Current guidelines and practices in cardiac rehabilitation within Europe.* Leuven: Acco.

Visser, M.C., Koudstaal, P.J., Erdman, R.A.M., Deckers, J.W., Passchier, J., van Gijn, J., & Grobbee, D.E. (1995). Measuring quality of life in patients with myocardial infarction: a feasibility study of four questionnaires in The Netherlands. *Journal of Epidemiology and Community Health*, **49**, 513-517.

Wagner, E.H., LaCroix, A.Z., Grothaus, L.C., & Hecht, J.A. (1993). Responsiveness of health status measures to change among older adults. *Journal of the American Geriatric Society*, **41**, 241-248.

Wallston, K.A., Wallston, B.S., & DeVilis, R. (1978). Development of the Mulitidimensional Health Locus of Control (MHLOC) scales. *Health Education Monograph*, **6**, 161-170.

Ware, J. & Sherbourne, C. (1992). The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, **30**, 473-483.

Ware, J.E. (1995). What information do consumers want and how will they use it? *Medical Care*, **33**, JS22-JS30.

Watson, D. & Tellegen, A. (1985). Towards a consensual structure of mood. *Psychological Bulletin*, **98**, 219-235.

Wells, G., Tugwell, P. Kraag, G.R., Baker, P. Groh, J., & Redelmeier, D.A (1993). Minimum important difference between patients with rheumatoid arthritis: The patient's perspective. *Journal of Rheumatology*, **20**, 557-5650.

Wenger, N.K., Froelicher E.S., Smith, L.K., et al. (1995).*Cardiac Rehabilitation.* Clinical Practice Guideline No. 17, Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research and National Heart, Lung and Blood Institute. AHCPR Publication No. 96-0672.

Wenger, N.K., Mattson, M.E., Furberg, C.D., & Elinson, J. (1984). Assessment of quality of life in clinical trials of cardiovascular therapies. *American Journal of Cardiology*, **54**, 908-913.

White, K. (1967). Improved health statistics and health services systems. *Public Health Reports* , **82**, 847-54.

Wiebe, S., Rose, K., Derry, P., & McLachlan, R. (1997). Outcome assessment in epilepsy: Comparative responsiveness of quality of life and psychosocial instruments. *Epilepsia*, **38**, 430-438.

Wilkinson, L. and the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. Guidelines and Explanations. *American Psychologist*, **54**, 594-604.

Williams, R.B., & Chesney, M.A. (1993). Psychosocial factors and prognosis in established coronary artery disease. The need for research on interventions. *Journal of the American Medical Association*, **270**, 1860-1861.

Wilson, I.B., & Cleary, P.D. (1995). Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *Journal of the American Medical Association,* **273**, 59-65.

World Health Organisation. (1947). The constitution of the world health organisation. *WHO Chronicle*, **1**, 13.

World Health Organisation. (1993). *Needs and action priorities in cardiac rehabilitation and secondary prevention in patients with coronary heart disease*. Geneva: WHO Regional Office for Europe.

Wright, J.G. (1996). The minimal important difference: Who's to say what is important? *Journal of Clinical Epidemiology*, **49**, 1211-1222.

Wright, J.G., & Young, N.L. (1997). A comparison of different indices of responsiveness. *Journal of Clinical Epidemiology*, **50**, 239-246.

Wrzesniewski, K. (1977). Anxiety and rehabilitation after myocardial infarction. *Psychotherapy and Psychosomtics*, **27**, 41-46.

Zigmond, A., & Snaith, R. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica,* **67**, 361-370.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Adjective Check List*. Educational and Industrial Testing Service, San Diego, CA.

Zuckerman, M., Lubin, B., & Rinck, C. (1983). Construction of new scales for the multiple affect adjective checklist. *Journal of Behavioral Assessment*, **5**, 119-129.

Multidisciplinary cardiac rehabilitation studies:

1. Engblom, E., Hämäläinen, H., Lind, J., Mattlar, C-E, Ollila, S., Kallio, V., Inberg, M. and Knuts, L-R. (1992) Quality of life during rehabilitation after coronary artery bypass surgery. *Quality of Life Research* **3**: 207-213.
2. Hillers, T., Guyatt, G., Oldridge, N., Crowe, J., Willan, A., Griffith, L. and Feeny, D. (1994) Quality of life after myocardial infarction. *Journal of Clinical Epidemiology* **47**: 1287-96.
3. Oldridge, N., Streiner, D., Hoffman, R. and Guyatt, G. (1995) Profile of mood states and cardiac rehabilitation after acute myocardial infarction. *Medicine and Science in Sports and Exercise*: **27**: 900-05.
4. Mattlar, C-E., Engblom, E., Vesala, P., Vänttinen, E. and Knuts, L-R. (1991) The proportion of patients with cognitive impairment after coronary artery bypass surgery: An 8-month follow-up study. *Psychotherapy and Psychosomatics* **55**: 145-150.
5. Newton, M., Murtrie, N. and McArthur, J. (1991) The effects of exercise in a coronary rehabilitation programme. *Scottish Medical Journal* **36**: 38-41.
6. Davis, C., Hare, D., Rushford, N., Burrows, G. and Buxton, B. (1995) Specific psychological intervention reduces depression and pain after urgent coronary surgery - a prospective randomised study. *Circulation* **92** (Suppl 1): 491 (Abstract).
7. Turner, L., Linden, W., van der Wal, R. and Schamberger, M. (1995) Stress management for patients with heart disease: a pilot study. *Heart and Lung* **24**: 145-53.
8. Denollet, J. (1993) Emotional distress and fatigue in coronary heart disease: the Global Mood Scale (GMS). *Psychological Medicine* **23**: 111-21.
9. Denollet, J. and Brutsaert, D. (1995) Enhancing emotional well-being by comprehensive rehabilitation in patients with coronary heart disease. *European Heart Journal* **16**: 1070-78.

10. Denollet, J. (1994) Health complaints and outcome assessment in coronary heart disease. *Psychosomatic Medicine* **56**: 463-74.

11. Denollet, J. (1993) Sensitivity of outcome assessment in cardiac rehabilitation. *Journal of Consulting and Clinical Psychology* **61**: 686-95.

12. Erdman, R., Duivenvoorden, H., Verhage, F., Krazemier, M. and Hugenholtz, P. (1986) Predictability of beneficial effects in cardiac rehabilitation: a randomised clinical trial of psychosocial variables. *Journal of Cardiopulmonary Rehabilitation* **6**: 206-213.

13. van Dixhoorn, J., Duivendoorden, H., Pool, J. and Verhage, F. (1987) Psychic effects of physical training and relaxation therapy after myocardial infarction. *Journal of Psychosomatic Research* **34**: 327-37.

14. Lewin, B., Robertson, I., Cay, E., Irving, J. and Campbell, M. (1992) Effects of self-help post-myocardial-infarction rehabilitation on psychological adjustment and use of health services. *Lancet* **339**: 1036-40.

15. Dracup, K., Moser, D., Marsden, C., Taylor, S. and Guzy, P. (1991) Effects of a multidimensional cardiopulmonary program on psychosocial function. *American Journal of Cardiology* **68**: 31-34.

16. Oldridge, N., Guyatt, G., Jones, N., Crowe, J., Singer, J., Feeny, D., McKelvie, R., Runions, J., Streiner, D. and Torrance, G. (1991) Effects on quality of life with cardiac rehabilitation after acute myocardial infarction. *American Journal of Cardiology* **67**: 1084-1089.

17. Oldenburg, B., Martin, A., Greenwood, J., Bernstein, L. and Allan, R. (1995) A controlled trial of behavioral and educational intervention following coronary artery bypass surgery. *Journal of Cardiopulmonary Rehabilitation* **15**: 39-46.

Exercise only cardiac rehabilitation studies

18. Porcari, J., Sierzant, M., Bird, J., Herrewig, T. and Olson, P. (1991) Changes in anxiety and depression levels of CABG patients following a clinic or home-based phase II cardiac rehabilitation program. *Journal of Cardiopulmonary Rehabilitation* **11**: 314 (Abstract).

19. Blumenthal, J., Emery, C. and Rejeski, W. (1988) The effects of exercise training on psychosocial functioning after myocardial infarction. *Journal of Cardiopulmonary Rehabilitation* **8**: 183-193.

20. Worcester, M., Hare, D., Graeme Oliver, R., Reid, M. and Goble, A. (1993) Early programmes of low and high intensity exercise and quality of life after myocardial infarction. *British Medical Journal* **307**: 1244-1247.

21. Giese, H. and Schomer, H. (1986) Lifestyle changes and mood profile of cardiac patients after a rehabilitation program. *Journal of Cardiopulmonary Rehabilitation* **10**: 255-263.

Cardiac rehabilitation studies with no data available:

22. Bär, F., Höppener, P., Diederiks, J., Vinken, H., Bekkers, J., Hoofd, W.v.h., Appels, A and Wellens, H.J.J. (1988) Cardiac rehabilitation contributes to the restoration of leisure and social activities after myocardial infarction. *Journal of Cardiopulmonary Rehabilitation* **12**: 117-125.

23. Barr Taylor, C., Houston-Miller, N., Ahn, D., Haskell, W. and DeBusk, R. (1986) The effects of exercise training programs on psychosocial improvement in uncomplicated postmyocardial infarction patients. *Journal of Psychosomatic Research* **30**: 581-587.

24. Bertie, J., King, A., Reed, N., Marshall, A. and Ricketts, C. (1992) Benefits and weaknesses of a cardiac rehabilitation programme. *Journal of the Royal College of Physicians in London* **26**: 147-151.

25. Fridlund, B., Högstedt, B., Lidell, E. and Larsson, P. (1991) Recovery after myocardial infarction. *Scandinavian Journal of Caring Sciences* **5**: 23-32.

26. Holmbäck, A., Säwe, U. and Fagher, B. (1993) Training after myocardial infarction: Lack of long-term effects on physical capacity and psychological variables. *Archives of Physical Medicine and Rehabilitation* **75**: 551-4.

27. Kugler, J., Dimsdale, J., Hartley, H. and Sherwood, J. (1990) Hospital supervised vs home exercise in cardiac rehabilitation: Effects on aerobic fitness, anxiety, and depression. *Archives of Physical Medicine and Rehabilitation* **71**: 322-325.

28. Mc Hugh Schuster, P., Wright, C. and Tomich, P. (1995) Gender differences in the outcomes of participants in home programs compared to those in structured cardiac rehabilitation programs. *Rehabilitation Nursing* **20**: 93-101.

29. Oldenburg, B., Allan, R. and Frastier, G. (1989) The role of behavioural and educational interventions in the secondary prevention of coronary heart disease. In *Clinical & Abnormal Psychology*, ed. P. Lovibond, P. Wilson, pp. 429-439. Elsevier Science Publishers B.V., North-Holland.

30. Uniken Venema-van Uden, M., Zoeteweij, M., Erdman, R., van den Berg, G.M., Smeets, J.S., Weeda, H.W.H., Vermeulen, A., van Meurs-van Woezik, H. and Ebink, C. (1989) Medical, social and psychological recovery after cardiac rehabilitation. *Journal of Psychosomatic Research* **33**: 651-656.

31. van Dixhoorn, J. (1994) Significance of breathing awareness and exercise training for recovery after myocardial infarction. In *Clinical Applied Psychophysiology*, ed. J.G. Carlson, A.R. Seifert, N. Birbaumer, pp. 113-132. Plenum Press, New York.

32. van Dixhoorn, J. and Duivendoorden, H. (1989) Breathing awareness as a relaxation method in cardiac rehabilitation. In *Stress and Tension Control 3. Stress Management*, ed. F.J. McGuigan, W.E. Sime, J Macdonald Wallace, pp. 19-36. Plenum Press, New York.

APPENDICIES

APPENDIX 1: Summary details of intervention studies assessing psychosocial outcomes before and after multi-factorial cardiac rehabilitation intervention

APPENDIX 2: Summary details of intervention studies assessing psychosocial outcomes before and after exercise –only cardiac rehabilitation intervention

APPENDIX 3: Letter of invitation to comparison group from physician or consultant cardiologist

APPENDIX 4: Letter included in each questionnaire envelope outlining the study

APPENDIX 5: Demographic and clinical status questionnaire

APPENDIX 6: Duke Activity Status Index (DASI)

APPENDIX 7: Hospital Anxiety and Depression Scale (HADS)

APPENDIX 8: Health Complaints Scale (HCS)

APPENDIX 9: Quality of Life Index-Cardiac Version III (QLI-CVIII)

APPENDIX10: Spielberger State-Trait Anxiety Inventory Form X-1 (SSAI)

APPENDIX 11: MacNew Quality of life after Acute Myocardial infarction Questionnaire (MacNew QLMI)

APPENDIX 12: Cardiac Depression Scale (CDS)

APPENDIX 13: Global Mood Scale (GMS)

APPENDIX 14: Heart Patients Psychological Questionnaire (HPPQ)

APPENDIX 15: MOS 36-item short –form health survey (MOS SF-36)

APPENDIX 16: Global rating of change questionnaire

APPENDIX 17: Comparison of the response rate of each of the study centres

APPENDIX 18: Comparison of the response rate for each questionnaire at each centre over time

APPENDIX 19: Comparison of baseline demographic details for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires

APPENDIX 20: Responsiveness statistics for each subscale

APPENDIX 21: Percentage of responsiveness scoring at the lowest 10% and top 10% range of scores possible for each instrument at Time 1.

APPENDIX 1: Summary details of intervention studies assessing psychosocial outcomes before and after multi-factorial cardiac rehabilitation intervention

Summary details of intervention studies assessing psychosocial outcomes before and after multi-factorial cardiac rehabilitation interventions.

Measure & Study Number	Design	N	Population	Intervention	Group	Effect size
Beck Depression Inventory						2mths 8mths
(21 items) 1	RCT Two groups t1= pre-CABG	194 102 rehab 92 hospital	CABG	Exercise & psycho- social rehabilitation vs. Hospital-based treatment	Rehabilitation Hospital treatment	0.01 0.21 0.13 0.09
2†	Stratified RCT Two groups t1= pre-rehabilitation	201 99 rehab 102 usual care	MI	Exercise & psycho- social rehabilitation vs. Community care	Rehabilitation Community care	8wks 4mths 8mths 12mths 0.19 0.38 0.24 0.26 0.05 0.15 0.25 0.28
note: 13 item version						
3	Stratified RCT Two groups t1= pre-rehabilitation	187 93 rehab 94 usual care	MI	Exercise & psycho- social rehabilitation vs. Community care	Rehabilitation Community care	8wks 12mths 0.20 0.27 0.00 0.26

							2mths	8mths	30mths	60mths
	4†	RCT t1= pre-rehabilitation	167 95 rehab 72 control	CABG	Exercise & psycho-social rehabilitation vs. ControI	Rehabilitation Control	0.01 0.13	0.25 0.07	0.21 0.10	0.21 0.14
							10wks			
	5	RCT t1= pre-rehabilitation	22 12 rehab 10 control	MI	Exercise & psycho-social rehabilitation vs. usual care (control)	Exercise Control	0.64 0.19			
Cardiac Depression Scale							8wks			
(35 item version)	6	Random allocation Pre-posttest Three groups	90 30 per	Coronary surgery	Psychological care vs. Non specific care	Psychological care Non specific care Usual care	1.28 0.04 0.14			
Note:All groups received standard rehabilitation		t1= pre-intervention	group		vs. Usual care					

General Health						8wks
Questionnaire-30 7	Random allocation	24	MI or CABG	Stress management	Stress management &	
(30 items)	Pre-posttest	SM+ER 18		(SM) & exercise	exercise rehabilitation	0.80
	Two groups	ER 6		rehabilitation (ER)	Exercise rehabilitation	0.18
	t1= pre-intervention			vs.		
				Exercise		
				rehabilitation (ER)		
Global Mood Scale					Negative affect	3mths
(20 items) 8	Pre-posttest	120	MI	Exercise & psycho-	Rehabilitation	0.42
	Two groups	60 rehab	CABG	social rehabilitation	Control	0.08
	t1=pre-rehabilitation	60 control	PTCA	vs.	Positive affect	
				Usual care	Rehabilitation	0.64
					Control	0.17

					Negative affect	3mths
9	Pre-posttest	170	MI	Exercise & psycho-	Rehabilitation	0.60
	Two groups	85 rehab	CABG	social rehabilitation	Control	0.02
	t1= pre-rehabilitation	85 control	PTCA	vs.	Positive affect	
				Usual care	Rehabilitation	0.63
					Control	0.14
Health Complaints					Somatic complaints	3mths
Checklist (24 items) 10	Pre-posttest	120	MI	Exercise & psycho-	Rehabilitation	0.55
	Two groups	60 rehab	CABG	social rehabilitation	Control	0.08
	t1=pre-intervention	60 control	PTCA	vs.	Cognitive complaints	
				Usual care	Rehabilitation	0.45
					Control	0.15
					Somatic complaints	3mths
9	Pre-posttest	170	MI	Exercise & psycho-	Rehabilitation	0.48
	Two groups	rehab 85	CABG	social rehabilitation	Control	0.11
	t1= pre-intervention	control 85	PTCA	vs.	Cognitive complaints	
				Usual care	Reabilitation	0.48
					Control	0.08

Heart Patients					Well-being	3mths
Psychological	Pre-posttest	120	MI	Exercise & psycho-	Rehabilitation	0.55
Questionnaire (40 items)[11]	Two groups	rehab 60	CABG	social rehabilitation	Control	0.03
	t1= pre-intervention	control 60	PTCA	vs.	Disability	
				Usual care	Rehabilitation	0.87
					Control	0.15
					Well-being	3mths
9	Pre-posttest	170	MI	Exercise & psycho-	Rehabilitation	0.72
	Two groups	rehab 85	CABG	social rehabilitation	Control	0.13
	t1= pre-intervention	control 85	PTCA	vs.	Disability	
				Control	Rehabilitation	0.67
					Control	0.05

					Well-being	6mths	60mths
12	Pre-posttest	57	MI	Exercise & psycho-	Rehabilitation	0.34	0.16
	Random allocation	rehab 27		social rehabilitation	Control	0.12	0.16
	Two groups	control 30		vs.	Disability		
	t1= pre-intervention			Usual care	Rehabilitation	0.26	0.19
					Control	0.11	0.02
					Despondency		
					Rehabilitation	0.05	0.09
					Control	0.08	0.38
					Social inhibition		
					Rehabilitation	0.21	0.25
					Control	0.00	0.24

					Well-being	6wks
13	Pre-posttest	137	MI	Exercise training &	Exercise &	
	Random allocation	relax 66		relaxation &	relaxation	0.63
	Two groups	exercise 71		breathing therapy	Exercise	0.15
	t1= pre-intervention			vs.	Disability	
				Exercise training	Exercise &	
					relaxation	0.29
					Exercise	0.13
					Despondency	
					Exercise &	
					relaxation	0.15
					Exercise	0.07
					Social inhibition	
					Exercise &	
					relaxation	0.03
					Exercise	0.04

Hospital Anxiety &					Anxiety	6wks	6mths	12mths
Depression Scale [14]	Pre-posttest	176	MI	Exercise & psycho-	Rehabilitation	0.74	0.60	0.71
(14 items)	Stratified Random	rehab 88		social rehabilitation	Control	0.25	0.15	0.10
	allocation	control 88		vs.	Depression			
	Two groups			Usual care	Rehabilitation	0.61	0.42	0.52
	t1= pre-intervention				Control	0.03	0.19	0.06
Multiple Affect Adjective					Anxiety	6mths		
Checklist (132 items) [15]	Pre-posttest	141	MI	Exercise & psycho-	Rehabilitation	0.50		
	Two groups	rehab 41	CABG	social rehabilitation	Control	0.00		
	t1= pre-intervention	control 100		vs.	Depression			
				Control	Rehabilitation	0.57		
					Control	0.00		

Profile of Mood States					Tension	8wks	4mths	8mths	12mths
(65 items)	Stratified, random	201	MI	Exercise & psycho-	Rehabilitation	0.37	0.44	0.41	0.51
2†	allocation	99 rehab		social rehabilitation	Community care	0.35	0.29	0.56	0.63
	Two groups	102		vs.	Depression				
	t1=pre-rehabilitation	control		Usual community	Rehabilitation	0.19	0.21	0.21	0.24
				care	Community care	0.10	0.06	0.29	0.31
					Anger				
					Rehabilitation	0.18	0.17	0.18	0.32
					Community care	0.04	0.09	0.14	0.17
					Vigour				
					Rehabilitation	0.60	0.42	0.29	0.58
					Community care	0.31	0.43	0.24	0.58
					Fatigue				
					Rehabilitation	0.34	0.30	0.23	0.25
					Community care	0.19	0.20	0.23	0.33
					Confusion				
					Rehabilitation	0.26	0.28	0.17	0.30
					Community care	0.06	0.14	0.35	0.31

					Tension	10wks
5†	RCT	22	MI	Exercise & psycho-	Exercise	0.80
	t1= pre-rehabilitation	12 rehab		social	Control	0.17
		10 control		vs.	Depression	
				usual care	Exercise	0.54
					Control	0.62
					Anger	
					Exercise	0.46
					Control	0.55
					Vigour	
					Exercise	0.18
					Control	0.13

					Fatigue	
					Exercise	0.56
					Control	0.18
					Confusion	
					Exercise	0.38
					Control	0.62
					TMD	
					Exercise	0.57
					Control	0.87
Psychosocial Adjustment to Illness Scale (46 items) 15	Pre-posttest Two groups t1= pre-intervention	141 rehab 41 control 100	MI CABG	Exercise & psycho-social rehabilitation vs. Control	Total Exercise Control	6mths 0.67 0.18

Quality of Life after					Total	8wks	4mths	8mths	12mths
Myocardial Infarction	Stratified, random	201	MI	Exercise & psycho-	Rehabilitation	0.99	1.13	1.14	1.16
(26 items) 2†	allocation	99 rehab		social rehabilitation	Community care	0.62	0.81	0.91	1.12
	Two groups	102		vs.	Self-Esteem				
	t1=pre-rehabilitation	control		Usual care	Rehabilitation	0.54	0.63	0.65	0.69
					Community care	0.35	0.49	0.55	0.73
					Restriction				
					Rehabilitation	1.24	1.39	1.43	1.47
					Community care	0.89	1.06	1.12	1.26
					Fatigue/physical				
					Rehabilitation	0.34	0.30	0.23	0.25
					Community care	0.19	0.20	0.23	0.33
					Emotion				
					Rehabilitation	0.34	0.46	0.42	0.46
					Community care	0.04	0.20	0.35	0.53
					Confidence				
					Rehabilitation	1.24	1.40	1.49	1.49
					Control	0.87	1.09	1.19	1.36

16†	Stratified, random allocation Two groups t1=pre rehabilitation	201 99 rehab 102 control	MI	Exercise & psycho-social rehabilitation vs. Usual care	<u>Limitations</u> Rehabilitation Control <u>Emotions</u> Rehabilitation Control <u>Total</u> Rehabilitation Control	<u>8wks</u> 1.05 0.75 0.81 0.44 0.99 0.63	<u>4mths</u> 1.14 0.89 0.95 0.63 1.12 0.83	<u>8mths</u> 1.13 0.93 0.99 0.74 1.14 0.93	<u>12mths</u> 1.17 1.09 1.01 0.94 1.18 1.13
Note: Same study as above but analyis of questionnaire using different factors									
<u>Spielberger State Anxiety Inventory</u> (20 items) 2†	Stratified, random allocation Two groups t1=pre-rehabilitation	201 99 rehab 102 control	MI	Exercise & psycho-social rehabilitation vs. Usual care	Rehabilitation Community care	<u>8wks</u> 0.59 0.16	<u>4mths</u> 0.65 0.31	<u>8mths</u> 0.45 0.46	<u>12mths</u> 0.55 0.43
3	Stratified, Random allocation Two groups t1= pre-rehabilitation	187 93 rehab 94 usual care	MI	Exercise & psycho-social rehabilitation vs. Usual care	Rehabilitation Control	<u>8wks</u> 0.56 0.11	<u>12mths</u> 0.56 0.44		

						6wks		
13	Pre-posttest Random allocation Two groups t1= pre-intervention	137 relax 66 exercise 71	MI	Exercise training & relaxation & breathing therapy vs. Exercise training	Exercise & relaxation Exercise	0.26 0.06		
Symptom Checklist 90 (90 items)					Depression, Anxiety & Somatisation	4mths	8mths	12mths
17	RCT Two groups t1 = pre-intervention	86 rehab 43 control 43	CABG	Exercise & psycho- social rehabilitation vs. Usual care	Rehabilitation Control	0.49 0.41	0.53 0.43	0.45 0.26

Note: psychosocial measures listed alphabetically
[†]Unpublished information from the study provided by authors;
RCT: randomised controlled trial; MI: myocardial infarction; CABG: coronary artery bypass graft; CHD: coronary heart disease

APPENDIX 2: Summary details of intervention studies assessing psychosocial outcomes before and after exercise –only cardiac rehabilitation intervention

Summary details of intervention studies assessing psychosocial outcomes before and after exercise only cardiac rehabilitation interventions.

Measure & Study	Design	N	Population	Intervention	Group	Effect size
Beck Depression Inventory						8wks
(21 items)	Pre-posttest	36	CABG	Clinic-based	Clinic based	0.71
18	Two groups	27 home		exercise	Home based	0.37
	t1= pre-intervention	9 clinic		vs.		
				Home-based		
				exercise		
Center for Epidemiological						12 wks
Studies-Depression (20items)	Random allocation	61	MI	High intensity	High intensity	0.14
19	(t1=pre-intervention)	high 33		exercise	Low intensity	0.15
		low 28		vs.		
				Low intensity		
				exercise		

Institute for Personality and					Depression	4mths 12mths
Ability Testing Depression	Pre-posttest	188-173	MI	High intensity	High intensity	0.04 0.06
Scale 20*	Random allocation			exercise	Low intensity	0.10 0.03
	Two groups			vs.		
	t1= pre-intervention			Low intensity		
				exercise		
Perceived Stress Scale						12wks
(14 items) 19	Random allocation	61	MI	High intensity	High intensity	0.19
	Pre-posttest	high 33		exercise	Low intensity	0.03
	Two groups	low 28		vs.		
	t1= pre-intervention			Low intensity		
				exercise		

Profile of Mood States					Tension	10mths
(65 items)	21	Pre-posttest	29	MI	Exercise	Rehabilitation	0.25
		Two groups	rehab 17		rehabilitation	Control	0.16
		t1=pre-rehabilitation	control 12		vs.	Depression	
					Non-exercise	Rehabilitation	0.13
						Control	0.08
						Anger	
						Rehabilitation	0.26
						Control	0.04
						Vigour	
						Rehabilitation	0.00
						Control	0.57
						Fatigue	
						Rehabilitation	0.11
						Control	0.19
						Confusion	
						Rehabilitation	0.10
						Control	0.07

Spielberger State Anxiety							12wks
Inventory		Random allocation	61	MI	High intensity	High intensity	0.00
(20 items)	19	Pre-posttest	high 33		exercise	Low intensity	0.02
		Two groups	low 28		vs.		
		t1= pre-intervention			Low intensity		
					exercise		
							8wks
	18	Pre-posttest	36	CABG	Clinic-based	Clinic based	1.73
		Two groups	27 home		exercise	Home based	0.87
		t1= pre-intervention	9 clinic		vs.		
					Home-based		
					exercise		
							4mths 12mths
	20*	Pre-posttest	200-173	MI	High intensity	High intensity	0.29 0.22
		Random allocation	high		exercise	Low intensity	0.14 0.10
		Two groups	100-83		vs.		
		t1= pre exercise	low		Low intensity		
			100-90		exercise		

Symptom Checklist 90						Anxiety	12wks
(90 items)	19	Random allocation	61	MI	High intensity	High intensity	0.07
		Pre-posttest	high 33		exercise	Low intensity	0.29
		Two groups	low 28		vs.	Phobic Anxiety	
		t1= pre-intervention			Low intensity	High intensity	0.16
					exercise	Low intensity	0.06
						Depression	
						High intensity	0.21
						Low intensity	0.07
						Hostility	
						High intensity	0.12
						Low intensity	0.07
						Grand symptom	
						index	
						High intensity	0.04
						Low intensity	0.16

UCLA Loneliness Scale						12wks				
(20 items) 19	Random allocation	61	MI	High intensity	High intensity	0.14				
	Pre-posttest	high 33		exercise	Low intensity	0.09				
	Two groups	low 28		vs.						
	t1= pre-intervention			Low intensity						
				exercise						

Note: psychosocial measures listed alphabetically.

† Unpublished information from the study provided by authors;

* Baseline standard deviation calculated from baseline standard error

RCT: randomised controlled trial; MI: myocardial infarction; CABG: coronary artery bypass graft; .CHD: coronary heart disease

APPENDIX 3: Letter of invitation to comparison group from physician or consultant cardiologist

Dear ,

I am writing to invite you to participate in a national study looking at quality of life and well-being in patients after a cardiac event (heart attack or cardiac surgery). These types of questions will be used to assess how well people are doing after a cardiac event and to identify what specific types of help and support are needed. We hope to learn from your experiences in order to help us provide a better service for others in the future. We would very much appreciate your help in our efforts to study quality of life of cardiac patients.

The study involves completing two questionnaires which ask about your health and well-being. Shortly after agreeing to participate, you will be posted a form which is to be completed and returned to the main study centre at Beaumont Hospital. You will then be asked to fill in one more form 6 weeks after filling in the first form.

If you are interested in participating in this study, then simply fill in the attached form and return it in the enclosed stamped addressed to the main study centre at Department of Cardiology, Beaumont Hospital, Dublin 9. All the postal costs involved in this study will be paid for by the researchers. Participation in the study is voluntary and refusal to participate will not affect your medical care.

We guarantee that the information given will be kept confidential. If you have any enquiries regarding any aspect of the study, then please feel free to contact Mr. David Hevey (main researcher with the project; tel. 01) 809 2451) at the Department of Cardiology, Beaumont Hospital, Dublin 9.
Thank you for your consideration and we look forward to hearing from you.

Yours sincerely,

APPENDIX 4: Letter included in each questionnaire envelope outlining the study

Dear ,

CARDIAC REHABILITATION STUDY

We are currently studying different measures to assess quality of life in patients doing cardiac rehabilitation. In the future, such questions will assess how health improves on cardiac rehabilitation programmes. This will be used to identify what specific types of help and support are needed. The information will also be used to demonstrate the benefits of cardiac rehabilitation so that rehabilitation services may be available to all who would benefit. We hope to learn from your experiences so that we can improve the service further in the future. We would very much appreciate your help in our efforts to evaluate these measures.

We ask you to fill in the attached form and return it in the enclosed envelope at the next rehabilitation session. We will then ask you to fill in one more form at the end of your rehabilitation programme.

The information given is confidential to the researchers and will be used only to evaluate the measures. If you have any comments or questions, we are glad to be of help during your hospital visits. Please feel free also to contact Mr. David Hevey (researcher with the project; tel. 01) 809 3231) at the Department of Cardiology, Beaumont Hospital, Dublin 9.

It is important for future patients and for the development of cardiac rehabilitation services that we are able to demonstrate its benefits and we very much appreciate your time in helping us to do this.

Thank you for your consideration and we look forward to hearing from you.

Yours sincerely,

APPENDIX 5: Demographic and clinical status questionnaire

Ref No.:

DEMOGRAPHIC INFORMATION

Name: _____
First name Surname

Address: _____

_____

Telephone: _____

Sex: ☐ Male ☐ Female **Date of birth** ☐☐ ☐☐ ☐☐☐☐
day month year

Today's date: ☐☐ ☐☐ ☐☐☐☐

Marital status: ☐ Married ☐ Single ☐ Widowed
☐ Other (please explain) _____

Occupation: _____

Occupational Status: ☐ Full-time ☐ Part-time ☐ Retired
☐ Unemployed ☐ Other(please explain) _____

Education: ☐ Primary ☐ Completed 2nd level ☐ Third level

Residential Location: ☐ City ☐ Town ☐ Village ☐ Rural
☐ Other (please explain) _____

Do you live: ☐ Alone
☐ With others (please specify) _____

_____

How many weeks is it since your most recent cardiac event (heart attack, coronary artery bypass, coronary angioplasty, etc.)? _____

Most recent cardiac event ☐ MI (heart attack) ☐ CORONARY ARTERY BYPASS
☐ ANGIOPLASTY ☐ OTHER (please explain) _____

Cardiac history before this event: ☐ No previous events ☐ ANGIOPLASTY
☐ MI (heart attack) ☐ CORONARY ARTERY BYPASS
☐ OTHER (please explain) _____

DATE OF YOUR FIRST CARDIAC REHAB SESSION: ☐☐ ☐☐ ☐☐☐☐
day month year

We would now like to ask you some questions concerning your general medical history.

Please tick the appropriate boxes:

Do you have any of the following chronic diseases:

Asthma ☐
Chronic obstructive pulmonary disease ☐
Congestive heart failure ☐
Diabetes ☐
Renal disease, or kidney damage ☐
Rheumatoid arthritis ☐
Hypertension (high blood pressure) ☐
Alzheimers or another form of dementia ☐
Cirrrhosis, or serious liver damage ☐
Leukemia or lymphoma ☐
Cancer ☐
Mental health problems ☐

Have you ever had a:

Stroke ☐

Surgery ☐ (reason) _____

_____

_____

Please list all medications that you are currently taking:

Medication	Pupose
1)	
2)	
3)	
4)	
5)	
6)	

APPENDIX 6: Duke Activity Status Index (DASI)

s scale is designed to identify what physical activities you feel you are capable of performing. We are not interested in whether u usually do these activities but rather what you think you could do at the moment. Please tick the most appropriate box.

	Yes	No
. Can you take care of yourself, that is, eating, dressing, bathing or using the toilet ?		
Can you walk indoors, such as around your house ?		
Can you walk a street or 2 on ground level ?		
Can you climb a flight of stairs or walk up a hill ?		
Can you run a short distance ?		
Can you do light work around the house like dusting or washing dishes ?		
Can you do moderate work around the house like vacuuming, sweeping floors, or carrying in eries ?		
Can you do heavy work around the house like scrubbing floors, or lifting or moving heavy iture ?		
Can you do yardwork like raking leaves, weeding or pushing a power mower ?		
. Can you have sexual relations ?		
. Can you participate in moderate recreational activities like golf, bowling, dancing, doubles tennis hrowing a ball or football ?		
. Can you participate in strenuous sports like swimming, singles tennis, football, basketball or ng ?		

APPENDIX 7: Hospital Anxiety and Depression Scale (HADS)

is questionnaire is designed to help us know how you feel. Read each item and place a firm tick in the box posite the reply which comes closest to how you have been feeling **in the past week**.
n't take too long over your replies; your immediate reaction to each item will probably be more accurate than ng thought-out response.

el tense or wound up:

- Most of the time..... ☐
- A lot of the time..... ☐
- Time to time, occasionally..... ☐
- Not at all..... ☐

I feel as if I am slowed down:

- Nearly all the time..... ☐
- Very often..... ☐
- Sometimes ..... ☐
- Not at all ..... ☐

ill enjoy the things I used to enjoy:

- Definitely as much..... ☐
- Not quite so much..... ☐
- Only a little..... ☐
- Hardly at all..... ☐

I get a sort of frightened feeling like "butterflies" in the stomach:

- Not at all..... ☐
- Occasionally..... ☐
- Quite often..... ☐
- Very often..... ☐

et a sort of frightened feeling as if mething awful is about to happen:

- Very definitely and very badly..... ☐
- Yes, but not too badly..... ☐
- A little, but it doesn't worry me..... ☐
- Not at all..... ☐

I have lost interest in my appearance:

- Definitely..... ☐
- I don't take as much care as I should ☐
- I may not take quite as much care ... ☐
- I take just as much care as ever..... ☐

an 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..... ☐

I feel restless as if I have to be on the move:

- Very much indeed..... ☐
- Quite a lot..... ☐
- Not very much..... ☐
- Not at all..... ☐

rrying thoughts go through my mind:

- A great deal of the time..... ☐
- A lot of the time..... ☐
- Sometimes, but not too often..... ☐
- Only occasionally..... ☐

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..... ☐

el cheerful:

- Not at all..... ☐
- Not often..... ☐
- Sometimes..... ☐
- Most of the time..... ☐

I get sudden feelings of panic:

- Very often indeed..... ☐
- Quite often..... ☐
- Not very often..... ☐
- Not at all..... ☐

n sit at ease and feel relaxed:

- Definitely..... ☐
- Usually..... ☐
- Not often..... ☐
- Not at all..... ☐

I can enjoy a good book, radio or TV programme:

- Often..... ☐
- Sometimes..... ☐
- Not often..... ☐
- Very seldom..... ☐

APPENDIX 8: Health Complaints Scale (HCS)

HCS

Name: ..... Sex: ..... Age: ..... Date: .....

Below are a number of problems and complaints that ill people often have. Please read **each item** carefully and then **circle the appropriate number** next to that problem. Indicate how much each problem has **bothered you lately**. Please use the following scale to record your answers.

0= NOT AT ALL 1= A LITTLE BIT 2= MODERATELY 3= QUITE A BIT 4= EXTREMELY

---

ately, how much were you bothered
y the following **specific** problems :

A1	Sleep that is restless or disturbed	→ 0 1 2 3 4
A2	Tightness of the chest	→ 0 1 2 3 4
A3	Feeling that you are not rested	→ 0 1 2 3 4
A4	Fatigue	→ 0 1 2 3 4
A5	Trouble falling asleep	→ 0 1 2 3 4
A6	Inability to take a deep breath	→ 0 1 2 3 4
A7	Stabbing pain in heart or chest	→ 0 1 2 3 4
A8	Feeling exhausted without any reason	→ 0 1 2 3 4
A9	Shortness of breath	→ 0 1 2 3 4
A10	Pain in heart or chest	→ 0 1 2 3 4
A11	Feeling weak	→ 0 1 2 3 4
A12	Feeling you can't sleep	→ 0 1 2 3 4

Lately, how much were you bothered
by the following **general** problems :

B1	The idea that your bad health is the biggest problem in your life	→ 0 1 2 3 4
B2	Not being able to work fluently, also with hobbies	→ 0 1 2 3 4
B3	Being afraid of illness	→ 0 1 2 3 4
B4	The idea that you were able to take on much more work formerly	→ 0 1 2 3 4
B5	Feeling blocked in getting things done	→ 0 1 2 3 4
B6	The idea that you have a serious illness	→ 0 1 2 3 4
B7	Feeling you are not able to do much	→ 0 1 2 3 4
B8	The idea that something serious is wrong with your body	→ 0 1 2 3 4
B9	Feeling you are no longer worth as much as you used to be	→ 0 1 2 3 4
B10	Feeling despondent	→ 0 1 2 3 4
B11	Worrying about your health	→ 0 1 2 3 4
B12	Thinking that all your worries would be over if you were physically healthy	→ 0 1 2 3 4

APPENDIX 9: Quality of Life Index-Cardiac Version III (QLI-CVIII)

rt 1: For each of the following, please tick the box that best describes how satisfied you are with at area of your life. There are no right or wrong answers.

OW SATISFIED ARE OU WITH:	Very unsatisfied	Moderately unsatisfied	Slightly unsatisfied	Slightly satisfied	Moderately satisfied	Very satisfied
Your health?						
The health care you are receiving?						
The amount of chest pain (angina) that you have?						
Your ability to breathe without shortness of breath?						
The amount of energy you have for everyday activities?						
Your physical independence?						
The amount of control you have over your life?						
Your potential to live a long time?						
Your family's health?						
) Your children?						
) Your family's happiness?						
) Your relationship with your spouse/significant other?						
) Your sex life?						
) Your friends?						
) The emotional support you get from others?						
) Your ability to meet family responsibilities?						

OW SATISFIED ARE OU WITH:	Very unsatisfied	Moderately unsatisfied	Slightly unsatisfied	Slightly satisfied	Moderately satisfied	Very satisfied
) Your usefulness to others?						
) The amount of stress or worries in your life?						
) Your home?						
) Your neighbourhood?						
) Your standard of living?						
) Your job (IF EMPLOYED) or go to next question						
) Not having a job (IF UNEMPLOYED)						
) Your education?						
) Your financial independence?						
) Your leisure time activities?						
) Your ability to travel on vacations?						
) Your potential for a happy old age/retirement?						
) Your peace of mind?						
)Your personal faith in God?						
) Your achievement of personal goals?						
) Your happiness in general?						
)Your life in general?						
) Your personal appearance?						
) Yourself in general?						
) The changes in your life t you have had to make cause of your heart problem e. changes in diet, physical ivity and/or smoking?)						

rt 2: For each of the following, please tick the box that best describes how important that area of e is to you. There are no right or wrong answers.

OW IMPORTANT TO YOU	Very unimportant	Moderately unimportant	Slightly unimportant	Slightly important	Moderately important	Very important
Your health?						
Health care ?						
Being completely free of chest pain?						
Being able to breathe without shortness of breath?						
Having enough energy for everyday activities?						
Your physical independence?						
Having control over your life?						
Living a long time?						
Your family's health?						
) Your children?						
) Your family's happiness?						
) Your relationship with your spouse/significant other?						
) Your sex life?						
) Your friends?						
) The emotional support you get from others?						
) Meet family responsibilities?						

OW IMPORTANT TO OU IS:	Very unimportant	Moderately unimportant	Slightly unimportant	Slightly important	Moderately important	Very important
) Being useful to others?						
) Having a reasonable amount of stress or worries?						
) Your home?						
) Your neighbourhood?						
) A good standard of living?						
) Your job? EMPLOYED) or go to next question						
) To have a job? (IF UNEMPLOYED)						
) Your education?						
) Your financial independence?						
) Leisure time activities?						
) The ability to travel on vacations?						
) having a happy old age/retirement?						
) Peace of mind?						
)Your personal faith in God?						
) Achieving your personal goals?						
) Your happiness in general?						
) Being satisfied with life?						
) Your personal appearance?						
) Yourself?						
) The changes in your life t you have had to make cause of your heart problem e. changes in diet, physical ivity and/or smoking?						

APPENDIX10: Spielberger State-Trait Anxiety Inventory Form X-1 (SSAI)

number of statements which people have used to describe themselves are given below. Read ach statement and then please tick the appropriate box to the right of the statement to indicate ow **you feel right now, that is, at this moment** . There are no right or wrong answers. Do not pend too much time on any one statement but give the answer which seems to describe your eelings best.

	Not at all	Somewhat	Moderately so	Very much so
I feel calm				
I feel secure				
I am tense				
I feel strained				
I feel at ease				
I feel upset				
I am presently worrying over possible misfortunes				
I feel satisfied				
I feel frightened				
0 I feel comfortable				
1 I feel self-confident				
2 I feel nervous				
3 I am jittery				
4 I feel indecisive				
5 I am relaxed				
6 I feel content				
7 I am worried				
8 I feel confused				
9 I feel steady				
0 I feel pleasant				

APPENDIX 11: MacNew Quality of life after Acute Myocardial infarction Questionnaire (MacNew QLMI)

Ve would now like to ask you some questions about how you have been feeling **DURING THE LAST 2 VEEKS.**

lease check the box ☐ that matches your answer

. In general, how much of the time during the last 2 weeks have you felt frustrated, impatient or angry?

ALL OF THE TIME..... ☐
MOST OF THE TIME..... ☐
A GOOD BIT OF THE TIME..... ☐
SOME OF THE TIME..... ☐
A LITTLE OF THE TIME..... ☐
HARDLY ANY OF THE TIME.... ☐

. How often during the last 2 weeks have you felt worthless or inadequate?

ALL OF THE TIME..... ☐
MOST OF THE TIME..... ☐
A GOOD BIT OF THE TIME..... ☐
SOME OF THE TIME..... ☐
A LITTLE OF THE TIME..... ☐
HARDLY ANY OF THE TIME..... ☐
NONE OF THE TIME..... ☐

. In the last 2 weeks, how much of the time did you feel very confident and sure that you could deal with your ear problem?

ALL OF THE TIME..... ☐
MOST OF THE TIME..... ☐
A GOOD BIT OF THE TIME..... ☐
SOME OF THE TIME..... ☐
A LITTLE OF THE TIME..... ☐
HARDLY ANY OF THE TIME..... ☐
NONE OF THE TIME..... ☐

. In general how much of the time did you feel discouraged or down in the dumps during the last 2 weeks?

ALL OF THE TIME..... ☐
MOST OF THE TIME..... ☐
A GOOD BIT OF THE TIME..... ☐
SOME OF THE TIME..... ☐
A LITTLE OF THE TIME..... ☐
HARDLY ANY OF THE TIME..... ☐
NONE OF THE TIME..... ☐

How much of the time during the past 2 weeks did you feel relaxed and free of tension?

- NONE OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- ALMOST ALL OF THE TIME..... ☐
- ALL OF THE TIME..... ☐

How often during the last 2 weeks have you felt worn out or low in energy?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

How happy, satisfied, or pleased have you been with your personal life during the last 2 weeks?

- VERY DISSATISFIED, UNHAPPY MOST OF THE TIME..... ☐
- GENERALLY DISSATISFIED, UNHAPPY..... ☐
- SOMEWHAT DISSATISFIED, UNHAPPY..... ☐
- GENERALLY SATISFIED, PLEASED..... ☐
- HAPPY MOST OF THE TIME..... ☐
- VERY HAPPY MOST OF THE TIME..... ☐
- EXTREMELY HAPPY, COULD NOT HAVE BEEN MORE SATISFIED OR PLEASED..... ☐

In general how often during the last 2 weeks have you felt restless, or as if you were having difficulty trying to calm down?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

9. How much shortness of breath have you experienced during the last 2 weeks while doing your day-to-day physical activities?

- EXTREME SHORTNESS OF BREATH..... ☐
- VERY SHORT OF BREATH..... ☐
- QUITE A BIT OF SHORTNESS OF BREATH..... ☐
- MODERATE SHORTNESS OF BREATH..... ☐
- SOME SHORTNESS OF BREATH..... ☐
- A LITTLE SHORTNESS OF BREATH..... ☐
- NO SHORTNESS OF BREATH..... ☐

10. How often during the last 2 weeks have you felt tearful, or like crying?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

11. How often during the last 2 weeks have you felt as if you are more dependent than you were before your heart problem?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

12. How often during the last 2 weeks have you felt you were able to do your usual social activities, or social activities with your family?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

3. How often during the last 2 weeks have you felt as if others no longer have the same confidence in you as they did before your heart problem?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

4. How often during the last 2 weeks have you experienced chest pain while doing your day-to day activities?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

5. How often during the last 2 weeks have you felt unsure of yourself or lacking in self-confidence?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

6. How often during the last 2 weeks have you been bothered by aching or tired legs?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

17. During the last 2 weeks, how much have you been limited in doing sports or exercise as a result of your heart problem?

- EXTREMELY LIMITED..... ☐
- VERY LIMITED..... ☐
- LIMITED QUITE A BIT..... ☐
- MODERATELY LIMITED..... ☐
- SOMEWHAT LIMITED..... ☐
- LIMITED A LITTLE..... ☐
- NOT LIMITED AT ALL..... ☐

18. How often during the last 2 weeks have you felt apprehensive or frightened?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

19. How often during the last 2 weeks have you felt dizzy or light-headed?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

20. In general during the last 2 weeks, how much have you been restricted or limited as a result of your heart problem?

- EXTREMELY LIMITED..... ☐
- VERY LIMITED..... ☐
- LIMITED QUITE A BIT..... ☐
- MODERATELY LIMITED..... ☐
- SOMEWHAT LIMITED..... ☐
- LIMITED A LITTLE..... ☐
- NOT LIMITED AT ALL..... ☐

21. How often during the last 2 weeks have you felt unsure as to how much exercise or physical activity you should be doing?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

22. How often during the last 2 weeks have you felt as if your family is being over-protective toward you?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

23. How often during the last 2 weeks have you felt as if you were a burden to others?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

24. How often during the past 2 weeks have you felt excluded from doing things with other people because of your heart problem?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

25. How often during the past 2 weeks have you felt unable to socialise because of your heart problem?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

26. In general, during the last 2 weeks how much have you been physically restricted or limited as a result of your heart problem?

- EXTREMELY LIMITED..... ☐
- VERY LIMITED..... ☐
- LIMITED QUITE A BIT..... ☐
- MODERATELY LIMITED..... ☐
- SOMEWHAT LIMITED..... ☐
- LIMITED A LITTLE..... ☐
- NOT LIMITED AT ALL..... ☐

27. How often during the last 2 weeks have you felt your heart problem limited or interfered with sexual intercourse?

- ALL OF THE TIME..... ☐
- MOST OF THE TIME..... ☐
- A GOOD BIT OF THE TIME..... ☐
- SOME OF THE TIME..... ☐
- A LITTLE OF THE TIME..... ☐
- HARDLY ANY OF THE TIME..... ☐
- NONE OF THE TIME..... ☐

APPENDIX 12: Cardiac Depression Scale (CDS)

This questionnaire consists of a number of statements about the way you feel at present.
Next to each statement there is a rating scale from 1 to 7 for you to indicate how much you agree or disagree with the statement.

Strongly disagree1 2 3 4 5 6 7Strongly agree

Please indicate how strongly you agree with each statement by circling one of the numbers on the scale.
For example, A score of 4 would indicate that you neither agree nor disagree with the statement, a score of 1 that you strongly disagree, and a score of 7 that you strongly agree.

EXAMPLE

Strongly disagree 1 (2) 3 4 5 6 7 Strongly agree

This indicates that you quite strongly disagree with the statement.

THERE ARE NO RIGHT OR WRONG ANSWERS.
PLEASE ENSURE YOU HAVE COMPLETED ALL ITEMS

HECK TO MAKE SURE YOU HAVE ANSWERED ALL QUESTIONS	STRONGLY DISAGREE	STRONGLY AGREE	
I have dropped many of my interests and activities.....	1 2 3 4 5 6 7 None dropped	All dropped	
My concentration is as good as it ever was.....	1 2 3 4 5 6 7 Very poor concentration	Excellent concentration	
I can't be bothered doing anything much.....	1 2 3 4 5 6 7 Keen to do things	Can't be bothered	
I get pleasure from life at present.....	1 2 3 4 5 6 7 No pleasure	Great pleasure	
I am concerned about the uncertainty of my health.....	1 2 3 4 5 6 7 Not concerned	Very concerned	
I may not recover properly...	1 2 3 4 5 6 7 Will recover completely	Will not recover	
My sleep is restless and disturbed.....	1 2 3 4 5 6 7 Not restless	Very restless	
I am not the person I used to be	1 2 3 4 5 6 7 Just the same	Completely different	
I wake up in the early hours of the morning and cannot get back to sleep.....	1 2 3 4 5 6 7 Never wake	Always wake	
0. I feel like I'm living on borrowed time.....	1 2 3 4 5 6 7 Unlimited time	Very much on borrowed time	
1. Dying is the best solution for me.....	1 2 3 4 5 6 7 No solution	Best solution	
2. I feel in good spirits.....	1 2 3 4 5 6 7 Very poor spirits	Excellent spirits	
3. The possibility of sudden death worries me.....	1 2 3 4 5 6 7 Not at all	Very worried	
4. There is only misery in the future for me.....	1 2 3 4 5 6 7 No misery	Only misery	
5. My mind is as fast and alert as always.....	1 2 3 4 5 6 7 Slow and inattentive	Very fast and alert	

6. I get hardly anything done.....	1 2 3 4 5 6 7 Everything done Nothing done	
7. My problems are not yet over.....	1 2 3 4 5 6 7 All problems over Still major problems	
8. Things which I regret about my life are bothering me.....	1 2 3 4 5 6 7 Absolutely no regrets Great regrets	
9. I gain just as much pleasure from my leisure activities as I used to.....	1 2 3 4 5 6 7 No pleasure at all Very great pleasure	
0. My memory is as good as it always was.....	1 2 3 4 5 6 7 Very poor memory Excellent memory	
1. I become tearful more easily than before.....	1 2 3 4 5 6 7 Not at all tearful Very easily tearful	
2. I seem to get more easily irritated by others than before.....	1 2 3 4 5 6 7 Never irritated Very easily irritated	
3. I feel independent and in control of my life.....	1 2 3 4 5 6 7 No independence Completely independent	
4. I lose my temper more easily nowadays.....	1 2 3 4 5 6 7 Never lose temper Lose it very easily	
5. I feel frustrated.....	1 2 3 4 5 6 7 Not at all frustrated Extremely frustrated	
6. I am concerned about my capacity for sexual activity.....	1 2 3 4 5 6 7 No concern at all Grave concern	

APPENDIX 13: Global Mood Scale (GMS)

GMS

Name: ..... **Sex:** ..... **Age:** ..... **Date:** .....

Below are a number of words that describe different feelings and emotions. Please read **each item** carefully and then **circle the appropriate number** next to that word. Indicate to what extent **you have felt this way lately**. Please use the following scale to record your answers.

0 = NOT AT ALL **1** = A LITTLE BIT **2** = MODERATELY **3** = QUITE A BIT **4** = EXTREMELY

To what extent have you felt this way lately :

1	*Wearied*	→ 0 1 2 3 4	11	*Listless*	→ 0 1 2 3 4
2	*Active*	→ 0 1 2 3 4	12	*Tired*	→ 0 1 2 3 4
3	*Worn out*	→ 0 1 2 3 4	13	*Enterprising*	→ 0 1 2 3 4
4	*Dynamic*	→ 0 1 2 3 4	14	*Relaxed*	→ 0 1 2 3 4
5	*Bright*	→ 0 1 2 3 4	15	*Insecure*	→ 0 1 2 3 4
6	*Helpless*	→ 0 1 2 3 4	16	*Sociable*	→ 0 1 2 3 4
7	*Hard-working*	→ 0 1 2 3 4	17	*Cheerful*	→ 0 1 2 3 4
8	*Feeble*	→ 0 1 2 3 4	18	*Fatigued*	→ 0 1 2 3 4
9	*Lively*	→ 0 1 2 3 4	19	*Weakened*	→ 0 1 2 3 4
10	*Physically weak*	→ 0 1 2 3 4	20	*Self-confident*	→ 0 1 2 3 4

APPENDIX 14: Heart Patients Psychological Questionnaire (HPPQ)

NSTRUCTIONS

On the following pages you will find statements such as:

	Corrrect	?	Incorrect
"I dread getting up early".....			

You should tick the answer which applies to you NOWADAYS.

	Corrrect	?	Incorrect
You can choose between three possibilities.....			
Tick correct if in your opinion the statement does apply to you nowadays.....	√		
Tick incorrect if in your opinion the statement does not apply to you nowadays			√
Tick ? only if you really can not answer correct or incorrect		√	

In all cases, what is wanted is your own impression.

There are no wrong answers. All answers are equally good provided they indicate what your own impression is.

Please do not omit to answer any of the questions.

PLEASE TICK THE MOST APPROPRIATE BOX.

		Correct	?	Incorrect
01	I quite often have vague stomach complaints			
02	I feel at ease nowadays			
03	If it's cold and windy outside I hardly ever leave the house			
04	I`m sometimes so restless that I can`t sit still			
05	I avoid busy streets			
06	If i`m going to meet a lot of people I get nervous			
07	I feel fit nowadays			
08	I don`t like speaking to a group of people I don`t know			
09	I could do a lot more work formerly			
10	I worry about the chance of having a heart attack (again)			
11	I`m easily irritated			
12	I don`t have enough stamina			
13	I worry about the future			
14	I feel happy nowadays			
15	I find it rather annoying that I cry more nowadays			
16	I used to be capable of a lot more			
17	I prefer working alone to working with a number of other people			

		Correct	?	Incorrect
18	In a group I`m usually leader			
19	I feel healthy			
20	I tire faster than I think is normal			
21	I feel tight in the chest quite often			
22	I worry quite a lot that I`ll be worse off financially			
23	I usually feel cheerful			
24	I feel calm nowadays			
25	I like giving instructions			
26	So far I`ve suceeded in almost everything that I wanted			
27	I feel comfortable and pleasant nowadays			
28	I`m shy in the company of strangers			
29	I feel relaxed nowadays			
30	If I had to live my life over again I wouldn't change anything			
31	I don`t like having a lot of people around me			
32	I sleep well nowadays			
33	I`m often out of sorts without knowing why			
34	I feel up to anything			
35	I feel relaxed when I`m going to take a journey			
36	I have a lot of self-confidence nowadays			
37	I can accept that I`m a heart patient			
38	Things often go wrong if I have to do something quickly			
39	I`m interested in a lot of things			
40	I tire quickly even if I don`t do anything out of the ordinary			
41	I feel self-assured nowadays			
42	I don`t like the idea of doing heavy work			
43	I feel physically healthy			
44	I often feel listless			
45	I feel content nowadays			
46	I often feel down			
47	I`m often nervous			
48	I get out of breath quickly			
49	I still feel quite capable of taking part in sports			
50	I have little influence on other people			
51	Once in a while I`m so bad-tempered that there`s no pleasing me			
52	I feel good nowadays			

APPENDIX 15: MOS 36-item short –form health survey (MOS SF-36)

This survey asks for your views about your health. This information will help to monitor how you feel and how well you are able to do your usual activities.
Answer every question by marking the answer as indicated. If you are unsure about how to answer a question, please give the best answer you can.

1. **In general** would you say your health is:

(please circle the most appropriate answer)

Excellent..... 1
Very good..... 2
Good..... 3
Fair..... 4
Poor..... 5

2. **Compared to one year ago,** how would you rate your health in general now?

(please circle the most appropriate answer)

Much better now than one year before..... 1
Somewhat better now than one year ago..... 2
About the same as one year ago..... 3
Somewhat worse than one year ago..... 4
Much worse now than one year ago..... 5

3. The following items are about activities you might do during a typical day. Does **your health now limit you** in these activities? If so, how much?

(Please tick the most appropriate box on each line)

ACTIVITIES	Yes, Limited A lot	Yes, Limited A Little	No, Not Limited At All
a. **Vigorous activities**, such as running, lifting heavy objects, participating in strenuous sports			
b. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, playing golf			
c. Lifting or carrying groceries			
d. Climbing **several** flights of stairs			
e. Climbing **one** flight of stairs			
f. Bending, kneeling, or stopping			
g. Walking **more than a mile**			
h. Walking **several blocks**			
i. Walking **one block**			
j. Bathing or dressing yourself			

4. During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of your physical health?**

(Please tick the most appropriate box on each line)

	YES	NO
a. Cut down on the **amount of time** you spent on work or other activities		
b. **Accomplished less** than you would like		
c. Were limited in the **kind** of work or other activities		
d. Had **difficulty** performing the work or other activities (for example, it took extra effort)		

5. During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious) ?

(Please tick the most appropriate box on each line)

	YES	NO
a. Cut down on the **amount of time** you spent on work or other activities		
b. **Accomplished less** than you would like		
c. Didn't do work or other activities as **carefully** as usual		

6. During the **past 4 weeks**, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbours, or groups?

(Please circle the most appropriate answer)

Not at all..... 1
Slightly..... 2
Moderately..... 3
Quite a bit..... 4
Extremely..... 5

7. How much **bodily** pain have you had during the **past 4 weeks**?

(Please circle the most appropriate answer)

None..... 1
Very mild..... 2
Mild..... 3
Moderate..... 4
Severe..... 5
Very severe..... 6

8. During the **past 4 weeks,** how much did pain interfere with your normal work (including both work outside the home and housework)

(Please circle the most appropriate answer)

Not at all..... 1
Slightly..... 2
Moderately..... 3
Quite a bit..... 4
Extremely..... 5

9. These questions are about how you feel and how things have been **during the past 4 weeks**. For each question, please give the answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks**-

Please tick the most appropriate box on each line)

	All of the time	**Most of the time**	**A good bit of the time**	**Some of the time**	**A little of the time**	**None of the time**
a. Did you feel full of pep?						
b. Have you been a very nervous person?						
c. Have you felt so down in the dumps that nothing could you cheer you up?						
d. Have you felt calm and peaceful?						
e. Did you have a lot of energy?						
f. Have you felt downhearted and blue?						
g. Did you feel worn out?						
h. Have you been a happy person?						
i. Did you feel tired?						

10. During the **past 4 weeks**, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

(Please circle the most appropriate answer)

All of the time..... 1
Most of the time..... 2
Some of the time..... 3
A little of the time..... 4
None of the time..... 5

11. How TRUE or FALSE is **each** of the following statements for you?

Please tick the most appropriate box on each line)

	Definitely True	**Mostly True**	**Don't Know**	**Mostly False**	**Definitely False**
a. I seem to get sick a little easier than other people					
b. I am healthy as anybody I know					
c. I expect my health to get worse					
d. my health is excellent					

APPENDIX 16: Global rating of change questionnaire

Since you began the cardiac rehabilitation programme, has there been any change in your overall quality of life?

Please mark the most appropriate box below.

- ☐ A very great deal worse
- ☐ A great deal worse
- ☐ A good deal worse
- ☐ Moderately worse
- ☐ Somewhat worse
- ☐ A little worse
- ☐ Almost the same, hardly any worse at all
- ☐ No change
- ☐ Almost the same, hardly any better at all
- ☐ A little better
- ☐ Somewhat better
- ☐ Moderately better
- ☐ A good deal better
- ☐ A great deal better
- ☐ A very great deal better

APPENDIX 17: Comparison of the response rate of each of the study centres

Comparison of the response rates in the study sample (n = 786)

Cardiac Rehabilitation Centre	Number of questionnaires administered per centre	Time 1 N (% completion)	Time 2 N (% completion)	Percentage of overall sample
1	435	298 (69)	276 (63)	35.1
2	75	62 (83)	53 (71)	6.7
3	75	56 (75)	43 (57)	5.5
4	75	49 (65)	39 (52)	5.0
Comparison centre				
5	630	347 (55)	293 (48)	38.4
6	135	77 (57)	66 (43)	7.1
7	45	22 (49)	16 (37)	2.2

APPENDIX 18: Comparison of the response rate for each questionnaire at each centre over time

	Time 1 N (% completion)	Time 2 N (% completion)
CARDIAC REHABILITATION		
White		
Centre 1 (150)	88 (59)	81 (54)
Centre 2 (25)	22 (88)	21 (21)
Centre 3 (25)	18 (72)	15 (60)
Centre 4 (25)	16 (64)	11 (44)
Total	144 (64)	128 (57)
Yellow		
Centre 1 (150)	113 (75)	106 (71)
Centre 2 (25)	19 (76)	15 (60)
Centre 3 (25)	21 (84)	16 (64)
Centre 4 (25)	15 (60)	12 (48)
Total	168 (75)	149 (66)
Blue		
Centre 1 (150)	97 (65)	89 (59)
Centre 2 (25)	21 (84)	17 (68)
Centre 3 (25)	17 (68)	12 (48)
Centre 4 (25)	18 (72)	16 (64)
Total	153 (68)	134 (60)
COMPARISON GROUP		
White		
Centre 4 (210)	111 (53)	98 (47)
Centre 5 (45)	23 (51)	19 (42)
Centre 6 (15)	7 (47)	6 (40)
Total	141 (52)	123 (46)
Yellow		
Centre 4 (210)	113 (54)	89 (42)
Centre 5 (45)	29 (64)	26 (58)
Centre 6 (15)	6 (40)	5 (33)
Total	148 (55)	120 (44)
Blue		
Centre 4 (210)	123 (59)	106 (50)
Centre 5 (45)	25 (56)	21 (47)
Centre 6 (15)	9 (60)	5 (33)
Total	157 (58)	132 (49)

Response rate for each questionnaire at each centre over time

Questionnaire (number administered)	Time 1 N (% completion)	Time 2 N (% completion)
CARDIAC REHABILITATION		
White		
Centre 1 (150)	88 (59)	81 (54)
Centre 2 (25)	22 (88)	21 (21)
Centre 3 (25)	18 (72)	15 (60)
Centre 4 (25)	16 (64)	11 (44)
Total	144 (64)	128 (57)
Yellow		
Centre 1 (150)	113 (75)	106 (71)
Centre 2 (25)	19 (76)	15 (60)
Centre 3 (25)	21 (84)	16 (64)
Centre 4 (25)	15 (60)	12 (48)
Total	168 (75)	149 (66)
Blue		
Centre 1 (150)	97 (65)	89 (59)
Centre 2 (25)	21 (84)	17 (68)
Centre 3 (25)	17 (68)	12 (48)
Centre 4 (25)	18 (72)	16 (64)
Total	153 (68)	134 (60)
COMPARISON GROUP		
White		
Centre 4 (210)	111 (53)	98 (47)
Centre 5 (45)	23 (51)	19 (42)
Centre 6 (15)	7 (47)	6 (40)
Total	141 (52)	123 (46)
Yellow		
Centre 4 (210)	113 (54)	89 (42)
Centre 5 (45)	29 (64)	26 (58)
Centre 6 (15)	6 (40)	5 (33)
Total	148 (55)	120 (44)
Blue		
Centre 4 (210)	123 (59)	106 (50)
Centre 5 (45)	25 (56)	21 (47)
Centre 6 (15)	9 (60)	5 (33)
Total	157 (58)	132 (49)

APPENDIX 19: Comparison of baseline demographic details for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires

Comparison of baseline demographic details for those completing only Time 1 questionnaire with those completing both Time 1 and 2 questionnaires.

Variable	Time 1 only (n= 125)	Time 1 and 2 (n = 786)	P value
Age	Mean (SD) 62.17 (9.73)	Mean (SD) 62.48 (10.35)	t-statistic = 0.31, w/903 df, p = 0.76
Gender	N (%)	N (%)	χ^2 = 0.74, w/1df, p = 0.39
Male	101 (80.8)	594 (75.57)	
Female	24 (19.2)	192 (24.43)	
Marital Status	N (%)	N (%)	
Married	93 (74.4)	625 (79.5)	
Single	16 (12.8)	66 (8.4)	χ^2 = 3.41, w/3df, p=0.33
Widowed	14 (11.2)	81 (10.3)	
Other/ Missing	2 (1.6)	14 (1.6)	
Occupational status	N (%)	N (%)	
Full-time	29 (23.2)	261 (33.2)	
Part-time	13 (10.4)	65 (8.3)	
Retired	64 (51.2)	334 (42.5)	
Unemployed	3 (2.4)	27 (3.4)	χ^2 = 10.38, w/5df, p=0.06
Homemaker	7 (5.6)	69 (8.8)	
Other/missing	9 (7.2)	30 (3.8)	
Education	N (%)	N (%)	
Primary	55 (44)	327 (41.6)	χ^2 = 7.85. w/3 df, p = 0.049
Secondary	40 (32)	303 (38.6)	
Third level	18 (14.4)	123 (15.7)	
Missing	12 (9.6)	33 (4.1)	
Residential area	N (%)	N (%)	
City	54 (43.9)	340 (43.3)	
Town	18 (14.6)	174 (22.1)	χ^2 = 5.99, w/3df p = 0.11
Village	21 (17.1)	89 (11.3)	
Rural/ missing	32 (25.6)	183 (23.3)	
Living	N (%)	N (%)	
Alone	15 (12)	88 (11.2)	χ^2 = 3.06, w/2df, p = 0.22
With others	107 (85.6)	692 (88)	
Missing	3 (2.4)	6 (0.8)	
Event	N (%)	N (%)	
MI	26 (20.8)	138 (17.6)	
CABG	60 (48)	386 (49.1)	χ^2 = 1.94, w/3df, p = 0.59
PTCA	22 (18.4)	174 (22.1)	
Other/missing	16 (12.8)	88 (11.2)	
CHD History	N (%)	N (%)	
None	77 (61.6)	428 (54.4)	
PTCA	19 (15.2)	147 (18.7)	
MI	10 (8)	67 (8.5)	χ^2 = 2.35 w/4df, p = 0.67
CABG	3 (2.4)	24 (3.1)	
Other/missing	16 (12.8)	120 (15.3)	

Co-morbidity	N	(%)	N	(%)	$\chi^2 = 0.005$, w/1df,
None	41	(32.8)	239	(30.4)	p = 0.92
Co-morbidity	88	(67.2)	497	(59.6)	
Group	N	(%)	N	(%)	$\chi^2 = 3.56$, w/1df,
CR	54	(43.2)	411	(52.3)	p =0.06
Control	71	(56.8)	375	(47.7)	
Questionnaires	N	(%)	N	(%)	$\chi^2 = 6.00$, w/5df,
White CR	16	(12.8)	128	(16.3)	p=0.31
Yellow CR	19	(15.2)	149	(19.0)	
Blue CR	19	(15.2)	134	(17.0)	
White comparison	18	(14.4)	123	(15.6)	
Yellow comparison	28	(22.4)	120	(15.3)	
Blue comparison	25	(20.0)	132	(16.8)	

APPENDIX 20: Responsiveness statistics for each subscale

Instrument	ES	SRM	RE	tuley	Guyatt
hcs cog	0.18	0.24	0.13	0.41	0.89
hcs som	0.14	0.17	0.07	0.24	1.83
hads dep	0.17	0.25	0.17	0.42	1.14
hads anx	0	-0.02	0.00	-0.01	1.47
qli health	0.03	-0.01	0.00	0.05	0.98
qli socio	0.07	0.04	0.01	0.12	1.09
qli psy	0.01	-0.09	0.03	0.03	1.28
qli family	-0.07	-0.10	0.03	-0.12	1.26
qli total	-0.01	-0.11	0.02	0	1.05
eds	-0.02	0.01	0.00	-0.01	0.31
stai	-0.06	-0.07	0.01	-0.08	0.67
qlmi e	-0.11	-0.17	0.01	-0.11	0.89
qlmi p	-0.03	-0.04	0.00	-0.08	0.63
qlmi s	-0.06	-0.05	0.00	-0.08	0.66
qlmi t	-0.06	-0.10	0.03	-0.08	0.72
gms pa	0.59	0.62	1	0.79	0.94
gms na	0.30	0.39	0.40	0.43	0.58
hppq s	-0.07	-0.04	0.01	-0.06	0.79
hppq d	0.30	0.27	0.32	0.56	0.66
hppq w	-0.06	-0.10	0.03	-0.06	0.80
hppq f	0.26	0.29	0.27	0.43	0.99
sf pf	0.03	0.12	0.04	0	0.54
sf rlp	0.64	0.44	0.58	0.51	0.91
sf rlw	0.36	0.38	0.08	0.53	0.57
sf e	0.15	0.14	0.20	0.15	1.32
sf ewb	0.24	0.22	0.15	0.25	0.05
sf swb	0.30	0.35	0.35	0.39	1.21
sf p	0.21	0.33	0.34	0.27	0.67
sf gh	0.06	0.06	0.01	0.11	0.92

Responsiveness statistics for each instrument

Instrument	ES	SRM	RE	tuley	Guyatt
hcs cog	0.18	0.24	0.13	0.41	0.89
hcs som	0.14	0.17	0.07	0.24	1.83
hads dep	0.17	0.25	0.17	0.42	1.14
hads anx	0	-0.02	0.00	-0.01	1.47
qli health	0.03	-0.01	0.00	0.05	0.98
qli socio	0.07	0.04	0.01	0.12	1.09
qli psy	0.01	-0.09	0.03	0.03	1.28
qli family	-0.07	-0.10	0.03	-0.12	1.26
qli total	-0.01	-0.11	0.02	0	1.05
cds	-0.02	0.01	0.00	-0.01	0.31
stai	-0.06	-0.07	0.01	-0.08	0.67
qlmi e	-0.11	-0.17	0.01	-0.11	0.89
qlmi p	-0.03	-0.04	0.00	-0.08	0.63
qlmi s	-0.06	-0.05	0.00	-0.08	0.66
qlmi t	-0.06	-0.10	0.03	-0.08	0.72
gms pa	0.59	0.62	1	0.79	0.94
gms na	0.30	0.39	0.40	0.43	0.58
hppq s	-0.07	-0.04	0.01	-0.06	0.79
hppq d	0.30	0.27	0.32	0.56	0.66
hppq w	-0.06	-0.10	0.03	-0.06	0.80
hppq f	0.26	0.29	0.27	0.43	0.99
sf pf	0.03	0.12	0.04	0	0.54
sf rlp	0.64	0.44	0.58	0.51	0.91
sf rlw	0.36	0.38	0.08	0.53	0.57
sf e	0.15	0.14	0.20	0.15	1.32
sf ewb	0.24	0.22	0.15	0.25	0.05
sf swb	0.30	0.35	0.35	0.39	1.21
sf p	0.21	0.33	0.34	0.27	0.67
sf gh	0.06	0.06	0.01	0.11	0.92

APPENDIX 21: Percentage of responsiveness scoring at the lowest 10% and top 10% range of scores possible for each instrument at Time 1.

Percentage of respondents scoring at lowest 10% and top 10% range of scores possible for each instrument at time one.

	Cardiac rehabilitation	**Comparison group**	
Instrument (range of scores) (Lowest 10% of possible scores) (Highest 10% of possible scores)	Percentage of respondents	Percentage of respondents	Average percentage
HCS som (0-48) (0-4.8) (43.2 – 48)	 25 0	 24.39 0	 24.7 0
HCS cog (0-48) (0-4.8) (43.2 – 48)	 9.4 2.3	 17.9 1.6	 13.7 2
HADS anx (0-21) (0-2.1) (18.9 – 21)	 19.5 0.8	 30.9 0	 25.2 0.4
HADS dep (0-21) (0-2.1) (18.9 – 21)	 33.6 0	 42.5 0	 38.1 0
QLI health (0 –30) (0 –3) (27 – 30)	 0 24.2	 0 22.8	 0 23.5
QLI socio (0 –30) (0 –3) (27 – 30)	 0 35.1	 0 35.8	 0 35.5
QLI psy (0 –30) (0 –3) (27 – 30)	 0.8 39	 0 35.8	 0.4 37.4
QLI fam (0 –30) (0 –3) (27 – 30)	 0.8 62.5	 0 66.7	 0.4 64.6
QLI total (0 –30) (0 –3) (27 – 30)	 0 18.7	 0 18.2	 0 18.5
CDS (26-182) (26 - 41.6) (166.4 – 182)	 7.3 0	 8.1 0	 7.7 0
STAI (20-80) (20 – 26) (74 – 80)	 38.2 0	 32.4 0	 35.3 0
QLMI e (1-7) (1 - 1.6) (6.4 - 7)	 0 15.8	 0 12.2	 0 14.4
QLMI p (1-7) (1 - 1.6) (6.4 - 7)	 0 13.3	 0 12.9	 0 13.1

QLMI s (1-7)			
(1-1.6)	0	0	0
(6.4-7)	24.8	24.3	24.6
QLMI t (1-7)			
(1-1.6)	0	0	0
(6.4-7)	14.5	16.4	15.5
GMS na (0-40)			
(0 – 4)	26.1	3.8	14.95
(36-40)	1.5	0	0.8
GMS pa (0-40)			
(0 –4)	0.7	39.2	20
(36 –40)	2.2	0	1.1
HPPQ s (6-18)			
(6-7.2)	8.2	0	4.1
(16.8 – 18)	3.7	6.1	4.9
HPPQ d (10 –30)			
(10 – 12)	22.4	25.8	21.1
(28 - 30)	5.2	0	2.6
HPPQ w (12 – 36)			
(12 – 14.4)	6.7	4.5	5.6
(33.6 – 36)	26.1	51.5	38.8
HPPQ f (12 – 36)			
(12 – 14.4)	3.7	12.1	7.9
(33.6 – 36)	6.7	0	3.35
SF pf (0-100)			
(0-10)	1.5	1.5	1.5
(90 – 100)	6.4	19.7	13.1
SF rlp (0-100)			
(0-10)	59	18.2	38.6
(90 – 100)	7.5	39.4	23.5
SF rle (0-100)			
(0-10)	34	22	28
(90 – 100)	31.3	58	44.7
SF e (0-100)			
(0-10)	0.7	3	1.9
(90 – 100)	9	10.6	9.8
SF ewb (0-100)			
(0-10)	0	0	0
(90 – 100)	23.9	39.4	3.2
SF swb (0-100)			
(0-10)	1.5	0	0.8
(90 – 100)	27.6	46.2	36.9
SF p (0-100)			
(0-10)	3	1.5	2.3
(90 – 100)	36.6	50	43.3
SF gh (0-100)			
(0-10)	3	0	1.5
(90 – 100)	14.9	16.7	15.8