

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Meta-analysis of moderators of psycho-oncology therapy effectiveness:

"It's the sick who need a doctor".

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

at Massey University, Palmerston North

New Zealand.

Heather Adele Heron

2009

Abstract

Reviews conflict regarding the effectiveness of psycho-oncological therapies at reducing patient anxiety, depression and general distress, suggesting that unknown factors are moderating trial results. This meta-analysis investigates the moderating impact of a large range of socio-demographic, psychological, medical and therapy factors using published and unpublished data from 146 prospective controlled trials, including non-random designs.

Preliminary analyses of trial design quality features exposed 2 moderators: recruit screening for psychological distress or history, and the nature of the control condition. These structured a 2 x 2 matrix used to conduct substantive analyses. Admitting only recruits with established baseline distress was found to predict greater effect size, as was excluding patients with a history of distress. Main effects for patients with baseline distress compared with untreated controls, were medium-strong at $g = 0.52 - 0.70$. Evidence of varying strengths indicated that patients who were older, of lower income, male, single, or suffering from cancer sited elsewhere than breast produced higher effect sizes. Data also highlighted particular stages in the cancer journey: re-entry to normal life at the end of medical treatment, recurrence, and distant disease spread.

Findings suggest that risk and distress screening should be employed by both clinicians and researchers. Researchers should also re-direct attention away from unscreened middle class early stage breast cancer patients, towards more vulnerable socio-demographic and medical groups. The potential of using survivors and indirect therapies to effectively and efficiently reach vulnerable groups deserves exploration. Reviewers need to take into account the 2 trial design moderators discovered, and should include non-random controlled trials which may have more access to particularly vulnerable groups because some past conclusions were confounded by the co-variation of study design with sampling characteristics.

Preface

The subtitle, “It’s the sick who need a doctor” is drawn from the gospel of Matthew 9:12 where Jesus says that His attentions are for the needy rather than for those who consider themselves invulnerable. It was chosen to reflect the theme that emerges from this research that therapy effectiveness tends to be moderated by deficit – medical, social, economic and psychological need – and the call for more focus on these factors in research and practice.

I came to the study of psychology late, having originally trained in law and then ‘retired’ to raise a family and help my mother cope with the advanced stages of Parkinson’s disease. I suffered an episode of depression, and later another, during my care-giving interlude, and learned about the multi-generational causes and effects of family dysfunction. Having ‘done the practical’, when my children were older, I enrolled in a university course in psychology.

My depression-lowered immunity levels exposed me to glandular fever, which took me out of study for six months, but on my return I was thrilled to be given the opportunity to do this piece of research. However, my weakened system invited more illness, and seven months into it, the research ‘went live’ when I was diagnosed with early stage breast cancer. I now had the opportunity to ‘do more practical’, but this time I was greatly assisted by what I had learned about the trajectory of emotional experience for cancer patients and about the medical treatment of the disease. I shed few tears as I faced my new identity as a cancer patient, a mastectomy, and the trials of chemotherapy, and kept on with the research between medical appointments and as ‘chemo-brain’ allowed.

Through the series of diagnostic tests, surgery and more surgery, chemotherapy and its side effects, oncology appointments, mammograms, heart scans and blood tests.... I became very aware of how privileged my personal circumstances were and how they buffered me from much of the stress that others suffer with this disease. I was in a loving marriage, and my children were now in their late teens – quality social support. My husband had a good job and my research work could flex around the needs of my treatment – financial security, occupation, a bolster to my self-esteem. I had a reasonable lay person’s understanding of the disease, its treatment and hospital structures and procedures, plenty of brochures were available as well as ready access to a specialist nurse, and I knew about the emotional side of the disease – information. I knew how to press for the strategic advice and services I needed in order to reduce stressful uncertainty and make well informed treatment choices – advocacy. I had caught the disease early – favourable prognosis. I lived near a well equipped hospital in an OECD welfare state – comprehensive free and timely medical treatment. I had long attended to the existential issues of life – the existential crisis posed by this life threat was minimised. And I did not have close personal experience of the toll that the disease can take on those that die from it – no fear-filled memories.

I could not imagine anyone better placed to deal with the experience than me. I was bolstered and buffered on every side. Despite my history of depression, I was not distressed by the experience, other than in quickly passing moments of fear and grief as I adjusted to my new identity, body image, and expectations of the future. I felt no need

for psycho-oncological support, but that too would have been available to me – free of charge - if I had. Ultimately, the experience boosted my appreciation and enjoyment of life and sharpened my focus on what matters.

There can't be many cancer patients who are so well supported. My comparative wealth and education set me apart from the majority of people in my own country, let alone those in lower and middle income countries. Many of the supports that I enjoyed turned upon this socio-economic position. Most cancer patients are battered by a succession of losses, practical difficulties and decisions for which they are ill-equipped.

In designing interventions, psychologists often do not realise how fundamentally important socio-demographic and economic factors are. The focus is on the type of therapy and how it is delivered rather than who it is delivered to and what their life circumstances are. I hope that the present research helps correct that focus, turning attention away from middle class people with favourable prognoses like me – where the majority of research has been directed in the past - to those who have gaps and pressures where they need supports.



The writer and her supervisors, Dr Don Baken, left, and Dr Shane Harvey, right, when research started mid 2007. Don is a psycho-oncologist and researcher, and Shane directs the Massey psychology clinic. This photo was taken for the funder, the Cancer Society of New Zealand, but was then used in requesting unpublished data from primary study authors...and to apparent great effect as half of them responded. “A good looking team”, one author said!

Acknowledgements

My supervisors, Dr Shane Harvey, Director of the Psychology Clinic here at Massey Palmerston North, and Dr Don Baken, lead researcher with the Massey / MidCentral Health Psycho-oncology Service. I have awarded you 'best supervisors ever' presents the last two Christmas's for good reason: competent, reliable, direct but kind, generous teachers. If wise teachers make learning a joy, then excellent teachers make their students want to be like them. Your input has caused me to alter the direction of my career and you have modeled for me the attitudes and standards that I would like to take into it. This research has been a fabulous life and learning experience for me, and it all started when you were willing to take a punt on someone whose health was not one hundred percent and who didn't measure up to the requirements of others. Thanks. I hope to 'pay it forward' in due course. Let me know if you ever need a reference!

Dr Kevin Ronan, former head of the Massey Psychology Clinic, since moved to Australia, mentor. Another who loyally supported me regardless of the 'downs' of my failures and health adventures. I've said it before but will say it again, Kevin: Your encouragement has been a lifeline without which I may have given it away long ago.

My husband, David. You have paid the bills faithfully for the many years that it has taken to get to this point with my studies – lets not count them - and I haven't finished yet. Then there has been all the technical support and the times that you have encouraged me to stay with it when I felt like quitting. Big hug. Promise to stick by you if you ever lose your marbles.

Harvey Jones, technical support, Massey School of Psychology. You are a patient and generous man Harvey!

Anne Hall and the team at Massey Library Document Supply, detectives extraordinaire! It is the primary study search that undergirds the quality of this research, and it was you who assiduously ferreted reports out from all over the world. What fabulous people to have on the team!

Thanks also to the many primary study researchers who generously provided additional effect size data. I hope you find the results useful. And to Dr Borenstein of CMA in the US for his generosity in statistical support.

Finally, thanks to the Cancer Society of New Zealand for funding the supervision and disbursements associated with this research project, and to the Lovell & Berys Clark Trust for their scholarship support during this phase of my masters study.

Heather Heron
July 2009

Brief Table of Contents

1. INTRODUCTION	1
The disease and its prevalence	2
The emergence of psycho-oncology	2
Psychological impact of the disease.....	3
Moderation of distress and intervention effectiveness.....	5
Types of intervention for distress.....	22
Psycho-oncological services	25
Why meta-analysis?	31
The present meta-analysis in context.....	34
2. METHOD.....	41
Guiding principles.....	41
Domain criteria.....	41
Search strategy	47
Coding.....	50
Study quality	53
3. ANALYSIS	57
Statistical matters	57
Groupings.....	62
4. RESULTS: PRELIMINARY ANALYSES	67
External validity	67
External validity summary	78
Internal validity	78
Internal validity summary	92
5. RESULTS: MAIN EFFECTS.....	94
Anxiety.....	94
Depression.....	96
Distress.....	97
Placement in the literature.....	98
6. RESULTS: THERAPY CHARACTERISTICS	100
Therapy type.....	100
Therapy trajectories.....	118

Therapy type conclusion	119
Therapy components	119
Therapy component conclusion	128
Delivery mode, dose and therapist variables.....	128
7. RESULTS: PATIENT CHARACTERISTICS	135
Socio-demographic moderators	135
Socio-demographic variables conclusion.....	144
Baseline distress and other screening.....	146
Baseline distress: Conclusion.....	150
Medical variables	151
Medical variables conclusion.....	163
Studies of more vulnerable patient groups.....	164
8. THEORETICAL MECHANISMS	165
Main effects.....	165
Therapy components	166
9. GENERAL DISCUSSION.....	167
Preliminary analyses	167
Substantive analyses.....	171
Conclusion	195
References	204
Appendices	on disc inside back cover

Full Table of Contents

1. INTRODUCTION	1
The disease and its prevalence	2
Medical treatment.....	2
The emergence of psycho-oncology	2
Distress intervention research	3
Psychological impact of the disease.....	3
Prevalence of psychopathology and symptoms	4
Nature of psychopathology and symptoms	4
Moderation of distress and intervention effectiveness.....	5
Conceptualising cancer distress	6
Baseline distress	8
Moderation by medical factors.....	9
Cancer site.....	9
Disease and treatment stages.....	10
Moderation by socio-demographic variables	11
Gender.....	11
Age.....	12
Race.....	13
Education and income.....	13
Marital status.....	14
Moderation by psychosocial resources	14
Social support.....	15
Distress moderation.....	15
Intervention and peer support groups.....	16
Prediction of therapy effectiveness.....	17
Self-efficacy	18
Coping.....	19
Self-esteem.....	20
Types of intervention for distress.....	22

Moderation by therapy variables.....	23
Therapy type.....	23
Therapy delivery	25
Psycho-oncological services	25
Financial burden on health system.....	25
Use of psychological services	26
Screening for psychological distress.....	28
Assessment.....	28
Screening approaches.....	29
Why meta-analysis?	31
An instructive synthesis	31
Its method and findings.....	31
Method critique	32
The present meta-analysis in context.....	34
Previous meta-analyses	35
General psychosocial intervention meta-analyses	35
The need for moderator meta-analysis.....	39
Hypotheses	40
2. METHOD.....	41
Guiding principles.....	41
Domain criteria.....	41
Types of study	42
Basic design	42
Nature of control	42
Study size	43
Types of participant.....	43
Types of intervention	43
‘Psychological’ intervention	43
Non-physical focus.....	44
Professionalism of therapist.....	45

Types of outcome	45
Outcome constructs	45
Scales and psychometrics.....	46
Theoretical moderators.....	46
Type of report.....	47
Search strategy	47
Sources	47
Search procedure	48
Keywords	48
Search yields	49
Coding.....	50
Coding instrument and procedure	50
Coding items	51
Procedure and reliability	51
Study quality	53
Approaches to date.....	53
Sensitivity analyses	54
The present approach	55
Allocation to conditions and attrition.....	56
3. ANALYSIS	57
Statistical matters	57
Computational model.....	57
Effect size index.....	57
Adjustment for baseline differences.....	58
Significance reporting and level.....	58
Multiple analyses	59
Omnibus effect sizes	59
Homogeneity testing	59
Intention to treat data	60
Missing data	60
Outliers.....	61
Fail-safe N.....	61

Closing of dataset.....	61
Groupings.....	62
Unit of analysis and decision rules.....	62
Assessment time points.....	62
Therapy type groupings.....	65
4. RESULTS: PRELIMINARY ANALYSES.....	67
External validity.....	67
Sampling bias.....	67
Distribution of effect sizes.....	68
Fail-safe N.....	70
Outliers and windzoration.....	71
Outlier summary.....	73
Publication.....	73
Study size.....	74
Possible confounding.....	75
Study size conclusion.....	76
Original language.....	76
Representativeness.....	76
OECD membership.....	77
External validity summary.....	78
Internal validity.....	78
Allocation to conditions.....	79
Comparison of designs.....	79
Possible confounding.....	82
Confidence in design report.....	84
Concealment of allocation to groups.....	84
Attrition.....	84
Nature of control condition.....	85
Blindness to condition.....	87
Participants.....	87

Therapists and assessors.....	88
Screening/floor effect.....	88
‘Screened in’ studies.....	90
‘Screened out’ studies.....	90
External validity and screening out.....	90
Targeting and screening out.....	91
Simultaneous screening.....	91
Decision.....	91
Treatment replicability/standardisation and fidelity	92
Internal validity summary	92
Dealing with the design confounds.....	93
5. RESULTS: MAIN EFFECTS	94
Anxiety.....	94
Depression.....	96
Distress.....	97
Placement in the literature.....	98
6. RESULTS: THERAPY CHARACTERISTICS	100
Therapy type.....	100
Main effects.....	101
Comparison with earlier works	103
Therapy types and combinations.....	103
Descriptive summary	105
Education.....	105
Relaxation	106
CBT.....	107
CBT detail: Depression.....	107
Expressive-supportive therapies.....	111
Expressive-support detail: Depression.....	111
Expressive-supportive v. CBT.....	114
Specific therapy combinations	114
Non-professional.....	114

Indirect therapies	115
Other.....	118
Therapy trajectories.....	118
Therapy type conclusion	119
Therapy components	119
Anxiety.....	120
Education.....	120
Relaxation	120
CBT	121
Expressive-support.....	122
Education.....	123
Relaxation	123
CBT	124
Expressive-support.....	125
Distress.....	125
Education.....	125
Relaxation	126
CBT	127
Expressive-support.....	127
Therapy component conclusion	128
Delivery mode, dose and therapist variables.....	128
Delivery mode variables	131
Dose variables	132
Therapist variables	133
7. RESULTS: PATIENT CHARACTERISTICS	135
Socio-demographic moderators	135
Age.....	135
Race.....	137
Gender.....	138

Main effect	138
Education level.....	140
Occupation / income	140
Main effect	140
Educational therapies	141
CBT	142
Conclusion	142
Marital status.....	143
Main effect	143
Terminal phase	143
Expressive-supportive therapy	144
Conclusion	144
Socio-demographic variables conclusion.....	144
Baseline distress and other screening.....	146
Baseline distress: Main effect	146
Baseline distress: Therapy type finding	147
Baseline distress: Breakouts within primary studies.....	147
Simultaneous screening.....	147
CBT	149
Screening out patients with psychological history.....	150
Baseline distress: Conclusion.....	150
Medical variables	151
Cancer site	151
Cancer prognosis	155
Cancer stage	156
Distant spread issues	157
Medical protocol	158
Medical treatment stage	159
Specific issues	161
Medical variables conclusion.....	163

Studies of more vulnerable patient groups.....	164
8. THEORETICAL MECHANISMS	165
Main effects.....	165
Therapy components	166
9. GENERAL DISCUSSION.....	167
Preliminary analyses	167
Baseline distress	169
Screening out and simultaneous screening	170
Nature of control condition	171
Substantive analyses.....	171
Main effects – Are psycho-oncological therapies successful?.....	171
Therapy type – What does what?	172
Education.....	172
Relaxation	175
CBT	177
Expressive-support.....	179
Non-professional	181
Indirect	182
Other.....	182
Therapy delivery and theoretical mechanisms – How?	183
Delivery variables	183
Psycho-oncologist role	183
Delivery and psychological background.....	184
Theoretical mechanisms.....	185
Patient characteristics - And for whom?	186
Socio-demographics.....	186
Age.....	186
Race.....	187
Gender.....	188
Education and income.....	190

Marital status and social support.....	191
Baseline distress	191
Distress screening.....	191
Medical characteristics.....	193
Cancer site.....	193
Prognosis, cancer stage, medical protocol.	194
Medical treatment stage.	195
Conclusion	195
Study limitations and strengths	197
Constraints imposed by the data	197
Rarefication of sample.	197
Choice of variables.....	198
Choices made relating to method.....	199
English language bias.....	199
Illness specific measures.....	199
Method approach and execution	200
Application.....	200
Practice recommendations	200
Research recommendations.....	201
Closing reflection.....	203
References	204
Appendices	on disc inside back cover
Appendix A. Physical focus studies	
List 1. Studies set aside after initial electronic screen	
List 2. Studies set aside after screening full text	
Appendix B. Foreign language studies that could not be used	
Appendix C. Studies that could not be traced	
Appendix D. Studies with insufficient sample size	
Appendix E. Studies with unusable outcome data	
Appendix F. Reports that were only abstract length	
Appendix G. Uncontrolled studies and studies with inadequate controls	

- Appendix H. Study that used inadequate measures
- Appendix I. Studies outside the research domain
- Appendix J. Studies coded for analysis
- Appendix K. Outline of research questions
- Appendix L. Coding instrument
- Appendix M. Code record sheet
- Appendix N. Studies of more vulnerable groups
- Appendix O. Relegated Results

Table of Figures

Figure 4-1. Funnel plot, anxiety	69
Figure 4-2. Funnel plot, depression	69
Figure 4-3. Funnel plot, distress.....	70
Figure 4-4. Allocation to conditions, anxiety.....	80
Figure 4-5. Allocation to conditions, depression	80
Figure 4-6. Allocation to conditions, distress	81
Figure 4-7. Study design confound matrix.....	93

Table of Tables

Table 1-1. Earlier meta-analyses, summary of therapy type effect sizes.....	23
Table 1-2. Previous meta-analyses, summary of main effects.....	35
Table 4-1. Outliers	71
Table 4-2. Sampling bias, publication.....	74
Table 4-3. Sampling bias, study size.....	75
Table 4-4. Sampling bias, study size, pairwise comparisons for heterogeneity	75
Table 4-5. Large studies, confound proportions	76
Table 4-6. Nationality, comparison by OECD membership	77
Table 4-7. Allocation to conditions.....	80
Table 4-8. Study design, pairwise comparisons for heterogeneity	81
Table 4-9. Allocation to conditions, anxiety, confounding variables	83
Table 4-10. Allocation to conditions, depression, confounding variables.....	83
Table 4-11. Nature of control condition.....	85
Table 4-12. Nature of control condition, pairwise comparisons for heterogeneity	86
Table 4-13. Patient blindness, all categories.....	87
Table 4-14. Patient blindness, where 'not reported' merged with 'not blind'.....	87
Table 4-15. Screening at recruitment	89
Table 4-16. Screening at recruitment, pairwise comparisons for heterogeneity.....	89
Table 5-1. Anxiety, main and breakout effects	95
Table 5-2. Depression, main and breakout effects.....	96
Table 5-3. Distress, main and breakout effects	98
Table 6-1. Therapy types, main effects.....	102
Table 6-2. Therapy types, summary.....	106
Table 6-3. CBT as a therapy type, depression	108
Table 6-4. CBT as a therapy combination, depression	110
Table 6-5. Expressive-supportive as a therapy type, depression	112
Table 6-6. Expressive-supportive as therapy combinations, depression.....	113
Table 6-7. Indirect, descriptive information	116
Table 6-8. 'Screened out' studies, therapy trajectory	118
Table 6-9. Therapy components, anxiety.....	121
Table 6-10. Therapy components, depression.....	123
Table 6-11. Therapy components, distress.....	126

Table 6-12. Therapy delivery mode, dose, and therapist variables: Summary	129
Table 7-1. Age.....	136
Table 7-2. Gender.....	138
Table 7-3. Occupation / income	141
Table 7-4. Occupation / income, educational therapies	141
Table 7-5. Occupation / income, CBT	142
Table 7-6. Marital status, quartile contrast.....	143
Table 7-7. Marital status, expressive-supportive therapy	144
Table 7-8. Screening at recruitment	146
Table 7-9. Screening at recruitment, pairwise comparisons for heterogeneity.....	146
Table 7-10. Screening at recruitment, simultaneous broken out.....	148
Table 7-11. Simultaneous screening, pairwise comparison	149
Table 7-12. Cancer site	151
Table 7-13. Cancer site, screening level frequencies	153
Table 7-14. Cancer site, breast v. other single sites	154
Table 7-15. Cancer prognosis	156
Table 7-16. Cancer stage.....	156
Table 7-17. Distant spread, therapy type	157
Table 7-18. Existential and CBT therapy components for distant spread patients	158
Table 7-19. Medical protocol.....	159
Table 7-20. Medical treatment stage	160
Table 7-21. Medical treatment stage, therapy types	161
Table 7-22. Recovery stage, therapy types	162
Table 7-23. Medical treatment stage, relaxation.....	163
Table 8-1. Theoretical mechanisms	165

1. INTRODUCTION

‘Cancer’ is a word that makes people shudder. Implications can be far-reaching, and the spectre of death hovers over them all. Yet we all die eventually. With cancer, the dread comes from more than that – it's about the lengthy and invasive treatment and worries about disfiguring surgery and pain and social isolation and finances. Stark questions forcefully impose themselves, such as, "Will life ever be the same again?", "Will I suffer a painful and lingering death?" and, "How will the family cope?". Stress. Depression. Adjustment.

The present study is about the psychological therapies that help some people through the emotional turmoil accompanying their cancer diagnosis and treatment. It is about uncovering ways to make those therapies more effective. This implies accurate targeting of therapy to need. Discussing another field of psychological therapy, Gordon Paul put the ‘ultimate question’ in these famous words: “What specific treatment, by whom, is most effective for this individual with that specific problem, and under which set of circumstances?” (Paul, 1969, p. 62). He also pointed out that this question cannot be answered by any single intervention trial. Indeed. What is needed is a means of synthesising the knowledge gained from all the trial information available – making allowance for the different trial designs – in order to glean the factors that repeatedly point to greater effectiveness.

Such a tool is called meta-analysis. A meta-analysis is a statistical synthesis – a grand averaging – of the results of primary research. It is a flexible tool, which can be designed around particular research questions. Most commonly in psycho-oncology it has been used simply to measure the magnitude of the effect of a therapy, but it can also be focussed on particular features of therapy, or the ‘what... whom...which’ of Paul’s ultimate question. This allows comparisons to be made of the effects produced by such broader features built into primary study design. It can even be focussed on study design features *per se* – anything for which data are regularly reported. Though it has its limitations, it is ideally suited to gathering all the available data and uncovering factors that moderate the effectiveness of therapy.

In this chapter the context is described for just such an investigation. The need arises out of doubts as to the effectiveness of psycho-oncological therapies (Coyne, Stephen, & Palmer, 2006; Newell, Simon-Fisher, & Savolainen, 2002). The doubts derive from conflicting evidence, which itself suggests the invisible influence of moderating factors (Ross, Boesen, Dalton, & Johansen, 2002). Given the severity of the emotional suffering of some cancer patients, it is important that services be provided and not cut due to a lack of evidence of their effectiveness. But it is also important that precious health resources are conserved where they are not effective.

The disease and its prevalence

The term 'cancer' is used to encapsulate more than one hundred different diseases with different aetiologies and courses. In common, they have the mechanism of an abnormal lack of control responses in cell growth resulting in the development of tumours, and after spreading regionally, the spawning of 'metastases' takes hold in organs distant from the original locality (Knight, 2004).

Cancer is now New Zealand's leading cause of mortality, contributing to one in three deaths (Dr John Childs, National Clinical Director Cancer Control, Ministry of Health, Morning Report radio broadcast, Radio New Zealand, 12 September 2007; New Zealand Health Information Service, 2007). The Ministry of Health attributed the prominence of the disease to New Zealand's aging population and to improvements in the treatment of other life-threatening illnesses. In developed nations generally, cancer has been the second leading cause of death (Knight, 2004, citing World Health Organisation figures for 2002) with a mortality rate of one in two of those diagnosed with the disease (Holland, 2002). In 2002 there were 10.9 million new diagnoses worldwide, 6.7 million deaths, and 24.6 million persons living (up to three years post-diagnosis) with cancer (Parkin, Bray, Ferlay, & Pisani, 2005). People with lower socio-economic status are disproportionately burdened in terms of both cancer incidence and survival (Dalton et al., 2008).

The most commonly diagnosed and deadly type is lung cancer. Breast cancer is the second most common and, because it has a relatively favourable prognosis, the most prevalent (Parkin et al., 2005). This means that breast cancer patients have the numbers to organise themselves to exercise political clout. An example was seen recently in New Zealand with the political party that won the last general election delivering on a promise to fund a longer treatment period of an expensive drug known as Herceptin, against the position taken by the drug funding authority.

Medical treatment

Medical treatment for cancer began with surgery in the 1850s, and palliative radiation a decade into the new century. In the 1940s, nitrogen mustards, developed in World War II, were discovered to have anti-tumour action. Important discoveries were also made in the forties that led to the later development of endocrine therapy and immunotherapy (Holland & Lesko, 1990). This led to the development of chemotherapy in the 1950s, and the real break-through, combined modality treatment, in the 1960s (Holland, 1990b; Jacobsen, 1990).

The emergence of psycho-oncology

Until the 1960s, a cancer diagnosis was routinely 'kept mum' from American sufferers by doctors and families since it was virtually a death sentence and carried a heavy social stigma (Holland, 1990b, 2002). However, successful medical treatment and modern understanding of the aetiology of cancer altered the conversational acceptability of the cancer diagnosis. Combined with a surge in interest in dying and increasing demand for

accountability from the medical profession (Freidenbergs et al., 1981-1982), this new optimism created a freedom to talk which set the stage for psychiatrists, psychologists, nurses and social workers to begin to address the psychosocial sequelae of the disease. Psychologists also became involved in issues relating to prevention, early detection and medical help seeking, and compliance with medical treatment as evidence emerged that cancer was related to environmental carcinogens, in particular, the habit of smoking (Holland, 1990b).

In 1975 a handful of researchers came together for the first U.S. national conference in psycho-oncology, and the field was 'born'. Jimmie Holland defines the field as dealing with (1) the psychological reactions of patients, their families, and staff to the various stages of cancer and its treatment, and (2) the psychological, social and behavioural factors impacting cancer cause and survival (Holland, 2002).

Distress intervention research

The nature of psychosocial therapy research with cancer patients has been largely a-theoretical. It has been driven by pragmatic considerations attendant to health administration and to a clinical population with urgent and serious needs whose interests are guarded by professionals from other disciplines who may lack an appreciation for psychological theory (Owen, Klapow, Hicken, & Tucker, 2001; Siegel, 1990). Knight (2004) explains further that patients often present with symptoms amenable to behavioural or cognitive-behavioural treatment, such as medical treatment aversion responses or depression, but they also often need to explore the meaning of what they are going through, express their feelings, and know the support of others, and so interventions may be designed to combine these elements in a group setting. Although education / information or anxiety treatments associated with particular medical treatments and their side effects may be delivered in isolation, many other treatments are delivered in such eclectic packages. Despite this pragmatism, there is no lack of theory that could be usefully tested, e.g. theoretical lines pertaining to adult development, reference groups, perceived personal control, and the impact of stressful life events (Siegel, 1990).

As a research tool, the 'kitchen sink approach' (Owen et al., 2001, p.226) thwarts long term research efforts by failing to isolate particular treatment elements and their effects, and preventing the identification of theoretical mechanisms of effectiveness. However, with sufficient numbers of studies, it should be possible to draw out the impact of some therapy and study features using a meta-analysis designed for that purpose – which is the aim of the present study.

Psychological impact of the disease

The depth and breadth of the cancer experience justify veteran oncology psychiatrist Alistair Cunningham's description of cancer as, "... primarily a psychosocial and existential experience" (Cunningham, 1988, p. 150). However, every patient's circumstances and experience is different, and for some the psychological sequelae are

more serious than for others. The following paragraphs provide background for the present study by describing the prevalence and nature of distress suffered.

Prevalence of psychopathology and symptoms

Psychosocial distress and psychopathology affecting cancer patients take varying forms, also varying over the course of the disease from symptom detection through diagnosis, medical treatment, survival, recurrence or disease progression, and palliative and terminal care. Studies of psychopathology prevalence vary widely in their findings. An early seminal study using formal psychiatric diagnoses found disorder in 47% of its randomly selected sample (Derogatis et al., 1983). Sixty-eight percent of these diagnoses were for adjustment disorders, mostly with depressed or mixed mood, but also with anxious mood. This translates into a prevalence rate in the base population of 32%. Disorders in which depressive affect played a prominent part (including adjustment disorder with depressed mood) accounted for 36% of diagnoses; and where anxiety was prominent, 16%. Together, depressive and anxiety disorders accounted for 85% of the diagnoses, amounting to 40% equivalent prevalence. These rates would presumably have been higher had not eligibility criteria for this study required a Karnofsky Performance Scale score of at least 50, excluding terminal patients who may be expected to suffer more distress associated with physical symptoms and grief.

Over the years there has been wide disagreement on prevalence rates, reflecting the disease-related and methodological complexity of research in this field (Jacobsen, Donovan, Swaine, & Watson, 2006; Trijsburg, van Kippenberg, & Rijpma, 1992; van't Spijker, Trijsburg, & Duivenvoorden, 1997) but most estimates agree upon rates above general population norms (Stanton, 2006). However, the prevalence of elevated distress among cancer patients may not be any higher than that among primary care patients: In a very large fairly recent study of patient records (n = 4496), half of which related to patients diagnosed in the previous 90 days, Zabora et al. (2001) found that, overall, 35% of cancer patients were clinically distressed. Meanwhile, a very large study of primary care patients (n = 18,489) found that 37% had clinically significant depressive symptoms (Herrman et al., 2002). However, this latter study itself portrays the difficulty in pinning down a distress prevalence figure since the six different countries that it drew data from produced rates ranging from 24 to 52%.

Nature of psychopathology and symptoms

As noted, psychological disorders tend to be adjustment and affective in nature (Derogatis et al., 1983). Underlying anxiety and depression can manifest in other common disorders too: A very large representative American household survey found a significantly increased likelihood of major depression, drug dependence, simple phobia, and agoraphobia in cancer patients (Honda & Goodwin, 2004). However, distress is often experienced as elevated but sub-syndromal symptoms. For example, anxiety may manifest as an increase in distractibility, absent-mindedness and loss of concentration, and depression as sadness, lowered self-esteem, persistent health concerns, uncertainty about the future and fears of death (Lovejoy & Matteis, 1997; Timms, 1990). Cancer

related depression has been described as similar to grief, but a pathologic response to the loss of certainty and normalcy in life (Lovejoy & Matteis, 1997). Sub-clinical depression symptoms in cancer patients include all of those which generally would, in the requisite numbers, qualify for the syndrome (Sellick & Crooks, 1999).

Such psychological symptoms affect and interact with marital and family support and circumstances including sexual relations, wider interpersonal relationships and social roles, work and finances, self-esteem, sleep, pain, activities of daily living, existential issues and grief, and satisfaction and compliance with demanding medical treatment regimes. At the level of personal meaning, Chekryn (1984) found that women suffering a recurrence of cancer felt uncertainty, grief, injustice, fear, anger, and existential concerns, and that there was not a sharing of the meaning of the recurrence between the patient and her spouse. The depth of the experience can be socially isolating.

Specific types of cancer and their treatments carry their own peculiar psychological impacts in addition to those more generally felt. For example, breast cancer patients may suffer effects to their body image and sexual identity and relationships, while head and neck patients may face major blows to their identity and self esteem following publically apparent disfigurement and loss of some ability to communicate, on top of the functional implications that arise from these. Hence the range of psychosocial impacts is broad, covering affect, functioning and every social domain. The present study focuses on just one aspect - affect.

Long term psychosocial impacts on cancer survivors have become more salient as medical treatments have improved. Stanton (2006) lists general impacts on quality of life such as reduced vitality and ability to participate in and perform physical and social roles and relationships given ongoing limitations, and goes on to explain that survivors are commonly left without the sense of well-being that they previously enjoyed. Many have to cope with lingering uncertainty and fear of recurrence or ongoing financial difficulties. She also points out that many cancer survivors identify a positive side to the cancer experience, however, in the personal growth they experience: Enhanced life appreciation and spirituality, a review of priorities and improved relationships, and the adoption of health-promoting behaviours.

Moderation of distress and intervention effectiveness

Cancer patients differ in their emotional responses to the experience. That truism indicates the working of moderating influences, meaning that some patients will be exposed to less distressing circumstances than others, and some will cope with their distress better than others. One factor drawn into the picture of distress is psycho-oncology intervention, but there are also characteristics of the patient and the disease that have critical influence, and all of these factors interact with each other (Sellick & Crooks, 1999). This means that the effectiveness of psycho-oncological therapy can depend, to a greater or lesser extent, on these other moderating factors. This section discusses what is known about some factors that moderate cancer-related distress, and begins with a discussion of how cancer distress is conceptualised, and how that level of

distress itself impacts the effectiveness of therapy. Moderation by therapy type is discussed later, after some descriptive information about therapies is given.

Conceptualising cancer distress

Which factors are recognised and addressed by clinicians and researchers will depend on how cancer distress is conceptualised by them. For example, coming from a traditional bio-medical model the focus will be more on medical factors: the initial emotional impact of diagnosis, the implications of recurrence or progression of disease, side effects of medical treatment, and cerebral dysfunction (Greer, 1987). Application of such a conceptualisation to a particular case might look like this:

“... a woman with breast cancer ... has undergone mastectomy and is currently receiving multiple chemotherapy [and] develops a depressive illness. Among the possible contributory factors (aside from any genetic predisposition) are (a) fear of a painful, lingering death, (b) loss of the breast with consequent loss of sexual attractiveness and self-esteem, (c) the side-effects of cytotoxic drugs such as fatigue, nausea, vomiting and alopecia [hair loss], and (d) hypercalcaemia.” (Greer, 1987, p.268.)

Psychiatrist Jimmie Holland – ‘the mother of psycho-oncology’ – has taken a broader psychosocial view of cancer distress, locating the patient in their social context (Holland, 1990a). She recognises the stress inherent in medical treatments that can be demanding, toxic, and temporarily or permanently disfiguring and disabling, and also that inherent in the threat of a foreshortened lifespan, but she also pays attention to the impact of disruption, disability, and ongoing uncertainty on marital, family, social, and vocational life and financial resources. This broader approach to distress moderation brings into view many more points of vulnerability, and also resources that can strengthen the patient to cope. Holland’s conceptualisation (1990a) falls under three heads:

1. Socio-cultural, i.e. social attitudes and stigma about the disease
2. Patient-related variables, including
 - (a) intrapersonal, namely age specific developmental tasks that are threatened or disrupted by the disease or its treatment, and personality, particularly in relation to coping abilities, and
 - (b) interpersonal, namely the quantity, quality and appropriateness to the moment of social support available to the patient; and
3. Medical-related variables, including cancer type, stage, prognosis, dysfunction caused by symptoms, treatment, rehabilitation options, and the psychological management provided by the health care team.

Holland’s scheme translates into a picture of vulnerability that includes low socio-economic status, social and marital problems and lack of social and spiritual support, a history of psychological problems, substance abuse, high anxiety, denial, suppression,

and general lack of effective coping abilities, complex medical factors, and a negative perception of the medical team and treatment. Greer's case example (1987, above) appears in a different light altogether when rounded out by the inclusion of such psychosocial details:

..... The woman is 55 years of age and recently divorced. She has been a heavy drinker 'off and on' for the last 20 years. She is overweight, and was diagnosed with diabetes 3 years ago. Our case subject has a deep-set suspicion of hospitals and medical procedures which she attributes to the experiences of near relatives, and is generally negative about the treatment she has received for both major diagnoses. In both instances, there was considerable delay before she initially sought medical advice and was diagnosed because, she said, she "...had too much else on [her] plate at the time, and just needed the symptoms to go away."

More recently, Knight (2004) has conceptualised the matter from the perspective of psychosocial risk factors. He summarised the literature as implicating young age, concurrent and previous psychiatric disorder or substance abuse, other concurrent stress (e.g. recent bereavement or other illness), previous adverse experience with cancer, difficulty accepting change in body image or physical function, pessimism, an avoidant coping style or suppression of negative feelings, pre-existing relationship problems (e.g. in marriage) or lack of social support, lack of involvement in meaningful activities, treatment with cytotoxic drugs, treatment associated with visible physical deformity, physical symptoms at diagnosis, and poor prognosis. More succinctly, Andersen (1992) puts cancer distress down to the patient's accumulation of losses and the strength of their social support.

In looking at it this way, both Knight and Andersen are emphasising the vulnerabilities or strengths that might leave patients more or less resourced to cope with cancer as a major psychological stressor, trauma, hardship or burden. This is effectively an application of broader theories of coping and self-efficacy (Bandura, 1997; Lazarus & Folkman, 1984). It makes available to psycho-oncology much of what has been learned about the treatment of distress in general or other chronically ill populations (Sellick & Crooks, 1999). Indeed, cancer distress can be conceptualised as simply the response of normal people undergoing abnormal stress, whose depressive or anxious reactions can be treated in standard ways (Cunningham, 2000).

Although there is a wide range of factors that could moderate an individual's emotional response to cancer, a synthesis of the literature is limited to investigating those for which data are sufficiently commonly reported by primary studies. These tend to relate to medical and socio-demographic variables, with a little psychological data. This is unfortunate because there is evidence to suggest that psychosocial variables are the more influential, and this may be because they are the more direct. For example, in a large cross-sectional study of early stage breast cancer and depression (Bardwell et al., 2006) it was found that neither cancer-related variables (stage, time since diagnosis, treatment protocol, tamoxifen treatment) nor personal characteristics (age, marital

status, ethnicity, education) nor health behaviours, nor even physical functioning / symptom variables (though these had the greatest influence), retained statistically significant impact on depression prevalence once psychosocial variables were added to the model (social support / strain, optimism, ambivalence over negative emotional expressiveness, life events, sleep disturbance). The authors concluded that subjective psychosocial factors were the more important predictors of depression to attend to.

However, there are other objective factors that were not part of that study (e.g. gender, income, other cancer sites, advanced stage disease) which may be of considerable importance, and objective indicators can also point to subjective factors.

Baseline distress

The present study seeks to identify moderators of therapy effectiveness. The most important meta-analytic finding to date is that the level of distress that patients bring to intervention, that is, their baseline distress level, strongly tends to predict the effectiveness of the intervention (Sheard & Maguire, 1999). The majority of studies do not screen potential recruits for baseline distress, treating them – as Sheard and Maguire put it - ‘preventatively’. This finding showed a dramatic difference in outcome: Anxiety: screened, Hedges’ $g = 0.85$ ($n = 3$ studies), unscreened, 0.33 (5); depression, screened 0.94 (4), unscreened 0.16 (5). The differences here, in clinical terms, are between small and large effect sizes. Because of insufficient statistical power, this difference did not reach statistical significance, but it makes perfect sense and conforms to general clinical psychology practice of treating proven need. If there is no need – no deficit - then it will be difficult to improve the patient’s wellbeing – show benefit. This moderation can also be thought of as an artefact of measurement, i.e. a floor effect impacts the post-intervention wellbeing scores of patients who are not distressed to start with. The tendency was also noted in a major recent review of psycho-oncology intervention studies with the expression, “once again ‘the rich did not get richer’, but rather those most in need experienced the greatest improvements (Sherman et al., 2004, p.356).

This tendency for ‘the poor to get richer’ in therapy forms the central thesis of the present study and a hinge between factors that moderate cancer-related *distress* and factors that moderate psycho-oncological *therapy effectiveness*. Factors that denote psychological ‘poverty’ are those that contribute to loss and diminish support, in Andersen’s terms (Andersen, 1992). The following sub-sections describe the literature relating to some of these factors, but as already noted, a study that synthesises primary data is dependent on frequently occurring data types, so the focus here is narrowed mostly to medical and socio-demographic variables, with a glance towards theoretical mechanisms.

Moderation by medical factors

There is an established association of increased psychiatric morbidity with increased physical illness and hospitalisation generally (Clarke, 1998). Distress prevalence is likely to vary with cancer and treatment site and stage, and these factors have been proposed as proxies for distress and possible moderators of therapy effectiveness accordingly (Andersen, 1992).

Cancer site

Different cancer sites (primary tumour site) confront patients with different issues in terms of treatment options, treatment side-effects and prognosis. Some cancers raise sexual and, therefore, relational issues, since treatment affects sexually significant organs. Others imply poor prognosis since they are seldom found in time for effective curative treatment. Some are well known to be causally linked to the patient's health behaviours, giving rise to feelings of guilt and shame. Some have greater functional implications than others. This variety implies that different sites will produce different prevalence of distress. The very large Zabora study mentioned above (Zabora et al., 2001) found that the highest rates of distress were in sufferers of lung (43.4%) and brain (42.7%) cancers. The authors attributed this in part to the difficulty of treating cancers at these sites resulting in poor prognosis. Lung cancer diagnosis also commonly carries self-attribution guilt. The breast cancer distress rate was 33%, prostate 31%, and the lowest rate was for gynaecological cancers at 30%. Depression has been highly associated with particular cancers by a recent review (Massie, 2004): oropharyngeal (22-57% prevalence range shown in studies reviewed); pancreatic (33-50%); breast (1.5-46%); and lung (11-44%). As usual, wide variations in the prevalence rates are notable. However, the lower limit of these ranges is interesting, showing that even in the most favourable of prevalence studies at least 22% of oropharyngeal patients and 33% of pancreatic patients are 'depressed' by some clinically significant measure. The former group is characterised by functional and identity issues arising from the disability and disfigurement that accompanies treatment, and the latter by dismal prognosis.

The prevalence of distress amongst breast cancer patients is of particular interest because of the disproportionate frequency that this group is studied – as shall be seen. In a recent large intervention trial that is relevant because of its proximity to home, Australian researchers Kissane et al. (2003) found that 37.3% of their early stage breast cancer sample had a DSM-IV diagnosis at baseline: 9.6% of the sample with major depression; 27.1% with minor depression (adjustment disorder with depressed mood, dysthymia and depressive disorder not otherwise specified); and 8.9% with an anxiety disorder (adjustment disorder with anxious mood, generalized anxiety disorder, and panic disorder). The overall rate sounds high given the favourable prognosis of this diagnosis and that participation criteria excluded patients with prior cancer history or with psychotic illness, dementia or intellectual disability. On the other hand, a larger recent correlational study (n = 3,088) found depression prevalence for women with this

diagnosis was much the same as that in the general population (17%)(Bardwell et al., 2006). Indeed, a main finding of that study was that such impact as cancer-related variables (stage, time since diagnosis, treatment protocol, and tamoxifen treatment) had on depression lost statistical significance once psychosocial variables were added to the model. Early stage breast cancer has a particularly good prognosis (98.3% five year survival rate relative to the general population, in year 2000, National Cancer Institute, 2009) and is not, therefore, the frightening diagnosis that some cancers are. The fact that perhaps two thirds or more of early stage breast patients are not clinically distressed should have implications for psycho-oncology intervention and research.

Disease and treatment stages

Cancer and medical treatment stages are relevant because distress peaks and dips over the trajectory of the cancer experience (Stanton, 2006). The initial impact of diagnosis and treatment is highly disruptive and can be traumatizing, even to the point of diagnosable post traumatic stress disorder for a small percentage of patients. However, for most patients, such anxiety or depression symptoms as may initially be experienced decline markedly over the first year or two after diagnosis. This may be particularly the case for breast cancer patients (Edgar, Rosberger, & Nowlis, 1992; Epping-Jordan et al., 1999). Once the initial diagnosis stage is past, and patients enter into treatment and recovery, concerns tended to broaden beyond focus on the disease to its wider social and other implications (Freidenbergs et al., 1981-1982). Disease recurrence has already been mentioned as a time of particular distress (Chekryn, 1984), and, logically, advanced disease stage is also. Miovic and Block (2007) reviewed the literature and found that about 50% of patients at that stage had psychiatric disorder, most commonly adjustment disorders (11-35%) and major depression (5-26%). Presumably many of the remaining 50% suffer sub-clinical symptoms. Pain is often noted as a distress correlate (e.g. Tatrow & Montgomery, 2006), and may be of particular relevance during advanced stage (Lovejoy & Matteis, 1997). However, advanced stage does not necessarily correlate with physical symptoms, and it may be that physical symptoms are a critical moderator of distress and therapy effectiveness (Given et al., 2004; S. L. Manne et al., 2007). A recent systematic review of ovarian cancer studies found a relationship between heightened distress and more advanced disease, more physical symptoms, greater disability, shorter time since diagnosis and poorer quality of life (Arden-Close, Gidron, & Moss-Morris, 2008). Suicide risk is heightened for terminal stage patients as well as those with a history of affective disorder, who abuse alcohol, whose judgment is disinhibited secondary to the use of pain medication, who suffer poorly controlled symptoms (e.g. pain, fatigue) or from a lack of social support especially through the recent death of a friend or spouse (Lovejoy & Matteis, 1997; Sellick & Crooks, 1999). Many of these factors can converge for elderly patients.

Different types of medical treatment carry particular psychological risks. Marked degrees of apprehension, depression and social withdrawal have been noted in association with radiotherapy (Forester, Kornfeld, & Fleiss, 1985). Distress or

depression has been found to be predicted by having had primary chemotherapy, recurrence chemotherapy, and having pain treatment with opioids (Gotay et al., 2007). Many therapies target the period when patients are in radio- or chemo-therapy because of this. As has been mentioned, the psychosocial implications of disfigurement or disability resulting from surgery (e.g. mastectomy, head and neck surgery, radical prostatectomy) are also distressing, and are therefore another common target of intervention.

Moderation by socio-demographic variables

Whilst demographic characteristics are thought of individually, as moderators of distress they are actually very much linked to each other and to other moderating constructs and are usually discussed with them in the literature. For example, it has already been noted that elderly patients can accumulate medical predictors of distress, and it will be seen that they are also vulnerable to socio-demographic ones. Although artificial to do so, some teasing apart of these moderators is necessary for the purposes of discussing their influence.

Gender.

There appears to be difference in the way that men and women manifest cancer distress. In the very large community survey mentioned above (Honda & Goodwin, 2004) it was found that the rate of male cancer patients developing depression was 533% of that amongst non-patients, whereas there was only an 81% higher rate in females. Overall, cancer patients were nearly four times more likely to be clinically depressed. These figures derived from non-institutionalised people only, so presumably they minimize the effect that would be seen if data were collected from hospitalised patients also. It was suggested that differences in perceived personal control and social support may be the cause of the gender differential. The same study found that about 18% of male cancer patients suffered drug dependence compared with 3% of female patients. Among females, it also found only one disorder for which cancer patients had a statistically significant increased prevalence, namely social phobia, but that PTSD (post traumatic stress disorder) and agoraphobia came close. For men, the only such anxiety disorder was panic attacks. Something of a sex-typical emotional response is seen in this, with women tending to manifest anxiety, while men externalise depression in substance abuse. A similar dynamic was found in a large recent German study where being female proved to be a significant risk factor for anxiety disorders and being male, for addictive disorders. However, more broadly, it also found that having previous mental trauma was a risk factor for anxiety disorders; that being female and having poor physical functioning were risk factors for affective disorders; and that younger age and having distant metastases were risk factors for addictive disorders, (Krauss, Ernst, Kuchenbecker, Hinz, & Schwartz, 2007, German language, abstract only sighted).

There are very few psycho-oncology intervention studies that have broken out outcome analyses by gender. One that did was mentioned above (radiotherapy) and found that men were more distressed at baseline and also gained more from the intervention

(Forester et al., 1985). In their recent German meta-analysis, which included the consideration of a few moderating variables, Rehse and Pukrop (2003) found that men produced an effect size about twice that of women. From such outcomes the suggestion arises that psychological intervention is able to disproportionately meet psychosocial needs of men – perhaps needs to regain some vital sense of control or tap into social support - thereby providing them with a disproportionate gain from therapy. If a disproportionate gain to them was based on need – on deficit - then the flip side of that coin is to say that outcomes for women suffer from something of a floor effect – they have other needs, or other ways of meeting their needs, and therefore do not show such great benefit from therapy. An attempt to replicate Rehse and Pukrop's finding will be interesting.

Age.

Younger patients are consistently found to have higher prevalence of anxiety and depression, according to a recent systematic review (Jacobsen et al., 2006). This was confirmed in a recent meta-analysis of breast cancer findings which used 40 years as the cut-off to compare older and younger patients: A large effect associating younger age with dissatisfaction with body image was found, and moderate magnitude effects associated age with distress and emotional function (Reis, 2007, foreign unpublished thesis, abstract only sighted). The ovarian cancer review mentioned above (Arden-Close et al., 2008) found strong evidence of this same relationship, and a study among lower income women (Ell et al., 2005) also found that younger age (< 50 years) was significantly correlated with depression.

However, excluding the review, these studies are all of female patients with sex-specific cancers. Furthermore, there is a greatly disproportionate frequency of breast cancer patient studies in the literature. It can be expected that, as a matter of developmental stage, younger patients will feel losses that relate to sexual attractiveness and reproductive ability more acutely. In addition to this, cancer tends to be more vigorous and life-threatening in younger patients. It may be that the true factors at work are loss and grief, as suggested by Andersen (1992), and that when a more representative range of cancers is considered, the association of younger age with greater distress weakens. Supporting this suggestion, the very large Zabora study (Zabora et al., 2001) found only a weak inverse relationship between distress and age.

Patients at the other end of the age range have to cope with a great deal of loss also. Elderly patients suffer more co-morbidity and associated disability (Yancik, 1997), loss of social support and income. It has been suggested that age in itself neither predisposes nor protects from distress, but that age-associated factors like these are the real predictors (Snowdon, 2001). Alternatively, such factors have been conceptualised as mediating the relationship between age and distress (Mosher & Danoff-Burg, 2005).

The severity of pain and its interference with everyday activities is another important factor, and advanced disease is more common amongst the aged, since cancer is generally a disease of older age. In a recent study of 120 advanced patients

(Mystakidou et al., 2006) it was pain rather than demographic or other medical variables that predicted anxiety, and – contrary to findings noted above - older age and female sex predicted more depression. Marriages between older people tend to be particularly satisfying (Levenson, Carstensen, & Gottman, 1993), and a main effect of marriage protecting from distress has been found in this age group, with moderation by the quality of the marriage (in terms of perceived relational equity) (Hagedoorn et al., 2006). This means that older people have much to lose from spousal bereavement, especially as they suffer losses to their social network generally. It has also been noted that older people may respond differently to questioning about their symptoms and feelings, resulting in lower rates of disorder being recognised despite high rates of serious sub-syndromal depression in this population (Snowdon, 2001).

It may be that a U-shaped relationship will eventually be found between age and cancer distress, as subpopulations at each end of the age range may tend to be particularly vulnerable in terms of other predictive socio-demographic, medical and psychological moderators and mediators which represent accumulated losses and/or diminished social supports.

Race.

Disparities in the incidence and mortality of cancer patients on racial grounds suggest that non-white patients ought to suffer poorer physical and mental functioning. However, although a link between physical functioning and race has been found (Eton, Lepore, & Helgeson, 2001), a link with mental well-being has proven more elusive. The large recent American study by Zabora et al. (2001) mentioned above found only a very small increase in prevalence of distress amongst patients who were not Caucasian. Other studies have found no difference on the basis of race (Ell et al., 2005; Friedman et al., 2006). However, as will be seen in the present study, the great majority of intervention research participants are white, and it may require greater numbers of non-white participants in order to pick up a subtle disparity - as in the Zabora study. If that is the nature of the effect, it suggests that other predictors of distress are more directly responsible for moderation.

Education and income.

Although there appear to be no recent reviews on the impact of income or education on cancer distress or psycho-oncological therapy effectiveness, there are hints from a few individual studies that align with the thesis that more vulnerable patients – those with less educational and financial resources - are likely to suffer more distress and therefore find therapy more beneficial.

A weak inverse relationship between distress and income was found in the Zabora study (Zabora et al., 2001). Eton et al. (2001) found that higher education level was associated with better mental functioning. In another study by the same team (Lepore, Helgeson, Eton, & Schulz, 2003) an educational intervention was tested and a significant interaction was found such that only the less well educated men who

received the intervention improved their physical functioning. The researchers posited that poor uneducated minorities suffer more from cancer in terms of mortality and quality of life, and therefore they stand to gain more from educational interventions. Another study with a similar design arrived at a similar outcome, but this time regarding psychological distress outcome rather than physical functioning (Lerman et al., 1996): A significant education level by treatment interaction confirmed a slight rise in distress over time for the lower educated untreated group, while both higher educated groups and the treated lower education group became less distressed.

Although these findings make sense in terms of the deficit thesis, the evidence for this inverse relationship is far from abundant, and there is research that shows no such relationship (e.g. between education level and depression: Ell et al., 2005). Part of the problem may be that income and education extremes are not sufficiently represented in the populations being sampled. Most of the studies are from the United States where, though far from what many New Zealanders would consider an egalitarian society, welfare exists and illiteracy is low. In studies where poverty and illiteracy are more pervasive and more highly represented in study samples, strong effects from educational interventions could be expected – and have been found: Very strong effects with such samples were achieved from educational interventions with cancer patients from Egypt and Puerto Rico (Ali & Khalil, 1989; Corchado, 2006). However, drawing a conclusion from these studies is not so straightforward since it appears also that information about cancer was not as freely available in those societies as it is in the West.

In order for the inverse relationship between cancer distress and education level or income to become visible it may be necessary to turn to samples where lower status on these socio-demographics is more extreme, or else to use a very large sample, as the Zabora study did (Zabora et al., 2001). A meta-analysis with a large enough database to draw on may be able to detect an effect – one of the objectives of the present study.

Marital status.

The marriage relationship is a particularly influential source of social support for cancer patients, and is discussed in those terms later in this section. Marital status *per se* becomes an indicator of interest only because data indicating of the quality of the support offered in patients' marriages and data regarding the availability of other social support resources are seldom reported. In that sense, it becomes a proxy – though inadequate – for social support.

Moderation by psychosocial resources

Psychosocial resources – both internal strategies and environmental supports – mesh with medical and socio-demographic variables and yet, as noted, most psycho-oncological intervention research has been a-theoretical, rather than springing from theory-driven models that name theoretical mechanisms of action. Where psychosocial resources have been measured, it has usually been as outcomes rather than for the

purposes of mediator / moderator analysis (Owen et al., 2001). The present study will confirm that relatively few studies provide useful data on such mechanisms.

There are many constructs that could constitute mechanisms of intervention action. Owen et al. (2001) look over the literature and recite perceived health, locus of control, coping, denial, self-efficacy, self-esteem, personality, knowledge of the disease and / or its medical treatments, purpose in life, alienation, and life satisfaction. In the context of group interventions, Sherman et al. (2004) add reassuring social comparisons, emotional disclosure and deeper levels of emotional processing. Social support is another, although its effects may themselves be mediated by some constructs on the former list such as self-esteem, coping, and emotional disclosure.

Because of the dependence of the present study on regularly reported data, the lack of attention to this area limits investigation to only a few more commonly measured psychosocial resources / potential theoretical mechanisms. It was also necessary to group them quite broadly in order to gather sample size. The selection of constructs was: 1. Social support, proxied by marital status; 2. Self-efficacy, which was grouped with perception of personal (locus of) control, dispositional optimism and (the converse of) helplessness; 3. Coping style (as opposed to particular coping strategy), simplified to approaching or avoiding threats; and 4. Self-esteem, grouped with self concept. Only a brief discussion of these constructs and how they may modify distress and therapy effectiveness is justified in what follows. However, it will be noted that the theme of baseline deficit predicting intervention effectiveness recurs.

Social support

Distress moderation.

The adequacy of social support and psychosocial adjustment has been frequently linked in both general psychological and psycho-oncological literature (e.g. Eton et al., 2001). Social support may be available from the patient's natural networks – spouse, family, friends, workmates, and community groups - or from a therapeutic group or therapist. In the context of an interpersonal psycho-oncological therapy, Badger and colleagues have defined social support as affective, instrumental, informational, and appraisal support (Badger et al., 2005, p. 274). They attribute its effect substantially to the resources it makes available for working through the emotional impact of a stressor, and marshalling aid for making assessments of the threats it poses, obtaining helpful information, and adjusting roles and functions appropriately in order to handle it. However, the meaning and importance of social support to a particular person will vary according to individual circumstances, including economic circumstances (Coyne & De Longis, 1986).

Social support is known to co-vary with survival in relation to medical conditions generally (Andersen et al., 2004; Spiegel, 2002; Spiegel, Sephton, Terr, & Stites, 1998). People who are socially isolated or without confidants (and people with lower socio-economic status) are exposed with regard to all-cause mortality, while married cancer

patients and those with strong support from friends, relatives, neighbours and employers have been found to survive longer. Social support can also be vital to adjustment following invasive medical treatment. In the context of recovery from mastectomy, partner support has been seen as particularly important to patient psychological adjustment (Bultz, Speca, Brasher, & Page, 2000; Northouse, 1988).

Support will vary in quality and quantity between relationships. While most cancer patients feel that they have very supportive and satisfying social networks, marriage partners and other supporters are not always helpful. Some do not allow the patient to express his or her distress or may minimize the distress or push the patient to be 'positive' and cheerful. Some supporters will even withdraw from the relationship. From the patient's point of view, such responses are obviously unhelpful and are known as 'social constraint'. They arise from the supporter's belief that dwelling on the negative will be detrimental to the patient's health, or from the supporter's desire to avoid discomfort (Gottlieb & Wachala, 2007). It is also common for cancer patients themselves to withdraw out of consciousness of their increased need, illness stigmatization, and concerns about burdening others. In doing this they increase their isolation and distress (Leszcz & Goodwin, 1998).

Obviously, if a patient is married, that relationship is the first port of call for social support. However, it has been shown generally, and for cancer patients specifically, that the quality of the relationship is critical, with poor marriages producing poorer psychological well-being than singleness, but happy marriages producing the greatest sense of well-being (Hagedoorn et al., 2006; Williams, 2003). Marriage quality has been measured with reference to a number of dimensions including communication / social constraint, harmony / conflict, and equity of contribution. Cancer has the potential to put pressure on all of these aspects of the relationship. It has already been noted that marriages between older couples tend to be particularly satisfying (Levenson, Carstensen, & Gottman, 1993).

Intervention and peer support groups.

Interventions may aim to provide an effective social support network for a patient, or may attempt to strengthen the support networks the patient already has by means of relational skills training (e.g. assertive communication with medical staff, addressing difficult health-related communications with family and workmates, and making strategic use of present networks). Therapy peer support groups can help patients by filling a general need for social support and in other distinctive ways. They are thought to buffer against the stresses of living with the uncertainty of cancer by providing a resource for the appraisal of threats (stress and coping theory: Gottlieb & Wachala, 2007; Leszcz & Goodwin, 1998). They can also improve treatment adherence (Leszcz & Goodwin, 1998) and other disease related coping, reduce stress through emotional expression, and give members the opportunity to feel useful by helping others (Gottlieb & Wachala, 2007). Peer support can normalise patient experiences and feelings and instil hope by providing an opportunity for comparison with peers (social comparison

theory)(Gottlieb & Wachala, 2007). Support groups can support learning in that they provide a context for rehearsal, reinforcement, clarification and modelling of skills and information taught (mastery – self efficacy theory)(De Lorenzo et al., 2004; Gottlieb & Wachala, 2007).

It is not a given that therapy groups will provide a positive supportive effect for a given patient, and negative impacts are experienced by some. They deliver a form of support that is different from that which natural supporters provide, and may not be successful for some patients who need more than merely the support of peers and a passive facilitator (Helgeson, Cohen, Schulz, & Yasko, 2001). Group leaders see both positive and negative effects. They have reported, on the positive side, that groups can allow members to feel accepted, have the opportunity to hear different perspectives, feel cared for, have their experiences normalised and enjoy emotional release. On the negative side, they have reported that some group members can find open communication threatening, feel a sense of loss when a member leaves, feel overwhelmed or embarrassed, or can become too dependent on the group (Galinsky & Schopler, 1994). Gottlieb and Wachala (2007) add that some patients decline to attend support groups because they feel that they have sufficient or better avenues for information or support elsewhere and others do not feel comfortable with the group format or the requirement to disclose to a group of people. Some fear hearing others' distressing experiences or watching them deteriorate in health. Some members drop out due to personal clashes or boredom or the feeling that their personal needs can be met better elsewhere. They also note a concern about possible rebound effects following group termination. From the writer's personal experience, an additional deterrent from joining a group might be the amount of prominence in the patient's identity that the time and effort invested in group attendance would give to the cancer diagnosis.

Prediction of therapy effectiveness.

Although there is little evidence in the psycho-oncological intervention literature, there are indications that the amount or quality of the social support naturally available to a patient will moderate benefit from therapy. In their intervention study with 230 breast cancer patients, Helgeson, Cohen, Shulz, and Yasko (2000) found that whether an education group or peer discussion group intervention resulted in improved physical indicators depended on the presence and quality of support from patients' spouses and oncologists. Those who lacked support gained most from interventions, rather than remaining the same or slipping in their physical well-being, as did comparable controls. However, those who were well supported gained from the educational intervention but actually deteriorated from participation in the peer discussion. The authors attribute the gains of those without strong support to meeting a deficit, and, contrariwise, the deterioration of those who had felt well supported at baseline, possibly to an alteration of their perception of that support or to shaking patients out of a false concept of how supported they were by promoting discussion in their natural networks that transpired to be less helpful than expected. In a trial of a couple focused group intervention for

breast cancer patients it was found that emotional expression and processing as well as patients' perception of the supportiveness of their partners moderated therapy effectiveness (Manne, Ostroff, & Winkel, 2007b; Manne et al., 2005). And in a study of expressive writing, it was found that social constraint at home resulted in greater effect (Zakowski, Ramati, Morton, Johnson, & Flanigan, 2004). The importance of intervention meeting deficit is suggested again by these studies.

Social minorities might be thought of as vulnerable in terms of the social support provided by medical health systems, which are geared to the dominant culture. Walker and Podbilewicz-Schuller, 2005 ran a trial comparing education material for newly diagnosed patients in preparation for their initial multidisciplinary clinic visit. They found interactions whereby unmarried patients, patients with a history of mental health treatment, and racial minority patients showed lowered distress, improved general well-being, and greater satisfaction (respectively) with the treatment. The authors posited that the treatment may have compensated for some of the lack of support endured by unmarried patients, modelled a problem-focused non-passive mode of coping that was particularly helpful for patients with a mental health history, and reassured African American patients that they could expect caring and respectful treatment by the visual cues portrayed. In this example the interrelationship of demographic and theoretical moderators is seen. Meeting deficit is also the critical factor, but if the authors are correct, the importance of accurately identifying the exact nature of that deficit and meeting it is also seen.

It has been mentioned that a synthesis of literature is limited by the data available from primary literature, and that means that the present study cannot directly tap quality and quantity of social support in an ideal direct way. Marital status is one proxy that will be used, and others that might assist – though even more distant include whether therapy was delivered in group or individual mode, whether a significant other was present in therapy, whether communication or relationship skills were taught, and the intensity of the therapist / patient relationship.

Self-efficacy

Bandura (1997) theorised self-efficacy as confidence in one's ability to perform a particular task. 'Mastery'- this sense of confidence - is derived from successful direct experience, from vicarious experience (i.e. comparative observation of modelling by similar others), or from verbal persuasion, and is best conveyed in a controlled state of arousal. Self-efficacy in relation to being able to make a difference to one's own health is obviously relevant to cancer patients. Evidence has been adduced showing that it affects treatment adherence, quality of life, and physical and psychological symptoms (Lev, 1997). For example, a low level of a similar construct, dispositional optimism, at diagnosis has been found to predict anxiety or depression symptoms at six months, and was itself partially moderated by a coping style (see below) that was not active but disengaging (Epping-Jordan et al., 1999). Perceived personal control and optimism are constructs that are closely related to self-efficacy. Optimism was found to predict

distress or depression in a large sample of breast patients with first disease recurrence (Gotay et al., 2007) and emotional well-being in a study of minority race breast cancer patients (Friedman et al., 2006), while in that same study pessimism predicted mood disturbance. In the context of cancer, this collection of concepts is about feeling some measure of self-determination in what is an experience characterised by high-stakes uncertainty. There is some evidence that baseline deficit in them predicts greater effect from therapy. Lower baseline optimism predicted reduced prevalence of moderate depressive symptoms in a breast cancer sample who received a stress management group intervention (Antoni et al., 2001).

Self-efficacy theory was applied by Weber et al. (2007) in their intervention trial for prostate cancer patients. They describe the importance of self efficacy as a resource that buffers the association between cancer, physical dysfunction, symptom management, and depression, and designed their intervention to bring Bandura's means of gaining mastery to bear. Weber and his team paired radical prostatectomy survivors with men who had recently received this treatment to provide them with modelling, encouragement (persuasion), to normalise their situation, in order to provide skills and reduce fear (control arousal). The intervention was successful in both lifting self efficacy and lowering depression. Having a sense of personal control has been found to be especially important to men in the context of depression (Watts, 2002), and therefore intervening to enhance self-efficacy was meeting an area of particularly felt deficit.

Coping

Coping theory (Lazarus & Folkman, 1984) can be summarised quite briefly: Whether a threat is coped with or becomes an overwhelming stressor depends upon a person's primary appraisal of the intensity of a threat and his or her secondary appraisal of the availability of resources to meet it. Coping is thus a process involving interaction between the individual and their environment, and a person's typical way of responding to threats is known as their coping style (Brennan, 2001). Coping can improve through altered appraisals, increased personal resources, or more active behavioural or cognitive style in which the threat is 'approached', rather than 'avoided' (which includes passive acceptance, denial and escapism) and allowed to grow worse. Generally, patients who have an active style report more positive affect, and vice-versa, and interventions that promote active coping may therefore be helpful (N. W. Fawzy, 1995).

For example, information-seeking is a strategy to increase resources for coping. Summarising work by Weisman and colleagues (1976, 1976, and 1979) Freidenbergs et al. (1981-1982) noted that 'good copers' face facts, find something favourable, and then confidently comply with medical recommendations, while 'poor copers' use suppression, passivity, stoic submission and a variety of attempts to reduce tension. In their meta-analysis of psychological sequelae following cancer diagnosis, van't Spijker, Trijsburg, and Duivenvoorden (1997) found that coping style significantly related to the prevalence of distress in that a cluster consisting of confrontation, fighting spirit or optimism related positively with psychological adjustment, while a cluster consisting of

passive acceptance, helplessness, anxious preoccupation, avoidance, denial, feelings of loss of control, or fatalism were negatively related. Such findings confirm the importance of active coping behaviour and also the power of self-efficacy in moderating cancer related distress.

As alluded to, education or information are often seen as contributing to self-efficacy or coping. The provision of concrete information to cancer patients can be conceptualised as helping them with problem solving and coping via self regulation theory – the information provided redirects attention from emotional responses and sources of stress to objective features of the situation or symptoms and thereby to achieving functional outcomes, which is then productive in reducing the impact of the stressor (Allard, 2006). Early researchers Bloom, Ross, and Burnell (1978) cited Janis (1958 and 1978) who theorised that information about threatening events reduces stress because it provides the patient with a framework from which to appraise the danger, order his or her thoughts, feelings and behaviour about it, and begin imaginary rehearsal of how he or she might deal with it, thereby undertaking the ‘work of worrying’ in anticipation. Janis also proposed the importance of a continuous supportive relationship with a member of medical staff in order to provide the motivational and emotional prerequisites to make the receipt of information influential. Their intervention therefore involved information and support from an oncology counsellor who became part of the medical team and continued to be available to patients before, during and after medical care. Helgeson, et al. (2000) said that their educational intervention’s theme was control in that it provided information to reduce confusion and uncertainty and enhance the patient’s control over their experience of cancer.

There have been findings in the intervention literature that support the deficit thesis of the present study. Scheier et al. (2007) found that some early stage breast cancer patients benefited from their cancer education and coping skills group or nutrition education group interventions, and some did not. Those who benefited suffered initial deficits in dispositional optimism and in their social environment. They noted that their finding was in line with resource theories of stress and coping. The education intervention studies by Helgeson, et al. (2000) and Lepore et al. (2003) discussed above also found that patients who had less education benefitted more.

Self-esteem

The impact of self-esteem has proven to be illusive to pin down in the context of cancer. Katz, Rodin, and Devins (1995) provide the only review of the theory and research relating to self-esteem and cancer. They explain that self-esteem can be defined in different ways and globally or as multiple dimensions. Essentially the construct refers to evaluations of aspects of one’s self, and four broad emergent domains relate to body image, social place and likability, achievement, and ‘identification’ (an evaluation of one’s moral and spiritual beliefs and character). Self-esteem can be conceptualised as a causal or moderating factor, for example in coping, problem solving, treatment compliance or seeking out social support. It can also be seen as an outcome impacted

by life events, such as loss of health or employment, and social feedback. It is not to be assumed that stressful life events will ultimately result in lowered self-esteem however, as it may be that successful handling of the crisis has the effect of building self-worth, or that one dimension is bolstered (e.g. identification) while another is diminished (e.g. body image).

Clearly cancer poses challenges to self-esteem, and body image is the domain most consistently found to be negatively affected, with evidence from mastectomy, head and neck, and colon patients who have undergone major surgery. Patients also have to accept altered social position and roles including, possibly, interference with sexual relationship and increased dependency. Some are burdened by guilt, and self-esteem may also be lowered as a consequence of depression. There has been relatively little cancer patient research into these latter domains of self-esteem.

Because self-esteem can be conceptualised as a moderator of distress, some interventions target it directly, for example, therapies that focus on body image for women with breast cancer, and the 'Look good, feel better' programme offered to patients undergoing chemo- and radiotherapy by voluntary organisations. However, Katz et al. (1995) concluded that in fact consistent disturbances in self-esteem after cancer diagnosis have not been found, and nor has consistency been found in the relationships between self-esteem and other constructs, or as a result of intervention. A significant impediment that they explain is that cancer or illness specific measures do (did) not exist, and that the extant measures have other psychometric flaws relative to their dimensionality.

There is some more recent literature that provides hints about the dynamics of this construct, however. A study of ovarian cancer patients found that self-esteem largely mediated distress predicted by unsupportive family and friend behaviours (Norton et al., 2005). Another study also found that negative self-esteem was moderately associated with various aspects of social support, but that positive self-esteem was only weakly so (Schroevers, Ranchor, & Sanderman, 2003). These studies suggest that self-esteem is more readily stripped by negative social support than built by positive social support. Confirmation of an association between distress and depression on the one hand and a relatively high sense of a related quality, coherence (i.e. a view of their lives as comprehensible, manageable, and meaningful), was found in the study of breast cancer patients with first recurrence mentioned above (Gotay et al., 2007). And finally, the thesis that deficit predicts therapy benefit was supported by a finding from the education arm of an intervention trial for breast cancer patients (Helgeson et al., 2000) that those with low initial scores on a 'personal resources index' comprising self-esteem, body image, control, and uncertainty about illness, gained more benefit from the intervention on their physical well-being indicator. It would be helpful if the present meta-analysis could shed any light on this complex variable.

Types of intervention for distress

In concluding their psychopathology prevalence study, Derogatis et al. (1983) observed that since the bulk of diagnoses in cancer patients are affective they are potentially very treatable. Fawzy, Fawzy, Arndt and Pasnau (1995) surveyed the different types of psychosocial therapies, categorising them into educational, behavioural training (including treatments to reduce chemotherapy effects and treatments targeting general distress), individual psychotherapy, and group interventions. Education and behavioural training aim to increase patients' mastery and hope by arming them with useful information or with stress management tools. Psychotherapy tends to be one of two kinds: cognitive-behavioural psychotherapy which aims to strengthen patients' ability to cope by identifying and correcting unhelpful thoughts, feelings and behaviours and assisting with problem solving (these strategies are usefully and briefly described in Lovejoy and Matteis, 1997); and the provision of social and emotional support relevant to the cancer experience. Individual and group interventions overlap in content, and lay persons (volunteers, often survivors) administer some treatments. Also some research exists about treatments that are carried out indirectly, through intervention with others who are significant in the lives of patients - usually spouses or doctors. Treatments also vary in delivery form from one-off sessions or the provision of educational materials, through brief structured formats with sessions running over less than six weeks, and longer structured or unstructured formats, to therapies that may be undetermined as to both specific content and length.

Andersen (1992) attempts to conceptualise the appropriateness of particular types of treatment according to the degree of patients' risk of psychological distress by virtue of the nature and stage of their disease. The disease dimensions that she takes into account are: local as opposed to disseminated disease; intensity of medical treatment required; and favourable as opposed to dismal prognosis. Andersen posits that low or medium risk patients, that is, patients at diagnosis, treatment, or early recovery stage, tend to have a 'best fit' with treatments that are of crisis intervention or brief nature, that is early assessment, present-day focus, limited goals, therapist directed, and prompt. She considers that the mechanisms for success may be the same as for effectively intervening to cope with other stressors, namely, learn about the stressor, confront it with positive cognitive states, apply active behavioural strategies, make realistic appraisals as stress declines, and enhance self efficacy and feelings of control early in the adjustment process. She suggests that helpful to these ends are an emotionally supportive context to address fears and anxieties about the disease, information about the disease and its treatment, and behavioural and cognitive coping strategies, including relaxation training to lower 'arousal' and / or enhance sense of control.

For patients at high risk of psychological distress, Anderson's scheme envisages many of the same components, but with a shift of emphasis to specific death or quality of life issues. She suggests that the same underlying mechanisms (i.e. enhanced self-efficacy, control and realistic appraisals) are at work, but that for these patients, intense, high quality social support is required through a therapist, therapeutic group, or spouse in

order to strengthen these patients, in view of their worst fears having been realised. This quality of support will mean sustained therapy over time.

Whilst it is debatable that patients at time of diagnosis tend to be at only low or medium risk of psychological distress, F. I Fawzy (1999) agrees that more structured and briefer therapies tend to address stages in the cancer journey that require intense adjustment to new challenges. She tabulates a summary of which psychosocial treatments best address each stage in the cancer experience, but concludes that patients need to be given the strength offered by all three of the main therapy approaches, namely, education, cognitive and behavioural coping skills, and social support. Indeed, much of the intervention research applies integrative treatment approaches, as has been noted. In considering the literature on medical factors and appropriate therapy, it pays to keep in perspective that medical factors may not be the main predictors of patient distress, but only a catalyst. A broader psycho-social conceptualisation of cancer distress, as discussed previously, is appropriate.

Moderation by therapy variables

Therapy type

Therapy type is the potential moderator of intervention effectiveness that has received most attention in psycho-oncology meta-analyses, which have otherwise focussed on main effects. A summary of effect sizes relative to the different therapies and populations that have been analysed is set out below (Table 1-1). Direct comparison is problematic since categorisations vary, but a couple of points are noted: First, all of the therapy types investigated by Devine and Westlake (1995) produced moderate to large effects (Cohen's *d*'s of 0.40 – 0.74) sustaining their conclusion that there was no statistically significant difference in the performance of different therapies and therefore clinicians could choose between a wide range to suit individual circumstances. This is saying that therapy type is not an important moderator of intervention effectiveness. Second, there is a wide range between meta-analyses in the results produced by some therapy types. For example, CBT scores a high Hedges *g* of 0.81 for Cwikel, Behar, and Rabson-Hare (2000), extraordinarily high *g*'s of 1.99 and 1.20 for Osborn, Demoncada, and Feuerstein (2006) and a negligible magnitude *d* of 0.13 for Tatrow and Montgomery (2006). The implication from this is that other moderators are at work relating to the design of primary studies – what research designs or populations they admit, for example. Both of these points give the present analysis some idea of what to expect in findings, and the latter shows how important design decisions will be to the validity of conclusions drawn.

Table 1-1. Earlier meta-analyses, summary of therapy type effect sizes

Meta-analysis	Therapy	Anxiety	Depression	Distress
Devine & Westlake (1995)	Education only	d = 0.74	d = 0.50	-
	Muscle relaxation only	d = 0.60	d = 0.40	-

	Other relaxation / distraction only	d = 0.66	-	-
	Muscle relaxation with guided imagery	d = 0.62	d = 0.40	-
	Multiple behavioural interventions with relaxation	d = 0.59	-	-
	Education with other behavioural treatments	d = 0.46	-	-
	Education with behavioural or non-behavioural counselling	d = 0.52	-	-
	Expressive-supportive counselling	-	d = 0.66	-
Sheard & Maguire (1999)	Group psycho-education	g = 1.59	g = 0.94	-
	Relaxation	g = 0.21	g = 0.03	-
Cwikel, Behar, & Rabson-Hare (2000)	CBT	'psychological outcome' g = 0.81		
	All other treatments	'psychological outcome' g = 0.49		
Rehse & Pukrop (2003)	Educational treatments	'emotional adjustment' d = 0.96		
	Social support	'emotional adjustment' d = 0.58		
	Coping skills training (CBT)	'emotional adjustment' d = 0.48		
	'Psychotherapy'	'emotional adjustment' d = 0.58		
Tatrow & Montgomery (2006)	CBT for breast cancer patients	-	-	d = 0.13
Zabalegui, Sanchez, Sanchez, & Juando (2005)	Social support groups	z score = 0.71	z score = 0.63	-
Luebbert, Dahme, & Hasenbring (2001)	Relaxation training	d = 0.45	d = 0.54	-
Osborn, Demoncada, & Feuerstein (2006)	CBT for survivors	g = 1.99	g = 1.20	-

Therapy delivery

The dose and means of delivery of a therapy might be expected to influence the benefit received. Primary studies commonly experiment with different delivery packages, comparing one-off interventions with lengthy programmes, simple brochures with multi-media educational packages, delivery by nurses or social workers with delivery by psychologists, etc. For ethical reasons, or in order to address placebo effects, many studies provide some kind of attenuated treatment to controls. However, there is very little in the psycho-oncology synthesis literature about delivery factors. Writing about group interventions, the review by Sherman et al. (2004) note a tendency for long term interventions to be provided for patients with advanced disease. The meta-analysis by Rehse and Pukrop (2003) found a greater effect for interventions lasting more than 12 weeks. And the meta-analysis by Cwikel et al. (1999) found that interventions delivered by social workers performed poorly. However, given the great range in participant characteristics, interventions and study designs in this field, such broad generalisations are of questionable validity and utility.

Psycho-oncological services

Having discussed some of the moderating variables that will be investigated in the present study, a more general description of the field is returned to by way of further orienting the reader to the context of the study and salient issues. The following sections discuss psycho-oncological services in terms of their financial value to the health system, their users, and how patients are assessed for distress and screened for intervention.

Financial burden on health system

Besides the toll of psycho-oncological distress on the life quality of patients and their families, the financial burden for the health system is likely to be significant. The general association between illness, psychiatric co-morbidity, and health service use is established (Clarke, 1998). In their prospective study of the impact of psychological illness on medical care cost, Levenson, Hamer, and Rossiter (1990) found that general medical inpatients who screened as anxious and/or depressed accumulated hospital costs 35% higher than others, and stayed a median of 40% more time in hospital, despite averaging the same on an indicator of disease severity. They also noted that half of the patients reporting high levels of pain also had elevated levels of depression or anxiety. Many other studies of general medical patients have shown similarly that heightened distress results in greater costs to health services, and that psychological services can therefore more than offset the cost of delivering them (Carlson & Bultz, 2004). Indeed, part of the reason for treating psycho-oncological distress – some would say the reason (e.g. Redd, 1995) - is to support the work of oncologists by reducing distress and increasing medical treatment compliance.

Although medical cost offset studies are few in psych-oncology intervention research (Owen et al., 2001), there seems to be no reason why results found for other medical

patients would not generalise. Further, as cancer treatment and survivorship rates improve, the need to deal with cancer-related distress grows. Stanton (2006) has noted that survivors who suffer ongoing depression may be twice as likely to utilise inpatient services. Because the effective treatment of psycho-oncological distress is a financial priority, locating the factors that make therapy more effective is also a financial priority.

Use of psychological services

In New Zealand, dedicated psycho-oncology services are quite new. Cancer patients have had limited access to health or clinical psychologists working for district health boards (DHBs) in the past, and there has been a team of health psychologists working for the Cancer Society of New Zealand in Auckland, but it was only in February 2006 that the first DHB-contracted dedicated psycho-oncology service began. This service, staffed by clinical psychologists, is run jointly by the MidCentral Health DHB and the Massey University Psychology Clinic at Palmerston North, and remains the only such service to date. It advertises various services to assist cancer patients and their families adjust to and better cope with the cancer experience, ranging from help with physical symptoms such as pain and anticipatory nausea, through stress management and self-esteem building, to assisting with constructing meaning, purpose and wairua (spiritual strength) in the experience. Services are offered individually, to families or partners, or in groups with peers.

The usefulness of psycho-oncology services is greatly limited by the populations they reach. Statistics provided by the Palmerston North Psycho-Oncology Service for the three year period from its inception until 23 March 2009 (pers. com. G. Shirley, 24 & 25 March 2009) show the distribution of service use by gender, ethnicity, diagnosis, and treatment stage. The service was used more by women (395 referrals, i.e. 62% of the total) than men (239). It was used disproportionately less by Māori: Pākehā (New Zealanders of European descent) comprised 521 referrals; Māori, 63 (9.9%); Pasifika (New Zealanders of Pacific Island descent), 13; and 'other', 37. Data from the 2006 New Zealand census for the Manawatu-Wanganui region, which roughly equates to the area MidCentral Health serves, shows that Māori made up 42,288 of the 222,423 total population in the area, that is, 19% (New Zealand Department of Statistics, 2009). It was used disproportionately more by breast patients and disproportionately less by prostate patients: Of the 423 patients who reported cancer site, 109 were breast; 65 colorectal plus 3 stomach, 2 bladder and 1 appendix; 47 reported lymphoma or lymph glands; 28 lung and 3 bronchus; 15 head, neck or eye plus 12 oesophagus and 3 larynx; 25 gynaecological; 18 melanoma; 18 brain; 13 haematological; 11 prostate; 7 testes; and 41 with multiple or other sites. These site figures can be compared with New Zealand statistics for cancers causing death, to give some sense of proportionality with the distribution of serious cancer in the population. Statistics from 2005 (New Zealand Health Information Service, 2009) show that cancers of the lung, bronchus and trachea were the biggest killers at 18.2% of deaths, followed by colorectal at 15.3%, then breast at 8.2%, and prostate at 7.1%. Finally, patients in the medical treatment stage were overly represented compared with those at other stages: At diagnosis (24), in treatment

(309), post-treatment (40), recurrence (11), during palliative care (26)(43 not recorded). Family / whānau members rather than patients made up another 181 referrals.

The overrepresentation of women, Pākeha, and breast cancer patients in the Palmerston North Service figures will be no surprise to anyone who reads psycho-oncology literature. The ethnic imbalance occurs despite the deliberate employment of a Māori clinician by the service. Given that the service accepts patients by referral – including self-referral - the causes for these imbalances will be complex and include structures affecting the New Zealand health system and society in general. Another factor may be that patients tend to be more willing to receive information and support from medical rather than psychological professionals (Coyne, Stephen, & Palmer, 2006) perhaps because of social stigma associated with mental health. The challenge of righting such disproportionalities is faced all over the developed world. The high proportion of those patients in treatment as opposed to other stages of the experience may reflect not only relative levels of distress but also avenues of access to the service, and gives further cause for reflection, particularly since the period soon after the end of treatment as the patient tries to re-engage with normal life has been noted as a particularly vulnerable time (Freidenbergs et al., 1981-1982). Are services reaching those who most need them or to whom they can be most beneficial? Or are other factors driving the distribution of services?

Overseas literature shows that there is a mismatch between psychological morbidity amongst cancer patients and their actual receipt of services. In a study of cancer outpatients at four hospitals in more affluent parts of Sydney, Australia (Pascoe, Edelman, & Kidman, 2000), it was found that 75% of those experiencing clinically significant levels of anxiety or depression had not received any psychological treatment. A recent American study of low income female cancer patients (Ell et al., 2005) found that only 12% of the women with major depression and 16% of those with dysthymia reported currently receiving medication for it, and far smaller percentages reported seeing a mental health counsellor or attending a support group. An analysis of the very large representative American National Health Interview Survey conducted around 1999 (Hewitt & Rowland, 2002) found that among cancer survivors, mental health service use was greater by younger patients, females, and whites. Unmet need for services existed particularly among the young and those without health insurance, and there was still unmet need amongst females. Cost was found to be a barrier, and referral suggested as a key determinant of service use. It was also suggested that social stigma was an issue for many, as was access for low income and rural patients.

Information as to the availability of services and awareness of the need for them may be other important barriers to access. In a large recent German study of breast cancer survivors (n = 1083, Mehnert & Kock, 2008) it was found that less than half of the sample felt adequately informed about psychosocial support services. Lower education, depression, and older age were associated with feeling ill-informed. Those who had used services were younger, more frequently single or divorced, and suffering progressive disease. Most interestingly, only 23% of those who were classified in the

study as suffering anxiety, depression or PTSD, expressed the subjective need for support. Obviously, if patients suffering clinical levels of distress are not themselves aware of their need, any system that relies upon them raising the issue in order to obtain assistance will fail. If depression itself predicts patient ignorance of services, then for patients suffering that morbidity there is a double barrier to accessing services. The overall rates of needy patients not receiving services suggest a desperate need for better systemic methods of connecting patients with psychosocial services. If the present study can identify populations for which therapy is particularly effective, systemic changes could be made that more effectively connect with those patients.

Screening for psychological distress

Assessment

Psychological distress is traditionally approached through the constructs of anxiety and depression. These constructs have long been part of psychiatric diagnostic categorisation and measurement, and in the present study both are conceptualised as dependent variables. Because of the need for measures administered to seriously ill patients to be brief, however, and because of the differing research goals represented by the wide range of psycho-oncology studies, clinical criteria and interview tools are not always – in fact, not often - used. Most of the commonly used tools nonetheless have recognised psychometric properties. They include the Distress Thermometer (DT), the Profile of Mood States (POMS), the Symptom Checklist (SCL-90-R or BSI), the Psychological Adjustment to Illness Scale (PAIS), and the Medical Outcomes Study Short Form 36 (SF-36), in relation to overall psychological well-being as well as for their individual construct scales. For specific anxiety and depression measurement, they include the Hospital Anxiety and Depression Scale (HADS), the State Trait Anxiety Inventory, (STAI), the Impact of Events Scale (IES), the Beck Depression Inventory (BDI-II), the Centre for Epidemiological Studies Depression Scale (CES-D). Gotay and Stern (1995) review most of these, together with other measures, for their suitability for use with cancer patients.

A validity issue arises in relation to those measures with items that register somatic distress symptoms due to their similarity with some physical symptoms associated with cancer or its treatment. Vegetative symptoms as lethargy, sleep problems, weight loss, weakness, headaches, backaches and gastrointestinal disorders are common to depression or anxiety on the one hand and progressing disease or treatment side-effects on the other (Barsevick, Sweeney, & Haney, 2002; Lovejoy & Matteis, 1997; Siegel, 1990). Somatic items can particularly affect professional-report measures, such as the Hamilton Rating Scale for Depression (HRSD), where the rater is dependent on the physical signs and symptoms that he or she observes. But they also affect self-rated instruments such as the oft used BDI (Love, Grabsch, Clarke, Bloch, & Kissane, 2004). Kissane et al. (2003) mention the need for measures to be valid as regards cancer related grief also, and uses instruments like the Hospital Anxiety and Depression Scale and the Monash Interview for Liaison Psychiatry accordingly. Another way to step around the

problem was demonstrated by Weber et al. (Weber, 2001; Weber et al., 2004; Weber et al., 2007) who used the Geriatric Depression Scale which discriminates out somatic symptoms.

This issue can make a great difference to outcome scores. In a study by Savard et al. (2006) an illustrative range of measures of depression were taken immediately after a CBT intervention: The HRSD, which is comparatively dependent on somatic signs as mentioned above, yielded a very large effect size of Hedges' $g = 1.10$; the BDI, which has some somatic items, yielded a moderately large g of 0.685; but the HADS depression subscale, which is designed especially for use with physically ill populations, yielded a small g of 0.222. The difference is dramatic, with the illness-specific measure picking up a much smaller effect size than the other two, and the measure that is most dependent on outward physical signs generating the greatest effect size. A study that administered both CBT and relaxation therapy illustrates the problem from another slant: Houby (1987) simultaneously administered both the BDI and a form of it adjusted by the removal of somatic items, yielding g 's of 0.861 and 0.469 respectively.

The interference of somatic symptoms in the assessment of cancer-related distress is further complicated by the possibility that particular psycho-oncological therapies have more effect on physical symptoms, such as problem solving CBT or relaxation therapies. Given et al. (2004) found that the impact of their CBT intervention on depression was moderated by the intervention's impact on physical symptoms. The combination of such an intervention with a somatically laden assessment tool could convey quite a misleading impression of changes in the psychological state of patients. When assessing cancer patients for distress, although the whole person should be the focus, the medical context cannot be forgotten.

Screening approaches

It has been known since Derogatis et al. (1983) found that psychological disorder was diagnosable in less than half of their cancer patient sample that clinical distress is not an inherent part of the disease. This finding would have shaken a myth to the contrary at large at the time (Sellick & Crooks, 1999) and which may still justify the lack of distress screening evident in most intervention research studies. Yet even patients with very serious cancers may not be clinically distressed. Goldberg and Wool (1985) found that both the lung cancer patients in their intervention trial, and their spouses, were well adjusted at baseline.

On the other hand some clinicians have expressed the fear that more patients anguish in silence than is recognised (Cunningham, 1988; Greer, 1987; Sellick & Crooks, 1999). Indeed, it has been found that many patients do not spontaneously disclose emotional distress, except perhaps by indirect nonverbal cues, and that many cases do go undetected by medical staff and untreated (Maguire, Hopwood, Tarrier, & Howell, 1985). Australian researchers Newell, Sanson-Fisher, Girgis, and Bonaventura (1998) found that only 17% and 6% respectively of patients classified as clinically anxious or

depressed were identified as such by their oncologists. In another revealing study, an experiment was intended to test a nurse counselling intervention for mastectomy patients. The intervention proved unsuccessful at reducing distress, but the regular monitoring before and after surgery and later at home led to 76% of the intervention group being referred for psychiatric help compared with only 15% in the control group, resulting in a greatly reduced rate of psychiatric morbidity (12% v. 39%)(Maguire, Tait, Brooke, Thomas, & Sellwood, 1980). Such evidence, together with that adduced above concerning service use, suggest that great numbers of patients do indeed suffer in silence.

Worden and Weisman (1984) explain that three approaches have been taken towards distress and the offering of psychological treatments to cancer patients in the past. One has been to offer them to everyone on the assumption that all will need them. However, this assumption has not been sustained. Another has been to 'wait and see' who will develop clinically significant distress and then offer help, with obvious disadvantage. And the third is to try to identify those most vulnerable to distress in advance and offer them preventative therapy. Clearly this approach would serve both economic and psychological ends if it was possible to do it well.

Medical visits provide a regular convenient opportunity to detect distress, but there are recent studies that show that they are not used optimally. A recent review (Kruijver, Garssen, Visser, & Kuiper, 2006) identified that staff were not able to adequately detect distress and suggested that time pressure and a "unidisciplinary work attitude and manner" (p.175) were causes. More integration of psychosocial awareness and information in and through medical staff was indicated, and psychosocial checklists were endorsed as helpful in both detecting and communicating issues. It was suggested that after the completion of primary medical treatment, when the patient was progressing out of initial 'survival mode', would be the best time to screen for distress during a routine visit, and the importance of this time has already been noted in the present paper. A very recent study of the practices and attitudes of cancer professionals in the United Kingdom with regard to distress screening (Mitchell, Karr, Coggan, & Herdman, 2008) found that most attempted to detect mood disorder at least 'regularly', but 62% relied upon their clinical skills alone for this, while 30% attempted to use one, two or three brief questions (e.g. from the PHG2) and only 6% used a formal questionnaire. One of these latter two methods were regarded as optimal by professionals, with time and lack of training and confidence cited as primary barriers to screening. From this it can be seen that staff were willing but ill equipped to detect distress.

Clearly time and simplicity will be keys to the use of any screening tool by medical professionals. Efficiencies could be made by the prior identification of vulnerable demographic or medical groups of patients, or by the use of waiting room time or automated touch-screen computer technology presently being developed (e.g. Cull et al., 2001), but the importance of distress screening must be appreciated by staff as a

precondition (Kruijver et al., 2006). In fact, distress screening may itself transpire to be the key to effective – and efficient - psycho-oncological intervention.

Why meta-analysis?

The substantive context of the present study has now been discussed, and a broad range of possible moderators of therapy effectiveness have been identified, the most important of which may be baseline distress. At the outset it was noted that the way to confirm the influence of these is through synthesis of all the relevant available literature. There are various ways to do that, including narrative review, ‘vote count’ review, and meta-analysis. Narrative reviews already abound in this field, yet it is clear from the practice issues referred to in the preceding discussion that they have not been greatly influential in persuading professionals and administrators to adopt practices that are optimally effective at identifying and treating cancer distress. What is needed is ‘hard data’ – an empirical synthesis. There was a large systematic vote-count review published recently that caught attention. In the writer’s opinion, it was fundamentally flawed in its method and therefore drew spurious conclusions – dangerous, when services to so many suffering patients are at stake. However, it is instructive to examine that study because it highlights a number of pitfalls in undertaking empirical reviews of whatever kind, and also demonstrates why meta-analysis is the preferable tool for use in this field.

An instructive synthesis

In 2002 a systematic review was published that concluded that the evidence for the effectiveness of psycho-oncological interventions was so lacking that it could only make “tentative recommendations” in relation to any therapy type (p.580). That review was done by ‘vote-count method’ which means that the number or proportion of statistically significant results are tallied according to some formula. It was carried out by Newell, Simon-Fisher and Savolainen (2002) and published by the (U.S.) National Cancer Institute. Of 36 reviews in this field compared for their quality by Lepore and Coyne (2006), this was the only one that was graded ‘platinum plus’, and the authors duly concluded that the literature did not make out a strong case for the effectiveness of psycho-oncological therapies. However, for some who knew firsthand the effect of psycho-oncological therapies and the need for them, the Newell study was a winding blow. They could not agree with the findings, and questioned the method (Bredart, Cayrou, & Dolbeault, 2002).

Its method and findings

The Newell review gathered outcome results from 155 randomised controlled trials of psycho-oncological interventions dated prior to December 1998. No dissertation database was searched, although unpublished papers that became known to the authors through their contacts with researchers in the field were admissible. Outcome constructs covered the full range of intervention endeavour, namely, psychosocial outcomes, side-effects from medical treatment, immune response, and survival. Under the psychosocial head, the authors collected outcome findings taken with measures of

anxiety, depression, general affect, hostility, and stress / distress, together with a number of functional indicators that are not of interest to the present study.

The trial reports were graded against a researcher-derived measure of study design quality containing ten internal validity items drawn from Cochrane Collaboration recommendations, prerequisite to their inclusion in the final synthesis. Of the 129 trials with psychosocial outcomes, the authors excluded 87 (67%) on the grounds of a poor methodological quality score, and a further eight because of inadequacies with data. This left 34 trials presenting results on psychosocial constructs; just 25 on the emotional outcome variables noted above.

In relation to each trial and intervention strategy, the authors recorded simply whether outcomes were statistically significant or not. Where more than one measure had been administered in relation to the same outcome variable, they recorded a positive ‘vote’ only when more than half of the results were statistically significant. When all the votes for each construct measured were tallied against each particular intervention strategy – that is, in each cell of the summary table - different levels of recommendation were allocated according to these criteria: First, whether 75% of the outcome votes in a given cell were favourable (leading to a ‘tentative recommendation’), and second, whether, amongst those that crossed the 75% threshold, there were any cells that had at least three positive votes including at least one drawn from a trial which had obtained a superior method quality score (‘strong recommendation’). An intervention was said to show equivocal results worthy of further research when it produced some positive votes, but an insufficient proportion to cross the 75% threshold.

Using this scheme, the Newell review made tentative recommendations for interventions impacting psychosocial outcomes in relation only to music therapy for anxiety, unstructured counselling and music therapy for general affect, and non-therapist delivered interventions involving structured counselling for stress / distress. Cautions were raised, however, regarding the low numbers of studies contributing to these recommendations – which were either one or two in each case. No recommendation could be made relating to any hostility outcome at any time period, and the majority of treatment components could not be recommended for intervention against any emotional construct or for further research. A greater range of treatment components were said to have produced outcomes that were equivocal and should be further researched, but none in relation to hostility. No strong recommendations were made in relation to any treatment component or outcome construct at any time point.

Method critique

As mentioned, this dismal picture of psycho-oncological interventions was latched onto by some (e.g. Coyne et al., 2006; Lepore & Coyne, 2006) who have commended the Newell team for an extensive search and the method quality demands they made of primary studies. However, the result that was arrived at, i.e. that no therapy could be more than tentatively recommended, *and that a single music therapy study gave rise to this highest recommendation*, should have sounded the alarm. In fact, as shall be seen

by comparison with the yield from the search undertaken for the present study, the Newell search was not very productive, and the lack of any search of databases for unpublished literature was a flaw. Regarding method quality criteria, the approach chosen was, firstly, to assume that all studies that were not randomised were of no value, and second, to devise a set of further criteria from a bio-medical model, without validity or reliability proven for the practical and ethical context of psycho-oncology. The most obvious problem was the requirement that treatment-providers and patients be blinded (Bredart et al., 2002). Method quality rating is an exercise in psychometrics, and as such, this approach lacked construct validity. As the authors themselves concede, their system also confused the reporting of the required quality features with their existence. The fact that their criteria were inappropriate was surely flagged when only one study achieved a 'good' rating according to the criteria applied. Instead of taking the hint, however, the authors complained of the poor method quality in the field and excluded 87 randomised trials from contributing to the analyses, raising a large question-mark over the validity of conclusions derived.

There were other important flaws in the design of the Newell synthesis. The authors did not justify their treatment delivery and content component categorisation either empirically or theoretically. The result is an arbitrary list that includes some categorisations that are very broad (e.g. 'therapist delivered') or amorphous (e.g. 'improving self-esteem / self-image'), others that are narrow and specific (e.g. audiotape delivered, hypnosis, music therapy), and others that seem to overlap (e.g. audio tape delivered, relaxation training, guided imagery / visualisation, hypnosis, music therapy). The authors refer repeatedly to the challenge presented by the wide variety of components covered by research trials in this field, but failed to address the issue in a systematic way.

Of further concern is the mixing of study purposes in the collation of outcome data. For example, a general affect outcome from a trial with the primary purpose of treating nausea and vomiting anticipatory to chemotherapy would have been collated in the same cell as the result from a study where affect was the primary target of the therapy simply because the two studies had an outcome construct in common, that is, without regard to the fact that the primary purpose of one intervention was to treat physical symptoms and the other to treat psychological symptoms. This inappropriate synthesis must have compromised the quality of data combined in a given result cell, adding numbers to the denominator but making it more difficult for the proportion of significant results to exceed the 75% threshold required for a given treatment strategy to gain a tentative recommendation of effectiveness. It also strips portions of the results tables that relate to more widely used outcome constructs (in particular, most emotional outcomes) of their interpretability and utility since the purpose of the interventions contributing data to them cannot be discerned.

The proportional approach that the Newell team used for analysing cell results (treatment component x outcome construct) favoured cells with few hits. It has been noted that the only (tentative) recommendations reached in relation to emotional

constructs had no more than two entries in each cell. This was the case notwithstanding that 41 of the entries had denominators in double figures. Indeed, from the whole overall results table (224 cells containing entries, Table 4, p.570), only 15 tentative recommendations could be made, and only one of these was drawn from a cell that had more than two hits. This meant a favouring of narrow treatment categorisations - like music - and, contrary to usual findings, the favouring of late follow up times, for which there was also little data. Hence, it appears that the proportional threshold rule led to recommendations that were negatively associated with the size of the denominator in the relevant table cell – the less researched a particular therapy is, the better the chances that results will cross the threshold. This produces conclusions based on a scheme completely contrary to what most clinical psychologists consider to be an ‘evidence based approach’ (e.g. Andrykowski & Manne, 2006; Institute of Medicine of the National Academies, 2008). All in all the authors have failed to realise how their weightings and categorisations have failed to fit the data and have distorted their conclusions.

There are other data selection and conceptual criticisms that could be levelled at this review, but enough has been said along those lines. The final and most important point concerns the choice of the vote-count tool for the purpose of drawing the conclusions that Newell and associates came to. ‘Vote-count’ relies upon the finding of statistical significance, itself a function of sample size. The authors acknowledge that type II errors are more likely in samples of less than 50, and that this imposes a limitation on their study. As shall be seen, for all but breast cancer and mixed cancer site samples, less than 50 is a practical reality for much psycho-oncological research (Bredart et al., 2002; Ross et al., 2002). The tool used to synthesise such data must be sensitive enough to combine them in a way that will extract what they may teach us. The vote-count review tends to register small sample outcomes as negative, simply because small samples struggle to reach statistical significance. When combined with the proportionality system that Newell used, the result is particularly misleading. Another layer of bias is added when sample size is itself systematically related to moderators of therapy effectiveness, as shall be seen in the present study. Meta-analysis is a finer tool, better suited to combining the smaller *n* research that characterises much of this field, and by-passing such bias. All in all, Newell et al. (2002) contains many lessons for data synthesis.

The present meta-analysis in context

The present meta-analysis is part of an iterative and ongoing process between primary trials and research syntheses that seek to develop the field from establishing general effects of intervention towards refinement of effectiveness through the study of possible moderators and mediators. It is appropriate to briefly survey features of previous meta-analyses in this field in order to place the present study the context of this development. This final section begins with a list of intervention meta-analyses, and then surveys those that are relevant to the domain of interest with the aid of a main effects summary

table. The chapter closes with a statement of the need for the present meta-analysis and of its hypotheses.

Previous meta-analyses

Although there are now hundreds of psycho-oncological intervention reports in the literature, and a great many reviews, there are only a few meta-analyses. In English, there are three that address psychosocial interventions generally, in relation to adult patients: Cwikel et al. (2000); Devine and Westlake (1995) and Meyer and Mark (1995). A number of others focus on specific sub-populations of, or treatments for, the adult population: Tatrow and Montgomery (2006), cognitive-behavioural treatments for breast cancer patients; Edwards, Hailey, and Maxwell (2004), women with metastatic breast cancer; Zabalegui et al. (2005), nursing and cancer support groups; Luebbert et al. (2001), relaxation treatments for acute non-surgical patients; Osborn et al. (2006), treatments for ‘survivors’. Some important meta-analyses focus on particular outcome spheres: Devine (2003), pain; Sheard and Maguire (1999), anxiety and depression; Rehse and Pukrop (2003), quality of life; and both Chow, May, and Harth (2004) and Smedslund and Ringdal (2004) investigate the impacts of psychosocial treatments on patient survival. Graves (2003) offers the only theory based analysis (social-cognitive theory), and there are two studies specific to children, the first-mentioned general, and the second specific to pain interventions to assist with invasive procedures: Pai, Drotar, Zebracki, Moore, and Youngstrom (2006) and Scholz (1999). The present meta-analysis is concerned with psychosocial interventions for distress in adult patients only. Studies relating to children are regarded as distinct because children respond differently to the cancer experience.

General psychosocial intervention meta-analyses

Features of the meta-analyses with domains and outcome constructs similar to the present study are now discussed. Details of their design and main effect findings are summarised in Table 1-2, below.

Table 1-2. Previous meta-analyses, summary of main effects

Study, design types sampled; and trial quality assessment/control method	Anxiety	Depression	Distress
Meyer & Mark (1995) published randomised trials only; no other control of quality	‘emotional adjustment’, unbiased $d = 0.31$ (45)		
Devine & Westlake (1995) nonrandomised controlled trials and pre-post designs included; sensitivity analysis for internal validity	unbiased $d = 0.56$ (68)	unbiased $d = 0.54$ (48)	‘mood’ unbiased $d = 0.45$ (30)
Cwikel et al. (2000) published controlled† studies only; no other control of quality	‘psychological outcome’ $g = 0.62$ (40)		
Rehse & Pukrop (2003) published controlled studies only; graded for internal validity	‘emotional adjustment’ $d = 0.65$ (37)		

Sheard & Maguire (1999) published and unpublished controlled studies, graded for internal validity	Screened and unscreened	$g = 0.36$ (18)	$g = \text{approx}$ 0.20 (20)	-
	Screened*	$g = 0.85$ (3)	$g = 0.94$ (4)	-
	Unscreened*	$g = 0.33$ (5)	$g = 0.16$ (5)	-

Figures are effect sizes with number of studies contributing in brackets.

Screening was defined by Sheard & Maguire to include patients at risk of or suffering significant distress, and is hence broader than the definition of ‘screening in’ in the present study, which did not include mere risk, and entirely distinct from the present ‘screened out’ category, which, together, are in the present study referred to as ‘screened’ studies.

* Authors report power insufficient for these results to reach $p < 0.05$ significance but came close.

†The authors do not specify controlled studies, but this seems implicit from the context.

Meyer and Mark (1995) is the most often referred to meta-analysis in this field. It allayed the fears of some early reviewers that too many interventions were not achieving statistically significant results with the conclusion that psychosocial interventions were effective, although the effect size found was small. The authors tested for the impact of screening-in only distressed patients, and for whether there was an interaction around Andersen’s low, medium and high patient psychological distress risk levels (Andersen, 1992). They found no statistically significant difference, but a tendency for effect size to increase with distress risk level. There was also no difference in the effects produced by different therapy types (refer earlier table, Table 1-1). It is noted that it appears that the authors (refer html p. 6 of 17) used an effect size from each study averaged over all assessment time points reported, which will have dragged their findings down and blurred the meaning of their result, already blurred by the breadth of the outcome construct ‘emotional adjustment’.

Meyer and Mark ruminated over the clinical significance of the modest effect sizes they had obtained, and drew a comparison with a meta-analysis by Matt (1989) which replicated Smith and Glass’s famous meta-analysis of psychotherapy outcome studies (1977) and achieved an effect size of $d = 0.35$. They noted that even small effects can make big differences in terms of the percentage of successful outcomes, and they wondered whether a ceiling effect produced by substantial proportions of well adjusted participants had robbed the results of their potential variance. This perceptive observation was taken up later by Sheard and Maguire (1999, below) who also commented on the lack of focus in Meyer and Mark brought about by the breadth of their outcome constructs and their choice to average together all measures of each construct regardless of psychometric quality. Sheard and Maguire also made note of their lack of unpublished studies.

Meyer and Mark thus made an important beginning to meta-analysis in this field. They were able to establish that psycho-oncology ‘works’, but – more importantly for present purposes - came up with some perceptive insights: Baseline distress may moderate therapy effectiveness; Anderson’s risk factors – essentially to do with the seriousness of the disease – may too; and particular therapy types may not.

Devine and Westlake (1995) examined the impacts of psychosocial (‘psychoeducational’) treatments for anxiety, depression, mood, knowledge relating to cancer, and a number of physical outcomes. They included a broader range of trial designs and used preliminary sensitivity analyses to check whether differences in internal validity grades that they assigned to studies based on the strength of their designs made any difference to effect sizes – and found they did not. They confirmed Meyer and Mark’s finding that psycho-oncological interventions ‘work’, finding that around 90% of studies produced positive effects (although not all reached statistical significance) and by finding considerably higher moderate magnitude overall effect sizes. Like Meyer and Mark they also found that effect sizes for different therapy types did not differ significantly from each other (as mentioned earlier – refer Table 1-1).

The primary purpose of the Cwikel et al. (2000) study was to compare vote-count and meta-analysis methodologies for use by the social work discipline, but it is of interest for several findings: It confirmed the moderate magnitude main effect sizes found by Divine and Westlake; found that medical treatment phase was probably the most effective time to administer interventions ($g = 0.51$) relative to new diagnosis phase (0.35), and late/terminal phase (0.59, n.s.); and found that cognitive-behavioural treatments outstripped the effectiveness of the grouping of all other therapy types (refer Table 1-1). They cited Hunter and Schmidt (1985) to note that effect sizes in the 0.20 - 0.40 range – smaller than the 0.62 they found - are typical of psychological interventions generally.

It might be suggested that the Cwikel and Divine meta-analyses arrived at stronger main effects because they incorporated trials of inferior design. However, the purpose of the sensitivity analyses in Devine was to establish empirically whether ‘inferior’ studies were in fact so. As the analysis showed no significant difference, the implication is, rather, that Meyer and Mark excluded research without empirical justification, and that their effect sizes are therefore artificially low. Their lack of unpublished studies could have also contributed. The present study takes a lead from the Devine approach to test the dataset against various validity criteria in order to understand its dynamics and ensure that the most data empirically justifiable is used for substantive analyses.

Rehse and Pukrop (2003) undertook a meta-analysis of studies in German and English language. A median split performed around internal design quality rankings showed that the better quality studies were more likely to detect effects than those with poorer methodology, but not by much ($d = 0.70$ compared with 0.59). Their substantive analyses again found a moderate strength main effect, but also a significantly higher effect for educational therapies over other types (refer Table 1-1) and an effect twice the

size for men (1.01) than women (0.52). A moderator that stood out as the only one to remain statistically significant when the impact of all other moderators was controlled was intervention duration of at least 12 weeks.

Finally, Sheard and Maguire (1999) used an internal validity rank to divide their controlled study sample into those regarded more and less reliable, and examined only the two clinical constructs, anxiety and depression, removing outliers (more or less than two standard deviations). They found that removing less reliable studies made the effect size for anxiety lift considerably, and for depression, drop slightly. The main effects tabled (Table 1-2) use all studies, but exclude outliers. The authors noted that the effect size of 0.36 for anxiety represented approximately the difference to be expected between a normal sample and a sample of general medical and surgical inpatients, according to the norming information for one of the anxiety measures commonly used in the studies. For depression the effect size was of negligible clinical significance.

Sheard and Maguire's most important finding – and for present purposes the most important of any of these meta-analyses – was of a disparity in effect sizes produced by studies that screened potential patient recruits on the basis of clinically significant psychological distress (or high risk), compared with those that recruited 'preventatively' i.e. on the basis of mere cancer diagnosis, self-referral, or being thought suitable by an oncologist. The great majority of studies took the latter course, and, the authors said, treated patients who may neither have needed nor benefited from intervention. The difference in results was dramatic, producing strong effect sizes from the studies that screened (refer Table 1-2), and the authors noted that this heterogeneity came close to statistical significance, missing only because of low power. Other differences they found were that therapies of longer duration were generally more effective, as were more experienced therapists, and group education outstripped other broad categories of intervention (refer Table 1-2), but note that most other categories were not listed there because they were defined by delivery mode rather than therapy content).

In summary, these meta-analyses were mainly about establishing main effects in order to establish the worth of psycho-oncological intervention. However, they did begin to reach for moderators. Indications were found that anxiety may be more amenable to therapy than depression (Sheard & Maguire, 1999); that distressed patients probably produce significantly larger effects (Meyer & Mark, 1995; Sheard & Maguire, 1999); that men receive more benefit than women (Rehse & Pukrop, 2003); that longer therapies are more effective (Rehse & Pukrop, 2003; Sheard & Maguire, 1999); more experienced therapists are more effective (Sheard & Maguire, 1999); and during active treatment may be the best time for therapy (Cwikel et al., 2000). Conflicting results emerged regarding the most effective types of therapy (in two, education proved best: Rehse & Pukrop, 2003; Sheard & Maguire, 1999), and one, CBT (Cwikel et al., 2000)), and it may be that this is not really a moderator (Devine & Westlake, 1995).

Other meta-analyses have addressed particular aspects of the domain of present interest. Some of these attempts suffered from the small number of available studies of the requisite quality, and some produced results of such a size as to raise doubts as to their validity. However, they raise the suggestions that metastatic breast cancer patients may be difficult to treat effectively (Edwards et al., 2004), support groups can be effective (Zabalegui et al., 2005), relaxation therapies are useful for patients undergoing chemotherapy (Luebbert et al., 2001), and that CBT can be very effective for long term survivors (Osborn et al., 2006).

The need for moderator meta-analysis

All in all, the field is still left uncertain as to the effectiveness of psycho-oncology due to considerable disparities in main effects (Coyne et al., 2006; Lepore & Coyne, 2006). This variety of findings itself strongly suggests the influence of moderators that are structuring the whole of the data of this field. The most important moderator that has been identified – important in terms of its practice implications and its ability to explain other moderating effects – could well be whether or not patients are screened for distress at entry to intervention. That is a finding of a ceiling or floor effect, which is a matter of trial design. What is needed now is analysis that scopes a comprehensive range of trial design features in order to confirm that finding and check for other trial design moderators that structure the dataset. Such an investigation ought to include analysis of whether RCTs produce effect sizes of significantly different magnitude to other controlled studies, so that it is known whether these other studies can safely be included in substantive analyses.

These preliminary investigations will provide a sure foundation from which to address Paul's 'ultimate question' (Paul, 1969) – the 'who, what and which' of substantive moderators of therapy effectiveness. The need for such investigation has been reiterated over the years by many 'big names' (e.g. Cunningham, 1988; S. L. Manne et al., 2007; Redd, 1995; Ross, Boesen, Dalton & Johansen, 2002; Silberfarb, 1982). Clinicians need moderator research to inform the design of more effective therapy protocols, and theorists need it for proving pointers as to underlying mechanisms of effectiveness and to promote new lines of thought and investigation. Administrators need it to inform service delivery decisions. Enough ground work has been done by previous meta-analyses to allow a comprehensive effort to be undertaken now, confirming or negating earlier findings and reaching to analyse untouched constructs.

Meta-analysis is a powerful and flexible tool, combining small n results that otherwise lack statistical significance, testing replication, and putting a figure on the size of the effect, as Kline (2004) advocates science should attempt to do. Although sometimes criticised for lumping together too many aspects of interventions - causing detail to drop from visibility – such a challenge of categorisation and construct validity faces most scientific endeavour. The gathering of multiple operationalisations of a particular construct can also be seen as a major strength of meta-analysis, increasing construct validity (Hall, Tickle-Degnen, Rosenthal, & Mosteller, 1994). By combining different

groups of data around different variables, meta-analysis also permits the exploration of theoretical relations not explored before, and though results may not be definitive because of limits to causal inference, they may be quite “provocative” and useful in guiding the direction of future research (Hall et al., 1994, p.21).

Hypotheses

The aim of the present study is to produce an understanding of the structural dynamics underlying the data in this domain, and on the basis of that understanding, to produce information about factors that moderate intervention effectiveness that will be useful to clinicians, researchers, and administrators. The only previous finding that is considered to be fairly firm is that screening for patient distress at baseline causes higher effect size. Confirming this finding is a major aim of this study. It is expected that such confirmation will have implications for the finding of other moderators, that is, that subpopulations who are more distressed for one reason or another – socio-demographic or medical – will tend to produce higher effect sizes. As to therapy characteristics, it is expected that more therapy that is more intensive and delivered by more experienced professionals will produce better results. Beyond these general directions, the investigation is considered exploratory in nature.

2. METHOD

Meta-analysis is simply an analysis of analyses. It is an experiment in which primary studies are the subjects and there are independent and dependent variables, the same as in any other. Sampling is critically important, as are the various aspects of validity that sustain conclusions drawn from the results. This chapter steps through the process of ‘setting up the experiment’, leaving detail relating to analysis for the next chapter.

The meta-analysis method requires the domain of the study to be established precisely as it is the basis for sampling primary studies. Then as much primary data as possible are collected – ideally, though this is seldom possible, all of it – and it is coded for features relevant to the research questions. Finally, the data are synthesised by way of analysis. At some point in the process – either at domain definition or before the analyses are conducted, or at both stages – the quality of the design of primary studies / differences in their features has to be addressed. In the case of the present research, that issue was of particular importance because the inconsistencies in previous meta-analytic results suggested that unknown design features were influencing results. The issue therefore demanded a series of sensitivity analyses that were as thorough as the analyses performed on substantive moderators.

The chapter begins with a statement of principles used to guide the research, progresses through the steps of domain definition, primary study collection, and coding, and concludes with an explanation of the approach taken to deal with primary study design.

Guiding principles

Before any planning relating to the detailed parameters or method of this study began, two principles of approach were agreed between the writer and two clinician researchers: first, the research would be oriented toward clinical usefulness, and, second, the most use possible would be made of extant data, subject to reasonable practical scientific quality assurance. In formulating these principles, the practical and ethical realities of the clinical context from which the data derived were acknowledged.

Domain criteria

The first task in designing a meta-analysis is to tightly define its domain (Lipsey & Wilson, 2001). This involves making both sweeping conceptual and detailed practical choices. Decisions are guided by research questions as well as principles of research design quality and practical constraints associated with the meta-analysis method. In the present case, the domain was defined according to types of study, participant, intervention, and outcome.

Types of study

Basic design

It was decided that only intervention trials involving a control group – whether naturalistic or experimental – would be collected. Uncontrolled pre/post analyses are heavily exposed to internal validity threats, and in the present context a dominant concern would be confounding due to maturation over the trajectory of the cancer experience (refer Introduction chapter). The contextual reality that ethical boards do not always allow randomisation with this doubly clinical population was taken into account in making the choice to include quasi-experimental (natural) and ‘pseudo’-random designs (defined as those with a strategy to defend against selection effects but which were not randomised by individual participant, e.g. those that randomised by week of entry into the study) as well as randomised clinical trials (RCTs). It was decided to admit this broad range of controlled studies to begin with, and then to test empirically for any impact of randomisation on effect sizes by way of sensitivity analysis. If a distinction was not clear, there would be neither need nor justification to exclude data from the quasi-experimental and pseudo-random designs. Nonetheless, naturalistic studies using self-selected groups (or ‘refuser controls’) were not included on account of the known selection effect associated with motivation.

Only prospective designs were admitted – retrospective designs were considered too unreliable since they rely upon memory for the measurement of the dependent variable, and correlational designs cannot not speak to causation.

Nature of control

Effect size is an index of contrast – in the present case, between intervention and control group scores assessed after therapy. In order to heighten variance, and so that the meaning of contrasts was clear, it was decided to admit only studies with controls that were essentially untreated. Therefore studies that used ‘treatment as usual’ or wait-list / no treatment groups were admitted, and also studies that used attention or treatment element placebos. Studies that compared only with groups that received alternative treatments or ‘straw’ treatments were excluded. The likelihood that studies that used placebos or treatment as usual controls would produce less contrast, i.e. smaller effect sizes, was taken into account by coding them differently in preparation for sensitivity analysis. Where there was doubt as to whether a comparison group was receiving an alternative treatment or a treatment element placebo, the researcher’s intention as to the nature of the comparison group guided the decision. However, this intention was discerned not only from the terminology used by the researcher, but from the substantive nature of the comparison being drawn. For example, in the Cunningham and Tocco study (Cunningham, 1987; Cunningham & Tocco, 1989) the ‘control’ group received 6 weekly 2 hour sessions of nurse led problem solving, information sharing and ventilation of feelings. This substantial intervention was regarded as an alternative

treatment for the purposes of the present meta-analysis and therefore the study had to be excluded for lack of control data.

Control group participants had to be drawn from a comparable cancer patient population. Studies using a control group that was not fairly comparable for some reason - healthy controls, 'refuser' controls as mentioned, controls from a different treatment setting (e.g. outpatients where the intervention group were inpatients) or having a different level of distress (e.g. Worden & Weisman, 1984) were excluded.

Study size

Studies were also required to include at least five participants in each of the treatment and control conditions. Although bias resulting from study size was taken into account somewhat by use of Hedges *g* as the effect size statistic (refer Analysis chapter), and it is technically possible to derive a standard deviation from a study of as few as three participants, this rule was decided upon in order for data spread to have some substantive meaning (c.f. the meta-analysis by Devine, 2003).

Types of participant

Qualifying participants were adult cancer patients only, that is, 17 years of age or older (nearly all studies had an entry age of at least 18 years). Children's appreciation of and psychological reactions and needs relating to cancer as a stressor differ from adults and it is therefore conventional to treat and study them separately.

The term 'cancer patient' was defined to include those who had been diagnosed with cancer, including survivors of more than five years, and patients who were being advised of their diagnosis at the time of the intervention. Studies of interventions administered during pre-diagnostic procedures (e.g. biopsies) or designed for 'at risk' populations (e.g. the daughters of breast cancer patients) were not included. Studies that included patients of other diagnoses (e.g. 'surgical patients') were included only if sub-group data relating exclusively to cancer patients could be extracted. Some studies collected psychological data from patients' significant others or third parties who were themselves the direct recipients of the intervention (e.g. medical staff or caregivers) but, again, only outcome results, sample sizes and demographic data relating to the cancer patients could be included in this research.

Types of intervention

'Psychological' intervention

Studies that delivered psychological, psychosocial, or psycho-educational treatments were included. Whilst the provision of information was counted in, the line was drawn to exclude art, drama, massage, reflexology, yoga, aromatherapy, pet therapy, prayer or treatments that were focused mainly on any of these. Likewise therapies with the main focus on training or education in exercise or nutrition were excluded. This decision was made not because such interventions were considered of less value but because they are

outside the conventionally accepted domain of ‘psychological’ treatments. Mindfulness-based therapies were counted in, although not when expressed through art. Studies that included a psycho-therapeutic component but which focused essentially on a type of intervention that was excluded were excluded. A study by Coyne et al. (2003), where the intervention was merely the provision of a more easily understood informed consent form, was excluded, but studies of interventions that oriented patients to cancer treatment facilities or provided information relating to the disease or its treatment were retained.

Therapies that amounted essentially to the provision of nursing or medical care were also excluded, as were studies where treatment included a pharmacological component, an implantable device or the provision of increased medical attention, unless that part of the treatment was made available in an equivalent way to the control group. This meant that ‘comprehensive’ psychological treatments, such as that assessed by Steel, Nadeau, Olek, and Carr (2007) which included educational, cognitive-behavioural, supportive-expressive, and pharmacological components, could not be included. Nurse advocacy to other medical care providers was considered a nursing intervention rather than a psychological intervention, and therefore excluded (e.g. Ambler et al., 1999). Studies of ‘collaborative care’ between disciplines were also excluded, although a distinction was made between those and studies that delivered a qualifying intervention through a multidisciplinary team. Again, these exclusions were made not because the interventions were thought to lack value, but because they did not isolate psychological interventions as such.

Interventions that amounted to mere psychological monitoring or screening by or for medical staff were excluded, again, because they are not therapy as such. However these studies are of particular value for the targeted application of psychological therapies, so for the reader’s interest it is noted that three studies excluded on this ground were Boyes, Newell, Girgis, McElduff, and Sanson-Fisher (2006), Maguire, Tait, Brooke, Thomas, and Sellwood (1980), and McLachlan et al. (2001).

Non-physical focus

Interventions that focused primarily on the physical effects of cancer (e.g. pain, sleep disturbance, fatigue) or its treatment (e.g. the nausea and vomiting commonly associated with chemotherapy, or physical limitations following surgical resection affecting sexual organs) were also excluded. Such a distinction was applied by Sheard and Maguire (1999) in their meta-analysis, but not by Newell, Simon-Fisher, and Savolainen (2002) in their systematic review, in the latter case, confusing the interpretation of findings. There is a considerable body of literature in this area, particularly relating to the aversive effects of chemotherapy, and it was set aside as a fairly discrete subset in order to make the study manageable given available resources. It is acknowledged that the judgment of where the primary focus of a study lies involves an element of arbitrary judgment, so lists of those studies set aside on this ground are appended for examination (**Appendix A, list 1**). Readers who are interested in this field

are referred to a meta-analysis by Devine (2003) and to reviews by Carey and Burish (1999) and Redd, Montgomery, and DuHamel (2001). Studies with length of survival as their primary outcome were retained to the extent that they theorised distress as a necessary moderator or mediator and provided useable data.

Professionalism of therapist

Studies of interventions administered by non professionals (e.g. peer counsellors) were included.

Types of outcome

Outcome constructs

The dependent variables targeted by this study are the classic clinical psychological constructs anxiety and depression, together with the more loosely defined construct 'psychological distress'. 'Adjustment' was considered, but was dismissed as too vague, with measures of it often tapping functional rather than emotional indicators. Hostility and anger were excluded because it is not clear whether a lift or a decline in scores would signify an improvement of the patient's psychological condition. Cancer patients are sometimes characterised as 'too nice' and needing to express anger (Thomas et al., 2000).

The admission of 'distress' as an outcome was a decision made to accord with the principle of relevance to the clinical field, since a great variety of measures of emotional well-being are used. Distress is a construct that overlaps with anxiety and depression, e.g. Carlson and Bultz (2004, p.838) say "emotional distress refers to problems such as anxiety, depression, and fears around the cancer experience". Although the construct is relatively loose and includes a substantial overlap with the more precise constructs, these factors could be taken into account in interpretation of results, and its use was advantageous in introducing more n and variance to the dataset.

In discussion with two clinician researchers, and with an eye to the NCCM Clinical Practice Guideline in Oncology (National Comprehensive Cancer Network, 2005, p. DIS-2) it was decided that the construct would be defined by the following features:

- An emotional experience
- May be part of or in itself a journey
- Is unpleasant
- May be evidenced by negative affect or poor emotional functioning
- May be marked by elevated arousal – at the course beginning anyway – and may lead to burnout and loss of vitality
- May be seen as indicating that a person is in a situation that is taxing beyond their coping resources (Lazarus & Folkman, 1984), or is a challenge to the

person's existence / identity / or some other characteristic of their life that is of great importance to them

- May itself obstruct adaptive coping
- Is multi-factorial in nature, including other more specific constructs such as anxiety and depression
- Its intensity can be graded on a continuum
- May be regarded as the opposite of 'emotional well-being'

Scales and psychometrics.

Flowing from this, it was decided that appropriate scales for measuring distress would include multi-factorial measures of emotional well-being such as the POMS, and of psychological symptoms such as the BSI. These have anxiety and depression subscales, but others as well, and are often used with the cancer population. Measures of affect, such as the Affect Balance Scale (ABS), were also included, and also measures of anxiety and depression where a score for those constructs separately was not made available, such as a total score from the HADS. Acceptable anxiety and depression measures included both commonly used screens (such as the subscales of the HADS, the BDI, or the Tension / Anxiety or Depression / Dejection subscales of the POMS) and diagnostic tools. Illness or cancer specific measures were preferred, to the exclusion of non-specific measures when both were available.

Only instruments that provided interval, rather than categorical, level data could be used, and established reliability was required, i.e. a commonly known published instrument or one that was substantiated by a Cronbach's alpha of at least 0.70 based on data in the study report. The preferable indices drawn from reports were means and standard deviations / errors but estimates could also be derived from *t*-test *p* values (if unspecified, in accordance with common practice it was assumed that a 5% type I error rate and a two tailed test were used).

Theoretical moderators

Studies were also coded for three potential theoretical moderators, self esteem, self-efficacy, coping style. Self esteem was defined to include self concept and was operationalised by such measures as the Rosenberg Global Self Esteem Scale, the Tennessee Self Concept scale (TSC), the Coppersmith Self Esteem Inventory, and the Adult Self Perception Profile. Self efficacy included perception of control and dispositional optimism / sense of coherence, and was often measured using the Mental Adjustment to Cancer Scale (MAC), Helplessness or Fighting Spirit subscales, the Mental Health Locus of Control Scale or the Life Orientation Test (LOT). Coping style was distinguished from specific coping strategy, so that the construct measured was whether the behaviour was adaptive or not. Common measures were the MAC Avoidance subscale and the Dealing with Illness Coping Inventory.

Type of report

Abstract length reports were not adequate because of the present focus on moderator data. Studies for which only such a brief report could be obtained, or which provided an English abstract but were otherwise in a foreign language, could not be used. Unfortunately resources did not permit the translation of foreign language studies. It is acknowledged that this may have introduced a bias to the dataset, since foreign language studies that are reproduced in English journals are often ‘good news stories’. This potential bias was also investigated empirically as part of preliminary analyses.

Search strategy

Having settled the domain, a search for primary studies was undertaken. As the strength and integrity of any meta-analysis depends upon the thoroughness of the search, care was taken in the selection of search tools and keywords, and much effort was put into the search itself. The process is described in this section.

Sources

Both electronic and manual searches were undertaken. Electronic databases were selected from those to which access was available to represent both general and specific content areas, and published and unpublished literature. PsychInfo was selected as the main psychological resource, MedLine as the main medical resource, and CINAHL for nursing studies. The Cochrane Central Register of Controlled Trials (‘CENTRAL’) was also searched. Unpublished theses were searched for using ProQuest (Europe and North America), Theses Canada, Index to Theses (Great Britain and Ireland), and the Australasian Digital Theses Programme (ADTP).

Citations were also derived from the primary study lists of all previous relevant meta-analyses (Cwikel, Behar, & Rabson-Hare, 2000; Devine & Westlake, 1995; Edwards, Hailey, & Maxwell, 2004; Luebbert, Dahme, & Hasenbring, 2001; Meyer & Mark, 1995; Rehse & Pukrop, 2003; Sheard & Maguire, 1999; Tatrow & Montgomery, 2006; Zabalegui, Sanchez, Sanchez, & Juando, 2005) and selected recent reviews (Campbell, Phaneuf, & Deane, 2004; Dunn, Steginga, Rosoman, & Millichap, 2003; Goodwin, 2005; Gottlieb & Wachala, 2007; Institute of Medicine of the National Academies, 2008; Jacobsen, Donovan, Swaine, & Watson, 2006; Kissane et al., 2004; McPherson, Higginson, & Hearn, 2001; Newell et al., 2002; Ross, Boesen, Dalton, & Johansen, 2002; Sellick & Crooks, 1999; Smith, Richardson, Hoffman, & Pilkington, 2005).

A very small number of papers were also picked up incidentally in the course of the research, for example, suggested by an author when a request for additional data was made.

A more extensive search including the specialized cancer database CancerLit and the grey literature database SIGLE would have been preferable, but access to these databases was not possible. An exhaustive search may also have involved contacting leading researchers in the field, advertising with relevant professional associations, and hand-searching particularly relevant journals. These efforts were not undertaken

because the increment in return was thought unlikely to justify the effort and expense. Furthermore, the number of papers already retrieved threatened to stretch resources beyond capacity and constituted the largest known collection of like studies to date.

Search procedure

The electronic search was run in late October 2007 and then repeated in July 2008 to update it to the end of June. (A full list of electronically located citations is available from the writer.) Citations and abstracts were downloaded to bibliographic referencing software (EndNote version 7) and then examined for relevance to the domain. The reference lists of the meta-analyses and selected reviews were also examined. Citations for those papers that were clearly or possibly within the domain were listed for retrieval of the full text. These papers were then screened in detail for usable effect size data, the presence of an acceptable control, prospective design, sufficient reliability of outcome measures, English language text, non-physical focus, reliable measures, and the remaining domain criteria. Unsuitable studies were discarded under these heads.

Where usable effect size data were not reported by an otherwise qualifying study, and it was dated 1990 or more recently, attempts were made to contact the author by email to request the required data, using the address provided in the article or derived from an internet search. One follow-up email was sent about a month later. Thirty-two of 61 such requests were generously responded to. Whilst it is admitted that this procedure may have introduced a small bias in favour of recent studies, it was thought much more important to obtain data where possible.

Keywords

A set of keywords for the electronic searches was derived from examination of those used by earlier reviews and meta-analyses and from consideration of the present domain. It is noted that the list of keywords used in this search was considerably longer than others seen, incorporating not only generic terms relating to intervention type, but also many specific terms. These specific intervention keywords appear to have been successful at locating a much larger range of studies for synthesis than has been collected previously.

The keywords used were those grouped under headings describing their purpose below, with groupings entered in different search boxes (where possible, otherwise grouped by parentheses) linked by the Boolean operator 'AND' so that the resulting hits registered at least one keyword from each group. A wildcard (* or \$) was used to catch varying suffixes.

Keyword groups:

Context:

cancer OR neoplas* OR oncolog*

Independent variable:

psychotherap* OR psycholog* OR psychosoc* OR counsel* OR nondirective OR psychodynamic OR psychoanalytic OR interpersonal OR supportive-expressive OR “social* support*” OR “emotional* support*” OR “support group” OR psychoeducat* OR psycho-educat* OR informat* OR individual OR behavior?ral OR cognitive OR cognitive-behavior?ral OR relax* OR hypno* OR meditat* OR mindfulness-based OR desensiti* OR imagery OR visualization OR coping OR self-help OR electromyography OR bibliotherapy OR existential OR death OR spiritual OR “stress management” OR “problem solving” OR “mutual aid” OR “peer support” OR “support group” OR biofeedback OR “communication training” OR “written emotional disclosure” OR “acceptance and commitment therapy” OR “patient educat*” or “client educat*”

Study type:

treatment* OR intervention* OR therap* OR trial*

Dependent variable:

anxiety OR depression OR distress

Fields searched included title, subject / heading / key-concept words, table of contents and abstract wherever available, but not full text. The date limitation of 1975 (refer Introduction chapter for the significance of this date as the beginning of this field) – June 2008 was applied, but few other limiters as it was unclear what criteria were used by each database in assigning them, and how reliably they were assigned.

Search yields

Detail of the yields of these searches and the disposition of references is given here. This follows, in substance, the recommendation of the QUOROM statement on reporting meta-analyses (Moher et al., 1999). Yields from the initial (October 2007) search were: 588 records from Cochrane CENTRAL; 817 from PsychInfo; 281 from MedLine; 259 from CINAHL; 25 from Index to Theses; 12 from Theses Canada; and 634 from ProQuest. The attempt to search ADTP was obstructed by a failure in the facility but after much effort 3 studies were obtained from it. Excluding the ADTP search, the total number of records examined from these initial database searches was thus 2616. After screening and removal of multiple records of the same study, 296 of these plus two from the ADTP search were retained for full text examination. A further 114 full texts were obtained after examination of the meta-analysis and review reference lists, and another 15 were discovered incidentally or suggested by authors.

The search was updated for the period 2007 to June 2008 on 3 July 2008, yielding 61 records from CENTRAL, 239 from PsychInfo, 64 from CINAHL, 496 from Medline, 76 from Proquest, 4 from Index to Theses, and 0 from Theses Canada totalling 942 (ADTP was still not operational). Some of these records overlapped with those already obtained, but a further 15 qualified for full text examination, bringing the total obtained in hard copy to 442.

The final disposition of reports from the 442 sought for final screening was:

- 20 foreign language, not retrieved (**Appendix B**)
- 10 could not be traced for retrieval (**Appendix C**)
- 20 physical primary focus (**Appendix A, list 2**)
- 3 insufficient sample size (**Appendix D**)
- 51 inadequate outcome data (and authors could not be reached) (**Appendix E**)
- 3 abstract length (**Appendix F**)
- 63 were uncontrolled or used an inadequate control (**Appendix G**)
- 1 used inadequate measures (**Appendix H**)
- 86 otherwise outside the research domain (**Appendix I**)
- 176 used for coding (**Appendix J**).

Multiple or longitudinal reporting meant that the 176 studies used for coding comprised 146 actual studies. Where a thesis was published, the published report of the study was used in order to keep costs down, so long as it provided sufficient outcome data. Note also that some reports fell into more than one discard category but are listed under only one.

Coding

Coding instrument and procedure

An outline of specific research questions (**Appendix K**) was developed after reading most of the meta-analyses and reviews listed above (under Search strategy, Sources), with especial attention to moderation issues raised by Andersen (1992) and Sherman et al. (2004), and to primary study method quality (discussed further below). A coding instrument (**Appendix L**) was drafted to operationalise these questions into a series of multiple choice items to ask of each study, which would provide answers in the form of nominal or interval level data. The parameters of categories were informed by the reviews read, and were refined with the help of two clinician researchers. The instrument was then piloted and adjusted. Coding was then undertaken by the writer. Detail of the content of coding items and the steps taken to ensure reliability in coding is now provided.

Coding items

The exact detail of coding items is seen in Appendix L, but an overview of item content is presented here:

Primary study quality

External validity issues

Publication type; original language; representativeness of sample; and nationality

Internal validity issues

Strategy used to assign participants to conditions; allocation concealment; attrition balance between groups (selection threat); reason for attrition; blinding; screening for baseline distress

Construct validity

Illness-specificity of outcome measure; source of report (e.g. self-report or professional); replicability / standardisation of treatment protocol; therapy fidelity

Statistical validity

Equivalence of groups at baseline

Independent variables i.e. possible moderators of therapy effectiveness

Patient variables

Socio-demographic: Level of education; occupation / household income; marital status

Medical: Cancer and treatment stages and types; cancer type prognosis

Therapy and control group variables

Nature of the control group; general therapy type(s) and specific therapy components; therapy recipient, means and setting of delivery; dose intensity and flexibility; therapist discipline, experience and involvement with the patient; homework.

Dependent variables i.e. data for the calculation of effect size

Main outcomes: anxiety, depression, distress

Theoretical moderators: self esteem, self-efficacy, coping style

Procedure and reliability

The coding process was designed to ensure the best possible validity and reliability within our resources, conscious that it was an exercise in psychometrics that undergirded the credibility of research conclusions. The coding instrument was drafted to contain considerable interpretative detail, reviewed by two clinician researchers, pilot

tested on seven studies, and then adjusted before use. The writer, who had completed two undergraduate and one post-graduate course in psychological research, coded all the studies. Where there was uncertainty as to correct coding, this was discussed with the clinician researchers for consensus. The coder read item detail from the full coding tool, and wrote response codes on a score sheet against item identifiers (**Appendix M**).

Having one coder made inter-rater reliability irrelevant but intra-rater reliability the appropriate test of consistency as drift in coding interpretation was the issue. There is no standard, to date, of what is an acceptable reliability rate in meta-analysis (Cochrane Collaboration, 2006, part 7.8; Lipsey & Wilson, 2001), but various strategies were designed to ensure consistency of coding, including the taking of test-retest consistency rates.

The full texts that were obtained were screened for the second time and, if appropriate, coded, in batches of 50. After each batch, two coded studies from the first batch were selected randomly and recoded, and then disqualified from being selected for recoding again. Among the fourteen studies recoded, test-retest reliability percentages ranged from 77.8 to 95.9, with an average of 87.6. Inconsistently coded items were examined for cause, and where it seemed to be a systematic issue of coding interpretation, explanatory detail was added to the coding tool and all previous studies were recoded on that item to ensure consistency. It was observed that most often the recoding rather than the original coding was of poorer quality since the recoding task did not hold the coder's full attention. The true reliability of the original coding was therefore higher than stated above. A list was kept of adjustments made to the coding tool and coding was again reviewed against it again when the process was almost complete.

It will be noted that ensuring reliability required a good deal of backworking through coded studies. More backworking resulted from modifications made to the coding instrument to more accurately or fully express the content of primary studies. There was thus a strong iterant character to the coding process and to the development of the coding schedule itself in order to accommodate the nature of the literature being coded, as it became apparent (Wortman, 1994). One adjustment to the coding schedule as a result of this was the addition of a coding response to accommodate studies that 'screened out' potential study recruits who had a history of psychological problems. To the writer's knowledge, this categorisation has not been made before, and it proved to generate some intriguing results. The iterant approach taken thus allowed the dataset itself to directly shape analyses.

Note that some coding items sought the 'predominant' characteristic of the study sample (e.g. the predominant education level). If there was an even or bimodal spread of the characteristic levels, then the level that was approximately the average was chosen. Further, some judgment calls based on inference were required. For example, an item regarding the researchers' procedure for concealing allocation to groups, an inference of robustness was taken from a report that some participants who had indicated intention to treat withdrew from the study due to dissatisfaction with their

allocated condition. At times it was difficult to know whether to infer a specific response, or to code ‘unreported / unclear’, bearing in mind that the more substantive coding that can be made, the more data are available and the more useful an analysis can be. This was a particular issue in relation to the blinding of participants and therapists. However, some kind of explicit pointer was sought, rather than making a guess from the ‘feel’ of the study or from common practice.

Study quality

Approaches to date

Various approaches have been taken towards ensuring the quality of the design of primary studies contributing to past meta-analyses in this field. Some researchers have assumed the ‘gold standard’ quality of published randomised trials and excluded all other studies (e.g. the oft referred to study by Meyer and Mark, 1995). Such an approach effectively weights the value of excluded studies at nil and the value of each RCT equally (Rosenthal, 1995). This defaults the task of judging quality in favour of the selection made by publishers – who have other objectives – and places blind and absolute faith in the words ‘randomised’ and ‘controlled’ to protect against all other internal validity threats.

Other meta-analysts in this field have used ‘the Jadad scale’ to grade for quality (Jadad et al. (1996, e.g. Chow, May, & Harth, 2004; Osborn, Démoncada, & Feuerstein, 2006; Zabalegui et al., 2005). This is a three item scale asking whether a primary study is randomised, double blind, and describes drop-outs, awarding a point for each and an extra point for each of the first two items if a description of the method was provided. On its face, this scale confuses the importance of the reporting of design features with the actual presence of those features, when in fact the quality of the reporting can make no difference to the validity of results. Further, it was expressly designed for medical studies where pain is an outcome or analgesia is an intervention, and so includes an emphasis on blinding which is not valid in psycho-oncology. (Chow et al. modified the scale by removing two of the five items in recognition of this, which did not leave much of the scale!)

Other researchers have shown a more thoughtful approach to addressing validity threats, such as Rehse and Pukrop (2003) and Sheard and Maguire (1999) who report scoring method quality against validity criteria suggested by Cook and Campbell (1979). However, these researchers could not provide psychometric validation of their *ad hoc* measure and did not have the publication space to make its details transparent. Any aggregate scoring system will be challenged for validity since it throws up the issue of the relative importance (weighting) of some items over others.

Because of such difficulties as these, the Cochrane Collaboration (Cochrane Collaboration, 2006, para. 6.11) conclude that until more is known about the impact of design features on outcomes and about the measurement of design quality, quality scoring can be misleading and should be avoided. Their advice is that since there is in

fact no ‘gold standard’ study design, no particular scoring system can be validated so it is best to deal with the issue in a way that is simple and transparent.

Sensitivity analyses

In his seminal speech on meta-analysis and reply to critics, Gene Glass (1976, 1978) advocated the treatment of primary study design features as a purely empirical matter. He said that the decision to discard data because the study design was considered weak ought to be based on empirical fact judged in relation to the particular dataset rather than judged *a priori*, and noted that in his experience such justification is likely not to be forthcoming:

“... whether weaknesses of method are a genuine concern in integrating a particular collection of studies is an a posteriori question of fact about which it is useless to generalize or speculate.... The sensible course to follow is to describe – in quantitative terms – features of designs and correlate them with the study findings: the obtained relationships will reveal how important matters of design are and precisely what to do about them” (1978, p.3)

When a student of his found that design did correlate with results, Glass said she, “did the sensible thing: covaried on design quality when comparing [treatments]”(1978, p.3). He went so far as to say that the only regret that he and Smith (Smith & Glass, 1977) had regarding their – now famous – meta-analysis of psychotherapy was that they arbitrarily excluded pre-post designs rather than including them and testing for the impact of this feature: “... we may have thrown away much good data”(1978, p.3).

Wortman (1994) took a lead from this approach in how he tried to defend meta-analysis conclusions from biases manifesting from the four basic threats to validity identified by Campbell and colleagues (Campbell & Stanley, 1966, and Cook & Campbell, 1979) – external, internal, construct, and statistical conclusion validity. He said that research quality is the product of relevance – construct and external validity – and acceptability – internal and statistical conclusion validity. Specific threats thrown up by the former two might be defended against largely through setting the domain of the meta-analysis, so as to include only the particular participants, times and locations that are appropriate and only those studies that use measures that operationalise constructs in the way intended. The latter two validities may be guarded in part by domain criteria – e.g. by the exclusion of uncontrolled repeat measures designs because they are open to too many internal validity threats – and also by care to make only appropriate comparisons and groupings in the analyses. In addition he suggested the performance of sensitivity analyses – the empirical comparisons that Glass was talking about - to see whether and to what extent some possible threats materialise in the particular dataset. He did not advocate going as far with sensitivity analyses as Glass appears to have, but does strongly advocate their use as one tool in a systematically considered strategy to defend a meta-analysis from biased conclusions.

Greenhouse and Iyengar (1994) and the Cochrane Collaboration (2006, para. 6.10) also promote the use of sensitivity analyses to aid understanding of the dynamics underlying

a dataset. As for Glass (1976, 1978) and Wortman (1994), their orientation towards method quality draws from the dataset itself to empirically judge the relative impact of design features. This insistence that there be a scientific justification for excluding data fits very nicely with the guiding principles decided upon when the present research was started, which included making the most use possible of available data, subject to reasonable practical scientific quality assurance in the practical and ethical context of the dataset.

The present approach

The approach taken in the present meta-analysis was informed by these views. The threats to validity identified by Campbell and colleagues as discussed by Kirk (1995), Pedhazur and Schmelkin (1991) and also by Wortman (1994) were worked through in the psycho-oncology context. They were then taken into account, as far as could be anticipated, in: 1. The domain of the study (e.g. the exclusion of studies using measures with inadequate reliability so as to strengthen statistical conclusion validity, and of uncontrolled repeat measures studies because of their vulnerability to maturation effects); 2. The way that groupings were organised (e.g. groupings of assessment time points, refer Analysis chapter); or 3. The coding of design features, to enable sensitivity analyses to be performed and decisions about handling the dynamics discovered to be made before substantive analyses were undertaken. The particular threats that are addressed by these analyses are briefly discussed in the relevant results sections.

In line with the guiding principles of this research, a particular approach to handling the outcomes of sensitivity analyses was decided upon. The aim was to retain data if possible in order to maximise variation in the dataset from which moderator patterns could be discovered. The establishment of accurate effect sizes, *per se*, was not the primary purpose of this work, and may not be a sensible objective anyway given the great variety in this field in terms of therapies and patient characteristics. The purpose was to see *patterns of relativity* – effect sizes that reflected significantly more and less impact of different levels of particular moderators. In these circumstances it made more sense to divide data along lines of heterogeneity discovered, but retain all groupings for analysis, than to exclude some data as ‘inferior’. The important thing was to make the presence and influence of the design features explicit.

This approach to research quality is messy compared with the arbitrary dispatching of studies according to pre-selected criteria. It involves making many choices based on the empirical evidence, and sometimes those choices are not clear cut. Further, it is necessary to end up with a scheme for dealing with quality that is manageable in a practical sense, and serves the purpose of the research. In the present case, that required not only retaining as much variation as possible within the dataset, but also limiting the number of divisions within the dataset to retain enough *n* to make given analyses worthwhile. This meant, again, that the emphasis would be on substantial patterns revealed in the preliminary validity analyses, rather than on trying to accommodate every last finding of heterogeneity or homogeneity. The question was not, “Is this *the*

scientific approach?”, since no such standard exists, but, “Is this *a* reasonably justified - and practical – scientific approach?”. No doubt other scientists would make different calls on some issues, and that is acceptable, provided they are justified.

Allocation to conditions and attrition

The decision to count in non-randomised studies requires fuller explanation. As Wortman (1994) points out, sensitivity analyses around this variable can show that these studies produce higher effect sizes which may be theoretically attributed to selection effects. He also notes that they can produce lower effect sizes. In the report of their meta-analysis of the impact of psycho-educational interventions on length of hospital stay, Divine and Cook (1983) observe that differential attrition between groups can result in selection effects despite initial randomisation. Their response was to initially include both designs and conduct a sensitivity analysis in order to clarify the dynamics of their particular dataset. In the context of psycho-oncology this approach makes even more sense because attrition due to illness and death is a particular issue, and, as mentioned earlier, the ethics and practicalities of the clinical context can make true randomisation unacceptable. On the positive side, non-randomised studies are often advantaged in terms of their generalisability to the clinical setting, so it would be doing the field a disservice to exclude them without proven justification. The approach taken in the present research was to admit non-random (but controlled) designs and conduct a sensitivity analysis not only of what means was used to allocate patients to conditions, but also of the balance of attrition between conditions, thereby tackling the possibility of selection effects from both ends.

On the subject of attrition, it is noted that the initial time point from which the attrition rate was measured was the pre-treatment test, where there was one. This was selected instead of the point of intention to treat because for some studies there was a considerable time lag between the two while for others they were simultaneous, giving them an advantage in terms of participant retention. Also, the rationale behind the usual favouring of intention to treat analyses was not applicable in relation to this issue. The usual desire is to take account of the acceptability of the treatment to participants – marked by their retention in or dropout from the study - in the final assessment of a treatment’s success, but in the current context the aim was to watch for bias from dropout caused by illness during therapy. Dropout during any lag between the registering of consent and pre-testing due to dissatisfaction with group allocation or pre-therapy illness, was not of concern here.

3. ANALYSIS

This chapter explains a series of decisions supporting the statistical integrity of the study. The first section steps through a number of technical statistical issues relevant to the meta-analytic method. The second explains some influential prioritising and categorising choices, and is therefore the more important.

Statistical matters

Analyses were conducted using the computer programme Comprehensive Meta-Analysis™ (CMA) version 2.2.046 (Borenstein, Hedges, Higgins, & Rothstein, 2007).

Computational model

Random effects was the computational model selected because of the wide range of therapy, sample and study styles being synthesised, meaning that an unknown number of true effects were expected to underlie any given effect size (Hedges, 1994). When subgroup comparisons were made, the software automatically switched to a mixed effects model because the assumption then became that study level variance was both systematic (known and dealt with in designing the levels of the variable that is being compared) and random (unknown). That switch explains a small difference in summary effect size that results from whether it is calculated as a unity, or produced as a by-product of a subgroup comparison. The practical effect of using a random effects model with this dataset - as opposed to a fixed effects model, which is only suitable for highly homogeneous studies - is that smaller studies carry more weight (because it is assumed that their variance is to a greater extent a reflection of true effects rather than random causes) and therefore, because they tend to have higher effect sizes, it lifts the mean effect size point estimate (the 'effect size') somewhat.

Effect size index

The effect size was drawn from a comparison of intervention and control group assessment scores at post-intervention, without adjustment for pre-intervention scores. This comparison was selected, having established that a control group was necessary, for the practical reason that it was expected that insufficient pre-intervention data would be regularly available to allow it to be built in to effect size calculations. The handling of baseline non-equivalence is described in the next subsection.

Hedges' g was selected as the effect size index for this comparison. It is the usual contemporary choice, and adjusts Cohen's d (standardised difference in means) to slightly favour studies with tighter variance, so that larger studies tend to be given more weight. A main effect size of d (or g) = 0.20 is 'small' by Cohen's calculations (Cohen, 1988) i.e. only 14.7% of equal-sized equally varying normal distributions for the two groups does not overlap; of 0.50 is 'medium' i.e. 33% of the combined area of their distributions does not overlap; and of 0.80 is 'large' i.e. 47.4% non-overlap.

CMA allows data for the computing of effect size estimates to be entered for this comparison from several parameters / statistics: n's, means and standard deviations; Cohen's d; *t*-values; and *p*-values. The most commonly used was the combination of n's, means and standard deviations, which produces the best estimate of effect size. For a few studies ANOVA *p*'s were used where baseline outcomes were the within groups factor, and treatment group was the between group factor (2 x 2 factors). On the rare occasion where means and standard deviations were reported but without n's, and the design and other reporting features of the study suggested similar group sizes, the total n was divided to enable the study to be used.

Adjustment for baseline differences

Because this meta-analysis uses post-intervention data only, it was recognised that results would be vulnerable to distortion up or down from large baseline differences between groups unless those rows of data affected could be excluded from analyses. Data were therefore tagged where statistically significant baseline differences were reported. For all but a few preliminary analyses (noted in the Preliminary analyses results chapter), such data were filtered out. There were a few studies where information regarding baseline differences and any statistical adjustment made was not supplied or no baseline assessment was taken. Data from these were retained in the set, erring on the side of inclusion. On the other hand, there were studies that statistically adjusted results for more than outcome differences at baseline, picking up differences in socio-demographic or medical variables as well, which ought to have strengthened their outcome data.

Significance reporting and level

Meta-analyses often do not report *z* score statistical significance because it is not relevant given the combined sample size from the number of studies analysed. The index of primary interest is one of magnitude rather than probability, i.e. effect size. In the present study, however, the attempt to account for structural confounds and the search for moderators required the dicing of data into frequencies that were quite small at times. *Z* score statistical significance then became relevant, and for consistency it is reported throughout (Rosenthal, 1995).

That said, null hypothesis testing has many limitations which are often overlooked as researchers and readers lazily rely on ' $p < 0.05$ ' to tell them whether an experiment 'worked'. At some length, Kline (2004) explains the theoretical and statistical drawbacks in relying on statistical significance, saying that the practice encourages mindless dichotomous thinking. He says that the emphasis in deciding upon the value of research outcomes should be on effect size, confidence intervals, replication (as the best way to deal with sampling error), and substantive - in the present case, clinical - significance.

The object of the present research was to detect moderating factors or trends that might confirm or prompt other research or practice. As a research tool, meta-analysis can produce powerful synthetic evidence, but it is entirely dependent on the availability and

quality of data presented in primary studies and on the categorisation and averaging of data. These three elements can easily cause subtle effects to be crushed out of view. Yet it was important for the present research to bring into visibility patterns that may become of interest in the field.

In view of these considerations, it was decided to table confidence intervals when needed, as well as regularly tabling effect sizes and precise Q statistic p 's, and noting z score p levels at both < 0.05 and < 0.10 (two tailed), and n sizes (number of studies – going to the issue of replication). The 0.10 alpha was chosen to help with the detection of subtle effects, and sometimes effect sizes that did not reach significance even at this level are discussed. By taking this approach, it was hoped that a more rounded – rather than dichotomous - picture of what the data were saying could be gleaned.

Multiple analyses

Many analyses were produced in the course of this research, raising the issue of 'random significance'. However, analyses were planned (to the extent possible given the need to accommodate results from preliminary analyses of study design features) *a priori* (Hedges, 1994), they generally used different independent variables, and important results are linked by a coherent theme, as shall be seen.

Omnibus effect sizes

For some of the substantive moderator analyses, the fine dicing of the dataset that was necessary meant that n would be very low and of no value if outcome constructs (anxiety, depression, and distress) were addressed separately. In order to preserve n , and given the relatively close relationship of the outcome constructs, an 'omnibus' outcome was often used to compute analyses, i.e. effects from the three outcomes were averaged together for any study that reported on more than one of them, and it was that effect size that contributed to the analysis. Although this loosened the interpretability of the outcome construct, the 'omnibus' results were considered theoretically adequate for the limited purposes to which they were put, i.e. detecting moderation based on patient characteristics (socio-demographic and medical).

In fact, the omnibus effect size is slightly weighted in favour of anxiety or depression constructs where, in relation to a particular study, effect size data had been entered for either or both of these outcomes as well as distress, and the former constructs were measured using subscales of the measure of the latter. This is because the anxiety or depression subscale is effectively counted twice – once on its own and once as a contributor to the distress measure - before the overall average is calculated. This weighting slightly strengthens the construct in favour of the more definite outcomes of anxiety and depression.

Homogeneity testing

In the search for moderators of effect size, a statistic that tests the statistical significance of differing variance associated with the effect sizes produced by different levels of a given variable is a vital tool. The finding of a significant difference between levels

indicates systematic heterogeneity in the data beyond that which would be expected from mere random variance, i.e. the formal moderation of effects by that variable, also known as heterogeneity. The statistic generally used – and used here – is Q statistic p , with alpha's reported at the same levels for the same reasons as for z scores (see above). Homogeneity testing is also the tool used for preliminary sensitivity analyses of study design features.

Note that while the presence of a statistical heterogeneity indicates moderation, the absence of it does not end the matter. Rebutting the conception that a line of homogeneity tests should only proceed so long as significance is found, Rosenthal (1995) points out that as significance tests, they are a function of the magnitude of both effect size (including variance) and sample size and says that the conclusion that variability is due to sampling error does not mean that you should not investigate moderators. On the contrary: “scientific progress can be defined in terms of scientists’ continually reducing the magnitude of sampling error by increasing their understanding of moderator variables” (electronic html p.9 of 23). The attempt to take such a balanced view of the limitations of statistical testing and the value of theory was made in the present study.

Intention to treat data

Data presented in research reports were generally based on ‘intention to treat’ rather than completion of the treatment course (‘effect of intervention’). Where both sets of data were presented - which was infrequently - the intention to treat data were preferred, but otherwise the data that were available was used. As commented by Edwards, Hailey, and Maxwell (2004), it is appropriate to include studies that did not report on an intention to treat basis owing to the nature of this particular population with its potential to be lost to follow-up.

Missing data

There were 146 studies in the final dataset, generating 614 unique rows of data with main outcomes (anxiety, depression, distress). Of these, 17 rows had insufficient data from which to calculate an effect size point estimate in that they reported only a p level of > 0.05 or > 0.10 . These rows were drawn from 7 studies. A further 45 rows were never entered due to gaps in reporting, and it may be assumed that these rows contained non significant data, although this gives little clue as to their effect sizes since p values are greatly influenced by sample size. Since there were a further 71 known reports (51 with insufficient data and 20 in a foreign language) that were also unusable and their impacts on analyses totally inestimable, and given that many reports of mere non-significance did not even state an effect direction, it was not worthwhile trying to calculate what range of impact the 17 rows may have had upon main effect sizes.

Two rows from one study reported only a p level of < 0.05 or < 0.005 , and the conservatively chosen values 0.049 and 0.0049 (respectively) were substituted, which enabled this study be included in the analyses. These two rows of data could have been

discarded, as the '> 0.05' rows were, but the desire to include as much variance as possible within the dataset made such a conservative estimate preferable.

Outliers

Outliers can have a considerable impact on mean effect size point estimates which are an average of the effect sizes from the studies in a given selection. Because the objective of a meta-analysis is to obtain summary data representative of a set, individual study result outliers – which are by definition unrepresentative - had to be dealt with. Effect sizes found in earlier meta-analysis (refer Introduction, Tables 1-1 and 1-2) fell in the range of 0.10 – 1.00, using Cohen's *d* or Hedges *g*. Limits of $g = 2.00$ or -2.00 were therefore considered reasonable to define outlier results and were set *a priori*. For analyses other than those conducted at the outset to investigate sampling bias, the outliers were windsorized, i.e. they were brought into line with the closest study effect size that was not an outlier (the target effect sizes were themselves all part of clusters rather than being isolated extremes). Choosing this strategy rather than simple deletion allowed outlier studies to be retained for their contribution to moderator analysis without undue concern about their impact on mean effect sizes.

Fail-safe N

Fail-safe N's were reported as part of an initial assessment of sample bias but not in relation to the many other effect sizes reported. The purpose of fail-safe N is to provide some indication of how robust a main effect size is, whereas the present study is about relative effects – moderators. The decision not to report them routinely was taken because in the present context its value was considered insufficient to justify the effort and space it would require. In any case, a large fail-safe N does not redeem the risks to conclusion validity that result from a poor search for studies, arbitrary method quality criteria, or unreliable coding. These are the issues of substance, and it is for the reader to judge how much weight can be put on an effect size accordingly. As mentioned above, *z* score statistical significance was reported for all substantive results. In the writer's opinion, this is a better understood and more useful statistic to consider in relation to small *n* conclusions.

Closing of dataset

As a matter of scientific integrity, it was resolved that no corrections would be made to the dataset once analysis was underway other than if erroneous transfer of data from the coding schedule to the programme was discovered. Having said that, an exception was made when it became necessary to check the coding of studies that 'screened in' for distress to determine which of them also 'screened out' for psychological history, and it was found that one study had been coded as screening in when it should have been coded as screening out. Given the significance that this difference made to the group of analyses that divided studies by type of screening (rather than simply screened *v.* unscreened) because of the low *n* involved in the screened in category and the theoretical importance of the results, the change was made. This single instance of recoding was not open to distortion from the writer's developing awareness of the

dynamics affecting the dataset, because no element of judgement was required in coding this item – the miscoding had been a simple error in noting the response number.

Groupings

As previously mentioned, decisions about how to group important variables, such as assessment time points and therapy types, have important implications for the construct and statistical conclusion validities of a meta-analysis. It is therefore important to provide a detailed description and rationale.

Unit of analysis and decision rules

Each study contributed only one effect size to each mean (or ‘summary’) effect size point estimate (usually referred to simply as the ‘effect size’, ‘result’ or ‘score’ for the variable). Where a study had used multiple assessment tools for the outcome being measured, effect sizes for each of those were averaged. Where a study had more than one treatment or control group, one had to be selected to represent the study in analyses where more than one treatment or control type was admitted (which was most of them) to avoid distortion to the results due to the statistical dependence of data. Along lines similar to those used by Devine and Westlake (1995) and (Sheard and Maguire (1999) in their meta-analyses in this field, the following decision rule was settled to enable the selection of one treatment group, and a preference coding was built into the dataset accordingly. The treatment group selected would be the one that was:

1. More psychologically substantial e.g. Berglund et al. (2007) had two treatment arms which both were provided with education and expressive-support, but one also provided group physical training – the latter arm was preferred because of the morale boost that it could have provided
2. More structured e.g. CBT was chosen over unstructured expressive-support
3. The researcher’s prediction of which would be more effective

It should be borne in mind that this meant that in analyses where more than one treatment type was admitted (e.g. analyses of socio-demographic moderators as opposed to analyses of different therapy types) CBT studies were more heavily represented.

There were only two studies that had two control groups. Because our aim was to maximise the contrast between treatment and control, the wait list or treatment as usual control was chosen over the attention placebo or treatment element placebo control.

Assessment time points

Assessment time points that were comparable across studies had to be selected. As with other issues of coding item design, clues as to what categories might fit the data were gleaned from the reading of previous meta-analyses and reviews in this field, then clinical relevance was discussed with two clinician researchers. Only the time points described as anchored to the completion of the intervention (below, group ‘A’) were anticipated in the original coding tool, however. While they netted by far the most data, as coding proceeded it was found necessary to develop other time point anchors to

accommodate the nature of certain classes of study. The three types of anchors and the particular time points associated with them are listed below:

- A.** Assessment points anchored to the completion of the intervention:
1. immediately post intervention
 2. short term follow up (up to 1 month post intervention)
 3. medium term follow up (more than one and up to six months post intervention)
 4. long term follow up (more than six and up to twelve months post intervention)
 5. very long term follow up (more than twelve months post intervention)

B. Assessment points anchored to the beginning of the intervention, used *only* where no post intervention data were available e.g. open-ended or very lengthy treatment regimes:

6. early mid term (three to six months into the intervention period)
7. late mid term (six to twelve months into the intervention period)

C. Assessment points anchored to the patient's stage in a particular medical treatment regime (radio-therapy, chemo-therapy, bone-marrow transplant) and therefore varying a little by individual:

8. before the commencement of medical treatment (very few studies: typically one-off educational or relaxation therapies)
9. mid way through the medical treatment term
10. post medical treatment
11. more than a month after medical treatment completion

Each study was coded according to only *one* of these three groups of assessment points in order to avoid creating statistical dependence, e.g. no study had one row of data coded '1. immediately post intervention' (group A) and another coded '2. post medical treatment' (group C).

Where more than one assessment was taken within the periods specified 2 – 4 or 6 – 7, the rule was to choose the point that was more central in the period. If there was more than one for period 5, then the rule was to choose the point closest to 12 months after intervention.

For the purpose of most analyses, it was necessary to group results from several of these time points together to ensure sufficient n. This grouping was made between studies however, not within them (refer Unit of analysis, above). The most commonly used

grouping was called ‘early times’. This was designed primarily to catch the first post-intervention assessment, as is the usual case with meta-analyses in this field. Most of the data it caught was from time points 1 and 2 – immediately or shortly after intervention. There were a few studies, however, that did not start assessments until more than a month after therapy had finished, and it caught them also, provided the assessment was made before six months had elapsed. In discussion with one clinician researcher, time points from the other two anchor groupings were chosen that were considered reasonably comparable. These results were used to compute an early times effect size. A ‘late times’ grouping was similarly designed for use in examining the durability of effects. The specific groupings are described:

‘Early times’ assessment time points:

EITHER

The *first* reported assessment point from

1. immediately post intervention
2. short term follow up
3. medium term follow up

OR

6. early mid term

OR

The *first* reported assessment point from

8. before medical treatment
9. mid way through medical treatment

‘Late times’ assessment points:

EITHER

The *first* reported assessment point from

4. long term follow-up
5. very long term follow up

OR

7. late mid term

OR

The *first* reported assessment point from

10. post medical treatment
11. more than a month post medical treatment

Therapy type groupings

Therapy type was a moderator that posed particular challenges for categorisation because of its complexity in psycho-oncology, where there is a great range of therapies, and because of its clinical importance. For this reason attention is given to it here.

Care was taken to choose a classification across only one dimension, namely therapy content (rather than confusing the issue with the mode of delivery), and to keep the classification on a high and broad level, avoiding very specific therapy types (c.f. the categorisation in the controversial systematic review by Newell, Simon-Fisher, and Savolainen, 2002, as discussed in the Introduction). The meta-analysis by Meyer and Mark (1995) and reviews by Barsevick, Sweeney, and Haney (2002) and Cwikel and Behar (1999) particularly informed the categorisation, and each of these had fairly similar categories. Alterations made for the purposes of the present study were to split the cognitive-behavioural categorisation into relaxation and other cognitive-behavioural therapies, and to add small categories for non-professional and 'indirect' therapies. The final therapy categorisation and parameter description was as follows:

- 1. Education / information** provided by a professional regarding cancer, cancer treatments, facilities (including orientation tour), or adjunctive services, nutrition, exercise, coping strategies or symptom management; includes bibliotherapy or information provided by some technological means, but does not include active rehearsal of new behaviours
- 2. Relaxation focused cognitive-behavioural treatment**, i.e. counselling / training in the use of coping strategies that focus on relaxation or stress management, including progressive muscle relaxation, mindfulness-based stress reduction, guided imagery, meditation, hypnotherapy, diaphragmatic breathing, autogenic training, systematic desensitization, biofeedback, electromyography, distraction, music
- 3. Broadly focused cognitive-behavioural treatment** i.e. counselling / training in the use of coping strategies that focus on cognitive reappraisal or behaviour modification or reinforcement, such as cognitive restructuring / reappraisal, challenging negative thoughts, positive self talk, self monitoring of thoughts or skills taught, problem identification, problem solving, contingency management, goal/expectation setting, activity pacing, behavioural activation, pleasant activity scheduling, assertiveness / communication / relational skills training, disability management, emotional control and anger management, fighting disease, cathartic, active interpretation / reconstruction, and may include role play or modelling. Problem solving included a wide range of topics, e.g. loneliness and isolation, morale and self management, sexuality and contact, body self esteem and general mood, communication, body self image and social adjustment, existential plight, social alienation and self identity, emotionality and personal control, dysphoria and depression

4. **Non-directive professional counselling / psychotherapy** ('expressive-supportive therapies') i.e. interactive verbal interventions, including nondirective, psychodynamic, existential, emotionally supportive / expressive/ reflective regarding the disease, its treatment, prognosis, and recovery, disability or death, general or crisis intervention; no specific behavioural or coping skills are taught; includes social support by professionals, but excludes therapist reconstruction
5. **Non-professionally led support or counselling** e.g. survivor testimony, self-help groups, telephone counselling, the teaching of coping skills by lay persons
6. **Indirect intervention** i.e. the immediate target of intervention is someone other than the cancer patient (e.g. communications training or counselling directed at medical staff or spouse without the patient present) but with the intention of benefiting the patient. (Measures were taken on the patient.)
7. **Other** (which transpired to comprise mostly written emotional disclosure studies)

Note that if a study was categorised 'indirect', it could not also be scored under any other category, e.g. the counseling of a spouse without the presence of the patient would be coded 'indirect' but not also 'expressive-supportive'. This is because the delivery mechanism of this categorisation was quite distinct so it was not comparable with other studies in some respects.

4. RESULTS: PRELIMINARY ANALYSES

Before computing main effect sizes and exploring substantive moderator effects, it was important to screen the sample of studies represented in the dataset for any threats to external or internal conclusion validity – that is, study design features that might bias conclusions drawn from results - that were not already covered by domain criteria. Most of the ‘external validity’ analyses below are fairly standard to meta-analytic method, and test the dataset for sample selection that is not representative of the theoretically extant distribution. In that sense, most of these analyses are a test of the thoroughness and balance of the sample search. The ‘internal validity’ analyses are not so commonly performed, and were conducted to undergird an empirical approach to vetting primary study design quality (refer Method chapter). They also scoped the research dynamics underlying the dataset, which was useful, but not their main purpose.

The interesting analyses among these results are set out in this chapter. Detail of results that were considered to be of no consequence was relegated to **Appendix O** under the appropriate heading. Because there were so many analyses conducted in the course of this study, many involving detailed categorisations, some discussion of results is entered into as results are presented while the reader has such details freshly in mind. This leaves discussion of broader implications for the final chapter.

External validity

The potential threats to external validity addressed include the more ready accessibility of published research and research reported in English, the nationality or culture of the study samples, changes affecting research over time, and different sample recruitment strategies.

For all analyses ‘early times’ assessment points were used (refer Analysis chapter) meaning that the very few studies without any such assessment time points did not contribute to results in this chapter. As with all results in the present study, where more than one effect size was available on a particular outcome construct relative to a particular treatment group (e.g. two measures for anxiety were used), the results were averaged for use in analyses.

Sampling bias

Possible biases in the sampling of studies were investigated from a number of angles. Main mean effect sizes were computed for each of the three constructs - anxiety, depression and distress - without regard to whether there was a statistically significant difference between groups at baseline or any other quality issues which had yet to be determined. The main effects and subset effect sizes computed for these purposes therefore should not be read in absolute terms but only relatively to each other. More accurate main effect sizes were computed once preliminary explorations of the dataset were complete.

Distribution of effect sizes

Funnel plots illustrate effect size (Hedges' g) against study size - or its proxy, precision ($1/\text{standard error}$), as in the case of the software used for the present study - and provide a visual check of the distribution of the set of effect sizes (Lipsey & Wilson, 2001). If there is a bias against low precision studies yielding negative effect sizes (usually these are small studies, which often go unpublished), then a 'bite' will appear in the distribution at the bottom left. However, the converse should not be assumed because the appearance of such a bite could be a result of true effects being observed in the plotted data, e.g. that small studies tend to yield higher effect sizes because psychotherapeutic treatments administered in an intense context afford greater personal attention to participants, which is itself therapeutic. The rigour of the search for studies is an important factor in weighing the likelihood of such an explanation. Outliers were not windzorised for these analyses so that the full spread of raw data could be seen.

Funnel plots for each of the three main constructs appear below (Figure 4-1. Funnel plot, anxiety, anxiety; Figure 4-2. Funnel plot, depression, depression; Figure 4-3, distress). While the plot for depression appears very evenly distributed, there is something of a 'bite' evident in the bottom left of both the anxiety and the distress plots. This asymmetry is confirmed by Eggers' regression intercepts of 2.08 for anxiety, 1.48 for depression, and 2.17 for distress, where 0 indicates perfect symmetry (Egger, 1997). Visual inspection suggests that perhaps six to eight low precision (i.e. small n) studies would be needed with extreme negative effect sizes to balance the distributions.

Figure 4-1. Funnel plot, anxiety

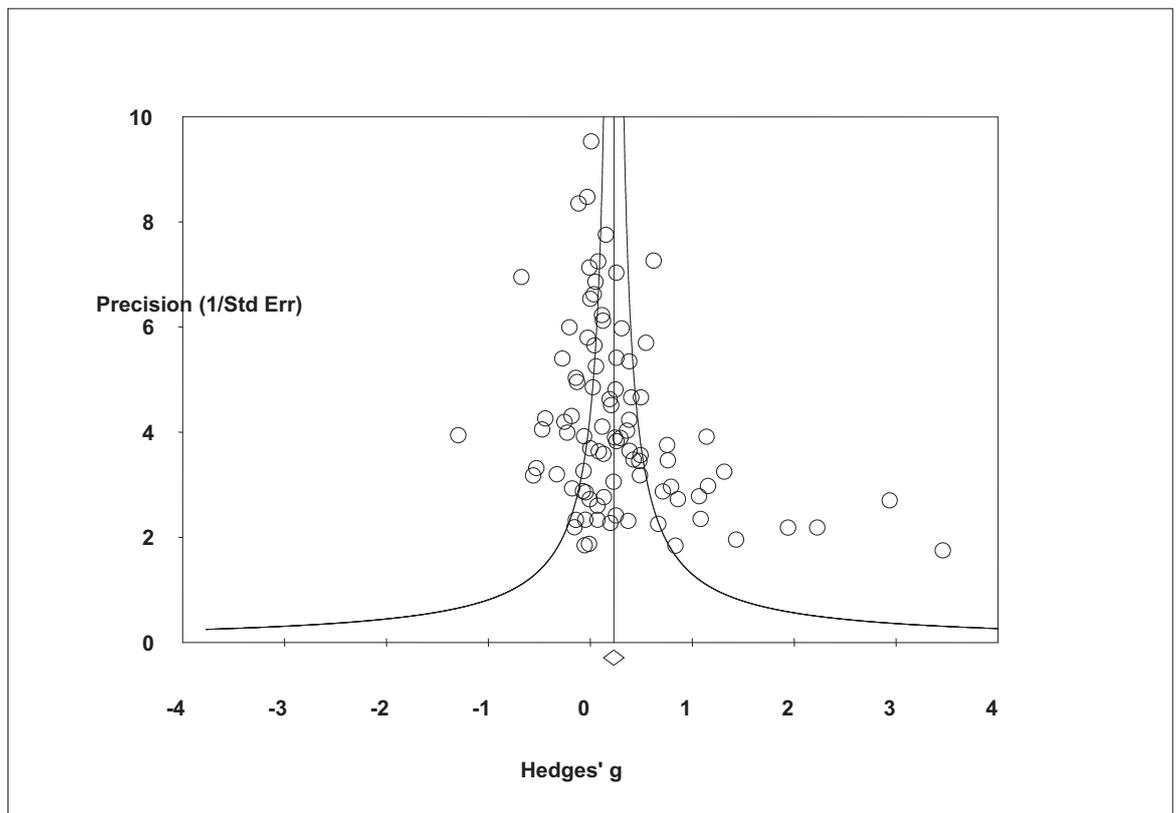


Figure 4-2. Funnel plot, depression

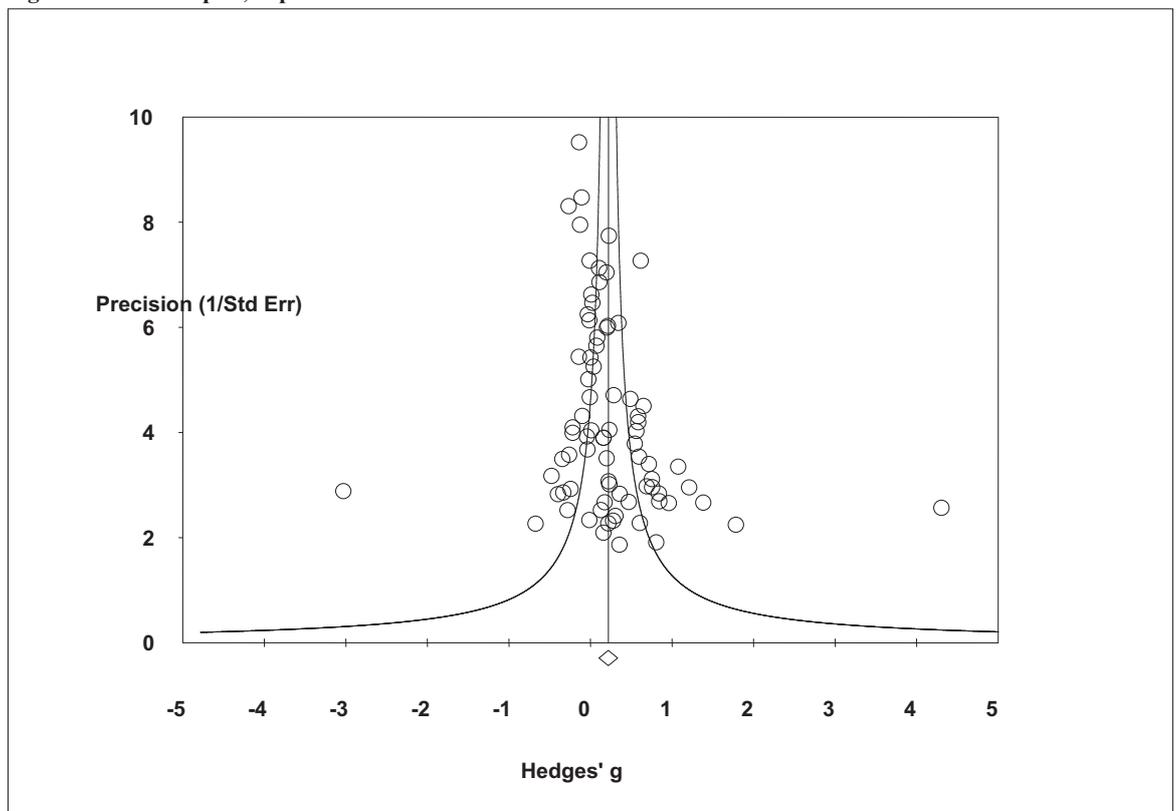
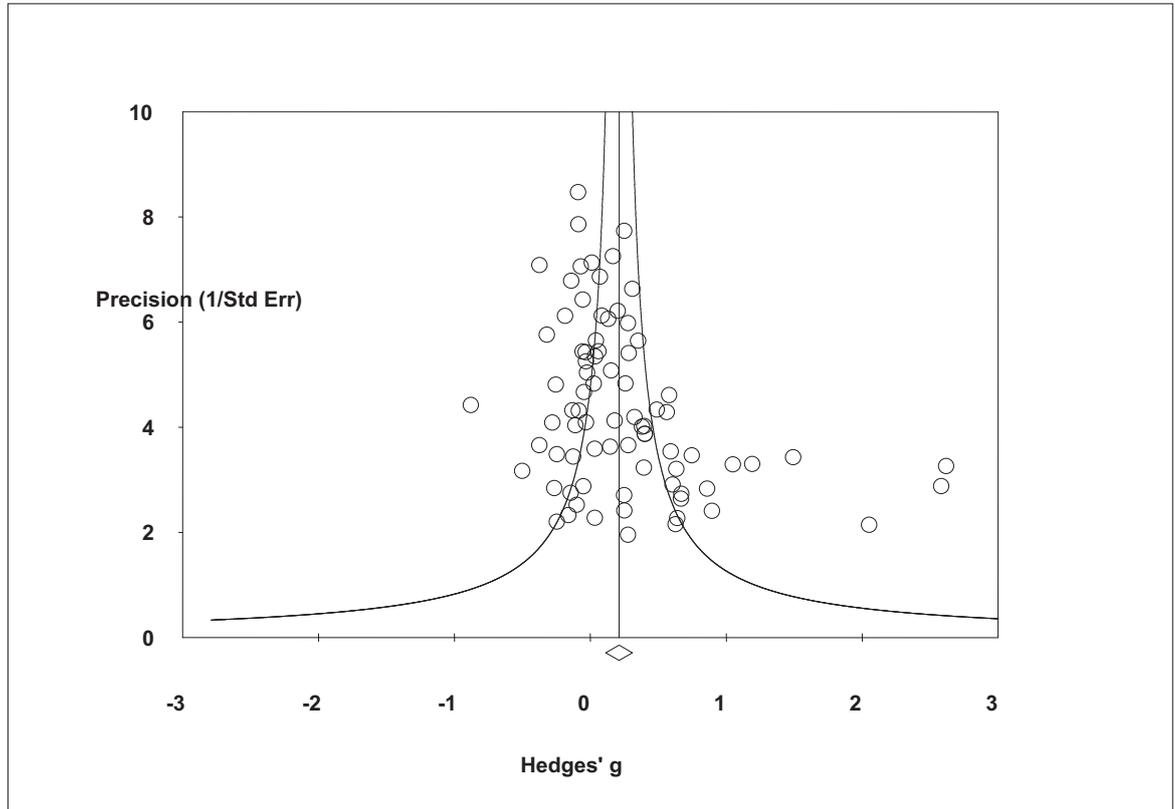


Figure 4-3. Funnel plot, distress



In interpreting this distribution it is asked how likely it is that six to eight small studies with extreme negative effect sizes were missed in the search, given its thoroughness? That consideration is weighed with the possibility that a true effect is at work of the nature mentioned, which would justify the slightly asymmetrical distribution. The latter explanation is supported by the comparison (further explained below) of published versus unpublished studies, which shows that they should not be regarded as distinct groups, and of studies of different sizes, which suggested studies should be distinguished on that ground. Also, as will be seen, depression outcomes often behave differently from those for anxiety or distress. Finally, it is asserted that the search for grey literature was sufficiently thorough, as it included extensive searches of three thesis databases plus the reference lists of 21 reviews and meta-analyses. It is reasonable to conclude that the 'bites' evident in the distributions displayed by the funnel plots are more likely to reflect true dynamics at work in the domain than any sampling bias.

Fail-safe N

Fail-safe N's calculated for these preliminary main effect sizes (0.23 anxiety, 0.22 depression, and 0.21 distress) were 1375, 909 and 1007 respectively. This means that given the nature of the distributions, it would take these numbers of fugitive studies reporting null findings to bring the p level relative to each effect size to an alpha exceeding 0.05 (Lipsey & Wilson, 2001). These are very large numbers, which could

be taken to indicate a robust dataset for main effects and support for the strength of the dataset for calculating more specific effects.

Outliers and windzoration

Outliers were identified at this stage and are tabled below (Table 4-1) with their respective effect sizes.

Table 4-1. Outliers

	Anxiety	Depression	Distress
Ali & Khalil (1989)	3.46		
Corchado (2006)	2.94		2.58
Trask, Paterson, Griffith, Riba, & Schwartz (2003)	2.23		
Badger, Segrin, Dorros, Meek, & Lopez (2007)		-3.03	
Nezu, Nezu, Felgoise, McClure, & Houts (2003)		4.30	2.61
Telch & Telch (1986)			2.05
Preliminary main effect sizes for comparison	0.23	0.22	0.21

It is worth setting out some detail of the features of these studies since they foreshadow some later moderation findings – both substantive and relating to study design features - and some crop up repeatedly in later discussion. Note that for all analyses following this description, outliers were windzored so that individual extreme cases did not distort effect size values.

Ali and Khalil (1989) use a sample of 30 bladder cancer patients from Egypt who underwent urinary diversion (stoma) surgery. The authors describe the social setting at the time as heavily stigmatizing of the disease, to the point where doctors never used the word ‘cancer’ and so never directly informed patients of their diagnoses, although terms such as ‘growth’ or ‘tumour’ were used. Despite this the authors’ investigations revealed that patients knew their diagnoses because of the mutilating surgery that they were obliged to undergo. None of the patients had had any formal education and most were illiterate peasant farmers. Standard medical care involved little information and patients bore the burdens of guilt and shame as well as fears concerning their prognoses, management of the stoma and its social impact, and financial worries. Unsurprisingly, the authors found that most of the patients were in “tremendous distress” (p.241) although pre-treatment measures were not reported other than to say there was no significant difference between groups. The one-off 30 minute pre-surgical intervention involved professionally and personally delivered education, delivered to the patient and their spouse together, explaining the surgical procedure and the stoma. It included the opportunity to express fears and a reassuring peer visit. The study was randomised.

Stand-out features include the high level of baseline distress suffered by the patients (although because this was not formally pre-tested, this was not recognised in coding),

the provision of information and reassurance that was highly relevant and otherwise very much lacking, personal delivery by people with credibility for the purpose, and delivery to both the patient and their significant other. It is to be further noted that the effect size that contributed to the preliminary 'early times' ($g = 3.46$) analysis was taken on the third post-operative day, but another assessment taken before discharge (approximately 12 days on) showed an even greater effect size, the highest of any in the dataset (6.17) as the anxiety of treated patients continued to drop at a faster rate than that of controls.

Corchado (2006) is another third world study, this time from Puerto Rico, and seems to share similarities with Ali and Khalil (1989) relative to its social context, notwithstanding that it is more recent. The sample of 60 patients were women about to undergo surgery for breast cancer and, again, the intervention comprised much more detailed information about the procedure than was normally provided. Social constrictions prevalent in the society are suggested by the author's note that only three of the eight hospitals approached were prepared to accommodate the study, citing patient privacy concerns and surgeon's disapproval as reasons. Again, the participants had had limited education (two thirds with high school or less) and were of limited financial means (only 5% employed full time).

The intervention consisted simply of a brochure setting out detailed information about what patients would experience during the surgery experience and regarding symptom management afterwards, including a large section of photographs showing theatre facilities and lumpectomy and mastectomy scars. So powerful was the impact of this information, that it actually reversed a baseline difference between groups (for distress, but not anxiety) that was statistically significant against the treatment group to become a difference in its favour of outlier proportions four weeks later. The distress outcome is filtered from later calculations because of the known distortion in the effect size caused by the baseline difference.

Trask et al. (2003) is a recent American study of 34 melanoma patients treated with three 50 minute sessions of CBT. Of note is that these patients were selected for their pre-therapy level of distress ($BSI\ GSI \geq 60$) and that, notwithstanding, they had no history of psychiatric problems and no stage IV disease. As will be seen (under Internal validity, screening / floor effect) each of these screens tends to lift effect sizes and the combination of both produces very high results. The therapy included relaxation, problem solving and cognitive challenging, and was tailored to individual need. CBT and relaxation will be seen to perform well for distressed patients. Interestingly, the dramatic impact displayed by anxiety scores ($g = 2.23$ at medium term, i.e. in the one to six months period) was not reflected in general distress scores, which produced a solid but unremarkable effect of 0.60.

Badger et al. (2007) is simply explained as the product of an unfortunate large (but non-significant) difference in baseline scores favouring the control group, despite randomisation. It also employed an attention placebo design, which can be expected to

generate smaller effect sizes than studies where the control receives no treatment at all since effect size is a measure of comparison.

Nezu et al. (2003) is another recent American study where participants were screened in for distress ($BSI\ GSI \geq 63$ and $HDRS \geq 14$) and out for other known psychological disorders before cancer diagnosis. The treatment, which was tailored, was again problem solving, and its delivery was highly intensive (ten 90 minute sessions with homework). The treatment group that is included in most analyses received the same therapy as a second treatment group which received therapy one-on-one, but patients had a significant other present. Both treatment arms produced high effect sizes on both depression and distress, but the one with the significant other present, more so, with improvement actually continuing over six and 12 months in the latter case, although no control was available for comparison at those times.

Telch and Telch (1986) is a well known outlier, also American, with two treatment arms. One provided intensive and individually relevant problem solving therapy (six 90 minute sessions plus homework delivered to 13 patients divided among three groups), but in combination also with relaxation, communication and feelings management training. This arm produced spectacular results, overcoming a statistically significant disadvantage to become a positive outlier. However, because the known baseline disadvantage was not adjusted for, this arm was filtered out of analyses. The second arm comprised similarly intensive professionally run group expressive-support. The treatment group once again overcame a disadvantage to produce a very strong effect size, but in this instance the baseline difference was small and non-significant and the effect size was not of outlier proportions, so these data transmitted through into the dataset used for substantive analyses without alteration. In this study, once again, patients were screened in for distress, this time using a structured interview.

Outlier summary

Some common themes emerge from this brief examination of the outliers, most notably high baseline distress, but also, intensively delivered and personally tailored problem solving for Americans, and cancer-treatment-specific information for patients from societies where this contrasted sharply with what was readily available. Such dynamics will be the focus of substantive moderator investigation after further preliminary examination of the dataset, this time for internal validity threats.

Publication

Comparisons of effect sizes yielded by published *v.* unpublished studies and large *v.* medium *v.* small studies were made to shed further light on the possibility and nature of any sampling bias. Such explorations indicate whether and what impact small unpublished studies that may have evaded the search may have had on overall effect sizes had they been found.

The results relating to publication status (published *v.* unpublished) against each outcome, noting the number of studies contributing to each effect size, and citing

relative Q statistic p 's which indicate statistical heterogeneity, can be seen in Table 4-2, below.

Table 4-2. Sampling bias, publication

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Published	0.21 (74)	0.11 – 0.31	0.21 (65)	0.11 – 0.30	0.16 (69)	0.07 – 0.25
Not published	0.24 (15)	-0.07 – 0.55	0.23 (12)	-0.08 – 0.53	0.33 (11)	0.04 – 0.62
Q statistic p	0.849		0.899		0.269	

ES = Hedges g effect size point estimate; n = number of studies in subset; CI's = confidence intervals.

The published studies comprised 121 journal articles and one other, while all 23 unpublished studies were Masters or PhD theses. Unpublished studies comprised about 18% of the 'early times' set, and produced point estimates 0.03, 0.02, and 0.17 higher than published studies for anxiety, depression and distress respectively. Had 'publication bias' been present, the reverse trend would have been evident, i.e. the detection of more published studies with higher effect sizes would have manifested (Lipsey & Wilson, 2001). Furthermore, the 95% confidence intervals associated with the unpublished studies completely engulf those of the published studies in every case, and their Q statistics p values do not approach significance. From this it is concluded that unpublished studies do not comprise a distinct subset, and the whole dataset does not suffer from publication bias.

Study size

The dataset broke into three subsets with similar frequencies when 'large' studies were defined as having a post treatment sample size of 100 or more; 'medium' 50-99; and 'small' 10 – 49. The emerging picture differs a little across constructs so they are examined one at a time (Table 4-3).

The effect sizes and confidence interval spreads for the small and medium subsets with anxiety outcomes are very similar and largely overlapping, indicating no difference arising out of this division of the data. However, the confidence intervals for the large sample subset is much tighter, despite fewer contributing studies, encroaching on only about one third of the spread of the other subsets, and setting a point estimate that is very much lower, at 0.07 as opposed to 0.32 and 0.27 for the small and medium subsets. The Q statistic p ' is exactly 0.05 indicating the statistical significance of this heterogeneity.

An even more extreme distinction is seen for distress, with the large study effect size a mere 0.05, while small studies yield 0.26, and medium, 0.33. The Q statistic p is 0.016 again confirming heterogeneity between these subsets.

However, for depression the effect sizes and confidence intervals are more aligned with each other, displaying more overlap and reducing the Q statistic p to non-significance

(0.326), although the difference between the effect size for large studies and medium or small ones is still considerable (0.11; 0.22; and 0.27, respectively).

Table 4-3. Sampling bias, study size

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Large	0.07 (27)	-0.03 – 0.18	0.11 (25)	0.00 – 0.23	0.05 (32)	-0.05 – 0.15
Medium	0.27 (29)	0.06 – 0.47	0.22 (20)	0.02 – 0.41	0.33 (25)	0.15 – 0.52
Small	0.32 (32)	0.11 – 0.53	0.27 (29)	0.08 – 0.46	0.26 (21)	0.03 – 0.50
Q statistic <i>p</i>	0.050		0.326		0.016	

Notes as for Table 4-2.

What the Q statistic *p* levels tell us is that there is likely to be a true (systematic) effect – or more than one - associated with study size impacting the anxiety and distress data, beyond what is to be expected from mere within-study (random) variation. A pairwise comparison of each of the possible combinations of subsets on this variable was made in an attempt to pinpoint the locus of the heterogeneity indicated by the significant Q statistic *p* levels. Results are tabled below (Table 4-4).

Table 4-4. Sampling bias, study size, pairwise comparisons for heterogeneity

	Anxiety		Distress	
	Large	Medium	Large	Medium
Medium	0.105		0.008	
Small	0.035	0.693	0.109	0.633

Figures reported are Q statistic *p* levels. Absence of shading indicates significance at $p < 0.05$; light shading indicates approaching significance. Four studies that did not report post treatment n's were excluded from calculations. Other notes as for Table 4-2.

Formal heterogeneity (at $p < 0.05$) is shown between large and small studies (Q statistic $p = 0.035$) and for distress, only between large and medium sized studies ($p = 0.008$). However, if Q statistics that merely approach significance are taken into consideration as indicators of trend, then the tabled comparisons confirm a consistent break between large studies on the one hand, and small or medium sized ones on the other.

Possible confounding

Later analyses of internal validity issues explain these results. It will be seen that studies that did not apply pre-recruitment screens for distress or psychological history produced results systematically lower than those that did. It will also be seen that studies that used treated (placebo or treatment as usual) control comparison groups produced lower scores than those that used a no treatment (e.g. wait list) control. Both of those variables were fairly evenly spread in the dataset as a whole, but an examination of the frequencies in relation to large studies shows a disproportionate

concentration of the lower scoring levels of both. Refer to the frequency percentages set out below (Table 4-5).

Table 4-5. Large studies, confound proportions

	Anxiety	Depression	Distress
% unscreened	74	76	72
% treated control	70	72	66

Three studies that did not report the nature of the control group were excluded.

This co-variation of study size with particular study design characteristics means that larger studies are not simply larger than others – they tended to be designed differently in other important ways as well, confounding any moderation that might be predicted by size alone.

Study size conclusion

The issue of concern here is whether there is a sampling bias associated with study size. It is a normal expectation that larger studies will produce smaller effects, so a certain manifestation of this is not to be considered ‘bias’ as such. But with the present dataset it has been seen that two internal validity factors also heighten and confound the heterogeneity seen. It was concluded that the other two variables were more descriptive of the cause of bias, and size, as such, was not a variable to cause concern in interpreting results.

Original language

The last of the sampling bias checks was made between those studies which appeared to be originally written and then published in English and those for which this judgment could not be made (e.g. the population was from a non English speaking country and the author also was resident there, and no source report, such as a thesis, was sighted which was written in English). The rationale for this analysis is found in the presumption that studies written in a foreign language will generally be first published in a foreign language, that only a few of them will be republished in English, and those that are, as a set, will tend to show a bias towards positive results. However, such trend as was apparent in results was not statistically significant and seemed explained by the screening design feature mentioned in relation to study size. Result detail was therefore relegated to **Appendix O**.

Representativeness

Analyses were run to investigate a number of features relating to the representativeness – the generalisability – of the dataset. These included analyses around nationality (OECD membership and culture), publication year and sample selection strategy. Results regarding OECD membership are reported in detail below. Those for culture are not, but are detailed in the appendix for reasons mentioned under the OECD analysis. Year of publication (or release) was analysed by meta-regression and a slight

negative but non-significant trend was detected. Sample selection strategy was coded according to how well the sample represented the cancer patient population from which it was drawn: representative selection (random or consecutive); volunteer response to advertising; referral; or a mixture. The comparison of these levels produced non-significant Q statistics, and detail of these results and those relating to time are also relegated to the appendix.

Another generalisability issue presented by the dataset was the screening out of potential study recruits for psychological history. This was done presumably with the object of producing scientific homogeneity amongst the sample, but threatened the generalisability of results to practice contexts. Another screening procedure – screening recruits in for a proven level of distress – does not pose a threat to external validity since such screening occurs routinely in many practice contexts, but could threaten internal validity. The impact of both forms of screening are considered below in the section on internal validity.

OECD membership

An analysis was performed comparing results based on OECD membership as a rough proxy for national wealth. The dataset drew samples from the following countries, and OECD non-member countries are marked with an asterisk (*) (numbers are given where they exceed one): USA (78), UK (12), Canada (16), Australia and New Zealand (nine), Scandinavia (seven), Greece (four), Japan (four), Hong Kong* (four), Taiwan* (two), Italy (two), Germany (two), Israel* (two), Egypt*, Netherlands, Puerto Rico*, Spain, and Brazil*. No analysis was run in relation to depression because there was only one study in the non-OECD subset for that outcome.

Statistically significant and quite dramatic differences are evident in the results table below (Table 4-6). Non-OECD nations produced effect sizes of $g = 0.64 - 0.79$ higher than the generally wealthier OECD member nations. Although there were many fewer studies in the non-member subset, confidence intervals barely overlapped, producing differences that were highly statistically significant (Q statistic p 's: anxiety, 0.035; distress, 0.005).

Table 4-6. Nationality, comparison by OECD membership

	Anxiety		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's
OECD member	0.17 (83)	0.08 – 0.26	0.13 (72)	0.05 – 0.21
Non member	0.96 (6)	0.23 – 1.70	0.68 (8)	0.30 – 1.07
Q statistic p	0.035		0.005	

Notes as for Table 4-2.

These dramatic results warranted further investigation. The two internal design confounds identified later in this chapter and found to be fairly evenly distributed over the whole dataset were found to be unevenly distributed in these subsets. Out of the 15

studies from countries that were not OECD members, 11 screened potential recruits and 12 used no treatment control comparison groups. This meant that they had disproportionate frequencies of these predictors of significantly higher effect sizes. Since the internal design confounds are based on larger subsets than the frequencies yielded for OECD non-members, it was decided that those variables will be preferred as the basis for making distinctions in the main effect analyses. However, it is noted that these features of the non-OECD studies may indicate the harsh realities imposed by their poorer economic conditions - that generally only those patients with proven distress can be offered treatment, and that the usual comparison group – the ‘usual care’ condition for patients – is no therapy at all. The results are also an early hint at the role of deficit in predicting high effect size.

Studies were also compared on a rough estimation of whether the predominant culture in the country of sample was collectivist or individualist. However, in practical terms this meant only that two studies from Japan and two from Israel ‘swopped sides’ from the OECD analysis and made no appreciable difference to the outcomes. Details of this analysis were relegated to the appendix.

External validity summary

In the above section the dataset was surveyed for threats to its external validity. It began with testing for a number of sampling biases. Investigation revealed a possible gap in the sampling of small studies with extreme negative effects, but this appears merely to illustrate the nature of the domain. No publication bias was found, and the proportion of small studies in the set (about one third) is pleasing. Some distinction between the effect sizes produced between medium or small studies and larger ones is apparent but is not taken as evidence of sampling bias. A small and non-significant tendency towards higher effect sizes in studies which may have first been published in a foreign language reflects a disproportionate number of screened studies in that group and may be dismissed. In sum, it appears that a well balanced dataset has emerged from an assiduous search.

Generalisability across nationalities was also investigated. The number of studies from poorer non-OECD nations was relatively few, but they produced dramatically higher statistically significant mean effect sizes. These differences appeared to have their source in research design features that may have been the consequence of relative poverty. The design features, rather than OECD status, will be accounted for in the substantive analyses.

Internal validity

A number of variables that could threaten the internal validity of the dataset were investigated and are the topic of this section. They are variables that speak to the nature or quality of the design of the primary studies contributing to the dataset. As mentioned explained in the Methods chapter, this set of sensitivity analyses was necessary to provide an empirical rationale for the way that study quality would be handled, so as to

avoid excluding data on arbitrary grounds. Studies that did not adjust for a known statistically significant difference in baseline scores on an outcome measure were culled from the dataset before these analyses were undertaken, and were not used for any further analyses. In other respects, analyses were done using the same data as for the external validity investigations (and most of the later substantive analyses) that is, with windzorised outliers and ‘early times’ assessment points, and multiple measures of any one construct in any given study averaged.

Allocation to conditions

Comparison of designs

Studies were coded both for how participants were allocated to therapy or control conditions and for whether or not the assertion of that design was supported with specific information as to how it was executed. The levels of design were: 1. Random assignment after matching, stratifying, or blocking; 2. Random assignment, simple; 3. Pseudo-random assignment after matching, stratifying, or blocking (but there were no studies in this category); 4. Pseudo-random assignment, simple; 5. Non-random assignment, post-hoc matching (again, no studies); 6. Non-random assignment, other; and 7. Not reported / unclear (two studies for each of depression and distress). The ‘not reported’ studies were left out of these analyses. Results based on the coding for mere assertion of design type are dealt with first, in an analysis that includes the investigation of possible confounding. An analysis regarding whether an assertion of randomisation was substantiated with the reporting of detail as to the means used, then follows.

‘Pseudo-random’ was a categorisation for studies that could not strictly claim randomisation because, for example, participants were ‘randomised’ by week rather than individually, or were assigned alternately individually or in blocks to the different experimental arms. It was important to create this category due to the practical and ethical constraints in the field and the need to investigate whether or not this design really made a significant difference to outcome. Such designs are qualitatively different from naturalistic (‘non-random’) designs, which take advantage of already occurring phenomena which are inherently more vulnerable to confounding. This latter category included studies that had the experimental and control arms at different hospitals or which were divided by a single time point in the one hospital (typically, the first time period was used to recruit a standard care control group, then the new intervention was introduced to the standard medical package and the patients that received that comprised the therapy arm). But recall (Method chapter) that studies were excluded from the domain altogether where participants chose their own group assignment or there was evidence that the groups were not equivalent in nature.

Results of the study design analysis are set out below (Table 4-7, all outcomes; then figures showing forest plots for each outcome: Figure 4-4, anxiety; Figure 4-5, depression; Figure 4-6, distress). Both tables and forest plots are presented so as to

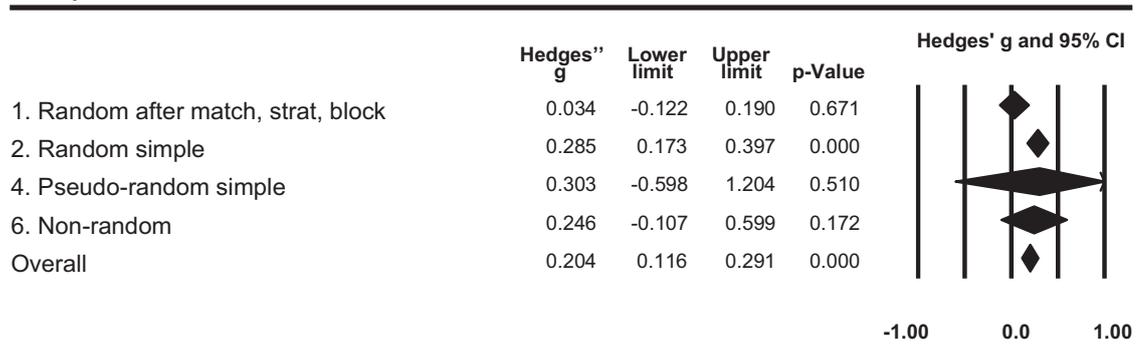
provide the clearest possible understanding of how the data lay on this contentious issue.

Table 4-7. Allocation to conditions

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Random strat	0.03 (16)	-0.12 – 0.19	0.06 (15)	-0.08 – 0.20	0.01 (11)	-0.09 – 0.11
Random simple	0.29 (55)	0.17 – 0.40	0.32 (43)	0.19 – 0.44	0.26 (50)	0.14 – 0.38
Pseudo-random	0.30 (5)	-0.60 – 1.20	-0.13 (3)	-0.39 – 0.14	-0.04 (7)	-0.25 – 0.16
Non-random	0.25 (7)	-0.11 – 0.60	0.26 (7)	-0.12 – 0.64	0.03 (3)	-0.25 – 0.30
Q statistic <i>p</i>	0.082		0.005		0.006	

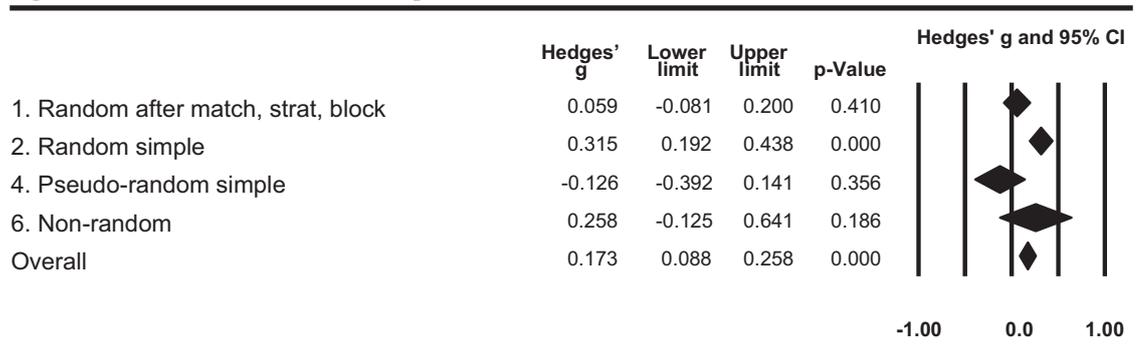
Random strat = Random, stratified. Other notes as for Table 4-2. Studies that did not report on this variable were excluded from the analysis.

Figure 4-4. Allocation to conditions, anxiety



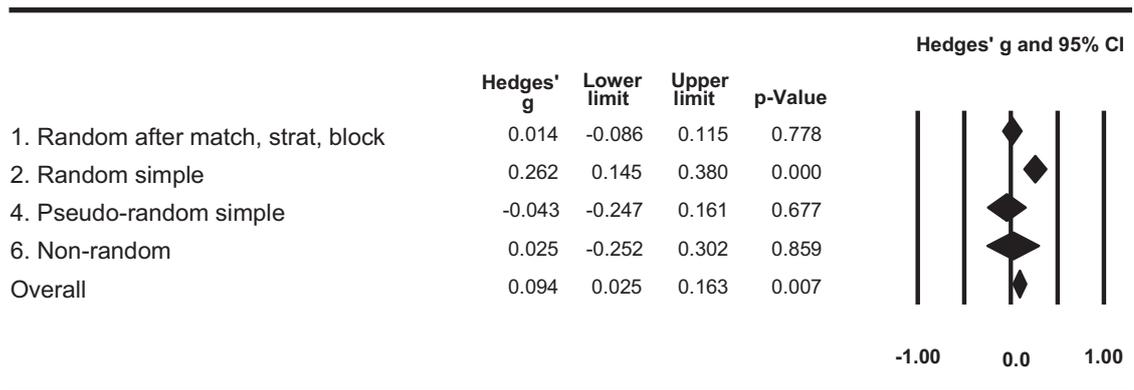
Outliers are windzorised. 'Early times' assessment time points. 'Not reported' studies and studies that did not adjust for known baseline significant differences are excluded.

Figure 4-5. Allocation to conditions, depression



Outliers are windzorised. 'Early times' assessment time points. 'Not reported' studies and studies that did not adjust for known baseline significant differences are excluded.

Figure 4-6. Allocation to conditions, distress



Outliers are windzorised. 'Early times' assessment time points. 'Not reported' studies and studies that did not adjust for known baseline significant differences are excluded.

An immediately striking feature of these results is that depression and distress yield statistically significant Q statistic p 's at $p < 0.05$ (0.005 and 0.006 respectively) and anxiety does so at $p < 0.10$ (0.082), meaning that the dataset is not to be regarded as the product of a unified effect but, rather, made up of distinct subsets apparently distinguished by study design. A second striking feature is that although both random stratified and random simple designs have quite tight confidence intervals, simple designs take a clear step to the right of about the distance of one confidence interval (nearly two standard errors) in every case, that is to say, they produce effect sizes that are clearly higher by a small effect size of about $g = 0.25$ for each outcome construct. The wider diamonds produced by pseudo-random and non-random designs are attributable to their small n 's (ranging from three to seven compared with n 's ranging from 11 to 16 for random stratified and 43 to 55 for random simple designs).

To identify which pairs of design types the statistically significant Q statistic p 's refer to (remembering that subsets with higher n 's attract significance with only modest effect size differences), pairwise comparisons were made, producing the following table of Q statistic significance levels (Table 4-8).

Table 4-8. Study design, pairwise comparisons for heterogeneity

	Anxiety			Depression			Distress		
	R strat	R sim	Pseudo	R strat	R sim	Pseudo	R strat	R sim	Pseudo
R sim	0.010			0.007			0.002		
Pseudo	0.564	0.970		0.230	0.003		0.618	0.011	
Non-R	0.281	0.836	0.908	0.339	0.782	0.107	0.943	0.122	0.696

Figures shown are Q statistic p levels. Absence of shading indicates significance at $p < 0.05$; light gray indicates approaching $p < 0.10$ significance. Studies for which the design type was not clearly reported ($n = 4$) are excluded from this analysis. N 's for each cell are reported in Table 4-9.

R strat = random, stratified designs; R sim = random, simple; Pseudo = pseudo random; Non-R = non-random

For all three outcome constructs, it is confirmed that the observed difference between random stratified and random simple designs is based on a true statistical underlying effect - Q statistic p 's are statistically significant at 0.010 for anxiety, 0.007 for depression, and 0.002 for distress. On the other hand, random stratified designs do not form a statistically distinct group from the theoretically weaker pseudo-random and non-random designs. Contrary to what might be expected given that result, random simple designs do prove to be distinct from pseudo-random designs, despite the smallish n 's of the latter group, in relation to depression and distress but not to anxiety where the pseudo-random n is just three (Q statistic p 's are statistically significant at 0.003 for depression and 0.011 for distress, but show almost perfectly homogeneity at 0.970 for anxiety). That outcome is undoubtedly influenced by the fact that the random simple subset has the highest n (55, 43, and 50 for anxiety, depression and distress respectively) while the largest subset for any other design and outcome construct is only one third as large (16, random stratified, anxiety). Finally, non-random designs approach statistical significance in relation to pseudo-randoms for depression, and simple randoms for distress, despite the low n 's in these subsets (seven and three, respectively).

What does this pattern of heterogeneity say? If random stratified designs are statistically distinct from random simples, then it would be expected that they would be statistically distinct from the weaker pseudo-random and non-random designs as well, but in no case does this prove to be so. Whilst this could be attributed to small subset sizes, there are reasonable numbers in the random stratified set, and near statistical significance was thrown up in comparisons of smaller subsets. Random simple designs show heterogeneity relative to more other subsets than any other design type, but again the pattern is far from complete. The distinction between random stratified and random simple designs, in particular, seems to pitch theory against empirical evidence. This confusion suggests that there is another element – or other elements – that are confounding the picture.

One implication that may be taken from this result pattern is that no strong or clear superiority of one design type has emerged, i.e. that a sensibly designed pseudo- or non-randomised study can produce results with similar effect sizes as a random or stratified random study in this domain. Such studies should not be excluded from reviews then, but can assist in providing added variation to a dataset.

Because there was no clear pattern of heterogeneity, the dataset was treated as homogeneous on this factor, and pseudo- and non-random designs were retained in it for substantive analyses.

Possible confounding

After investigations into internal design and substantive moderators had been completed, frequencies for three of the moderators found were collated to see whether they might shed light on the peculiar pattern displayed by the allocation to conditions variable. Those moderators are recruit screening and nature of the control comparison

group (details in this chapter) and cancer site (see the results for Patient characteristics: medical variables). Proportions of the lowest scoring level of each of these three moderators are tabled against the various study designs for the two more precise outcome constructs: anxiety (Table 4-9) and depression (Table 4-10).

Table 4-9. Allocation to conditions, anxiety, confounding variables

	% unscreened	% treated control	% breast site
Random, stratified	87.5	56.3	62.5
Random, simple	50.9	45.5	34.5
Pseudo-random	66.6	50.0	50.0
Non-random	42.9	57.2	28.5

Studies that did not report on the moderating variables were built into the percentages. They comprised: screening: 4 random, simple; nature of control: 3 random, simple; cancer site: 1 random, simple; 2 non-random. Other notes as per Table 4-2.

Table 4-10. Allocation to conditions, depression, confounding variables

	% unscreened	% treated control	% breast site
Random, stratified	86.6	66.6	53.3
Random, simple	48.8	41.9	37.2
Pseudo-random	75.0	50.0	25.0
Non-random	28.6	57.1	28.6

Studies that did not report on the moderating variables were built into the percentages. They comprised: screening: 4 random, simple; nature of control: 2 random, simple; cancer site: 1 non-random. Other notes as per Table 4-2.

These results take the mystery out of the pattern shown by the allocation to conditions results. The heterogeneity between the random stratified studies and the random simples is supported by a difference of 11 - 37 percentage points in the frequency of each of the low scoring confounds for each of the outcome constructs. Simply put, the random stratified results are confounded by having a disproportionately high share of the low scoring levels of these other moderating variables – they typically do not screen recruits, use treated control groups for comparison, and treat breast cancer patients. These features presumably accompany a desire to maximize sample size. The relatively good effect sizes yielded by the non-random studies are supported by relatively low proportions of breast cancer patients and, in relation to depression, unscreened studies. It may be that non-random designs would produce significantly smaller effect sizes than random designs if they had the same proportions on these moderators. However, we cannot know that.

Confidence in design report

Many simple random studies merely asserted their design rather than specifying the mechanism of randomisation. Some reviews, in an attempt to uphold high method quality standards, have excluded studies because of this reporting failure which contravenes CONSORT (*Consolidated Standards of Reporting Trials*, Moher, Schulz, & Altman, 2001) recommendations. An analysis was performed around whether or not those studies that asserted randomisation also specified a method i.e. whether the poorer reporting of some studies had implications for the overall quality of the study which was reflected in different results. This proved not to be the case however, with levels of this variable showing clear homogeneity. Detail of the analysis was relegated to the appendices (**Appendix O**).

Concealment of allocation to groups

The *Cochrane handbook for systematic reviews of interventions* (Cochrane Collaboration, 2006, para. 6.3) warns that failure by researchers to secure the process of allocating participants to conditions, so that that assignment is unknown to both participant and researcher until it is fixed beyond possible alteration, can cause a greater threat of bias than inadequacies in the randomisation technique itself. This is because either conscious or unconscious motivations of the participants or researchers can be brought to bear on the allocation, resulting in more patients with a particular characteristic in common ending up in one or other group. The possible impact of this phenomenon was tested in regard to all studies purporting randomisation, including pseudo-random designs. Studies were coded as either specifying their method of allocation concealment, merely asserting concealment, or not concealing the allocation process. Results for this analysis also showed homogeneity and were also relegated to the appendices.

Attrition

As explained in the Methods chapter, imbalanced attrition between the treatment and control groups was recognised as posing a possible threat to the validity of results. In their meta-analysis Divine and Cook (1983) excluded studies with an overall attrition rate of 15% as well as those with differential attrition between groups exceeding 10%. In the context of the present study, where participants all had a potentially lethal disease, the latter strategy was seen as the important one to protect against a possible selection effect. Studies were therefore coded for any imbalance in attrition between groups that exceeded 10% at each assessment point, and a heterogeneity analysis was conducted between studies that showed balance on this factor and those that did not. Studies were also coded for the reasons cited for attrition, categorised by predominance: Illness or death; Treatment side-effects; 'Other' (e.g. travelling to therapy was found to be inconvenient, or participant was lost to assessment due to an un-notified change of address); and mixed reasons.

Only one of the six analyses run with regard to these two aspects of attrition (reasons, anxiety) produced heterogeneity, and it was not considered sufficiently important to take into account in structuring further analyses, nor were the results sufficiently interesting to detail here. Refer to the results appendix.

Nature of control condition

Effect size is an index of contrast rather than of absolute value. In the present case, the outcomes being contrasted are those produced by the intervention group and the control group of each study. Therefore the nature of the control condition has as much bearing on effect size as the nature of the intervention condition, and the possible impact that differences in control condition type may have on effect size ought to be investigated.

Studies were coded according to the following responses to the question, “What is the nature of the control condition?”: 1. Participant receives nothing / is placed on a waiting list for the intervention / receives only the very minimal contact essential to administering the study; 2. Treatment as usual; 3. Attention placebo; 4. Treatment element placebo; and, ‘Unclear’. Because different authors and settings engender the use of different understandings of the term ‘treatment as usual’, this was defined to include medical protocols that appeared likely to contain some elements of the psychological intervention but in an undeveloped form. If it appeared that there was no element of the intervention in the normal medical protocol, then the study was coded ‘receives nothing’ regardless of whether or not the author used the term ‘treatment as usual’ to describe the control. To qualify for coding as a ‘treatment element placebo’, the control group had to have been deliberately provided with a partial form of the intervention (e.g. in one study, taped relaxation instructions with background music was provided by way of intervention, and the control was provided with a tape of the music alone).

It was expected that statistically significant differences would be produced from this analysis, and they were, for every construct (Table 4-11): anxiety, Q statistic $p = 0.016$; depression, 0.000; distress, 0.002. Pairwise comparisons of the different combinations were carried out to locate the particular relationships that were heterogeneous (Table 4-12).

Table 4-11. Nature of control condition

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Nothing	0.30 (42)	0.23 – 0.38	0.34 (32)	0.18 – 0.51	0.36 (36)	0.36 – 36.00
Usual	0.06 (26)	-0.01 – 0.13	0.16 (23)	0.05 – 0.28	0.05 (23)	0.05 – 23.00
Attn P.	-0.88 (2)	-1.29 – -0.47	-0.54 (2)	-0.90 – -0.17	-0.05 (2)	-0.05 – 2.00
Tx El P.	0.05 (10)	-0.06 – 0.17	0.07 (11)	-0.12 – 0.25	-0.02 (8)	-0.02 – 8.00
Unclear	0.21 (3)	-0.19 – 0.61	0.70 (2)	0.19 – 1.21	-0.17 (4)	-0.17 – 4.00

Q statistic <i>p</i>	0.016	0.000	0.002
-----------------------------	-------	-------	-------

Usual = treatment as usual; Attn P. = attention placebo; TxEl P. = treatment element placebo. Other notes as per Table 4-2.

Table 4-12. Nature of control condition, pairwise comparisons for heterogeneity

	Anxiety			Depression			Distress		
	Nothing	Usual	Attn P	Nothing	Usual	Attn P	Nothing	Usual	Attn P
Usual	0.005			0.083			0.000		
Attn P	0.428	0.235		0.111	0.000		0.103	0.664	
TxEI P	0.009	0.911	0.732	0.031	0.387	0.431	0.001	0.466	0.887

Figures reported are Q statistic *p* levels. Absence of shading indicates significance at $p < 0.05$; light gray indicates approximately 0.10 significance. Usual = treatment as usual; Attn P = attention placebo; TxEl P = treatment element placebo. Studies for which their status regarding this variable was unclear were excluded from this analysis.

The shading on Table 4-12 highlights an almost consistent heterogeneity between control conditions receiving ‘nothing’ and all others, if statistical significance at around $p = 0.10$ is accepted. The two comparisons of nothing with attention placebo that just miss 0.10 significance (depression and distress) can be explained by very low *n* in the attention placebo subsets (two in each - Table 4-11). The one comparison outside of a ‘nothing’ column that reaches significance (depression, attention placebo *v.* treatment as usual, Q statistic $p = 0.000$) is unexplained and again may be due to the few studies in the attention placebo subset (two).

The conclusion is that studies that employed a no treatment control group design generally elicited distinctly higher effect sizes than others, and that there is no statistically significant distinction between the other designs. This finding is theoretically readily understandable, given that the three more-than-nothing control types could, in practical terms, involve similar, and a mixture of, therapeutic components, whereas the ‘nothing’ control has just that – nothing of the nature of the therapy provided.

Although the pattern was not perfect in every cell of the pairwise comparison, the existence of the general pattern in the structure of the dataset is certainly substantial and has the potential to confound the results of later substantive analyses. It was therefore considered a structural variable that had to be accounted for in those analyses, and it was decided that this would be achieved by dichotomous display of results around its two levels or by exclusion of the lower scoring level, i.e. the group of ‘treated control’ studies (treatment as usual, attention placebo, and treatment element placebo studies). For the purposes of interpretation, results that used untreated control comparisons would provide a sense of the absolute magnitude of the effect of a therapy, whereas those that used treated controls would provide a sense of what a therapy could achieve

incrementally over conditions that could be provided as part of usual care or without specialised input.

Blindness to condition

Participants

Participant blindness to whether they are in the treatment or control condition is difficult or impossible to achieve in many psycho-oncological studies where it is obvious that the therapy is not part of a normal medical regime. It is likely that this is why the majority of studies did not report on the matter and were coded so. Notwithstanding, blinding and reporting on blinding have been features for which study designs have been quality-graded in literature syntheses in the past. It was therefore important to investigate whether there was any empirical justification for downgrading or excluding studies that did not keep participants blind. An initial investigation leaving ‘not reported’ as a distinct category (Table 4-13) was likely to be complicated by this limitation, and so another analysis was run (Table 4-14) on the assumption that studies that did not report on the topic were not patient blind, and removing the one study for which the issue was not applicable.

Table 4-13. Patient blindness, all categories

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Blind	-0.02 (10)	-0.23 – 0.18	0.00 (13)	-0.15 – 0.14	0.14 (9)	-0.06 – 0.33
Not blind	0.42 (23)	0.26 – 0.58	0.36 (24)	0.19 – 0.54	0.24 (21)	0.07 - 0.40
N/A	-	-	-	-	0.67 (1)	-0.05 – 1.39
N/R	0.18 (50)	0.05 – 0.30	0.17 (33)	0.05 – 0.29	0.14 (42)	0.03 – 0.26
Q stat p	0.003		0.007		0.424	

N/A = not applicable; N/R = not reported; Q stat p = Q statistic *p*. Other notes as per Table 4-2.

Table 4-14. Patient blindness, where 'not reported' merged with 'not blind'

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Blind	-0.02 (10)	-0.23 – 0.18	0.00 (13)	-0.15 – 0.14	0.14 (9)	-0.06 – 0.33
N/R plus Not blind	0.26 (73)	0.15 – 0.36	0.25 (57)	0.15 – 0.36	0.18 (63)	0.08 – 0.27
Q stat p	0.018		0.005		0.717	

Notes as per Table 4-13.

The heterogeneity of anxiety and distress effect sizes and the homogeneity of distress effect sizes found in the original analysis (Table 4-13, Q statistic *p*'s: anxiety, 0.003;

depression, 0.007, distress, 0.424) proved robust to the re-analysis (Table 4-14, anxiety, 0.018; depression, 0.005; distress, 0.717). The fact that this important apparent heterogeneity appeared in relation to anxiety and depression but not in relation to the distress outcome, and that so few studies reported patient blind designs, invited further investigation. Given that claims and counterclaims have been made about the practicality and ethics of blinding in psycho-oncology (e.g. Bredart, Cayrou, & Dolbeault, 2002, and Newell, Sanson-Fisher, & Savolainen, 2002), it was appropriate to go back to the studies themselves to try to understand whether the results may reflect confounding.

It was discovered that of the 16 studies claiming blindness, 11 had treated control conditions, meaning that a disproportionately high number of these studies (cf. approximately 50% overall, Table 4-11) fell into a category that yielded systematically lower effect sizes on that ground. (One of the five studies with an untreated control also appeared to have been coded 'blind' in error, meaning that the proportions were actually 4:11.) The reason for this is that patient blindness was usually achieved where the nature of the treatment tested allowed it to be slipped in to an already existing medical procedure by addition to doctor or nurse routines without it being apparent that anything new was being tried. Other than the eight 'treatment as usual' controls, there were also three studies that used a treatment element placebo. Logically, designs of this type produce the lowest effect sizes of all (Table 4-11).

Patient blind studies are therefore confounded by the nature of their control conditions and do not fairly represent the nature of most therapies in the domain, so it was decided that the heterogeneity found on this variable would be disregarded as spurious.

Therapists and assessors

Investigation into whether blind therapists or assessors in some studies but not in others caused heterogeneity could not be carried out because there was only one study that reported that the therapist was blind and only three that used assessors who were not participants (i.e. not self-report assessments). This would appear to confirm a comment by Kissane et al. (2003) in the context of a large cognitive-existential group study lasting 20 weeks, that the blinding of research assistants "is not methodologically possible in research of this type"(p.534).

Screening/floor effect

Because many studies in this domain recruit and treat patients who have not sought help for distress but merely on the basis of diagnosis and, perhaps, other medical variables, the risk is run that effect sizes may be limited by a floor on how much participants can actually lower their distress levels, i.e. they cannot improve much because they may already be quite psychologically well. On the other hand, as mentioned in the section on external validity, results might also be affected by the screening out of potential participants who have a history of psychological diagnosis or treatment.

Studies were therefore coded as follows: 1. ‘Screened out’ recruits with a psychiatric history or some particular present level of distress i.e. participants with history or current distress were excluded from participation; 2. ‘Screened in’ recruits for the same; 3. ‘Unscreened’ i.e. did not administer either screen. In the instance of a few studies that screened in for current distress and out for past psychiatric history, the code applied was ‘screened in’. Note that screening of this nature is to be distinguished from the administration of exclusion criteria designed simply to ensure the safety of participants or that they have the basic abilities needed to participate in the intervention. Commonly such exclusion criteria included cognitive impairment (dementia or intellectual disability) and current psychotic illness or suicide risk. Results are shown below (Table 4-15).

Table 4-15. Screening at recruitment

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Screened in	0.70 (6)	0.21 – 1.20	0.68 (8)	0.18 – 1.18	0.52 (7)	0.09 – 0.94
Screened out	0.34 (24)	0.17 – 0.51	0.24 (17)	0.06 – 0.42	0.32 (20)	0.15 – 0.48
Unscreened	0.11 (49)	0.00 – 0.23	0.12 (41)	0.02 – 0.22	0.05 (44)	-0.04 – 0.14
Not reported	0.39 (4)	0.08 – 0.69	0.46 (4)	0.15 – 0.77	0.66 (2)	0.29 – 1.02
Q statistic <i>p</i>	0.020		0.029		0.000	

Notes as per Table 4-2.

As all outcome constructs yielded highly heterogeneous significance levels (Q statistic *p*'s: anxiety, 0.020; depression, 0.029; distress, 0.000), further investigation was made of which pairs of categories were statistically distinct from each other (Table 4-16).

Table 4-16. Screening at recruitment, pairwise comparisons for heterogeneity

	Anxiety		Depression		Distress	
	Out	In	Out	In	Out	In
In	0.174		0.105		0.385	
Unscreened	0.033	0.023	0.250	0.031	0.006	0.034

Q statistic *p*'s are presented. No shading indicates significance at $p < 0.05$; light shading indicates p approximately 0.10. Studies that did not report clearly were excluded.

The pattern revealed was that unscreened studies were distinguished from both of the screened subsets, showing lower effect sizes in every case and statistical heterogeneity (Q statistic *p*'s < 0.05) in all but one pairwise comparison (depression, against screened out, 0.250, which had a smallish n of 8). The higher effect sizes consistently scored by studies that ‘screened in’ for distress did not significantly differ from the scores of those that ‘screened out’ for psychological history, but with two of three Q statistic *p*'s below 0.20 (anxiety, 0.174; depression, 0.105), it may be that with greater n in the ‘screened in’ category, significance would be attained.

‘Screened in’ studies.

In relation to the comparison between unscreened studies and those that ‘screened in’ for distress, this finding is consistent with the operation of a floor effect impacting the unscreened studies. By this is meant that a lack of proven distress at baseline has limited the scope for patients to improve with treatment. It also means that the inclusion of participants from the full spectrum of psychological neediness and capability in the unscreened subgroup has caused – in the averaging process - the ‘watering down’ or cancelling out of scores from individuals whose psychological distress did improve by those that did not or whose condition was better before therapy. In other words, if the majority of participants in a particular unscreened study were psychologically fairly well adjusted to start with, then they would likely benefit little from therapy and would drag down the average effect size of the whole sample.

Further, patients with psychological complexities who were in ‘unscreened’ studies may have found that a standardised protocol geared to the average patient was not helpful or even stressed them because their need was for individually tailored therapy regimes. During coding it was found that in fact all of the studies that ‘screened in’ did so based upon a particular level of baseline distress rather than upon psychological history. This could be expected to ensure that, in ‘screened in’ studies, participants were highly motivated for change, while the opportunity to participate in those studies was left open to patients with complexities in their psychological backgrounds (except in the few that screened both ways). A higher than normal proportion of the latter group would likely be represented, but this higher level of neediness could be anticipated and catered for in therapy design where researchers were anticipating recruits would be clinically distressed. Screened in studies might therefore be better designed to help patients that brought psychological complexity to therapy.

‘Screened out’ studies.

On the other hand, it was found that all of the studies in the ‘screened out’ subset excluded potential recruits because of mental health history rather than because of baseline distress. Except for the exclusion of more psychologically complex cases, these study samples would have been psychologically similar to the unscreened study samples. Given that complex cases can be expected to require extended and individualised programmes of psychotherapy in order to produce benefit, in the research context - where most studies deliver therapies that have been largely or fully predetermined in terms of content and duration (refer to Therapy characteristics results chapter, delivery) - removing such cases could be expected to lift the outcome mean. However, the mixture of distress levels in studies that ‘screened out’ patients with psychological complexity in their histories would also mean that a floor to improvement would continue to operate, leaving the final result between the low of the ‘unscreened’ group and the high of the ‘screened in’ group, as often happened.

External validity and screening out.

The screened out subset poses some theoretical complexity. In practice psycho-oncologists would not have the luxury of screening out patients with psychological

history, but, on the contrary, could expect that a disproportionate number of their clients would present with complexities of one kind or another. Employing a screening out procedure in that sense threatens the generalisability of results from such studies. But many therapies for cancer patients are not delivered by psychologists, and the nurses, social workers or counsellors / therapists who do deliver them likely screen out and refer more complex cases to psychologists or psychiatrists. In that sense, this screen poses no such threat.

Further, as has been mentioned, more complex cases would better suit and be more likely to receive longer and more tailored therapies in practice anyway, and so trial therapies should not be expected to transfer directly to them. Given that, the practice of screening complex cases out can be seen as producing more generalisable results than using an unscreened sample, if the psychological qualifications of the therapist better suit a manualised protocol.

Targeting and screening out.

Screening out also has the advantage of better defining the recipients for the purpose of designing and delivering particular therapies. This, again, should result in higher effects. Screening out can also generate a sample that is more homogeneous for scientific purposes in that the overall level of distress in the sample can be more confidently attributed to the cancer experience alone, largely excluding the likelihood that some of it is due to an interaction of that stressor with pre-existing psychological morbidities. This type of screening generates a 'purer' more cancer-specific experiment in terms of both the psychological state of the participants and the design and appropriateness of the therapy for treating cancer related distress.

Simultaneous screening.

Some studies applied both screens to potential recruits, producing a 'best of both worlds' sample in that their participants were currently distressed and therefore motivated for treatment, but were simultaneously unencumbered by psychological history. These studies set themselves up with ideal clients for yielding the largest possible effect sizes, and their results have contributed to the subset that has come out on top of the present analysis (the 'screened in' subset). Two outlier studies, Nezu et al. (2003) and Trask et al. (2003), applied both of these psychiatric screens before admitting recruits. (However, recall that for all of these calculations outlier scores were attenuated by windzoration.)

Decision.

The way that results fell around this variable presented a challenge for deciding how to cut the data for the substantive analyses. Statistical heterogeneity was found between those studies that screened in or out on the one hand, and those that did not screen at all on the other, but not between those that screened in and those that screened out. While homogeneous at 0.05 or 0.10, there is a consistent trend towards difference between the latter two subsets, which has a theoretical basis. However, the screened in and screened out subsets also have subtle theoretical commonality - in both cases the nature of the

sample in terms of the level and / or carcinogenic origins of psychological distress is better (if not fully) defined and visible, allowing better targeting of therapy and less flooring of improvements.

These commonalities and differences suggest the lumping together of the two screened subsets for the purpose of some analyses, and their separation for others. Simply being able to reduce the impact of flooring will allow effects that require a sensitive analysis to display variation and, perhaps, statistical heterogeneity. This empirical argument is a strong one in a meta-analysis that aims to expose moderators of effectiveness. Furthermore, ignoring the heterogeneity between the screened groups on the one hand and the unscreened group on the other is not attractive as it would result in the ‘drowning’ of the screened effect sizes in the much higher frequencies and very low effect sizes produced by the unscreened group. The consequences would be both greatly lowered mean effect size estimates and misleading effect size spikes when it happened that a greater proportion of screened studies contributed to a particular effect size – interpretation would be confounded. For both reasons, the outcomes would lose meaning and worth.

The strategy taken into the substantive analyses was therefore to divide studies on the basis of whether they screened recruits (either way) or not, in accordance with the empirical heterogeneity found and supported by the common theoretical thread of tailoring to distress type and level that has been proposed. The theoretical and empirical reservations in combining the two screened categories was kept in mind, however, and in analyses of therapy type where these issues were of particular significance, results were broken out by particular screening type to assure that they had the most meaning possible.

Treatment replicability/standardisation and fidelity

Studies were coded on the degree to which the therapy protocol was manualised, recorded, or described sufficiently for replication (fully; in part; not at all) and also whether therapy fidelity was monitored, and if so, whether a method was specified.

This was another instance of six analyses producing one heterogeneous result (replicability, distress) which was not considered important because it was isolated and also ran against the trend shown by the other outcomes. Details are in the results appendix.

Internal validity summary

In order to produce an empirical basis for assessing the quality of study design, the dataset was tested for statistical heterogeneity on a large range of possible threats to internal validity. By this process, it was found that several variables comprised statistically distinct levels. After further statistical investigations and taking into account theoretical and practical considerations, two of these variables were considered of sufficient importance to demand consideration in the conduct of later substantive moderator analyses: the nature of the control condition and the screening of study

recruits. Notably, the method used to allocate participants to conditions and whether or not participants were blinded to condition were variables not considered of such importance. This implies that earlier reviews have excluded or downgraded studies on these bases without need, impoverishing the datasets from which their conclusions were drawn.

Dealing with the design confounds

It was important that the differential impacts of the levels of the two variables with confounding potential be exposed in substantive analyses. One level of each could be excluded from the dataset to ensure that results were not confounded, but this was not necessary because no level signified poor quality design as such, but simply a different nature of design. Further, it was a priority to retain as much data as possible in the dataset.

It was decided to account for the potential confounds by aligning them with the axes of a 2 x 2 matrix, and draping all analyses over this frame (Figure 4-7). Some substantive analyses would use data from all of the quadrants of the matrix – with results each displayed separately – others, only from the two untreated control quadrants because of their greater variance, and some only from the single most sensitive quadrant. Which quadrants were used would depend upon the purpose of the analysis and the amount of data available: where there was a shortage of data, or they were thinly spread due to the number of levels of a given substantive variable, more quadrants would be used in order to verify any apparent pattern.

Figure 4-7. Study design confound matrix

	Untreated control	Treated control
Screened	Greatest variance – most sensitive quadrant	Medium sensitivity
Unscreened	Medium sensitivity	Least sensitive quadrant

5. RESULTS: MAIN EFFECTS

Having established heterogeneity on just two validity threats, both internal, analyses could proceed to the computation of main effect sizes. The focus of this meta-analysis was on moderators but it was necessary to investigate main effects in order to locate the study in the literature. Effect sizes that combine data from all therapy types on all demographics etc. for a given outcome are theoretically fairly meaningless in such a varied therapy field as psycho-oncology. However, they have been the main indices reported in past meta-analyses due to the newness of the field and the complexities of selecting and applying moderators. The present results will be presented by outcome. Earlier findings are discussed and tabulated in the Introduction chapter (Table 1-2) and will be referred to again at the end of this chapter.

The tabulating of results in this section is organised as follows: First a main effect (i.e. by psychological outcome) calculated without regard to the structural confounds, at early (i.e. usually the first assessment after intervention – refer Analysis) and late times (i.e. usually at least six months post intervention), and using a random effects computation model is tabulated. In the lower portion of the same table, a 2 x 2 matrix of effect size data accounts for the structure of the dataset around the two design variables, nature of control condition and whether recruits were screened. These results were computed using a mixed effects model (refer Analysis chapter for rationale). Tacked on top of the matrix but beneath the main effect index (in italics) is a breakout of the screened study results into three levels: screened out, screened in, and - a subset removed from the latter group - screened in and out. Although the screening levels did not differ significantly from each other in preliminary analyses, these effect sizes are included because they are of clinical interest. The Q statistic p for the breakout (italicised and bracketed) takes into account comparison with unscreened studies. The breakout is added in relation to untreated controls only because they provide the greater variance.

Note that because of the importance of the two structural confound variables, studies that did not report on either matter (n ranged between two and four over outcomes) were excluded from all mixed effects analyses that were based on the confound matrix both here and in later chapters, but were included for calculating the main effects at the top of the tables in this chapter. This accounts for any small difference in n. Note also that all calculations in this and following chapters exclude data where there was a known statistically significant difference between the treatment and the control group on the outcome variable at baseline and no statistical adjustment was made for it.

Anxiety

Table 5-1 sets out the main effect and breakout results for anxiety. The main effect at early times is small but statistically significant ($g = 0.23, p < 0.05$) and is retained (0.22, $p < 0.05$) through to late times. The main effect Q statistic p 's at both early and late times (0.000 and 0.001, respectively) show the great heterogeneity to be expected given

the two design moderators already found and the likelihood that substantive moderators will also be producing systematic variance. Screened studies produce effect sizes much higher than unscreened at early times, for both untreated and treated control comparisons (0.53 and 0.35 respectively, both $p < 0.05$), fading to about half magnitude at late times but from small n (3 for both). Unscreened effect sizes begin poorly but tend to remain static or even increase over time (untreated control: 0.33, $p < 0.05$ at early times to 0.28, $p = 0.114$, late; treated: -0.04, n.s. to 0.17, $p < 0.05$). Perhaps the effect of therapy on unscreened patients is to give them a small advantage that shows more over time than initially. Note also that at early times a trend favours screening both ways over screening in, over screening out, over no screening at all.

Note that a feel for the portion of the variance explained by the nature of the control condition (treated or untreated) is provided by the Q statistic p 's that remain once the structural breakout is performed. In this case, at late times strong homogeneity is shown at late times (Q statistic p 's of 0.857 and 0.850), meaning that a great deal of the systematic variance shown in the Q statistic p for the main effect at late times (0.001) is explained by the nature of the control condition. Where more heterogeneity remains, as with the early times results here (Q statistic p 's of 0.195 and 0.014), the implication is that status on the screening moderator is playing a greater role in the variance of the particular data sample. Of course, some other unknown moderating effect may also be at play. If there was a more even spread of n , this might suggest that screening plays a great role at early times but its influence fades by late times. However, there are insufficient screened data at late times to say.

The main effect of psycho-oncological intervention on anxiety is small but durable according to this statistical synthesis, and the impact of the two moderating design variables is readily apparent at early times.

Table 5-1. Anxiety, main and breakout effects

		Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
Main effect		0.23 (83)**	0.13 – 0.32	0.22 (30)**	0.09 – 0.35
Q statistic p		0.000**		0.001**	
Untreated control	<i>Screened in and out</i>	1.34 (2)**	0.21 – 2.46	-	-
	<i>Screened in</i>	0.73 (3)**	0.26 – 1.20	0.52 (1)**	0.18 – 0.87
	<i>Screened out</i>	0.37 (12)**	0.12 – 0.61	0.02 (2)	-0.35 – 0.38
	Screened (combined)	0.53 (17)**	0.29 – 0.76	0.24 (3)	-0.14 – 0.61
	Unscreened	0.33 (22)**	0.13 – 0.52	0.28 (7)	-0.07 – 0.63
	Combined	0.41 (39)**	0.26 – 0.56	0.26 (10)**	0.01 – 0.52
	Q statistic p	0.195 (0.163)		0.857 (0.143)	

Treated control	Screened	0.35 (11)**	0.07 – 0.63	0.13 (3)	-0.27 – 0.53
	Unscreened	-0.04 (28)	-0.18 – 0.10	0.17 (17)**	0.05 – 0.29
	Combined	0.04 (39)	-0.08 – 0.17	0.17 (20)**	0.05 – 0.28
	Q statistic <i>p</i>	0.014**		0.850	

ES = Hedges *g* effect size point estimate; *n* = number of studies in subset; CI's = confidence intervals. * statistically significant at $p < 0.10$; ** statistically significant at $p < 0.05$; two tailed.

Italicised figures are a breakout of screened study results, with 'screened in and out' data having been extracted from 'screened in' data. The Q statistic *p* for the breakout (italicised and bracketed) takes into account comparison with unscreened studies.

Depression

Table 5-2 shows that the main effect for depression at early times is small though statistically significant (0.22, $p < 0.05$), but fades to negligible magnitude at late times (0.11, $p < 0.05$). As for anxiety, near perfect heterogeneity is produced at both early and late times (Q statistic *p*'s 0.000 and 0.016 respectively), but while some of this is explained by the nature of the control condition, a good deal is left to be explained by screening (the main four matrix Q statistic *p*'s range from 0.175 to 0.599). Where the nature of screening is broken out (i.e. for untreated controls) Q statistic *p*'s reduce to near 0.10 statistical significance (early times, 0.093; late, 0.017) indicating that screening is a major source of variance for depression outcome.

In the untreated control rows, the screened studies produce a moderately strong effect at early times (0.42, $p < 0.05$) which is retained but with loss of considerable *n* and, probably as a consequence, statistical significance at late times (0.54, n.s., $n = 2$). The proportion of studies screening in (or in and out) for distress versus those that screened out for complexity has risen from what it was for anxiety (8/17 for untreated controls at early times, compared with 5/17). In the next chapter it will be seen that these are largely CBT studies which performed well for patients suffering distress at baseline.

The very low score for untreated control screened out studies at early times (0.12, n.s.) together with both unscreened scores (untreated control, 0.25, $p < 0.05$; treated, 0.08, n.s.) suggests that with no proven baseline distress, there is little point in intervention for depression. The only effect that remains of value at late times is for screened, untreated controls (0.54, n.s. but *n* of only 2) which (as will be seen) is produced by CBT studies.

The depression main effect is small and fades to negligible over six months. However, the latter finding may be due to the loss of *n* in distressed categories (screened in, and screened in and out) as this outcome shows particular sensitivity to baseline distress.

Table 5-2. Depression, main and breakout effects

	Early times		Late times	
	ES (<i>n</i>)	95% CI's	ES (<i>n</i>)	95% CI's

Main effect		0.22 (70)**	0.12 – 0.31	0.11 (20)**	0.02 – 0.20
Q statistic <i>p</i>		0.000**		0.016**	
Untreated control	<i>Screened in and out</i>	1.07 (4)**	0.30 – 1.85	-	-
	<i>Screened in</i>	0.51 (4)**	0.05 – 0.96	0.16 (1)	-0.18 – 0.49
	<i>Screened out</i>	0.12 (9)	-0.13 – 0.38	1.04 (1)**	0.35 – 1.72
	Screened (combined)	0.42 (17)**	0.13 – 0.71	0.54 (2)	-0.31 – 1.40
	Unscreened	0.25 (12)**	0.04 – 0.46	0.02 (4)	-0.14 – 0.19
	Combined	0.31 (29)**	0.14 – 0.48	0.04 (6)	-0.12 – 0.20
	Q statistic <i>p</i>	0.353 (0.093*)		0.241 (0.017**)	
Treated control	Screened	0.28 (7)**	0.01 – 0.54	0.00 (1)	-0.38 – 0.38
	Unscreened	0.08 (29)	-0.03 – 0.19	0.11 (13)**	0.00 – 0.21
	Combined	0.11 (36)**	0.00 – 0.21	0.10 (14)**	0.00 – 0.20
	Q statistic <i>p</i>	0.175		0.599	

Notes as for Table 5-1

Distress

The main effect for distress as an outcome is shown in Table 5-3 and is small at early times (0.18, $p < 0.05$) dropping to negligible after six months (0.11, $p < 0.05$). As for anxiety and depression, both of these results are underlain by heterogeneity (Q statistic p 's both 0.000). A good deal of this heterogeneity is explained by screening, if superimposed on a division by control type (early times untreated control Q statistic $p = 0.015$; treated, 0.070; late times untreated control screening breakout, 0.143) but at late times there is a sharp fall in screened n for both control types, rendering comparison across screening categories unreliable.

Twelve of 18 screened untreated control studies at early times screened out for history rather than in for distress or both. As seen previously, screened out studies scored considerably lower than those that recruited distressed patients, which in turn scored lower than those that screened in and out, but all of these scored higher than studies that were unscreened. The heterogeneity indicated by the breakout Q statistic p of 0.037 may, judging by confidence interval overlaps, divide the distressed patient studies from the unscreened ones, indicating the importance of baseline distress to intervention effectiveness once again. At late times the untreated screened is too low to sustain comment. Unscreened n drops by half at late times and effect size slips from low (0.19, $p < 0.05$), to negligible (0.11) and loses significance.

For treated controls no effect is evident overall (0.03, n.s.) and the only effect even approaching small magnitude is for screened patients at early times (0.17, $p < 0.10$).

This means that generally unscreened patients are receiving no benefit beyond they could expect from usual treatment or placebo for this outcome.

Screening is seen to be important to the distress outcome. The main effect is small at early times, fading to negligible over time.

Table 5-3. Distress, main and breakout effects

		Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
Main effect		0.18 (73)**	0.09 – 0.26	0.11 (26)*	-0.01 – 0.24
Q statistic p		0.000**		0.000**	
Untreated control	<i>Screened in and out</i>	1.08 (2)**	0.21 – 1.95	-	-
	<i>Screened in</i>	0.60 (4)**	0.26 – 0.95	0.52 (1)**	0.18 – 0.87
	<i>Screened out</i>	0.39 (12)**	0.14 – 0.65	0.02 (2)	-0.35 – 0.38
	Screened (combined)	0.53 (18)**	0.30 – 0.76	0.24 (2)	-0.06 – 0.54
	Unscreened	0.19 (17)**	0.03 – 0.35	0.11 (9)	-0.05 – 0.27
	Combined	0.30 (35)**	0.17 – 0.43	0.14 (11)*	0.00 – 0.28
	Q statistic p	0.015** (0.037**)		0.454 (0.143)	
Treated control	Screened	0.17 (8)*	0.00 – 0.34	-0.04 (1)	-0.42 – 0.35
	Unscreened	0.00 (25)	-0.09 – 0.08	0.07 (13)	-0.06 – 0.21
	Combined	0.03 (33)	-0.05 – 0.10	0.06 (14)	-0.07 – 0.19
	Q statistic p	0.070*		0.598	

Notes as for Table 5-1

Placement in the literature

Main effects over the three outcome variables are all small (around 0.20) at early times, and for depression and distress, fall away to negligible (around 0.10) by late times, though this may well be due to a disproportionate loss of n from the high performing categories that screened in or both ways. The very broad range in the breakout results (from 0.00 to 0.54) makes clear the major influence of the structural confounds (screening and nature of control). The two categories that screened in for baseline distress lifted results considerably at early times (results ranged from 0.51 to 1.34 over the three outcomes) but there were insufficient data available to sustain comment at late times.

Bearing these points in mind, and that usually such structural confounds have not been used to frame main effect analyses in the past, results from previous meta-analyses can be considered.

Most of the earlier meta-analyses used data from the first assessment point offered by each study, which is a similar approach to that taken in the present study for ‘early times’ calculations. The main effect sizes generated by the present study are generally a good deal lower, however. The earlier main effects ranged from 0.16 to 0.94 (refer Introduction, Table 1-2) but were mostly of moderate magnitude, whereas present main effects were small. These differences may reflect different ratios of screened / unscreened studies and studies with treated / untreated controls. Alternatively it may reflect the maturing of the field - having moved on from an initial stage typified by the publication of a few exciting results - or the size of the body of research now at large, the thoroughness of the search used in the present study, the inclusion of research designs that have been here empirically justified but were excluded by others on method quality grounds, different outcome constructs, or, likely, a combination of all of these. With many such advantages over earlier works, the present results should produce a more realistic picture of the state of the art.

The meta-analysis by Sheard and Maguire (1999) is closest in design to the present one, using anxiety and depression as outcome constructs and breaking out studies screened for risk of or actual psychological distress at baseline. Although their ‘screened’ category is not directly comparable, it is close to the present distress screened categories (screened in or both ways). Results show quite similar trends: main effect sizes for both outcomes are small (anxiety, 0.36; depression, 0.20), and a substantial difference is shown between the yields of screened (anxiety, 0.85; depression, 0.94) and unscreened (anxiety, 0.33; depression, 0.16) studies. Such small differences in magnitude as there are may be explained by the study quality design criteria applied by Sheard and Maguire, by more data bringing stability and conservatism to the present results, or by differing proportions of studies in structural confound categories.

6. RESULTS: THERAPY CHARACTERISTICS

Substantive moderator analyses begin with an exploration of various aspects of therapy type and delivery. First, main effects are briefly reported and then investigation is made concerning each of the seven broad types; the more frequent specific components that comprise therapy protocols; therapy trajectory; and variables relating to the means of therapy delivery including mode, dose and therapist variables. Where there is limited evidence of moderation – which is the norm in this chapter - summary tables are used to describe the field and detailed explication of data is relegated to **Appendix O**.

Therapy type

The seven therapy types categorised (refer Analysis chapter) included four main categories delivered professionally - professional education / information (education or 'e'); relaxation focused behavioural therapies (relaxation or 'r'); (other) cognitive-behavioural therapies (CBT or 'c'); and expressive-supportive or non-directive professional counselling / psychotherapy (expressive-support or 's') – and three smaller categories: non-professionally led support or counselling (non-professional or 'n'); indirect interventions (those that deliver benefit to the patient by intervening with someone else – a medical professional or significant other)(indirect or 'i'); and a residual category for 'other' therapies.

Therapy categorisation in this field is challenging not only because of the wide range of therapy types but also because of the frequent packaging of theoretically different elements together. In the present study, therapy protocols were categorised by two strategies - 'inclusively' and 'exclusively' – and both are drawn upon in analyses in this chapter.

The inclusive (or 'therapy type') approach categorises each study by *any and all* of the theoretical elements that comprise it. This means that a given study would be coded under more than one inclusively categorised therapy type, if its therapy protocol comprised more than one. For example, if a study delivered a problem solving therapy in a professionally run group structure in which expressive-support was expected to provide part of the therapeutic benefit, then it was coded for both CBT and expressive-supportive therapy and effect size estimates and data from it contributed to *both* of these inclusively defined therapy type analyses. Because this categorisation is statistically dependent (one study can contribute to more than one category), it does not allow statistical comparison across therapy types. Also, obviously, results are influenced by the contribution of studies using protocols that include components from therapy types other than the one of interest. However, this categorisation allows the accumulation of the most n, and with large n the variance contributed by other therapy types becomes less important. The inclusive categorisation of therapy type is therefore the most used throughout this study.

The exclusive (or 'therapy combination') approach categorises each study according to the particular combination of therapies in the protocol, meaning that each study falls into *only one* mutually exclusive category. For example, the problem solving group would be coded 'cs', i.e. 'c' for CBT and 's' for expressive-supportive therapy. Other studies might be coded with just one letter if there was only one therapy type that contributed to the protocol, but frequently studies were coded with three or four letters, signifying their particular eclectic mix of therapy types. This approach dramatically reduced the n available for any specific therapy combination, but by this means the therapy combinations that were most common could be identified and analysed as specific protocols. Also, this categorisation allowed the opportunity to examine results from protocols that used only the one type of therapy ('pure' protocols) without the variance introduced by protocols that also involved other types. For example, it allowed examination of results from studies that provided only CBT (studies coded 'c'). Such exclusively categorised results could be compared with the larger n inclusively categorised results of the same therapy type for a more rounded picture.

Main effects

Effect size estimates for the six therapy types (inclusive definition, and excluding the residual 'other' category) - education, relaxation, CBT, expressive-support, non professional and indirect therapies – are tabulated by psychological outcome in Table 6-1. These results are set out in order to provide data for comparison with earlier literature (refer Table 1-1 in the Introduction) and therefore take no cognisance of the two study design moderators identified in the Preliminary analyses results.

Main effects are generally small and significant at $p < 0.05$ at early times, most commonly waning over time. An interesting exception is the more durable, even slightly increasing, impact of education on anxiety and distress (anxiety, 0.13 early to 0.17 late; distress, 0.03 to 0.05). All the late times depression and distress effect sizes are of negligible magnitude and all but one loose statistical significance. It may be that the distress early times outcomes for education (very low at 0.03, n.s.) and relaxation (does not reach statistical significance at 0.22) are a little out of step with those for anxiety and depression because this outcome construct is relatively loosely defined.

A few particular results are noteworthy: Indirect therapies generate the largest effect sizes (anxiety, 0.50, $p < 0.05$; depression 0.64, $p < 0.05$; distress, 0.38, n.s., all early times) but n is small (3 in each case). Of the four conventional therapy types, relaxation produces the highest main effect at 0.31 for anxiety, and CBT follows with 0.28 for depression, both at early times and significant at $p < 0.05$. Non-professional results are more often negative than positive and none are sufficiently large and stable to reach statistical significance, but n is patchy and small (2 – 7).

Anxiety results are consistently large and stable enough to be statistically significant, and mostly at $p < 0.05$. It is striking that anxiety displays such consistent results across early and late times and across all therapy types except those that are non-professionally run, and that such consistency is not displayed in relation to depression or distress.

Table 6-1. Therapy types, main effects

	Early times					
	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Educ	0.13 (40)*	-0.01 – 0.28	0.15 (29)**	0.03 – 0.26	0.03 (33)	-0.07 – 0.13
Relax	0.31 (38)**	0.18 – 0.44	0.16 (29)**	0.03 – 0.29	0.22 (33)	0.7 – 0.37
CBT	0.18 (31)**	0.09 – 0.28	0.28 (30)**	0.13 – 0.43	0.19 (29)**	0.05 – 0.34
Exp-sup	0.23 (25)**	0.04 – 0.42	0.13 (23)	-0.03 – 0.29	0.18 (23)**	0.02 – 0.33
Non-prof	0.39 (4)	-0.37 – 1.16	0.27 (4)	-0.16 – 0.69	-0.03 (7)	-0.19 – 0.13
Indirect	0.50 (3)**	0.09 – 0.91	0.64 (3)**	0.21 – 1.06	0.38 (3)	-.07 – 0.83

ES = Hedges g effect size point estimate; n = number of studies in subset; CI's = confidence intervals. * = statistically significant at $p < 0.10$; ** = statistically significant at $p < 0.05$; two tailed. Shading indicates statistical significance at $p < 0.10$ is not met. Where there was uncertainty as to whether a study fitted a category, the study was excluded.

Early and late times are assessment time points defined in the Analysis chapter.

Educ = education; Relax = relaxation; CBT = cognitive behavioural therapy; Exp-sup = expressive-supportive therapy; Non-prof = non-professional counselling or support.

Table 6-1 continued

	Late times					
	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Educ	0.17 (13)*	-0.00 – 0.35	0.11 (9)*	-0.01 – 0.24	0.05 (15)	-0.07 – 0.18
Relax	0.16 (17)**	0.03 – 0.30	0.10 (10)	-0.03 – 0.23	0.06 (11)	-0.03 – 0.15
CBT	0.18 (11)**	0.02 – 0.34	0.10 (10)	-0.04 – 0.24	0.05 (9)	-0.07 – 0.17
Exp-sup	0.20 (11)**	0.05 – 0.35	0.08 (9)	-0.04 – 0.21	0.07 (13)	-0.09 – 0.24
Non-prof	-0.19 (3)	-0.46 – 0.08	-0.07 (2)	-0.40 – 0.25	-0.11 (5)	-0.39 – 0.12
Indirect	-	-	-	-	-	-

In sum, this glance at the impact of different types of therapy through a table of main effects raises interest in anxiety as a more general and enduring outcome construct and indirect therapies as particularly effective, at least at early times. It also highlights the poor performance of non-professional therapies, and the strong performance of indirect therapies.

Comparison with earlier works

Where do these therapy type findings fit against the literature? Direct comparison is not possible since categorisations vary, but some sense can be gain by reference to the summary table of treatment type results from earlier meta-analyses in the Introduction (Table 1-1).

Generally earlier effect sizes are higher than the overall early times outcomes found in the present study. Rehse and Pukrop (2003), which found Cohen's d 's ranging between 0.48 and 0.95 used a broader outcome ('emotional adjustment') and included German studies not included here. The low distress result produced by Tatro and Montgomery (2006) in relation to CBT for breast cancer patients ($d = 0.13$) may be the result of moderation by socio- and medical demographics, in that later results reported here show that women and breast cancer are the lowest scoring categories amongst these variables. The extremely high anxiety and depression scores in Osborn, Demoncada, and Feuerstein (2006) in relation to CBT for survivors (g 's of 1.99 and 1.20) may similarly be a result of patient characteristics – perhaps from the use of mixed cancer type samples. Other scores may result from particular frequencies of studies falling into one or other of the study design moderator categories recognised by the present study. The present meta-analysis, being the most recent, has had access to many more studies than any previous one, and most of those tabled (Devine & Westlake, 1995, being an exception, but it is more than a decade old) used more or less arbitrary methods for culling studies for research method quality. The present data should therefore be the most reliable.

A particular point to note is that where earlier studies have compared educational interventions to other types, education has done well (Rehse & Pukrop, 2003; Sheard & Maguire, 1999) whereas in the present study it does not. That distinction may be partially explained by the passing of time, with helpful information becoming more generally available to patients in recent years. More regarding education in the next section.

The most important point that arises out of the present study is that psycho-oncological intervention studies do not produce a body of results that should be regarded as homogeneous. The main effects table presented above is really of dubious value because it masks the impact of important moderators, creating a distorted impression. The analyses that follow take account of the two design feature moderators found in preliminary analyses, and identify more moderators of effect size.

Therapy types and combinations

The seven general therapy types coded were approached one by one, using the two definitions – therapy type (inclusive definition) and therapy combination (exclusive definition) – and several forms of analysis which are described below. However, few moderators were found in this set of analyses, so most of the detail was relegated to **Appendix O**. A summary table is provided instead (below, Table 6-2), with some

more interesting results reported in closer detail. After this section on therapy types, related sections follow regarding therapy trajectories and specific therapy components.

The different types of analysis that contributed to results reported in this section were as follows:

1. For the four main professional therapy types, inclusively defined, 2 x 2 matrices based on the study design confounds identified in preliminary analyses are used, i.e. whether or not potential study recruits were screened and the nature of the control condition. They include an overall effect size (mixed effects computational model) and sub-breakouts for screening type (screened out, screened in, and screened in and out) and combined distress outcome (screened in combined with screened in and out). It was theoretically important to be able to identify which of the screening categories was producing particular effects relative to different therapy types, so they were broken out further despite the statistical homogeneity of all screened studies regardless of screening strategy.
2. In separate analyses of therapy combinations (exclusive definition), only 'pure' protocols and those categories with an n of three for at least one outcome are presented. Because the number of exclusively defined categories is so great, spreading n so thin, analyses were not run in 2 x 2 matrix form, which would have reduced n in a given cell still further, but as a simple breakout by screening *or* nature of control. So, for each of these rows of data, the other confound is *not* taken into account, unlike the matrixed data. The same format for setting out results is used for non-professionally delivered therapies, also because of low n. While this approach means that any given result may be the product of an imbalanced frequency of the other structural confound, these results can still contribute to an overall picture of therapy type dynamics. Because therapy combinations (e.g. 'cs') can relate to more than one therapy type, data regarding multiple-type combinations appears more than once, under each of the individual types as they are dealt with.
3. Results for the low n and internally various indirect and other categories are tabulated in descriptive form as synthesis is not appropriate.

Note that whereas a random effects computation model was used for the main effects already reported and for 'overall' effect sizes in those that follow, the mixed effects model was used to make calculations involving the study design moderators (screening and nature of control condition) because inclusion of these moderators accounted for some systematic variation. The mixed effects model tends to slightly reduce overall effect sizes with this dataset, hence any discrepancy. Another source of minor discrepancy arises from the inclusion of studies that did not report on either of the study design moderators in the main effects analysis and in 'overall' effect sizes, but their exclusion from the substantive moderator analyses which rely upon the 2 x 2 confound matrix.

Descriptive summary

Table 6-2 (overleaf) provides a description of the four main professional therapy types results. The table allows comparison of different types of therapy for effect, which allows consideration of how they may complement each other in patient care. It also shows where data are sparse and non-existent. The table breaks out by baseline distress (combining studies that screened in for distress with those that screened both ways, except where marked *) and for psychological non-complexity. It is based primarily on inclusively categorised data, but allowance is made in the one case where exclusively categorised data produced quite a different result (refer note ‡). Early times data are presented at the top of each cell, with late times below. Notes run onto the following page.

Education

The table (Table 6-2) shows that anxiety is the main outcome against which educational interventions produce therapeutic effect. The part that anxiety plays underlying depression and general distress may explain the small impact seen against those outcomes also. There were relatively little data comparing with untreated controls, and unfortunately no data were available for patients distressed at baseline. Because lower confidence intervals from most education studies fall below zero, it is apparent that some patients are made more distressed by receiving more information, and for a lesser number that effect may be enduring. However, there is also a result (untreated control, unscreened, distress) that suggests the possibility of some small increase in the therapeutic effect of education over time.

A strong 'pure' education result gave rise to speculation - in the more detailed explication relegated to the appendix - that patients from societies where information about cancer is not freely available can be greatly relieved from anxiety by the frank provision of information about the disease and about medical treatments and procedures. Generally poor results may be a consequence of most studies having been conducted in Anglo-western societies where, in recent years, this type of information has been communicated quite freely.

Table 6-2. Therapy types, summary

Patient group		Education / Information			Relaxation			CBT			Expressive-supportive		
		Anx	Dep	Dis	Anx	Dep	Dis	Anx	Dep	Dis	Anx	Dep	Dis
Untreated control	DisB	-	-	-	✓ ?	✓ ?	✓ ?	✓ *	✓ *	✓ *	✓ ?	✓ ?	✓ ?
		-	-	-	-	-	-	✓ ?	✓ ?	✓ ?	-	-	-
	Non comp	✓	0 ?	0	✓	0	✓	✓ ?	✓	✓	✓ ?	0	0
		-	-	0 ?	0 ?	-	-	0 ?	✓ ?	-	0 ?	-	0
	Unsc	✓ ‡	✓	0	✓	✓ ?	✓ ?	0	0	0	✓	✓ ?	✓ ?
		0 ?	-	✓	✓	0	0	✓ ?	0	0	✓	✓ ?	✓ ?
Treated control	DisB	-	-	-	✓ ?	0 ?	0 ?	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-
	Non comp	✓	✓ ?	✓	✓ ? †	✓ ?	✓	✓	✓ ?	✓ ?	✓	✓ ?	-
		0 ?	0 ?	0 ?	✓ ? †	-	-	✓ ?	-	-	✓ ?	-	-
	Unsc	0	0	0	0 †	0	0	0	✓	0	0	0	0
		✓	0	0	✓ †	0	0	0	0	0	✓	0	0

Results at the top of each cell relate to early times, and at the bottom to late times. Outcomes: Anx = anxiety; Dep = depression; Dis = distress. Patient groups: DisB = distressed at baseline, i.e. ‘screened in’ and ‘screened in and out’ categories combined; Non comp = psychologically non-complex cases, i.e. ‘screened out’ for psychological history.

Effect sizes: 0 = negative, zero, or negligible range (Hedges' $g < 0.15$); ✓ = small range ($g = 0.15 - 0.35$); ✓ = moderate range ($g = 0.36 - 0.65$); ✓ = large range effect size ($g > 0.66$); ? = result is larger than ‘negligible’ (‘0’) and is non-significant ($p < 0.10$) or is supported by n of only one or two.

* These results are based on even, or near even, small frequencies of studies that screened in and that screened in and out, yielding a mean effect size that is strong. However, the studies that did not simultaneously screen out produced only a moderate 0.53 ($p < 0.05$, $n = 2$) against anxiety, a small 0.30 ($p < 0.10$, $n = 3$) against depression, and a moderate 0.45 ($p < 0.05$, $n = 3$) against general distress.

† These results are influenced by a disproportionately high number of studies that combine relaxation with other therapies, probably resulting in the attenuation of early and the exaggeration of late times effects.

‡ Whilst this effect size is moderate, the ‘pure’ education result is strong.

Relaxation

Relaxation therapies are designed to treat anxiety and it is against this outcome that best results are achieved. Studies were few for patients distressed at baseline but offered

promise that strong effects might be available, including against depression, somewhat surprisingly. Nothing is known about how effects endure over time for that patient group. Interestingly, samples lacking psychologically complex backgrounds enjoyed moderate range early times results against anxiety and distress (nothing against depression), but again there was a lack of *n* to show how durable these results were at late times. In the trajectory section of this chapter, however, this patient group shows a propensity to hold or even improve effect sizes over all outcomes through to medium term (one to six months). For unscreened patients there was a benefit from relaxation in the moderate range at early times for anxiety, and a small effect hung on at late times. Benefits attenuated predictably for depression and distress, along with *n*. There was evidence of similar effect dynamics against treated controls, but patchy *n* and some confounding of therapy combinations leave uncertainty (refer note † in the table).

CBT

The dynamics for CBT are quite distinctive as early times effectiveness is clearly moderated by simultaneous screening for baseline distress and psychological history. These studies produce *g*'s exceeding 1.00 against each outcome and untreated controls (although for each of anxiety and distress *n* is only two). Patients who are simply screened in for distress but not out for complexity, produce the next highest effect sizes, which are in the small to moderate range. Unfortunately there are very little late times data for distressed categories (i.e. simultaneously screened or screened in for distress) but some indication that effects may be maintained to small or moderate magnitudes. On the table (Table 6-3) effect sizes for both types of studies that screen for distress are combined, and it can be seen that the effects from those that screened both ways are so strong as to lift the average effect sizes into the large range.

For non-complex patients the early benefit is small, but there is again some evidence of increasing effect over time, pointed to also in the trajectory results in relation to depression and general distress. For unscreened patients no benefit is seen at early times. Data are so sparse against treated controls that overall patterns cannot be discerned with any certainty, but there is again some evidence of a small and lasting benefit for non-complex patients and no benefit for unscreened patient groups, except in relation to depression where the early times effect is small but not lasting.

Because CBT results do evidence moderation, they deserve closer examination. However, a similar dynamic is shown across all three outcomes so only the results for depression – the outcome for which this therapy was originally designed - are reproduced below from **Appendix O**.

CBT detail: Depression.

CBT produces a comparatively strong early times main effect of 0.28 ($p < 0.05$, $n = 30$), but this falls with *n* to 0.10 (n.s., $n = 10$) at late times (Table 6-1). However, these effects differ dramatically from those yielded by samples with baseline distress.

The study design feature matrix (Table 6-3) produces homogeneity on the screening dichotomy for all quadrants, but heterogeneity when screening is broken out further in the two quadrants (both untreated controls) that have samples screened in for distress (early times, Q statistic $p = 0.092$; late, 0.027). Overall effects are negligible and non-significant for all quadrants except for early times untreated control, where the effect is still small (0.25, $p < 0.05$, $n = 17$).

The early times untreated control quadrant is of great interest because it presents results for three studies that screened in and four that screened both ways. Whilst the confidence intervals for these sub-groups do have an overlap of about 0.30, the mean effect sizes are very different: simultaneously screened samples yield a very high 1.07 ($p < 0.05$), while studies that merely screened in yield a surprisingly small 0.30 ($p < 0.10$). The upper confidence interval of the screened in sub-group does not reach the average effect size when its results are combined with the simultaneously screened sub-group. The only one of these studies that survives into late times is a screened in study, which yields a non-significant and disappointing 0.16. The screened in and out result is so strong that it produces statistical heterogeneity when combined with the screened in result and compared with screened out and unscreened results (Q statistic $p = 0.092$). If the screened in and out confidence intervals are compared with others, it can be seen that 0.05 alpha heterogeneity would be likely against screened out and unscreened studies.

Screened out and unscreened untreated control samples yield negligible and non-significant effect sizes at early times with very similar confidence intervals. The unscreened result remains flat at late times, but the screened out row produces an anomalous very high effect size from its one remaining study.

For treated controls, which have no screened in samples, the effect size for the one screened out study is the same as for the nine unscreened ones: of negligible magnitude but significant at $p < 0.10$. The unscreened result is all that survives through to late times, where it fades further to 0.05 ($n = 5$) and loses all statistical significance.

This set of results is consistent with the pattern displayed by CBT for anxiety and general distress outcomes, except that early times anxiety results for untreated controls tend higher for the screened in and out category (1.34, $n = 2$) and screened in category (0.53, $n = 2$) lifting the distress combined result to the point that the screening breakout shows formal heterogeneity (Q statistic $p = 0.035$).

Table 6-3. CBT as a therapy type, depression

Depression		Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
Untreated	<i>Screened in and out</i>	1.07 (4)**	0.30 – 1.85	-	-
	<i>Screened in</i>	0.30 (3)*	-0.04 – 0.64	0.16 (1)	-0.18 – 0.49
	<i>Distress combined</i>	0.75 (7)**	0.20 – 1.31	0.16 (1)	-0.18 – 0.49

	<i>Screened out</i>	<i>0.07 (6)</i>	<i>-0.25 – 0.38</i>	<i>1.04 (1)**</i>	<i>0.35 – 1.72</i>
	Screened (combined)	0.42 (13)**	0.07 – 0.77	0.54 (2)	-0.31 – 1.40
	Unscreened	0.13 (4)	-0.17 – 0.43	0.05 (3)	-0.19 – 0.28
	Overall effect	0.25 (17)**	0.03 – 0.48	0.08 (5)	-0.14 – 0.31
	Q statistic <i>p</i>	0.220 (0.092*)		0.273 (0.027**)	
Treated control	<i>Screened in and out</i>	-	-	-	-
	<i>Screened in</i>	-	-	-	-
	<i>Distress combined</i>	-	-	-	-
	<i>Screened out</i>	<i>0.16 (1)</i>	<i>-0.34 – 0.66</i>	-	-
	Screened (combined)	0.16 (1)	-0.34 – 0.66	-	-
	Unscreened	0.16 (9)*	-0.02 – 0.34	0.05 (5)	-0.10 – 0.19
	Overall effect	0.16 (10)*	-0.01 – 0.33	0.05 (5)	-0.10 – 0.19
	Q statistic <i>p</i>	0.999 (0.999)		1.000 (1.000)	

Italicised results (screened in and out, screened in, screened out) are breakouts of the 'screened (combined)' result. 'Distress combined' is a combination of the two categories that screened in for baseline distress (screened in and out, screened in).

Q statistic *p*'s relate to the screened *v.* unscreened comparison; those in brackets and italicised relate to comparison of distress combined *v.* screened out (i.e. studies that excluded patients with a history of psychological distress, also referred to as 'non-complex' samples) *v.* unscreened.

Where there was uncertainty as to whether or not a treatment group fitted a category (treatment type, screening, or nature of control condition) that group was excluded from computations. The occasional smaller summative *n* in a confound breakout will be noticed as a result. Relevant notes from Table 6-1 apply.

This set of results is consistent with the pattern displayed by CBT for anxiety and general distress outcomes, except that early times anxiety results for untreated controls tend higher for the screened in and out category (1.34, *n* = 2) and screened in category (0.53, *n* = 2) lifting the distress combined result to the point that the screening breakout shows formal heterogeneity (Q statistic *p* = 0.035).

In Table 6-4 Table 6-4 results for the three CBT protocol combinations are presented. By itself, 'c' (the sample of studies with therapy protocols comprising CBT alone) presents the strongest overall effect sizes of any 'combination' result in this chapter: 0.50 (*p* < 0.05, *n* = 11), early times; 0.32, (*p* = 0.148, *n* = 3), late times. The screened results, untreated controls, are responsible for these impressive overall effect sizes. Screened studies produce a high 0.71 (*p* < 0.05, *n* = 7) at early times. The late times figure looks promising at 0.54 (but *n.s.*, *n* = 2) but from the previous table it is known to be made up of an anomalously high screened out score, and a poor screened in score. The unscreened early times result is only 0.17 (*p* = 0.10 precisely, *n* = 4).

Rcs comprises n that is too light to sustain comment, but it is noted that rc produces a small and non-significant early times overall effect (0.17, n = 7) that is bolstered by a medium strength result for screened studies of 0.39 ($p < 0.05$, n = 4). Unfortunately there are no late times data for screened studies using the rc combination. By comparison, unscreened results are around zero (early, -0.12, n = 2; late, 0.04, n = 4). Untreated control studies yield a modest non-significant 0.26 (n = 6), and the other confound rows are below zero. At late times n is sparse.

Results for c and for rc both demonstrate moderate to strong early times results for screened patients, but there is little evidence regarding the durability of this effect.

Table 6-4. CBT as a therapy combination, depression

Therapy combination	Confound breakout	Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
c	Overall	0.50 (11)**	0.20 – 0.80	0.32 (3)	-0.11 – 0.76
	Overall <i>p</i>	0.001**		0.148	
	Screened	0.71 (7)**	0.20 – 1.21	0.54 (2)	-0.31 – 1.40
	Unscreened	0.17 (4)	-0.03 – 0.38	0.07 (1)	-0.28 – 0.42
	Untx control	0.60 (8)**	0.15 – 1.05	0.54 (2)	-0.31 – 1.40
	Tx control	0.15 (2)	-0.05 – 0.35	0.07 (1)	-0.28 – 0.42
rc	Overall	0.17 (7)	-0.20 – 0.55	0.04 (3)	-0.13 – 0.21
	Overall <i>p</i>	0.256		0.676	
	Screened	0.39 (4)**	-0.31 – 1.08	-	-
	Unscreened	-0.12 (2)	-0.31 – 0.07	0.04 (3)	-0.13 – 0.21
	Untx control	0.26 (6)	-0.20 – 0.72	0.05 (2)	-0.29 – 0.38
	Tx control	-0.14 (1)	-0.35 – 0.07	0.07 (1)	-0.15 – 0.28
rscs	Overall	0.10 (3)	-0.31 – 0.51	0.20 (2)	-0.46 – 0.87
	Overall <i>p</i>	0.636		0.549	
	Screened	-0.10 (1)	-0.55 – 0.35	-	-
	Unscreened	0.21 (2)	-0.47 – 0.88	0.20 (2)	-0.46 – 0.87
	Untx control	-0.10 (1)	-0.55 – 0.35	-	-
	Tx control	0.21 (2)	-0.47 – 0.88	0.20 (2)	-0.46 – 0.87

Only 'pure' combinations (e.g. 'e' or 's' rather than 'es') and those with a frequency of at least three in any one of the three outcomes are tabled. Relevant notes from Table 6-3 apply.

Therapy combination codes: r = relaxation focused cognitive-behavioural treatment;; c = cognitive-behavioural treatment; s = expressive-supportive therapy / non-directive counselling / psychotherapy ('s' for supportive).

In sum, these tables portray CBT as capable of very strong results against depression at early times for patients who were screened both ways, but, surprisingly only a small effect for those who were merely screened in for distress (though n was only three). Unfortunately no late times data were available for the high scoring group and only one study for the screened in group. Compared with treated controls, unscreened studies still produced a small effect size, possibly because the form of treatment received by controls did not convey any significant feature of this very specialised form of therapy. However, the main point illustrated by the results in Table 6-3 is the moderation by simultaneous screening.

Expressive-supportive therapies

Expressive-supportive therapies produced spectacular results (exceeding $g = 1.00$) against all three outcomes at early times for patients distressed at baseline (who were *not* also screened out for psychological history) but n was only three in each case and no late times follow-up data were available. In the trajectory section of this chapter, some single study results at medium term showed retention of some benefit for both anxiety and depression, but not for distress. The pattern of therapeutic benefit for non complex samples was inconsistent between untreated and treated categories, at some points producing moderate magnitude results, and at others, no benefit. Unscreened patients gained a small and durable benefit generally, with the effect increasing over time against anxiety, but this benefit was nullified when compared with treated controls, with the exception of the small lift in the effect against anxiety after six months. Clearly more evidence relating to distressed patients is needed, including follow-up data.

The tables detailing depression outcome results were selected for reproduction from the relegated results appendix, in order to illustrate the moderation effect by simultaneously screened studies, and the small effect attained by unscreened studies. This set of results is again fairly representative of those for this therapy type, and depression was the outcome chosen to allow comparison with the CBT results. Further discussion of that comparison will follow.

Expressive-support detail: Depression.

Main effects for expressive-support as a therapy type against depression are very poor and do not reach statistical significance at either early (0.13) or late times (0.08)(Table 6-1). In the Table 6-5 matrix, the early times overall result for untreated control studies is of small to moderate magnitude (0.30, $p < 0.10$) but there is only one study for untreated controls at late times. Both overall effects against treated controls are negligible and non-significant.

All quadrants of the Table 6-5 matrix are statistically homogeneous on the comparison of screened with unscreened rows except treated control, early times, where the peculiar phenomenon of a very high scoring screened out study crops up (1.21, $p < 0.05$, Q statistic $p < 0.05$ against unscreened, -0.01, $n = 13$) and counters the trend of results in the untreated control quadrant (screened out, -0.08, $n = 3$; unscreened, 0.27, $n = 3$, both *n.s.*). More importantly, heterogeneity is also found in the breakout of screening types,

untreated controls, early times (Q statistic $p < 0.05$), due to the very high effect size generated by the two screened in studies (1.03, $p < 0.05$). This result contrasts with the screened out score in the same quadrant already mentioned (-0.08), and a small to moderate effect size for unscreened studies (0.27, $p = 0.173$). It also contrasts with the smallish CBT score for screened in untreated controls (0.30, $p < 0.10$, $n = 3$, Table 6-3).

The treated control unscreened result is well supported by n at early times (13) which it retains fairly well into late times (8), so should be fairly reliable. It shows that expressive-supportive therapy adds no value to placebo or treatment as usual against depression for unscreened patients at early or late times.

These results are consistent with those produced against anxiety and general distress outcomes, except that in the latter case the early times unscreened result for untreated controls reached 0.40 ($n = 5$, n.s.). Study frequencies for all breakout rows in the untreated half of the tables are light: distressed patients produce strong results but there is no incremental value over treated controls for unscreened patients.

Table 6-5. Expressive-supportive as a therapy type, depression

Depression		Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
Untreated control	<i>Screened in and out</i>	-	-	-	-
	<i>Screened in</i>	1.03 (2)**	0.55 – 1.51	-	-
	<i>Distress combined</i>	1.03 (2)**	0.55 – 1.51	-	-
	<i>Screened out</i>	-0.08 (3)	-0.38 – 0.23	-	-
	Screened (combined)	0.36 (5)	-0.17 – 0.89	-	-
	Unscreened	0.27 (3)	-0.12 – 0.66	0.16 (1)	-0.38 – 0.69
	Overall effect	0.30 (8)*	-0.01 – 0.62	0.16 (1)	-0.38 – 0.69
	Q statistic p	0.798 (0.001**)		1.000 (1.000)	
Treated control	<i>Screened in and out</i>	-	-	-	-
	<i>Screened in</i>	-	-	-	-
	<i>Distress combined</i>	-	-	-	-
	<i>Screened out</i>	1.21 (1)**	0.55 – 1.87	-	-
	Screened (combined)	1.21 (1)**	0.55 – 1.87	-	-
	Unscreened	-0.01 (13)	-0.17 – 0.14	0.09 (8)	-0.05 – 0.23
	Overall effect	0.05 (14)	-0.10 – 0.20	0.09 (8)	-0.05 – 0.23
	Q statistic p	0.000** (0.000**)		1.000 (1.000)	

Notes as for Table 6-3.

In Table 6-6 expressive-support in its pure form is seen to outstrip combinations with education and with relaxation and CBT at early times: main effects are 0.30, $p < 0.10$; 0.06, n.s.; and 0.10, n.s., respectively (but n for the rcs combination is only 3). For ‘s’ (expressive-support in its pure form) at early times, high scores are seen for screened and untreated controls (0.77, $p < 0.05$, in both cases) and very low scores for unscreened and treated control studies (0.05, in both cases). Frequencies are again too light for comment at late times. Both of these patterns are repeated for anxiety and distress outcomes.

There is generally insufficient n to sustain comment on the confound breakout data from therapy combinations with expressive-supportive, but for es, unscreened, early times ($n = 5$), the results are very poor.

Overall the data on this table are not encouraging, except for the scores produced by this therapy in its pure form for screened patients against untreated controls at early times.

Table 6-6. Expressive-supportive as therapy combinations, depression

Therapy combination	Confound breakout	Early times		Late times	
		ES (n)	95% CI's	ES (n)	95% CI's
s	Overall	0.30 (6)*	-0.02 – 0.63	0.09 (2)	-0.14 – 0.33
	Overall p	0.070*		0.438	
	Screened	0.77 (3)**	0.21 – 1.32	-	-
	Unscreened	0.05 (3)	-0.14 – 0.23	0.09 (2)	-0.14 – 0.33
	Untx control	0.77 (3)**	0.21 – 1.32	-	-
	Tx control	0.05 (3)	-0.14 – 0.23	0.09 (2)	-0.14 – 0.33
es	Overall	0.06 (6)	-0.39 – 0.52	0.10 (2)	-0.15 – 0.35
	Overall p	0.788		0.433	
	Screened	1.21 (1)**	0.55 – 1.87	-	-
	Unscreened	-0.13 (5)	-0.50 – 0.24	0.10 (2)	-0.15 – 0.35
	Untx control	0.54 (1)**	0.03 – 1.06	-	-
	Tx control	-0.03 (5)	-0.53 – 0.47	0.10 (2)	-0.15 – 0.35
rcs	Overall	0.10 (3)	-0.31 – 0.51	0.20 (2)	-0.46 – 0.87
	Overall p	0.636		0.549	
	Screened	-0.10 (1)	-0.55 – 0.35	-	-
	Unscreened	0.21 (2)	-0.47 – 0.88	0.20 (2)	-0.46 – 0.87
	Untx control	-0.10 (1)	-0.55 – 0.35	-	-
	Tx control	0.21 (2)	-0.47 – 0.88	0.20 (2)	-0.46 – 0.87

Therapy combination codes: e = professional education/information; r = relaxation focused cognitive-

behavioural treatment;; c = cognitive-behavioural treatment; s = expressive-supportive therapy / non-directive counseling / psychotherapy ('s' for supportive). Other notes as for Table 6-4.

In sum, this pair of tables shows that expressive-supportive therapy offers nothing to unscreened patients over what is available through treatment as usual or placebo in the short term, and a negligible non-significant benefit after six months. However, for patients presenting with distress at baseline there may be great value, but there are no late times data and n is light at early times. For non complex patients, the picture is confusing, with conflicting trends emerging between treated and untreated control subgroups.

Expressive-supportive v. CBT.

The very strong early times effects yielded by expressive-supportive therapies against all three constructs for patients screened in for baseline distress (and not simultaneously screened out – there were no data of that nature) cause the moderate magnitude CBT results for the same group (i.e. excluding patients screened both ways) to pale by comparison. The very high expressive-supportive results are each supported by only two studies, and the only follow up data are single study results at medium term together with hints of some durability of effect that emerge from results for unscreened patients, so more research with follow-up assessment is critical with this patient group. On the evidence available at present however, an effective strategy for treating distressed patients may be to start with expressive-supportive therapies, later working to consolidate and fix gains with CBT, but the suggestive evidence as to the durability of CBT may in fact be due to the non-complexity of some distressed samples treated with that therapy type.

Specific therapy combinations

In the detailed results concerning the four main therapy types presented in **Appendix O**, relaxation and CBT was the combination that showed the best results (excluding 'pure' combinations, i.e. protocols comprising only one therapy type) – a moderate magnitude impact against depression and general distress and a strong impact against anxiety - but there was no evidence available at late times. Expressive-support in combination with education was noted for its poor results, but it suffered a disproportionate frequency of studies from the low scoring confounds (unscreened and treated control comparisons) and may also be the result of confounding by cancer site (breast patient studies). The combination of relaxation, CBT and expressive-support proved particularly poor and raises a question about the theoretic consistency of that combination, but it is more likely that this result is also a result of confounding.

Non-professional

Effects for therapies administered by non-professionals are of limited interest because they are supported by such small n (two to seven). No main effect, however, reached statistical significance, and for four of six the effect size point estimate was actually negative (Table 6-1). However, the remaining two effect sizes (early times anxiety and depression) were in the small to moderate range showing that there are circumstances

under which non-professional interventions can work. Discussion of what these circumstances might be is left for the General discussion chapter, and further detail of results, including tables, is relegated to **Appendix O**.

Indirect therapies

There were few studies using an 'indirect' protocol, but it is an interesting category which shows exciting potential for future development, so the data are described in detail here.

It was difficult defining this category. The intention was to distinguish therapies that directly treated the patient from those that intervened with someone else – a significant other or medical professional – passing on the benefit indirectly. The difficulty arose where the intervention was not in the form of a typical therapy, but amounted to the training of medical professionals in communication skills. Two studies did this (Rutter, Iconomou, & Quine, 1996; Stewart et al., 2007) and to worthwhile effect. The line was drawn against another worthwhile study (Kristeller, Rhodes, Cripe, & Sheets, 2005) where oncologists were trained to administer a short series of questions that opened up discussion of how patients were coping spiritually and led to the offer of a referral if desired. In that case it was decided that because the content of the material was different from that which was normally conveyed by medical personnel, the questions amounted to an intervention with the patient, and the training of the oncologist would be considered 'therapist training'. The study was therefore coded expressive-supportive, of an existential nature, but remains of interest in this section because it shows how the influential position of medical personnel can be used to deliver powerful interventions very efficiently. The other two studies were coded 'indirect' because the personnel training merely taught better ways to communicate medical content that would normally be conveyed. The three remaining studies in the set intervened with the significant others of patients using usual therapy methods: education, expressive-supportive therapy, and written emotional expression.

Of the five indirect studies, one coded 'unclear' on the screening variable and consequently did not contribute to moderator analyses, though it did contribute to main effect calculations.

Main effects were: anxiety, 0.50 ($p < 0.05$), depression, 0.64 ($p < 0.05$), and distress, 0.38 (n.s.), all early times (Table 6-1). These are much the highest main effect sizes of any therapy type, but were each based on a frequency of only three.

There is no point in producing the standard tables for such small n data, but instead a descriptive table is provided, which includes effect sizes and confound information, and also includes the Kristeller et al. (2005) study (Table 6-7).

Table 6-7. Indirect, descriptive information

Study name and confound codes	ES (study n), assessment point, outcome construct	Descriptive notes
<i>Immediate therapy target: Significant other</i>		
Bultz, Specca, Brasher, Geggie & Page (2000), Unclear/Untx	0.71*(32) im anxiety	Large baseline differences were not tested for significance and were not adjusted for, and could explain the results. An education and support group for the significant others of breast cancer patients. Results for anxiety and depression fell only very slightly at mt, and were not reported for distress.
	0.84**(34) im depression	
	0.86** (34) im distress	
Goldberg & Wool (1985), Unsc/Untx	0.67 (20) im anxiety	Significant others of newly diagnosed lung cancer patients received 12 sessions of expressive-supportive therapy.
Rosenbaum (2006), Unsc/Tx	0.43 (15) st depression	Early stage breast cancer. Written emotional expression for husbands. Effect size rose from 0.2 at im.
<i>Immediate intervention target: Doctor</i>		
Rutter et al. (1996), Sc.out/Tx	0.23 (36) im anxiety	Doctors trained to convey information re upcoming treatment with structure and style that aided understanding and recall in two 45 minute workshops.
	0.69** (36) im depression	
Stewart et al. (2007), Unsc/Tx	0.15 (102) im distress	Intensive continuing education workshop in patient centred communication for family physicians, oncologists, and surgeons taking 6 hours rather than the usual 2.
Kristeller et al. (2005), Unsc/Tx	0.37* (111) st depression	Oncologists trained to make brief enquiries as to patients' spiritual wellbeing. Training took 2-3 hours, and the one-off delivery to patients added an average of 2 minutes to the time spent with controls. Scores rose from negligible or slightly negative (n.s.) at im. to the st. measure at 3 weeks.
	0.37* (111) st distress	

Assessment point codes: im = immediately after intervention; st = short term follow up, i.e. up to one month after intervention, but not immediately after; mt = medium term, i.e. 1 – 6 months after intervention.

Confound codes: screening status is presented first, then the nature of the control group after the slash: Unclear = screening status was unclear in the report; Unsc = unscreened; Sc.out = screened out for psychological history or complexity; Untx = untreated control comparison; Tx = treated control.

The therapies targeting significant others assume that relationships at home and the emotional and educational state of family members will interact with the distress level of the patient. Both the Goldberg and Wool (1985) and the Rosenbaum (2006) studies are expressive in nature, helping significant others to process their feelings. They produce moderate effect sizes (0.67 and 0.43, respectively) which did not reach significance because of low *n*. Given what this meta-analysis discovers about the ability of expressive-supportive therapies to favourably impact male patients, whether or not screened for distress (see socio-demographic results), the Rosenbaum study, which treated the husbands of breast cancer patients, is particularly interesting as a means of passing on benefit to patients. The Goldberg and Wool study, on the other hand, involved only 28% female patients, so most of the direct recipients of the therapy (spouses of male patients) would have been female. However, given the dismal prognosis of the cancer type involved (lung cancer) it is likely that these women would have been quite distressed at the time of the intervention, which shortly followed diagnosis, and the present study also shows that both non-sex-specific cancers and baseline distress predict effect impact (see Patient characteristics results, baseline distress and medical variables). It also finds (results above) that expressive-supportive therapies can have very strong beneficial effect on people who are distressed. Unfortunately neither the Rosenbaum nor the Goldberg study assessed at medium or later terms.

The Bultz, Specia, Brasher, Geggie and Page (2000) study is not discussed because it did not test the significance of or adjust for known large baseline differences in favour of the treatment group, calling the results into question.

None of these studies screened for the distress of either the patient or the significant other, and results may well have been even stronger had they done so.

In sum, the concept of benefiting a patient through a significant other has particular value where the patient cannot or will not receive help directly. In particular, these studies show the potential for helping men and late stage patients through their spouses.

The three doctor studies all illustrate how relatively brief training can yield worthwhile incremental gain to the full range of patients under their care. As mentioned, Rutter et al. (1996) and (Stewart et al. (2007) involved upskilling in medical communication while Kristeller et al. (2005) taught a means to briefly broach the issue of spiritual wellbeing. Further details of the interventions are tabled. All were successful, although the more focussed and directive interventions administered by Rutter et al. and Kristeller et al. were noticeably more so (0.69, $p < 0.05$ and 0.37, $p < 0.10$, against depression, respectively). Training in these latter two studies also took little time, and the fact that the Kristeller result improved at a later assessment for depression is attractive. None of these studies screened for patient distress either, but this type of intervention is designed to convey a benefit to all patients, with efficiency.

Other

The ‘other’ therapy category comprised three studies using written emotional expression, a particular therapy for prostate patients, and a Chinese holistic group therapy. Descriptive detail is provided in **Appendix O**.

Therapy trajectories

The four professional therapy types (education, relaxation, CBT, expressive-supportive, inclusively defined) were investigated more closely for the trajectories of their effectiveness over time. It will be recalled (Analysis chapter) that three sets of assessment time points were used in collating data: those anchored to the completion of the intervention; those anchored to the beginning of the intervention; and those anchored to the patient’s stage in a particular medical treatment. It was found that only assessment points anchored to the completion of the intervention and up until long term regularly yielded at least three studies upon which some analysis could be sustained, so only data from these assessment time points were tabulated. They were coded: 1. immediately post intervention; 2. short term i.e. up to 1 month post intervention; 3. medium term, i.e. more than one and up to six months post intervention; or, 4. long term, i.e. more than six and up to twelve months post intervention. Even so, n was mostly too sparse to be helpful. There was, however, one interesting thread of evidence that emerged.

In relation to studies that screened out for psychological history, some evidence was noted that effects may ‘fix’ or even lift over time, rather than fade away, as would normally be expected and as was the case for other data. This tendency was apparent over all outcomes for relaxation, and over depression and distress for CBT and expressive-supportive therapies. It may be that this trend also applies to samples that are screened both ways, resulting in very strong and lasting benefits. Unfortunately there were not sufficient data available from that subset of studies to glean any impression.

This thread of evidence would be interesting to follow up in future research, with important implications for practice and theory. The trajectory data for this one subset, against each therapy type and each outcome, and with cell frequencies of less than three shaded, are set out in Table 6-8 below.

Table 6-8. ‘Screened out’ studies, therapy trajectory

		Immediate ES (n)	Short term ES (n)	Medium term ES (n)	Long term ES (n)
Education	Anxiety	0.13 (1)	0.48 (1)*	0.27 (3)	-
	Depression	-0.26 (1)	-	0.09 (2)	-
	Distress	0.03 (1)	-	0.18 (3)	-0.50 (1)
Relaxation	Anxiety	0.25 (6)	0.95 (2)**	0.30 (6)	-0.02 (1)

	Depression	-0.24 (4)	-	0.10 (3)	-
	Distress	0.22 (6)*	1.13 (2)**	0.40 (7)**	-
CBT	Anxiety	0.09 (4)	0.02 (1)	0.03 (4)	-0.02 (1)
	Depression	-0.24 (3)	-	0.10 (3)	-
	Distress	0.13 (4)	0.26 (1)	0.26 (5)**	
Expressive -support	Anxiety	0.10 (3)	0.48 (1)*	0.05 (4)	-0.02 (1)
	Depression	-0.08 (3)	-	0.07 (2)	-
	Distress	-0.04 (2)	-	0.32 (3)	0.12 (1)

Assessment time points: Immediate = immediately post intervention; Short term = up to 1 month post intervention; Medium term = more than one and up to six months post intervention; Long term = more than six and up to twelve months post intervention.

Inclusive categorisation of therapy type was used. Results sustained by $n < 3$ shaded. Relevant notes from Table 6-1 apply.

Therapy type conclusion

In the therapy type results the ability of baseline screening for distress to lift scores – and of simultaneous screening for psychological complexity to lift them still further - is seen wherever data for those variables are available. The possibility is raised that psychological non-complexity may also help fix therapy effects. These are the most important observations from the above analyses, rather than any findings regarding particular therapy types.

Therapy components

Attention now turns to the detailed consideration of specific therapy components. Following this, the examination of therapy characteristics closes with examination of variables relating to delivery mode, dose and therapist specifics. That final section is dealt with by way of summary due to a lack of interesting findings.

More than 50 particular therapy components within the four main therapy types were coded. These categories were statistically dependent, i.e. any given study could belong to more than one component category and, indeed, most did. This meant that direct computed comparisons for heterogeneity were not possible. Further, effect sizes were exposed to random distortion by the other components of therapy that were packaged with the one under investigation in contributing studies. As with any other random variable, larger n provides protection. However, some components are regularly coupled with others, so an element of systematic variation is also involved. For these reasons, it should be borne in mind that general patterns built from higher n results are more likely to be reliable indicators of therapy component effectiveness rather than any individual result. Z-score statistical significance can also assist in interpreting these low n results.

Once again, the untreated control comparison dataset was used at early times. Studies with unclear status on screening, nature of the control, or the variable of substantive interest, were held out of analyses. Particular outcomes were relevant, and this set of results is presented by outcome. Because this reduced n, and because the need to take into account the structural confounds would reduce it further, a total frequency (across cells) of five was required of each therapy component to qualify for analysis. Results are tabled with shaded cells denoting frequencies of less than three, which are generally disregarded.

Anxiety

Attention is drawn to Table 6-9, below.

Education

It will be recalled that there were relatively few studies implementing education therapies that utilised an untreated control comparison, giving cause for reservation in interpreting present results. Three quite high effect sizes are reached for education components administered to unscreened patients (re cancer or cancer treatments, 0.74, $n = 4$; re facilities or services, 0.64, $n = 3$; re nutrition and exercise, 0.53, $n = 4$) but these do not reach statistical significance since their confidence intervals are very broad. However, they do indicate that this type of education can be very important in lowering anxiety for some patients. The only educational component results that do reach statistical significance are regarding coping strategies, but the effect sizes are modest (screened, 0.31, $p < 0.05$, $n = 6$; unscreened, 0.16, $p < 0.05$, $n = 8$).

The true value of education remains illusive in these results, although again hints are seen as to its potent potential for reducing anxiety.

Relaxation

Two particular relaxation therapies show their worth against this outcome, yielding moderate magnitude statistically significant results for screened patients: progressive muscle relaxation, 0.54 ($p < 0.05$, $n = 8$), and unspecified relaxation education or training, 0.58, ($p < 0.05$, $n = 4$). Note that diaphragmatic breathing approached significance for both screened and unscreened patients (0.48, $p = 0.129$, $n = 4$, and 0.13, $p = 0.159$, $n = 8$, respectively) producing a moderate magnitude result for screened patients. Stress management education or training reached 0.10 level significance, yielding a small effect for unscreened patients (0.17, $p < 0.10$, $n = 6$). Results for (guided) imagery and meditation are poor (screened, 0.17 and 0.06 respectively, both n.s.).

In sum the advantage for screened patients appears to be with progressive muscle relaxation, unspecified 'relaxation training' (from the context of the studies it is guessed that this is usually PMR), and deep breathing.

CBT

A number of classic CBT components reached moderate effect sizes and statistical significance for screened patients (cognitive restructuring, 0.56, $p < 0.10$, $n = 5$; challenging negative thoughts, 0.55, $p < 0.05$, $n = 6$; problem solving, 0.41, $p < 0.05$, $n = 8$; pleasant activity scheduling, 0.39, $p < 0.05$, $n = 4$). Assertiveness / communication training for screened patients returned a small result that approached significance (0.22, $p = 0.130$, $n = 5$) and self monitoring showed promise (0.85, n.s., only two studies) with goal / expectation setting less so (0.29, n.s.). For unscreened patients the best gains were the negligible to smallish statistically significant results from challenging negative thoughts, problem solving (0.17, $n = 7$; 0.16, $n = 5$; 0.16, $n = 5$, respectively, all $p < 0.10$).

These results confirm that CBT is not for unscreened patients, but there are a number of components that are valuable for treating the anxiety of screened patients.

Table 6-9. Therapy components, anxiety

Therapy component	Screened		Unscreened	
	ES (n)	95% CI's	ES (n)	95% CI's
Education re cancer or cancer treatments	0.30 (2)	-0.10 – 0.69	0.78 (4)	-0.19 – 1.74
Education re emotion and cancer	0.28 (2)	-0.08 – 0.63	0.31 (3)	0.03 – 0.60
Education re cancer facilities / services	-	-	0.64 (3)	-0.45 – 1.72
Education re nutrition and exercise	0.24 (1)	-0.26 – 0.74	0.53 (4)	-0.38 – 1.44
Education re coping strategies	0.31 (6)**	0.07 – 0.54	0.16 (8)**	0.00 – 0.31
Education re managing side effects	-	-	0.55 (3)	-0.72 – 1.82
Stress management education / training	0.24 (1)	-0.26 – 0.74	0.17 (6)*	-0.01 – 0.35
Progressive muscle relaxation	0.54 (8)**	0.26 – 0.83	0.09 (9)	-0.07 – 0.25
Relaxation education or training (unspecified type)	0.58 (4)**	-0.07 – 1.23	0.16 (4)	-0.19 – 0.52
Cue-controlled relaxation training	-	-	-	-
(Guided) imagery	0.17 (4)	-0.12 – 0.46	0.10 (5)	-0.08 – 0.28
Meditation	0.06 (3)	-0.31 – 0.43	-	-
Diaphragmatic breathing	0.48 (4)	-0.14 – 1.10	0.13 (8)	-0.05 – 0.31

Cognitive restructuring / reappraisal	0.56 (5)*	-0.11 – 1.23	0.09 (4)	-0.14 – 0.31
Challenging negative thoughts	0.55 (6)**	0.06 – 1.04	0.17 (7)*	-0.01 – 0.36
Positive self talk / imagine success	-	-	0.05 (3)	-0.20 – 0.30
Self monitoring	0.85 (2)	-1.23 – 2.93	-0.06 (1)	-1.11 – 1.00
Problem identification	0.50 (1)**	0.08 – 0.92	0.08 (2)	-0.19 – 0.36
Problem solving	0.41 (8)**	0.07 – 0.74	0.16 (5)*	-0.01 – 0.33
Goal / expectation setting / plan making	0.29 (3)	-0.24 – 0.83	0.16 (5)*	-0.03 – 0.35
Pleasant activity scheduling	0.39 (4)**	0.01 – 0.77	-	-
Assertiveness / communication education or training	0.22 (5)	-0.06 – 0.51	0.14 (5)	-0.05 – 0.33
Establishment or optimisation of use of social networks	0.14 (2)	-0.53 – 0.80	0.08 (2)	-0.54 – 0.70
Body image counselling	-0.06 (2)	-0.41 – 0.29	0.38 (1)**	0.01 – 0.75
Sex therapy / education re sexual intimacy / sexuality	-	-	0.23 (2)*	-0.03 – 0.49
Exp-sup re existential, spiritual, grief or death issues	0.87 (2)	-0.19 – 1.93	0.18 (6)	-0.09 – 0.44
Exp-sup re cancer experience, physical and psychosocial issues	0.47 (6)**	0.06 – 0.89	0.31 (6)**	0.09 – 0.53

Early times, untreated control data. Exp-sup = Expressive-supportive therapies. Cells with $n < 2$ are shaded. Other relevant notes from Table 6-1 apply.

Expressive-support

Under expressive-supportive therapies we see a strong result for existential issues with screened patients that just misses statistical significance despite being supported by only two studies (0.87, $p = 0.107$) but it plummets to a non-significant low impact for unscreened patients (0.18, $n = 6$). Expressive supportive therapies around more general topics relating to the cancer experience and its psychosocial impact produce a moderate and statistically significant result for screened patients (0.47, $p < 0.05$, $n = 6$) and a more modest significant return for unscreened patients (0.31, $p < 0.05$, $n = 6$).

The stark contrast between screened and unscreened results for existential issues is thought provoking. The difference is not statistically significant (Q statistic $p = 0.213$) but with more n in the screened cell it may well get there. Perhaps distressed patients particularly value the opportunity to talk through grief issues and develop spiritual

meaning in the midst of their distressing journey, whereas unscreened patients do not generally feel the same need.

Depression

Education

Table 6-10 portrays education as having no impact on depression at early times. Setting aside the moderate magnitude statistically significant result for unscreened patients re emotion and cancer since it is supported by only one study (0.54, $p < 0.05$), the only gain which even approached significance is a negligible to small outcome for screened patients who received education re coping strategies (0.14, $p = 0.149$). However, a similar benefit from this component was gained in relation to anxiety (see above).

Relaxation

Of the relaxation therapies the one hopeful result against depression was for unspecified relaxation training for screened patients (0.46, $p = 0.201$, $n = 3$) which may have reached statistical significance with more n . However, this result is not terribly helpful since the studies contributing to it did not provide specifics as to the nature of the therapy.

Table 6-10. Therapy components, depression

Therapy component	Screened		Unscreened	
	ES (n)	95% CI's	ES (n)	95% CI's
Education re cancer or cancer treatments	-0.26 (1)	-0.81 – 0.29	0.20 (1)	-0.36 – 0.76
Education re emotion and cancer	-0.26 (1)	-0.81 – 0.29	0.54 (1)**	0.03 – 1.06
Education re cancer facilities / services	-	-	0.01 (2)	-0.24 – 0.26
Education re nutrition and exercise	0.16 (1)	-0.34 – 0.66	-	-
Education re coping strategies	0.14 (7)	-0.05 – 0.34	0.14 (3)	-0.19 – 0.47
Education re managing side effects	-	-	-	-
Stress management education / training	0.16 (1)	-0.34 – 0.66	-0.01 (2)	-0.25 – 0.23
Progressive muscle relaxation	0.08 (5)	-0.19 – 0.35	0.02 (6)	-0.15 – 0.18
Relaxation education or training (unspecified type)	0.46 (3)	-0.25 – 1.17	0.09 (2)	-0.27 – 0.46
Cue-controlled relaxation	0.47 (1)	-0.26 – 1.20	-	-

training				
(Guided) imagery	-0.16 (4)	-0.58 – 0.25	0.01 (4)	-0.17 – 0.19
Meditation	-0.15 (2)	-0.52 – 0.23	-	-
Diaphragmatic breathing	-0.40 (2)	-0.93 – 0.12	0.02 (3)	-0.20 – 0.24
Cognitive restructuring / reappraisal	0.42 (8)**	0.02 – 0.82	0.13 (4)	-0.17 – 0.43
Challenging negative thoughts	0.28 (7)	-0.10 – 0.66	0.02 (3)	-0.23 – 0.27
Positive self talk / imagine success	-	-	-0.01 (1)	-0.28 – 0.26
Self monitoring	0.55 (3)	-0.33 – 1.43	0.01 (1)	-0.47 – 0.50
Problem identification	0.85 (3)*	-0.16 – 1.87	-	-
Problem solving	0.41 (11)**	0.02 – 0.80	0.13 (4)	-0.17 – 0.43
Goal / expectation setting / plan making	0.01 (2)	-0.36 – 0.38	0.19 (4)	-0.14 – 0.53
Pleasant activity scheduling	0.11 (4)	-0.10 – 0.33	-	
Assertiveness / communication education or training	0.08 (5)	-0.16 – 0.31	0.40 (3)*	-0.02 – 0.83
Establishment or optimisation of use of social networks	0.23 (2)	-0.45 – 0.91	0.01 (1)	-0.47 – 0.50
Body image counselling	-0.17 (2)	-0.52 – 0.18	-	-
Sex therapy / education re sexual intimacy / sexuality	-	-	-	-
Exp-sup re existential, spiritual, grief or death issues	0.69 (2)	-0.25 – 1.63	0.51 (2)**	0.04 – 0.97
Exp-sup re cancer experience, physical and psychosocial issues	0.41 (5)	-0.15 – 0.96	0.27 (3)	-0.12 – 0.66

Notes as for Table 6-9.

CBT

Classic cognitive therapy components generally performed well against depression for screened patients: Restructuring delivered a moderate and statistically significant effect size (0.42, $p < 0.05$, $n = 8$); challenging negative thoughts, a smaller effect size that approached significance (0.28, $p = 0.155$, $n = 7$); problem identification and problem solving returned strong and moderate significant results (0.85, $p < 0.10$, $n = 3$, and 0.41, $p < 0.05$, $n = 11$) respectively; but the moderate magnitude result for self monitoring did not reach statistical significance (0.55, $n = 3$). Interestingly, pleasant activity

scheduling and assertiveness / communication proved of no value to screened patients (0.11, $n = 4$, and 0.08, $n = 5$, respectively, both n.s.) although the latter did quite well for unscreened patients yielding a moderate and significant g of 0.40 ($p < 0.10$, $n = 3$). Otherwise, only goal / expectation setting / plan making showed any prospect of benefiting unscreened patients with a small non-significant outcome (0.19, $p = 0.252$, $n = 4$).

Perhaps the most interesting results from this selection are the strong showing of problem identification and problem solving for screened patients, and the benefit that assertiveness / communication training yielded for unscreened patients. The former components were identified as contributing to outliers in the chapter on external validity results, and seem particularly well suited to assisting depressed cancer patients. The latter finding comes as something of a surprise, though it is based on only three studies. The contrast with its screened result, a negligible and non-significant 0.08, is also interesting. How is it that unscreened patients found this type of training helpful whereas screened patients did not? Perhaps the screening applied was screening out for psychological complexity rather than in for distress. On the other hand, given this component's poor showing against anxiety and distress, the unscreened result could be a fluke caused by the kind of random variance mentioned at the beginning of this chapter.

Expressive-support

The expressive-supportive outcomes are a little better than they look as both screened results are not far from reaching statistical significance (re existential issues, 0.69, $p = 0.152$, $n = 2$; re cancer experience generally, 0.41, $p = 0.149$, $n = 5$). Again, it was the existential result that was higher, but this time, against depression rather than anxiety, the unscreened group also seemed to gain a moderate benefit (0.51, $p < 0.05$) although sustained by only two studies. A small but non-significant reduction in depression accrued to unscreened recipients of the more general therapy.

Distress

General psychological distress results are presented below (Table 6-11).

Education

Again, the data were constrained by the lack of education studies with untreated control groups. However, a moderate statistically significant effect resulted from small n for screened patients receiving education regarding nutrition and exercise (0.49, $p < 0.05$, $n = 2$) - although recall that studies that primarily focussed on these therapies were excluded from the meta-analysis domain so there will be much more extant data that are not taken into account here. Also, education re coping strategies produced a small effect for screened patients (0.26, $p < 0.05$, $n = 7$). Education re nutrition and exercise reached a similar significant result from small n for unscreened patients (0.51, $p < 0.05$, $n = 2$), so it may be that this form of education is useful generally, and has moderate strength of effect, which is notable for any educational intervention at early times. The only other results that approached significance were also both for unscreened patients,

namely education re cancer / treatments (0.28, $p = 0.159$, $n = 6$) and education re emotion and cancer (0.59, $p = 0.152$, $n = 3$). This latter result - of moderate magnitude, being supported by only three studies and applying to cancer patients generally - is worth noting.

Relaxation

The distress outcome from progressive muscle relaxation was the second highest well substantiated effect size for any of the relaxation therapies administered to screened patients (0.51, $p < 0.05$, $n = 8$), following on from a similar outcome against anxiety. Diaphragmatic breathing came through to top this ranking, as well as its performance against anxiety, with a moderately high statistically significant g of 0.63 ($p < 0.10$, $n = 4$). All other relaxation therapy results were non-significant, but the small outcome produced by imagery for screened patients approached significance level (0.22, $p = 0.162$, $n = 5$).

Table 6-11. Therapy components, distress

Therapy component	Screened		Unscreened	
	ES (n)	95% CI's	ES (n)	95% CI's
Education re cancer or cancer treatments	0.03 (1)	-0.51 – 0.58	0.28 (6)	-0.11 – 0.66
Education re emotion and cancer	0.03 (1)	-0.51 – 0.58	0.59 (3)	-0.22 – 1.39
Education re cancer facilities / services	-	-	-0.14 (3)	-0.40 – 0.11
Education re nutrition and exercise	0.49 (2)**	0.11 – 0.88	0.51 (2)*	-0.01 – 1.03
Education re coping strategies	0.26 (7)**	0.04 – 0.48	0.20 (5)	-0.22 – 0.62
Education re managing side effects	0.25 (2)	-0.60 – 1.11	0.07 (5)	-0.28 – 0.43
Stress management education / training	0.40 (1)	-0.11 – 0.91	0.11 (5)	-0.18 – 0.39
Progressive muscle relaxation	0.51 (8)**	0.24 – 0.79	0.06 (6)	-0.21 – 0.32
Relaxation education or training (unspecified type)	0.10 (3)	-0.39 – 0.58	0.11 (2)	-0.31 – 0.54
Cue-controlled relaxation training	0.67 (1)*	-0.08 – 1.41	-	-
(Guided) imagery	0.22 (5)	-0.09 – 0.52	0.09 (4)	-0.12 – 0.29
Meditation	-0.09 (1)	-0.54 – 0.37	-	-
Diaphragmatic breathing	0.63 (4)*	-0.05 – 1.31	0.16 (5)	-0.14 – 0.46

Cognitive restructuring / reappraisal	0.31 (7)**	0.02 – 0.61	-0.08 (4)	-0.36 – 0.19
Challenging negative thoughts	0.36 (7)**	0.13 – 0.59	-0.05 (2)	-0.69 – 0.59
Positive self talk / imagine success	-0.25 (1)	-1.14 – 0.64	0.17 (1)	-0.10 – 0.44
Self monitoring	0.39 (4)**	0.01 – 0.77	-0.11 (1)	-0.60 – 0.37
Problem identification	0.93 (3)**	0.28 – 1.57	-0.04 (3)	-0.43 – 0.34
Problem solving	0.46 (10)**	0.16 – 0.76	0.02 (6)	-0.16 – 0.20
Goal / expectation setting / plan making	-0.09 (1)	-0.54 – 0.37	0.10 (2)	-0.18 – 0.39
Pleasant activity scheduling	0.28 (3)	-0.06 – 0.63	-0.14 (1)	-0.86 – 0.57
Assertiveness / communication education or training	0.11 (5)	-0.14 – 0.35	-0.04 (5)	-0.35 – 0.28
Establishment or optimisation of use of social networks	0.23 (2)	-0.43 – 0.89	0.11 (2)	-0.31 – 0.54
Body image counselling	-0.12 (3)	-0.43 – 0.19	-0.08 (2)	-0.38 – 0.23
Sex therapy / education re sexual intimacy / sexuality	-	-	0.13 (4)	-0.18 – 0.44
Exp-sup re existential, spiritual, grief or death issues	0.89 (2)**	0.29 – 1.50	0.32 (3)	-0.43 – 1.08
Exp-sup re cancer experience, physical and psychosocial issues	0.53 (7)**	0.17 – 0.88	0.43 (4)	-0.22 – 1.09

Notes as for Table 6-9.

CBT

A familiar pattern was seen for CBT, with the classic components of restructuring (0.31, $n = 7$), challenging negative thoughts (0.36, $n = 7$), self monitoring (0.39, $n = 4$) and problem solving (0.46, $n = 10$) all yielding low to moderate statistically significant ($p < 0.05$) results against distress for screened patients, but no component producing any beneficial effect to speak of for unscreened patients. Problem identification once again produces the highest significant effect size, this time at 0.93 ($p < 0.05$, $n = 3$), and pleasant activity scheduling just misses significance producing a small 0.28 ($p = 0.109$, $n = 3$). Once again, assertiveness / communication training performs poorly, as do body image counselling and sex therapy.

Expressive-support

Expressive-supportive therapies also conformed to their established pattern, yielding higher and statistically significant results for screened patients, and the highest for the

existentially oriented therapy (0.89, $p < 0.05$, but n only two) compared with a moderate magnitude effect for the more generally oriented therapy (0.53, $p < 0.05$, $n = 7$). The unscreened results are small to moderate in magnitude, but not statistically significant, indicating broad variation in response (existential, 0.32, $n = 3$; general, 0.43, $n = 4$).

Therapy component conclusion

The education therapy components that cropped up repeatedly for providing greater impact were those regarding the disease and its medical treatment and regarding emotion and cancer. Often the better results were those for unscreened patients, but from earlier analyses it is known that there were no studies that screened in for distress so screening must all have been out for psychological history. The outcomes most impacted by this form of treatment are anxiety and distress rather than depression, indicating that these education components ought to form a mainstay of usual care delivered to all patients. Note also the effectiveness of education regarding nutrition and exercise against general distress, although that primary focus was outside the domain of this paper.

Of the relaxation therapies, progressive muscle relaxation, ‘unspecified’ relaxation, which probably denotes PMR in most cases, and diaphragmatic breathing are useful for reducing anxiety and general distress in screened patients, but not for depression. Other therapies of sufficient n to be analysed did not show worthwhile results, and unscreened patients were not well served by relaxation therapies.

The components of CBT that performed repeatedly over all three constructs, but for screened patients only, were cognitive restructuring, challenging negative thoughts, self monitoring, problem identification and problem solving. Pleasant activity scheduling had a lesser effect. Problem identification and solving may be the most valuable components. Assertiveness / communication training yielded almost consistently disappointing results.

Both expressive-supportive components proved effective for screened patients, with existential consistently taking the lead over general physical and psychosocial issues. A lesser benefit was available to unscreened patients.

Components contributing to indirect therapies, notably the training of medical personnel in improved communication, and ‘other’ therapy types, including written emotional expression, were discussed in some detail in the therapy types results chapter.

Delivery mode, dose and therapist variables

A large range of variables relating to therapy delivery mode, dose, and the qualities of therapists were coded for and analysed. Those variables were: Delivery mode: therapy recipient, delivery technology, and therapy setting; Dose: flexibility of session time, number of hours with the therapist; duration of therapy in weeks; number of therapy sessions; nature of the therapy work; homework expected; tailoring of therapy content;

Therapist: discipline; experience level; and level of involvement with the patient. For many of the dose variables data were collected in continuous form and had to be broken into categories for analysis. Such data were first examined for the shape of their distribution by stem and leaf plot, a few outliers were excluded, and then distributions were broken into upper and lower quartiles with one category collecting all the data in between. This was done to allow the extremes of the data to contrast. As usual, the untreated control portion of the dataset was used, at early times. However, for some of the dose variables (number of hours, weeks, and sessions) analyses were made that broke out assessment points in order to test whether greater dose resulted in more durable effect. The assessment points used in these cases were: immediate and short-term (up to one month after therapy); medium term (one – six months); and combined late times (more than six months).

Although there are a few interesting results among these analyses, by and large they did not produce statistical heterogeneity, showing either that they are not moderators of great importance or that researchers already have the balance of these factors about right relative to the type of therapy and characteristics of the patients delivered to. Since there are a great many analyses, they are not set out here, but are presented in detail in the relegated results appendix (**Appendix O**). That presentation includes consideration of confounding in relation to some results. Instead, a brief comment is provided here, together with a summary table, below (Table 6-12). The table describes the levels of categorisation used for each variable, as well as the conclusion reached for screened and unscreened subsets.

Note that in the table, the word ‘heterogeneous’, as opposed to ‘homogeneous’, is used when statistically significant differences were found between variable categories, and the weaker term, ‘trend’, when this was not established, but the results from one category appeared nonetheless quite different from others. The discussion following the table is structured under the three headings of delivery mode, dose and therapist variables.

Table 6-12. Therapy delivery mode, dose, and therapist variables: Summary

	Variable and coded levels	Screened	Unscreened
Delivery	<p>Therapy recipient</p> <p>1. Individual patients; 2. Individual patients accompanied by a significant other; 3. Group of patients; and 4. Group of patients with their significant others. Significant others were most commonly spouses</p>	Trend favours individual patient over patient plus significant other or group of patients	Negligible trend favours individual patient over patient plus significant other or group of patients. Indirect interventions show promise
	<p>Delivery technology</p> <p>1. In person; 2. By telephone; 3. By some other interactive technology (personal letter, email or interactive web site); or, 4. By non-</p>	Insufficient data for comparison	Homogeneity highlights the efficiency and accessibility of non-interactive technology

	interactive technology (i.e. by means that do not involve patient / therapist interaction, including printed material, video or audio recording, non-interactive web site, or snoezelen environment)		where appropriate
	<p>Therapy setting</p> <p>1. Inpatient or residential care; 2. Outpatient (including hospital or hospice outpatient clinics, professional premises, and community facilities); 3. The patient's home; or, 4. 'Split' i.e. the patient and therapist were in different settings, as for telephone delivery</p>	Homogeneity between inpatient and outpatient settings	Homogeneity between available categories implies that self directed or telephone therapy at home can be of similar value to outpatient therapy
Dose	<p>Flexibility of session time</p> <p>1. Limited; 2. Additional time available according to need; 3. Not applicable (e.g. self directed)</p>	Homogeneity implies efficiency from a limited format, but this result may be misleading	Homogeneity implies efficiency from a limited format, but this result may be misleading
	<p>Session hours with a therapist</p> <p>1. ≤ 4 hours; 2. 5 – 11 hours; 3. ≥ 12 hours; 4. Not applicable (self directed)</p>	Homogeneity found at immediate / short term but probably due to confounding by cancer site, or could result from appropriate hours being allocated for different therapy types. Effect confounded at medium term.	Homogeneity implies efficiencies could be obtained from lower doses or self-directed therapies, but possible confounding by cancer site was not investigated except for self-directed therapies. Effect evaporates by medium term.
	<p>Weeks of therapy sessions 1. 1 – 3 weeks; 2. 4 – 9 weeks; 3. 10 – 22 weeks; 4. Not applicable (e.g. printed or recorded material was provided to patients)</p>	Weak immediate / short term trend in favour of more weeks. Effect confounded at medium term.	
	<p>Number of therapy sessions</p> <p>1. 1 – 4 sessions; 2. 5 – 7 sessions; 3. 8 – 13 sessions; 4. Not applicable (non-interactive)</p>	Immediate / short term homogeneity suggests efficiencies from fewer sessions. Effect confounded at medium term.	
	<p>Nature of patient's therapy work†</p> <p>1. With homework / skill practice; 2. Without same</p>	Limited data but homogeneity suggests homework adds little to active participation	

	Frequency of homework 1. One-off (e.g. the reading of a literature pack provided at the first or only session); 2, Regularly expected (e.g. with sessions); 3. Irregularly expected; 4. Optional; and, 5. None at all	Homogeneity implies that regular homework (beyond a one off expectation) could be dispensed with, but not for relaxation therapies or CBT	Heterogeneous in favour of no/one-off homework
Therapist	Therapist discipline 1. Psychology; 2. Psychiatry; 3. Social work; 4. Counseling / trained therapist; 4. Nursing; 5. Multidisciplinary team; 6. Lay; 7 Mixed lay and professional team; 8. Not applicable (e.g. bibliotherapy)	Social workers do well	Trend favours psychologists, psychiatrists and counsellors / trained therapists
	Therapist experience level 1. Lay; 2. Students (including those with training); 3. Practitioners / professionals; 4. Mixed lay and professional; 5. Not applicable	Trend favours professionals over students	Homogeneous
	Therapist involvement 1. Minimal, i.e. one on one initial contact for setting up the research only; 2. Group, i.e. contact as part of group delivery and possibly initial one on one contact for set up purposes as well; 3. One-on-one, i.e. individually delivered therapies or those with both group and individual components; and 4. Intense, i.e. therapist was available at group or individual sessions and also beyond the normal session frame e.g. on crisis call.	Small trend favours one-on-one involvement	Homogeneity implies group involvement over one-on-one

Untreated control data, omnibus outcome, at early times unless stated as follows: immediate and short-term (up to one month after therapy); medium term (one – six months); and combined late times (more than six months).

‘Heterogeneous’, as opposed to ‘homogeneous’, is used when statistically significant differences were found between variable categories, and the weaker term, ‘trend’, when this was not established, but the results from one category appeared nonetheless quite different from others.

† This category originally coded for active (interaction, discussion, exercises) or passive (listening, reading) nature of work, plus whether homework was required, but there was only one study in the untreated control data that used a passive protocol, so that study was excluded and the analyses effectively became a comparison of the effect of the presence or absence of a homework / practice requirement.

Delivery mode variables

The first three categories tabled relate to delivery variables, and include who the therapy is delivered to, by what technological means, and in what setting. The summary word that appears repeatedly in relation to the latter two factors is homogeneity, i.e. in terms of therapy effectiveness, it matters little how the therapy is delivered or in what setting. There is a non-significant trend, however, in favour of the patient receiving the therapy

as an individual rather than as part of a group or with a significant other. This is very small in relation to unscreened studies, but greater for screened studies. Receiving the therapy individually is not the same as receiving it personally, however (refer delivery technology and therapist involvement), and includes receiving non-interactive instructional or other therapeutic material. There is also promise in the few studies that attempt to benefit patients by treating significant others instead, or by up-skilling doctors, and the latter can benefit the whole range of patients.

The fact that the results for means of delivery and setting are homogeneous (to the extent that data are available) gives scope for financial efficiencies and for flexibility in designing therapies to reach the socio-demographically and medically defined groups that need them most. It is helpful to know that in designing therapies, the focus can be placed on access for such people, rather than on a particular delivery means or setting. However, the whole truth of that proposition has yet to be tested as there were gaps in the data such that not all means of delivery or settings could be compared, but only in-person *v.* non-interactive delivery, and inpatient *v.* outpatient, for screened patients. Where non-interactive delivery produced enough *n* to merit attention (unscreened), its result was non-significant, indicating a broad range of individual difference around how therapies delivered by this form of technology are received.

Dose variables

‘Homogeneity’ is also the word commonly used in relation to dose variables. It was found in relation to flexibility of session time (limited or unlimited number of sessions) for both screened and unscreened patients. Efficiencies might be considered from fewer therapy sessions, but a trend at immediate / short term with screened patients favoured therapy spread over more weeks. Results for actual hours spent with a therapist were likely confounded by cancer site.

Statistically significant differences favouring no or one-off homework were found for unscreened patients. That finding is consistent with the notion that this group, which lacks proven distress, may not comply with demanding homework expectations and if the therapy depends on homework being done for its effectiveness, then it will not work. For screened patients, homogeneity was found between results for therapies that relied on homework and those that did not or required only a one-off effort, leading to the suggestion that homework (beyond a one-off effort) might be dropped for this heavily burdened group. An exception is noted in relation to CBT and relaxation where for screened patients (though not unscreened) skill practice is necessary and a trend in favour of homework was shown, consistent with the meta-analytic finding based on studies of general psychological clients by Kazantzis, Deane, and Ronan (2000).

Further in relation to dose, once again, homogeneous findings in relation to number of hours and sessions give rise to the opportunity for flexibility and efficiencies in how therapies are designed. The law of diminishing returns found in economics seems to be at work as most therapeutic ‘punch’ appears to be delivered within the first hours, sessions and weeks of therapy and incremental gains are smaller as input continues.

This is consistent with the finding of Howard et al. (1996 and Kopta, Howard, Lowry & Beutler, 1994, both as cited in Beutler et al., 2004) that the relationship between therapy intensity and benefit is - with general psychological samples - a negatively accelerating curve. That said, there was a non-significant trend in favour of more weeks for screened patients, suggesting that spreading therapy may be more beneficial than delivering it in a brief intensive dollop for this group. The Rehse and Pukrop (2003) finding of significantly greater effect for studies of more than 12 weeks duration was not replicated, however. Self-directed work emerges comparatively strongly again, though less so for unscreened patients.

There is some evidence that dose results co-vary with cancer site in that higher dose therapies tend to be delivered to breast cancer patients, a low scoring group. This evidence, together with such homogeneity as was found, suggests that reaching needy socio-demographic and medical groups should take prominence over dose in therapy design.

Before moving on to summarise therapist variables, it is important to note that all of these findings are drawn from a dataset that is heavily skewed towards the simple and easy end of the spectrum of patient needs. At primary level it was clear that most researchers were striving for scientific control and homogeneity by excluding patients with serious physical conditions additional to cancer. However, in 'real life' a great many cancer patients labour under such complications since cancer is generally a disease of old age. Further, a disproportionate number of studies were of women with early stage breast cancer or patients with other favourable prognoses who were in curative treatment. Such circumstances are a far cry from those suffered by the half of the patient pool who die of their diagnoses, and most of those patients who present to psycho-oncologists, whose need of support is because of the very multiplicity, complexity and severity of their burdens. The outcomes of the present analyses should therefore be regarded as a starting point for relatively simple cases, rather than the answer for all - or even typical - psycho-oncological patients.

Therapist variables

Therapist variables are statistically homogeneous on all three counts – discipline, experience level, and involvement – for unscreened patients, once again leaving much scope for design with the aim of efficiency and accessibility to needy socio-demographic and medical groups. However, in relation to screened samples there was a trend in favour of those disciplines that specialise in work of a more psychological nature (psychologists, psychiatrists, and counsellors / trained therapists – n for the latter two was very light) which may hint that such specialised skill is necessary to elicit results from patients who lack proven baseline distress. For screened patients there were trends towards more professional and one-on-one involvement by therapists. However, the professional need not always be in the most highly paid bracket, but with training and a type of therapy that lends itself to manualisation (e.g. problem solving or relaxation) could very effectively be a social worker. Screened results for the latter

discipline were higher than those for others, reversing the finding of Cwikel and Behar (1999). However, the types of therapies that this discipline delivered were more structured and simple than the mixture of protocols administered by some other disciplines, and two of the five studies that produced the high effect size for social workers were outliers that simultaneously screened recruits both in for distress and out for psychological history. Distressed patients with more complex histories ought to be referred to a psychologist or psychiatrist.

7. RESULTS: PATIENT CHARACTERISTICS

In this results chapter attention is turned to patient characteristics – patients' sociodemographic features, level of distress before intervention, and status on various medical criteria. Baseline distress has already received some attention as a screening criteria, but will be examined in further detail as a substantive moderator. As noted in the Introduction chapter, medical variables such as cancer and medical treatment type and stage have themselves been suggested as proxies for distress and therefore possible moderators of therapy effectiveness. Analyses in this chapter go to the main purpose of the present research and make some exciting and far reaching findings.

Generally, untreated control comparisons were used in analyses, with the omnibus outcome at early times in order to maximise n. However, treated control data were brought into results tables where it was necessary in order to muster enough n to inspect for a pattern. Data were broken out by whether or not recruits were screened, without regard to screening type (in or out or both) because the empirical - rather than the theoretical – homogeneity of the screening groupings was more important for the purpose of drawing out statistical effects against these variables of interest. Where specific therapy types were brought into analyses, the inclusive categorisation of therapy types was used to maximise n. Confidence intervals are reported only on important tables. Asterisks signifying z-score significance still provide an indication of variances on other tables.

Socio-demographic moderators

Broad socio-demographic variables, namely, age, race, sex, education level, occupation / income and marital status are investigated in this section. These factors are of interest because of their propensity to create either vulnerability to or resilience against psychological distress.

Age

The initial step in performing this analysis of the impact of age on therapy effectiveness was to split the continuous mean age data into younger and older groups. This was done after inspection of a hand-drawn stem-and-leaf plot which showed a fairly normal distribution with values ranging from 41 to 66 years, bi-modal at 51 and 54 years. The split was made at the approximate average and median, 52.99 years. Results are shown in Table 7-1.

Table 7-1. Age

	Screened				
	41-52 years		53-66 years		Q statistic <i>p</i>
	ES (n)	95% CI's	ES (n)	95% CI's	
All tx	0.56 (14)**	0.28 – 0.85	0.54 (14)**	0.35 – 0.73	0.888
Educ	0.14 (3)	-0.34 – 0.63	-0.03 (1)	-0.58 – 0.51	0.635
Relax	0.33 (7)*	-0.05 – 0.72	0.60 (9)**	0.32 - 0.88	0.270
CBT	0.45 (8)**	0.06 – 0.84	0.48 (8)**	0.21 – 0.74	0.908
Exp-sup	0.47 (5)**	0.06 – 0.88	0.37 (3)	-0.21 – 0.96	0.801

Untreated control data at early times. Where it was unclear whether a study fitted a certain category, it was excluded from the analysis.

ES = Hedges *g* effect size point estimate; *n* = number of studies in subset; CI's = confidence intervals. * statistically significant at $p < 0.10$; ** statistically significant at $p < 0.05$; two tailed. 95% CI's = 95% confidence intervals.

Inclusive categorisation of therapy types. All tx = all therapies, results averaged. Educ = education / information; Relax = relaxation training; CBT = cognitive behavioural therapy; Exp-sup = expressive-supportive therapy / non-directive counselling; Non-prof. = non-professional counselling or support.

Table 7-1 continued

	Unscreened				
	41-52 years		53-66 years		Q statistic <i>p</i>
	ES (n)	95% CI's	ES (n)	95% CI's	
All tx	0.11 (12)	-0.03 – 0.25	0.42 (16)**	0.19 – 0.65	0.022**
Educ	0.10 (5)	-0.08 – 0.28	0.52 (6)**	0.04 – 0.99	0.109
Relax	0.12 (5)	-0.15 – 0.40	0.21 (7)**	0.01 – 0.41	0.603
CBT	0.11 (7)	-0.09 – 0.32	0.13 (3)	-0.15 – 0.42	0.892
Exp-sup	0.13 (3)	-0.19 – 0.45	0.46 (4)**	0.03 – 0.88	0.225

For screened patients, with all therapies combined, there is no statistically significant difference in the effect sizes, which are moderate and statistically significant for both age groups (younger, $g = 0.56$, older, 0.54 , both $p < 0.05$, Q statistic $p = 0.888$). This homogeneity is robust to a breakout of therapy type (Q statistic p 's are non-significant and range from 0.270 for relaxation to 0.908 for CBT). This means that screening generally levelled effect between the two age groups – baseline distress or psychological non-complexity were generally more influential than age. However,

there is a non-significant trend in favour of older patients for relaxation (younger, 0.33, $p < 0.10$; older, 0.60, $p < 0.05$). The education result is meaningless because of low n.

For unscreened patients, combined therapies, the older group benefitted significantly more than the younger who gain non-significant benefit of negligible magnitude (older, 0.42, $p < 0.05$; younger, 0.11, n.s., Q statistic $p = 0.022$). This means there was no value from treating younger patients who did not fall into either of the screened categories. Though non-significant, the same trend is shown consistently through the four therapy types.

An interesting contrast appears for education (younger, 0.10, n.s.; older 0.52, $p < 0.05$), which almost reaches $p < 0.10$ significance for heterogeneity, and for expressive supportive therapies (younger, 0.13, n.s.; older, 0.46, $p < 0.05$), which begins to approach heterogeneity. For the latter result it appears that older unscreened patients enjoy much the same level of benefit that screened patients do (screened younger, 0.47, $p < 0.05$; older, 0.37, n.s.). The education result is eye-catching given the modest effect sizes produced by education in the therapy characteristics analyses. Results are very close and very poor for both unscreened groups for CBT. Earlier results show that distress screening is important for the effectiveness of that therapy.

The similarity of the screened effect sizes, in contrast with the difference between the unscreened effect sizes, suggests that screening (i.e. distress or non-complexity) is as important, or more important, than age in moderating effect. It may be that older patients generally suffer elevated distress and that explains the moderation seen in the unscreened sample. The reasons for this may include social isolation and poverty (further discussed below) as well as more serious and complex morbidity. Note that later analyses produced a non-significant trend in favour of more advanced cancer. Or it could be that older people are better able to take advantage of therapy because of greater life skills, similar to non-complex patients.

By way of summary, both age groups in the screened category receive similar levels of benefit from therapy, except for a tendency for older patients to benefit more from relaxation. Older unscreened patients benefit significantly more overall than their younger counterparts, particularly from education and supportive-expressive therapies. CBT is of negligible benefit to either unscreened group. Though older age predicts higher effect generally, distress or non-complexity are also important.

Race

The percentage of patients who were white was coded for each study. Again using the untreated control data, a quick survey of the dataset rows showed that: 48 reported 0% white; 33 a mixture comprising 59 - 79% white; 31, 80 - 89%; 63, 90 - 100%; and 111 rows did not report on the subject. Efforts were made to cut the data in a way that would heighten contrast, but there was such an imbalance of n favouring white patients that analyses were thwarted. Details are relegated to **Appendix O**. It would be helpful if more studies reported on the racial composition of their participants – and if more emphasis was placed on recruiting non-white patients.

Gender

Main effect

Studies were coded on the percentage of participants who were female. Also, where gender based breakout data were presented (two untreated control studies), these data were recorded and in these analyses was substituted for the whole study data.

A stem and leaf plot revealed nine studies with 0% female participants, including five that were not exclusively for prostate cancer patients. At the other end of the distribution, there were 38 studies with 100% female participation, including five that were not exclusively for breast or gynaecological patients. Overall, there were 83 studies in the set, 37 of which were for sex-specific cancers. Disregarding the ends of the distribution (0 and 100%), the central part of the distribution ($n = 36$) was fairly normally distributed, with no studies in the 10 – 19.9% or 90 – 99.9% female ranges.

In order to capture the extremes of the distribution for contrast, whilst still retaining enough n to conduct meaningful analyses, it was decided to exclude the middle block of studies, namely those with 40 – 69.9% female participation, leaving groups comprising mostly men (0 – 39.9% female) and mostly women (70 – 100% female) for analysis. Studies for patients with sex specific cancers were excluded from one calculation ('All tx B' on the table). Holding these studies out removes any bias relating to sex-specific issues. It will also be seen that breast cancer studies produce lower effects, and that of the screened breast cancer studies, all but one screen out rather than in or both ways, which also lowers results. However, therapy types analyses were broken out using data from all cancer sites combined. These results are presented in Table 7-2. Zero and 100% female contrasts were also separately computed, but because of confusion with and possible confounding by sex specific cancer samples, these results are of little interest and were relegated to **Appendix O**.

Table 7-2. Gender

	Screened				
	Men		Women		Q statistic p
	ES (n)	95% CI's	ES (n)	95% CI's	
All tx A	0.69 (7)**	0.39 – 1.00	0.39 (17)**	0.56 – 0.00	0.086*
All tx B	0.91 (5)**	1.22 – 0.00	0.76 (4)**	1.22 – 0.00	0.599
Educ	-	-	0.03 (3)	0.00 – 0.12	1.000
Relax	1.10 (3)**	0.00 – 5.44	0.32 (12)**	0.00 – 2.57	0.001**
CBT	0.36 (3)**	0.00 – 2.26	0.27 (10)**	0.00 – 2.83	0.634
Exp-sup	0.95 (1)**	1.55 – 0.00	0.24 (5)*	0.53 – 0.00	0.039**

Gender categories as defined in the text. All tx A = all therapy types combined, and all cancer sites; All tx B = all therapy types combined, but without data from sex specific cancers (prostate, breast, gynaecological). Q stat p = Q statistic p . Specific therapy type results derive from all

cancer sites (i.e. sex specific data are not removed). Other relevant notes as per Table 7-1.

Table 7-2 continued

	Unscreened				
	Men		Women		Q statistic <i>p</i>
	ES (n)	95% CI's	ES (n)	95% CI's	
All tx A	0.55 (7)**	0.96 – 0.00	0.21 (17)**	0.41 – 0.00	0.150
All tx B	0.55 (4)**	1.10 – 0.00	0.23 (4)	0.67 – 0.00	0.373
Educ	0.31 (3)	0.00 – 1.45	0.45 (5)	0.00 – 1.58	0.685
Relax	0.21 (2)	0.00 – 0.86	0.05 (8)	0.00 – 0.52	0.527
Exp-sup	0.89 (3)**	1.64 – 0.00	0.28 (5)	0.71 – 0.00	0.164

In the screened comparisons the difference reaches or approaches statistical significance in relation to all therapies combined when sex specific cancer data are included (men, 0.69; women, 0.39; Q statistic $p = 0.086$), and in relation to relaxation and expressive-supportive therapies (relaxation: men, 0.1.10; women, 0.32; Q statistic $p = 0.001$; expressive-support: men, 0.95; women, 0.24; Q statistic $p = 0.039$). The latter finding for men is based on only one study but is consistent with the unscreened comparison, at about three times the magnitude of the effect for women. Removing sex specific cancers reduces the difference in results greatly, probably due to the exclusion of breast studies, but the trend still favours men (men, 0.91; women, 0.76; Q statistic $p = 0.599$).

In the unscreened comparison men's effect sizes are about double those for women for both combined treatment comparisons ('All tx A': men, 0.55; women, 0.21; Q statistic $p = 0.150$; 'All tx B': men, 0.55; women, 0.23; Q statistic $p = 0.373$). The combined therapy contrast that includes sex specific data remotely approach significance with men producing twice the effect size of women (men, 0.55; women, 0.21; Q statistic $p = 0.150$). In the result for expressive-supportive therapies, men benefit to a magnitude about three times that of women, though again the breadth of variance is such that the difference only remotely approaches heterogeneity (men, 0.89; women, 0.28; Q statistic $p = 0.164$). The latter result may be due to women supporting each other in somewhat similar fashion more naturally, and it must be remembered that breast cancer (and other sex-specific) studies are included in this calculation. However, it is clear that men stand to benefit greatly from expressive-support.

In all but one of the tabled comparisons, higher effect sizes are attained by men. Both of the combined treatment results (with and without sex specific data) show that men obtain at least moderate benefit from therapy, regardless of screening – an important finding. For both unscreened results relating to women, effect sizes are small. When broken out by therapy type, men always obtain at least a small benefit from therapy, but

women did not in three of eight comparisons. The removal of sex-specific data decrease the gender heterogeneity but leaves the trend in tact, with unscreened men still obtaining twice the benefit from therapy generally. It may be that the lesser death threat associated with prostate and breast cancer has some moderating effect of its own. Overall, the pattern of the data conveys the impression that gender moderates the effectiveness of therapy, and indicates greater research emphasis on men in the future.

Education level

Studies were coded into four categories for the predominant level of education of their participants: 1. No formal schooling; 2. Primary / elementary; 3. Secondary / high school; 4. Tertiary / college (whether trade school or university, though usually it transpired to be the latter). Unfortunately, insufficient studies reported on this variable, thwarting analysis. Details of the attempts made, including attempts to examine the impact of this variable on education and CBT therapies, are therefore relegated to **Appendix O**.

Occupation / income

Main effect

Studies were coded on four levels for predominant level of occupation or income: 1. Very low: unemployed / very low income (< US\$25k); 2. Low: un/semi skilled labourer / homemaker / retired / low household income (US\$25 - 49k); 3. Medium: skilled labourer / self employed / medium income (US\$50 - 74k); and, 4. High: professional / business / managerial / high income (> US\$75k). The US dollar categorisations were lifted from one of the primary studies (Scheier et al., 2005), and allowance was made for studies before and after its date. Despite efforts to define this categorisation, in application it is nonetheless a rough estimate. Many studies did not report on this variable at all, or reported employment status instead (employed / unemployed, part-time, at home, or retired). This information was not useful except where a majority of participants were reported as at home or retired. There were no studies in the very low income category in any analysis. The number of studies available to contribute data for this analysis was consequently similar to that for education, so the same analysis strategy was used.

The table below (Table 7-3) presents data on the 2 x 2 matrix of study design moderators, using omnibus outcomes and combined therapies once again, and n is fairly evenly spread. It is immediately apparent that in every cell for which there were data, the effect size decreases as occupation / income rises. Only in the screened / untreated control quadrant does the progression reach statistical heterogeneity (at the 0.10 level - Q statistic $p = 0.068$), but the pattern of Q statistic p 's over the whole matrix accords with what would be expected of a moderator variable, i.e. lowest p where the higher scoring confounds collide (top left quadrant), highest in the diagonally opposite corner, and in between in the other two. The Q statistic p in the screened / untreated control quadrant is derived from comparison of adjacent levels of the variable of interest, and

could be expected to be lower if there were high income data available for comparison with the low income level. Given the n available, these patterns of effect size progressions and Q statistics display moderation perfectly.

Table 7-3. Occupation / income

		Untreated control		Treated control	
		ES (n)	95% CI's	ES (n)	95% CI's
Screened	Low	0.90 (4)**	0.40 – 1.40	0.59 (3)	-0.23 – 1.42
	Medium	0.32 (3)*	-0.05 – 0.69	0.12 (3)	-0.14 – 0.39
	High	-	-	-	-
Q statistic p		0.068*		0.289	
Unscreened	Low	0.43 (6)	-0.13 – 0.99	0.13 (3)	-0.15 – 0.41
	Medium	0.18 (6)	-0.06 – 0.43	-0.02 (4)	-0.21 – 0.18
	High	-0.07 (2)	-0.40 – 0.26	-0.04 (7)	-0.15 – 0.08
Q statistic p		0.267		0.549	

Relevant notes from Table 7-1 apply. Occupation / income levels are defined in the text.

Educational therapies

It could be hypothesised that patients in low income groups would also be more likely to suffer from a lack of information about the disease and its physical and psychological management and may therefore benefit more than others from educational interventions. The next analysis investigated this using data from studies of education therapies only (Table 7-4). N is sparse and there are several empty cells, but it is notable that such data that do exist generally keep to the pattern noted already, of effect sizes decreasing with a rise in occupation / income.

Table 7-4. Occupation / income, educational therapies

		Untreated control	Treated control
		ES (n)	ES (n)
Screened	Low	-0.42 (1)	0.88 (2)
	Medium	-	0.02 (1)
	High	-	-
Q statistic p		1.000	0.408
Unscreened	Low	0.43 (2)	0.13 (2)
	Medium	-0.06 (1)	0.00 (4)
	High	-	-0.04 (4)

Q statistic <i>p</i>	0.268	0.632
-----------------------------	-------	-------

Relevant notes from Table 7-1 apply. Occupation / income levels are defined in the text.

CBT

Finally, CBT is singled out for investigation on similar grounds, i.e. because it may be suggested that lower income patients face more complex problems when the burden of cancer is added to their limited financial means, and therefore stand to gain more from therapies that help address these problems (Table 7-5). N is low and particularly sparse under treated control, but again, over all quadrants we see consistent patterns of effect sizes falling away as occupation / income rises, and Q statistic *p*'s relating to each other as would be expected where there is moderation. The result in the screened x untreated control quadrant is heterogeneous (Q statistic *p* = 0.049) though, of course, based on very low n.

Table 7-5. Occupation / income, CBT

		Untreated control	Treated control
		ES (n)	ES (n)
Screened	Low	1.20 (2)**	0.27 (1)
	Medium	0.32 (3)*	0.19 (1)
	High	-	-
Q statistic <i>p</i>		0.049**	0.780
Unscreened	Low	0.63 (3)	-
	Medium	0.43 (2)	-
	High	-0.06 (1)	-0.06 (4)
Q statistic <i>p</i>		0.664	1.000

Relevant notes from Table 7-1 apply. Occupation / income levels are defined in the text.

Conclusion

Despite low n, a consistent inverse relationship between occupation / income and effect size is seen through main effect, education therapy, and CBT analyses. This moderating relationship was formalised by a finding of statistical heterogeneity in relation to screened participants and untreated controls in the main effect and CBT analyses. Given the limited n available, this is strong evidence that economic status moderates psycho-oncological therapy effectiveness generally.

Marital status

Main effect

As explained in the Introduction chapter, marital status is a deficient proxy for social support since it indicates nothing as to the quality of the relationship, in particular the social constraint or freedom enjoyed by the patient, but had to suffice since it is the only indicator for which data are regularly reported. Because of its limited value in representing the effective moderator, a muting of any association between marital status and therapy effectiveness was expected. Further distance from the real variable of interest would result from the need to use study average percentages as the basis for dividing the data, and the limited downward range in the distribution of the variable (nearly all studies had more than half of participants married).

The dataset (untreated controls only) was of fairly normal distribution. Studies had been coded for the percentage of married participants, and the spread ranged from 47% to 100% with a median at 73.4%. A comparison based on median split data lacked sufficient contrast to be of interest - although the trend was in the expected in direction – and was therefore relegated to **Appendix O**. A further comparison was made of studies in the upper and lower quartiles in an attempt to sharpen contrast and hoping that the spread of studies would be adequate over screened and unscreened breakouts (Table 7-6). This transpired to be so, and the outcome showed effect sizes of about double for the level with fewer married patients in both screened and unscreened groups, though the difference did not approach statistical significance (screened: over 79.9% married, 0.38; under 64.1% married, 0.64; *Q* statistic $p = 0.535$; unscreened: over 79.9% married, 0.22; under 64.1% married, 0.45; *Q* statistic $p = 0.561$). The lack of statistical significance was expected, due to the bluntness of the tool (as described above). That considered, the finding of about double the effect size from studies with fewer married patients is really very important.

Table 7-6. Marital status, quartile contrast

	Screened		Unscreened	
	ES (n)	95% CI's	ES (n)	95% CI's
Over 79.9% sample married	0.38 (5)**	0.70 – 0.00	0.22 (6)	0.56 – 0.00
Under 64.1% sample married	0.64 (4)	1.43 – 0.00	0.45 (5)	1.15 – 0.00
<i>Q</i> statistic <i>p</i>	0.535		0.561	

Relevant notes from Table 7-1 apply.

Terminal phase

Two specific questions were asked of the median split data (i.e. data split around 73.4% of participants in a study married), in the hope that there would be enough *n* in that dataset to provide some indication of trend. First, what impact does marital status have at terminal phase, when it may be argued that pressures from the disease (pain,

debilitation, and existential issues) cause most isolation and considerable distress? The recurring / palliative set was used for this, but, as might have been expected, there were no studies that had more than 73.4% of participants who were married or partnered in that set, so no analysis could be made. However, the lack of studies in that category itself highlights the social need of this group of patients and may help explain the higher effects they produce in later analyses.

Expressive-supportive therapy

The second question asked of the median split data was whether expressive-supportive therapies – themselves providing a form of social support - are more effective for unmarried patients (Table 7-7Table 7-7). Results for screened patients were hampered by lack of n in the under 73.4% married category. Disregarding the screened column therefore, the trend for unscreened patients is much the same as that found in the overall quartile comparison, i.e. the sample with more married patients produced an effect size about half of that produced by the sample with fewer married patients, although the difference does not approach statistical significance (over 73.4% married, 0.27; under, 0.46, Q statistic $p = 0.532$). The importance of expressive-support for unmarried patients is underlined by the fact that similar effect sizes from the comparison of combined therapy data were produced only by the sharper upper and lower quartile comparison above – not the median split comparison.

Table 7-7. Marital status, expressive-supportive therapy

	Screened	Unscreened
	ES (n)	ES (n)
Over 73.4% of sample married	0.35 (4)**	0.27 (4)
Under 73.4% of sample married	-0.12 (1)	0.46 (4)*
Q statistic p	0.078*	0.532

Relevant notes from Table 7-1 apply.

Conclusion

These analyses show a statistically small but theoretically important trend whereby marital status influences therapy effectiveness, both generally and especially for expressive-supportive therapies. There were insufficient data to make an analysis relative to terminally ill patients.

Socio-demographic variables conclusion

From these analyses it can be seen that socio-demographic variables have a considerable moderating influence on the effectiveness of psycho-oncological therapies. Age has been seen to moderate the effectiveness of therapy generally for unscreened patients,

such that older patients benefit to a moderate magnitude while for younger ones the benefit is negligible. This difference can be traced back to education and expressive-support rather than to relaxation and CBT where the outcomes are similar between age groups. The difference may relate to the non-significant trends later found relating to marital status and cancer stage such that unmarried and distant spread patients benefit more, and the statistically significant finding that patients with lower income (including retirement income) benefit more from therapy. Older patients can easily fall into more than one of these vulnerable categories. These factors may lift their overall distress, and it may well be that it is heightened baseline distress which is the true moderator at work.

The gender result was both interesting and complex. Gender produced formal heterogeneity for screened patients ($p < 0.10$) and approached it for unscreened patients, with men yielding considerably higher effect sizes. However, further analyses filtering out sex-specific cancers left no statistically significant contrasts, although the tendency still favoured men, and, of course, n was considerably reduced. Screened men were found to benefit considerably and significantly more than screened women from supportive-expressive and relaxation therapies, with gains of high magnitude. The difference that approached statistical significance for unscreened expressive-support also favoured men, who still achieved statistically significant ($p < 0.05$) and moderate magnitude results, despite their unproven distress. It may be that the overall trends favouring men indicate that they generally have more to gain from therapy, particularly expressive-supportive therapy, and that women are better at supporting themselves. The finding of high and statistically significant effect sizes ($p < 0.05$) for both genders in relation to non sex-specific cancers is important and hints at a finding relative to cancer type, yet to come.

Occupation / income level proved to be influential, showing a consistent inverse trend with effect size over all confound combination quadrants where data from all treatment types were combined, including formal heterogeneity in the screened / untreated quadrant. The same pattern was borne out in analyses of education and CBT therapies, where it could be argued that higher socio-economic groups would have more ready access to alternative similar support. These results are consistent with the expectation that poorer people are less buffered from the impacts of trauma and are therefore more likely to be distressed and likely to benefit more from therapy.

A small and non-significant trend was found in favour of unmarried / unpartnered patients. The weakness of this trend was expected given the inadequacy of marital status as a proxy for social support, so the trend – which was heightened in relation to expressive-supportive therapy - is actually quite theoretically important.

The socio-demographic factors to remember generally are therefore: older, poorer, single, and men – with greatest emphasis on economic status and least on marital status.

Baseline distress and other screening

Baseline distress: Main effect

Evidence of the importance of baseline distress to therapy effectiveness was presented with the preliminary analysis results. The two tables evidencing this are reproduced here for convenience (Table 7-8 and Table 4-16). That finding was that patients whose entry into the research therapy was conditional upon having a clinically significant level of baseline distress (‘screened in’ recruits) achieved significantly higher effect sizes than those admitted to unscreened studies, and scores that were also higher, though not significantly so, than those admitted to studies that ‘screened out’ recruits for history of psychological distress.

Table 7-8. Screening at recruitment

	Anxiety		Depression		Distress	
	ES (n)	95% CI's	ES (n)	95% CI's	ES (n)	95% CI's
Screened in	0.70 (6)	0.21 – 1.20	0.68 (8)	0.18 – 1.18	0.52 (7)	0.09 – 0.94
Screened out	0.34 (24)	0.17 – 0.51	0.24 (17)	0.06 – 0.42	0.32 (20)	0.15 – 0.48
Unscreened	0.11 (49)	0.00 – 0.23	0.12 (41)	0.02 – 0.22	0.05 (44)	-0.04 – 0.14
Not reported	0.39 (4)	0.08 – 0.69	0.46 (4)	0.15 – 0.77	0.66 (2)	0.29 – 1.02
Q statistic <i>p</i>	0.020**		0.029**		0.000**	

Both treated and untreated control comparison studies were used. ‘Screened in’ studies included those that simultaneously screened out. Studies that did not report clearly on this variable were excluded. Other relevant notes as per Table 7-1.

Table 7-9. Screening at recruitment, pairwise comparisons for heterogeneity

	Anxiety		Depression		Distress	
	Out	In	Out	In	Out	In
In	0.174		0.105		0.385	
Unscreened	0.033	0.023	0.250	0.031	0.006	0.034

Q statistic *p*'s are presented. No shading indicates significance at $p < 0.05$; light shading indicates p between 0.10 and 0.20. In = recruits were included (screened in) only if they showed clinically significant distress at baseline; Out = recruits were excluded (screened out) for history of psychological distress (including studies that screened both ways); Unscreened = recruits were not screened in either way.

However, for these calculations the ‘screened in’ categorisation included data from studies that simultaneously screened out for psychological history. It was seen in the therapy types results that the latter group can produce very high effect sizes. Further computations breaking apart data from these two screening levels were therefore undertaken in order to check whether distress screening on its own could sustain

heterogeneity from no screening. This proved to be so. Results are reported as part of the discussion of simultaneous screening, below (Table 7-10 and Table 7-11).

Baseline distress: Therapy type finding

Further evidence of the impact of baseline distress on effect sizes was apparent in the chapter on therapy type results where studies that screened in produced higher scores than those that screened out or that were unscreened in the untreated control / early times quadrant of every therapy type matrix in which screened in data appeared. Studies that screened in and out simultaneously produced even higher scores. There were no education therapy studies that screened in for distress, but examination of two particular studies suggested that baseline distress may predict better results for that therapy as well.

Baseline distress: Breakouts within primary studies

It was hoped that the issue could be assessed from another angle also, namely by using differential baseline distress breakout data provided by some studies. Unfortunately only four studies provided useable data of this nature, and after accounting for the two confounds that were known to structure the dataset, this did not leave sufficient n for analysis.

It was interesting to find, however, that two of the four studies reported effects in the expected direction and two in the opposite direction. One of the latter studies was Krischer, Xu, Meade, and Jacobsen (2007). In that study, the text relating to the distress breakout results is in conflict with the direction of the tabled results, suggesting a mislabelling of the table, and no inconsistency with the usual trend that more distressed patients produce higher effect sizes. The other study was Given et al. (2004) in which intervention patients with a variety of solid tumours undergoing chemotherapy were administered 10 sessions of CBT over 20 weeks. The higher distress group was defined by depression score, and the authors found not only that they were made more distressed by the burden of the therapy, but also that where improvement in depression scores was found overall in the whole intervention group, that improvement was mediated by improvement in non-affective-related physical symptoms rather than affective-related symptoms. The authors suggest that the pathway to lowering depression was therefore through non-affective symptom management, and also that patients who were more depressed at baseline may have found the burden of trying to learn and implement new symptom management strategies at the same time as coping with chemotherapy was too much and counterproductive. Therapists must take into account the amount of burden that depressed patients can carry.

Simultaneous screening

Two analyses were performed, using untreated control data, breaking out the effect of simultaneous screening in order to show the impact of this possible interaction (Table 4-16) and to test it for heterogeneity against other levels of the screening variable (Table 7-10). Using omnibus data there is sufficient n to comfortably reach heterogeneity between the screening levels on the effect size breakout table ($p = 0.24$, see the

continuation section of Table 7-8). Results from the three standard outcomes show overall heterogeneity for distress and depression, but at only the lower level for the latter ($p < 0.10$), and it is only approached for anxiety. The number of studies that screened both ways was very small (four in all, and only two under anxiety and distress outcomes) but the consistency of the pattern makes this phenomenon notable.

Table 7-10. Screening at recruitment, simultaneous broken out

	Anxiety		Depression	
	ES (n)	95% CI's	ES (n)	95% CI's
Screened in and out	1.34 (2)	0.21 – 2.46	1.07 (4)	0.30 – 1.85
Screened in	0.73 (3)	0.26 – 1.20	0.51 (4)	0.05 – 0.96
Screened out	0.37 (12)	0.12 – 0.61	0.12 (9)	-0.13 – 0.38
Unscreened	0.33 (22)	0.13 – 0.52	0.25 (12)	0.04 – 0.46
Q statistic <i>p</i>	0.163		0.093*	

Untreated control data. Other notes as per Table 7-1.

Table 7-10 continued

	Distress		Omnibus	
	ES (n)	95% CI's	ES (n)	95% CI's
Screened in and out	1.08 (2)	0.21 – 1.95	1.13 (4)	0.57 – 1.68
Screened in	0.60 (4)	0.26 – 0.95	0.59 (4)	0.22 – 0.95
Screened out	0.39 (12)	0.14 – 0.65	0.41 (20)	0.23 – 0.58
Unscreened	0.19 (17)	0.03 – 0.35	0.30 (33)	0.15 – 0.44
Q statistic <i>p</i>	0.037**		0.024**	

Again looking at the omnibus outcome results, the pairwise comparison table Table 4-16) locates the specific heterogeneous screening comparisons as between unscreened studies and those that screen in or both ways, and between those that screened out and those that screened both ways (all $p < 0.05$). This means that screening both ways produces results that are so high that they significantly differ from all other results except those that screen in for baseline distress – although that result also approaches formal heterogeneity ($p = 0.113$). This result is consistent with the assumption made originally (when studies that screened both ways were categorised 'screened in' for the purposes of preliminary analyses) that distress was the more influential factor in the apparent interaction. Studies that simply screen in for distress remain heterogeneous from unscreened studies ($p = 0.40$) on this smaller dataset, as for the whole dataset as

was used in the preliminary analyses. It is seen that the two levels that are least distinct from each other are the unscreened and those that screen out ($p = 0.347$).

Note also that having used only untreated control data, the screened out effect size shows homogeneity with both the screened in and the unscreened studies (omnibus Q statistic p 's of 0.166 and 0.347, respectively) meaning that this category could technically be grouped with either for statistical analysis of this subset of data, though it is, of course, closer to the screened in result. In the preliminary investigations based on the whole dataset this category was statistically homogeneous only with the screened in category. However, it is now seen that for the divided dataset it would be statistically valid to combine data from the screened out and unscreened categories, and for one analysis where the characteristics of the data call for it (cancer site, below) this is done.

Table 7-11. Simultaneous screening, pairwise comparison

	Anxiety			Depression		
	Screened in & out (2)	Screened in (3)	Screened out (12)	Screened in & out (4)	Screened in (4)	Screened out (9)
Screened in	0.326			0.218		
Screened out	0.099	0.186		0.023	0.152	
Unscreened	0.083	0.125	0.186	0.045	0.321	0.444

Q statistic p 's are presented. Numbers in brackets are study frequencies. No shading indicates significance at $p < 0.05$; light shading indicates $p = 0.05 - 0.10$.

Frequencies for unscreened are: omnibus, 33; anxiety, 22; depression, 12; distress, 17.

Table 7-11 continued

	Distress			Omnibus		
	Screened in & out (2)	Screened in (5)	Screened out (12)	Screened in & out (4)	Screened in (5)	Screened out (20)
Screened in	0.314			0.113		
Screened out	0.135	0.164		0.016	0.166	
Unscreened	0.047	0.005	0.183	0.005	0.040	0.347

CBT

In an untabled analysis, CBT studies that simultaneously screened both ways were tested for heterogeneity against those that screened only one way or the other. Omnibus data were used to maximise n and focus on the therapy effect rather than any particular outcome. The effect size produced by studies that screened both ways (untreated

controls, $g = 1.13$, $n = 4$) was compared with that for, first, those that simply screened in (0.42 , $n = 3$), and second, those that simply screened out (0.23 , $n = 10$). The first comparison produced a Q statistic p of 0.023 and the second, 0.003 , despite the low n involved. This means that CBT is particularly sensitive to simultaneous screening, since an interaction between the two types lifts the effect size to a level that is statistically significantly higher than either of the single types of screening, even on low n . It may be that CBT is a particularly demanding therapy, therefore the combination of distress-driven motivation and background psychological strength gives a particular advantage.

There was insufficient data to run a similar calculation for other therapy types, but it is noted from earlier analyses that the effect sizes from simply screening in for expressive-support were a good deal higher than for CBT, so it may be a therapy with a quite different dynamic in this regard.

Screening out patients with psychological history

In the therapy trajectory results a thread of evidence was detected suggesting that this type of screening may produce more durable effects. This was noted in relation to all outcomes for relaxation, and under depression and distress for CBT and expressive-supportive therapies. Also, it has just been noted that in combination with screening in for distress, screening out may cause a powerful interaction, at least in relation to CBT.

Although screening out was presumably done for the purpose of providing scientific homogeneity or participant safety (e.g. where the therapist was less highly trained or therapy was to be delivered in group mode), comparison of outcomes based on this variable would now be valuable for what it could tell us about the nature of psychopathology and optimal conditions for psychotherapy. It would be helpful if intervention study results were broken out and analysed around whether or not the patient had a history of psychological diagnosis.

Baseline distress: Conclusion

Screening for baseline distress alone – that is, not in combination with screening out for psychological history – generally produces higher results than those for unscreened studies. Omnibus data confirm this moderation. How much higher the results are may be influenced by therapy type and whether or not the sample was also screened out for psychological history. CBT studies that simultaneously screened out produced results that were significantly higher than those that only screened in for baseline distress, which may be due to the demanding and structured nature of this therapy. It is to be noted in relation to depression, that the study by Given et al. (2004) shows that outcomes for patients distressed at baseline are not universally positive. It appears that there can come a point where the burden of distress can become overwhelming rather than motivating in the face of a demanding therapy regime.

Medical variables

The impact of a range of disease and medical treatment variables was explored using the same approach as for other analyses in this chapter: cancer site; prognosis associated with site; cancer stage; medical protocol; and medical treatment stage.

Cancer site

Studies were coded by predominant cancer site of participants, namely breast, prostate, skin (melanoma), gynaecological, colorectal, mixed sites (i.e. where it would be misleading to assert the predominance of any one type over others) or other single type. The latter category comprised six studies: Hepworth (2004), head and neck; Katz, Irish, and Devins (2004), oral; Goldberg and Wool (1985), lung; Ali and Khalil (1989), bladder; Moynihan, Bliss, Davidson, Burchell, and Horwich (1998), testicular; and Lin, Tsang, and Hwang (1998), hepatocellular. However, only one of these (Goldberg & Wool, 1985) used an untreated comparison group and was therefore included in all of the analyses. In order to get the fullest picture of the dynamics in relation to this basic medical variable, effect sizes for both treated and untreated controls were therefore investigated. Results in the form of the 2 x 2 confound matrix are presented below (Table 7-12).

Table 7-12. Cancer site

	Predominant cancer site	Untreated control		Treated control	
		ES (n)		ES (n)	
Screened	Breast	0.29 (13)**	0.11 – 0.46	0.09 (4)	-0.10 – 0.28
	Prostate	0.37 (2)*	-0.03 – 0.76	-	-
	Melanoma	0.83 (2)	-0.34 – 2.00	0.29 (1)	-0.22 – 0.79
	Colorectal	1.10 (2)**	0.62 – 1.57	-	-
	Gynaecological	-	-	0.75 (1)**	0.18 – 1.31
	Mixed	0.80 (9)**	0.45 – 1.14	0.38 (5)*	
	Other	-	-	1.94 (1)**	1.09 – 2.79
	Q statistic <i>p</i>		0.005**		0.000**
Unscreened	Breast	0.20 (13)	-0.04 – 0.45	-0.04 (20)	-0.17 – 0.09
	Prostate	0.56 (3)	-0.17 – 1.28	0.18 (4)	-0.17 – 0.53
	Melanoma	-	-	-0.16 (2)	-0.88 – 0.56
	Colorectal	-	-	-	-
	Gynaecological	-	-	0.17 (2)	-0.09 – 0.43
	Mixed	0.30 (13)**	0.09 – 0.51	0.07 (14)	-0.09 – 0.22

	Other	0.67 (1)	-0.20 – 1.53	0.15 (1)	-0.71 – 1.01
	Q statistic <i>p</i>	0.629		0.639	

Cancer sites are by predominance within the primary study. Early times assessment points were used. Where it was unclear whether a study fitted a certain category, it was excluded. Other relevant notes as per Table 7-1.

An imbalanced distribution in favour of breast cancer studies is immediately apparent (breast studies, n over all four quadrants = 50, i.e. 44% of the tabled studies; mixed sites, n = 41, i.e. 36%; other individual sites, n = 22, i.e. 20%) and it will be seen that most breast studies were early stage.

Heterogeneity was found in the screened / untreated control quadrant (Q statistic $p = 0.005$) and so another analysis was run between the three sites with larger frequencies – breast, prostate and mixed sites - to try to locate the difference. The breast ($g = 0.29$, $p < 0.05$, $n = 13$) and prostate (0.37 , $p < 0.10$, $n = 2$) comparison yielded a non significant Q statistic p of 0.732, breast and mixed (0.80 , $p < 0.05$, $n = 9$) yielded a significant Q statistic p of 0.010, and prostate and mixed sites came close to significance with 0.108, locating the heterogeneity between breast and mixed. Note that the relatively poor effect size for prostate in this quadrant is based on low n and is not consistent with results in the screened quadrants. It would be interesting to see how prostate would relate to breast and mixed site categories if it had higher frequency.

The heterogeneity found in the screened / treated control quadrant (Q statistic $p = 0.000$) was investigated as between the two higher n categories, breast (0.09 , n.s., $n = 4$) and mixed (0.38 , $p < 0.10$, $n = 5$), with a result that did not reach heterogeneity but began to approach it (Q statistic $p = 0.182$). The Q statistic tabled will relate to breast cancer and one of the very small n results and is therefore meaningless. Heterogeneity was not found in the two unscreened quadrants.

Breast cancer yields the lowest results in three of four quadrants, with a negative result in the fourth (unscreened / treated control). Taken together, the consistency with which other cancer sites yield higher effect sizes than breast, the heterogeneity seen in the comparison of breast with mixed site studies in the screened x untreated quadrant, and the fact that the significantly higher scoring ‘mixed’ category is made up of no single predominant type and therefore is comprised mainly of cancers other than breast cancer, convey the impression that interventions targeting breast cancer patients are the least effective relative to those directed towards other sites. However, it was suspected that breast-predominant studies may be of a different nature than other site studies, and so the frequencies of the different levels of screening were checked. Frequencies are tabulated below (Table 7-13).

Table 7-13. Cancer site, screening level frequencies

		Breast			Mixed	Other single site	Not reported	Total	
		n	% relative to row(s)		n	n	n		
Untreated controls	Screened in & out	1	25	13	46	2	1	-	4
	Screened in	-	0			4	-	-	4
	Screened out	12	60			3	5	-	20
	Unscreened	13	39			13	4	3	33
	Totals	26	42			22	10	3	61
Treated controls	Screened in & out	-	-			-	-	-	0
	Screened in	-	-			1	-	-	1
	Screened out	4	36			4	3	-	11
	Unscreened	20	47			14	9	-	43
	Totals	24	44			19	12	0	55

Studies that did not report clearly on screening or nature of control were excluded.

The data in Table 7-13 show that the composition of the screened breast cancer studies was imbalanced in that only one screened in for baseline distress out of the whole dataset. The bias towards low scoring screened out and unscreened studies had to be addressed in order to enable a fair comparison with other studies, so it was decided to run a comparison that excluded all studies in the dataset that screened in or both ways. In order to sharpen the contrast between breast and other cancer sites, those studies that used mixed samples, which likely included a sizable proportion of breast patients although they were not in the majority, were also culled. To consolidate n from the remaining non-breast studies (i.e. all other predominantly single site studies), their data were combined. The power of the analysis was further intensified by combining screened out and unscreened data. This choice was made in the knowledge that those two levels of screening are close enough to lose their statistical heterogeneity when the dataset is divided by treated and untreated control (refer findings under simultaneous screening, above) and that frequencies of these two screening types were fairly evenly distributed between breast and other single site studies. Using both untreated and treated control data again, the analysis results are shown in Table 7-14.

Table 7-14. Cancer site, breast v. other single sites

Control type	Breast		All other sites		Q statistic <i>p</i>
	ES (n)	95% CI's	ES (n)	95% CI's	
Untreated	0.23 (25)**	0.07 – 0.39	0.57 (9)**	0.28 – 0.85	0.042**
Treated	-0.01 (24)	-0.12 – 0.10	0.24 (12)*	-0.03 – 0.51	0.090*

Studies that screened in or screened both ways are excluded. Mixed cancer site studies are excluded. Screened out data are combined with unscreened data. Other applicable notes as per Table 7-1..

These results show formal moderation by cancer site, with breast cancer patients returning significantly lower effect sizes than those with cancer at other single sites (Q statistic *p*, untreated control, 0.042; treated, 0.090). It is noted also that for untreated control data the effect size for breast studies is small and less than half that for other sites. For treated control data the breast study effect size is null, and the only non-significant result tabled, while the result for other single sites retains 0.10 level statistical significance, despite being supported by only half the n of the breast category. This means that compared with the benefit patients could expect to receive from placebo or treatment as usual, there is no impact on distress at all for non-complex or unscreened breast patients, while patients with cancer at other single sites would, on average, receive a small effect. Compared with no treatment controls, breast patients receive a small average effect while other patients receive a moderate one.

Since there are so many breast studies that sample only early stage patients, it was thought prudent to check the cancer staging of the studies used to compute this important finding. For studies with untreated control comparisons, the breast studies comprised 18 local, 1 distant, and 4 mixed stage studies, with 2 unreported, while other single site studies comprised 5 local, 1 mixed, and 3 not reported. This makes 72% of breast studies local, compared with 55% of other single site studies. For treated controls, breast: 17 local, 1 regional, 2 distant, and 2 not reported; other single sites: 5 local, 4 mixed, 3 not reported; 71% versus 42% are local. These calculations include the studies that did not report on stage, which had more impact on the calculations for other single sites because of their smaller n and their larger proportions of studies that failed to report. So, roughly three quarters of the breast studies were local spread (early stage) cancer, and roughly half of the other single site studies were.

In the analysis of cancer stage below, the unscreened untreated control data (which are the most comparable to the data mix used in the present untreated control analysis) shows a non-significant tendency for local spread studies to produce smaller effect sizes. This suggests the possibility of a small bias against breast studies in the above analysis. Whilst the cancer stage result is only a tendency, and there is not a large gap between the percentages of local spread studies in the untreated control data above, it would be wise to qualify the present finding with the note that most of the breast studies

that contributed to it were early stage, and not so many of the comparison single site studies were.

Other than the small bias that may result from differences in cancer stage, the above finding may reflect higher degrees of distress experienced by patients with other types of cancer, since breast cancer has a relatively good prognosis and, although it poses a threat to body image acutely felt by some sufferers, physical functioning may be relatively less affected, in the long term, by medical treatment. Or it may be that breast cancer is more socially understood and acceptable, and is often better catered for by psychological and voluntary services, its patients having greater numbers and political strength than other cancer patients. It may be that the numbers of women who get this disease are such that they are able to support each other to a degree that other cancer site sufferers cannot. This kind of support, often facilitated by voluntary organisations, can be quite substantial and of both an educational and expressive-supportive nature. Any or all of these factors may reduce the distress and social isolation that breast cancer patients experience on average and relative to other cancer sufferers, and therefore reduce effect sizes and create statistical contrast with data from other patients.

Referring back to Table 7-12, the mixed cancer site result in the screened / untreated control quadrant is eye catching ($g = 0.80, p < 0.05, n = 9$). However, it was found that the data contributing to this were bolstered by disproportionate numbers of screened in and simultaneously screened studies (n 's of 4 and 2, respectively) compared with breast and all other sites data which had only one each of these higher scoring studies (Table 7-13). There was only a one study bias of this nature in the screened / treated results quadrant, and the unscreened results are not affected. In those three quadrants the mixed results fall into the position that would be expected given that they probably include a good number of breast patients, i.e. between the breast effect size and those of other single sites.

Cancer prognosis

Whether the prognosis of the predominant cancer site of a study sample was favourable, guarded, or dismal according to five year survival rates (categories ex Andersen, 1992) supplemented with information from the American Cancer Society, 2007) was also coded and analysed. Note that this categorisation is not nearly as sharp as one based on individual patient prognoses, for which data were not available. Because so few studies appeared in guarded and dismal cells, treated control studies were again brought into the analysis. The results are set out below (Table 7-15).

The screened / untreated control quadrant produces statistical heterogeneity between the only two cells with n - favourable and guarded prognosis cells (Q statistic $p = 0.003$). Results in the unscreened / treated control quadrant come close to heterogeneity (0.105). However, both of these results are hampered by lack of n in the guarded category, and the complete absence of dismal prognosis studies, so can only be taken as suggestive. A consistent trend is not found in every quadrant, though effect sizes are lowest for patients with favourable prognoses in three of them.

This analysis is frustrated by low n in the guarded and dismal prognostic categories. This fact in itself suggests, however, a crying need for more research with those populations. If the focus on favourable diagnostic groups shown in the frequencies in this analysis (a ratio of 66 favourable prognosis studies to seven guarded and four dismal) is a true reflection of the focus of research in this field then the lack of guarded and dismal n is confronting, although it may be due to ethical restrictions.

Table 7-15. Cancer prognosis

	Cancer type prognosis	Untreated control		Treated control	
		ES (n)	95% CI's	ES (n)	95% CI's
Screened	Favourable	0.34 (17)**	0.17 – 0.51	0.31 (6)*	-0.05 – 0.67
	Guarded	1.10 (2)**	0.62 – 1.57	0.75 (1)**	0.18 – 1.31
	Dismal	-	-	0.37 (1)	-0.43 – 1.17
	Q statistic <i>p</i>	0.003**		0.437	
Unscreened	Favourable	0.25 (16)**	0.02 – 0.48	-0.02 (27)	-0.14 – 0.09
	Guarded	0.20 (2)	-0.28 – 0.68	0.22 (2)	-0.05 – 0.48
	Dismal	0.61 (3)**	0.19 – 1.03	-	-
	Q statistic <i>p</i>	0.299		0.105	

Prognoses groupings are by predominance within the primary study and categories were drawn from Andersen (1992) supplemented with five year survival rate information from the American Cancer Society (2007). Other relevant notes as for Table 7-1.

Cancer stage

Cancer stages were again based on Andersen (1992) and placed stage I or II (and stage III and *in situ* breast cancer) in the 'local' category, stage III and first recurrence for initially stage I disease under 'regional spread', and stage IV together with recurring regional disease and late or advanced disease under 'distant spread'. Studies that did not contain a predominance of any one stage were left out of the analysis. Results for treated controls were again calculated and tabled (Table 7-16).

Table 7-16. Cancer stage

Predominant cancer stage		Untreated control		Treated control	
		ES (n)		ES (n)	
Screened	Local focus	0.51 (16)**	0.29 – 0.74	0.32 (4)**	0.10 – 0.55
	Regional spread	-	-	-	-
	Distant spread	0.23 (4)	-0.15 – 0.61	0.37 (1)	-0.43 – 1.17
	Q statistic <i>p</i>	0.207		0.913	

Unscreened	Local focus	0.23 (13)*	-0.04 – 0.49	-0.01 (19)	-0.17 – 0.16
	Regional spread	-	-	0.10 (2)	-0.34 – 0.54
	Distant spread	0.52 (6)**	0.08 – 0.95	0.03 (2)	-0.28 – 0.35
	Q statistic <i>p</i>	0.266		0.889	

Stage groupings are described in the text and are based on (Andersen, 1992)*. Mixed studies were excluded from the analysis.

Early times assessment points were used. Where it was unclear whether a study fitted a certain category, it was excluded. Other relevant notes as for Table 7-1.

Once again results are challenged in terms of the distribution of n: only one quadrant picked up any regional spread studies and two had less than three distant spread. No quadrant showed statistical heterogeneity. In the untreated control column the effect sizes between screened and unscreened studies almost perfectly invert themselves (screened: local, 0.51, $p < 0.05$; distant, 0.23, n.s.; unscreened: local, 0.23, $p < 0.10$; distant, 0.52, $p < 0.05$) prompting pondering as to cause given that n is reasonable (in those cells that have any at all). Perhaps it has something to do with how distressed distant spread patients are coping with their prognosis, or it may be due to inadequacies in treating their particular needs, but it could just be a result of the grouping of screened in and screened out studies together. It is also noteworthy that unscreened distant spread patients received such substantial benefit from therapy. Based on an n of six, this result warrants attention from both clinicians and researchers as it again draws attention to patients suffering more advanced stage / poorer prognosis disease.

Distant spread issues

Despite low available n, some specific questions were asked of the data relating to distant spread patients. First, is any particular therapy type more effective for them (Table 7-17)?

Although only hints can be taken from an analysis with such low cell frequencies, it would appear that the full range of therapies is useful to distant spread patients and, it may be that relaxation therapies are of particular value (unscreened, $g = 0.58$, $p < 0.05$), although there is no way to test the significance of the different effect sizes of this statistically dependent data.

Table 7-17. Distant spread, therapy type

	Education	Relaxation	CBT	Expressive-supportive
	ES (n)	ES (n)	ES (n)	ES (n)
Screened	-	-0.12 (1)	0.28 (3)	-0.01 (2)
Unscreened	0.33 (4)	0.58 (4)**	0.38 (1)**	0.40 (3)**

Untreated control data. Other relevant notes as for Table 7-1.

Taking the opportunity to investigate further, might existential therapy be better than CBT, given that the prognostic fears of distant spread patients are rational (Edelman, Bell, & Kidman, 1999)(Table 7-18)?

From the statistically significant (all $p < 0.05$) small to moderate results accruing to challenging negative thoughts, problem solving and goal setting for unscreened patients, it seems that there are components of CBT that are as helpful to patients in advanced stages of disease as the existentially oriented expressive-supportive components that produce very similar results.

Table 7-18. Existential and CBT therapy components for distant spread patients

Therapy component	Screened	Unscreened
	ES (n)	ES (n)
Cognitive restructuring / reappraisal	0.28 (3)	-
Challenging negative thoughts	0.15 (2)	0.40 (2)**
Positive self talk	-	-
Self monitoring of thoughts	-0.12 (1)	-
Problem solving	-0.12 (1)	0.38 (1)**
Goal / plan making or expectation setting	0.15 (2)	0.40 (2)**
Activity pacing	-	-
Pleasant activity scheduling	0.15 (2)	-
Expressive-supportive re existential issues	0.23 (1)	0.40 (3)**
Appreciation of personal growth opportunities	-	0.40 (2)**

Untreated control data. Relevant notes as per Table 7-1.

In sum, unscreened results that distant spread patients can benefit well from therapy – the full range of types - and perhaps from relaxation therapies in particular, but this is clearly another subpopulation that deserves more research.

Medical protocol

The type and therefore arduousness of a patient’s medical protocol was another factor put forward as a possible predictor of distress by Andersen (1992). Based on her categories, studies were coded for this factor, and an analysis was performed using untreated controls (Table 7-19). The categories coded were: 1. Single treatment (surgery or radiation); 2. Combination treatment (surgery and radiation and/or chemotherapy); 3. Multiple treatments (combination plus symptom control); 4. Palliative treatment only; and, 5. Mixed protocols (no predominant group).

No particular pattern is apparent, and there is no heterogeneity between categories. One point to note is the low and non-significant scores for palliative patients in both screened and unscreened columns. Both have relatively low n (four studies each) so

could be challenged by future research, but at this stage it seems that there is a need for effective therapies for patients who are in this phase of medical treatment. Another point is the moderate magnitude results for therapies that address patients with a mixture of medical protocols (screened, 0.61, $p < 0.05$; unscreened, 0.40, n.s.). These are the strongest results amongst groupings with larger n, and they are interesting in that they counter the intuitive expectation that that protocol-specific therapy would be more effective since it could target issues relevant to that treatment protocol. Protocol-specificity does not seem to be an important predictor of effect size.

Table 7-19. Medical protocol

Predominant nature of protocol	Screened		Unscreened	
	ES (n)	95% CI's	ES (n)	95% CI's
Single treatment	1.03 (1)**	0.07 – 1.98	0.19 (4)	-0.24 – 0.62
Combination treatment	0.43 (12)**	0.13 – 0.72	0.11 (11)	-0.02 – 0.24
Multiple treatments	-	-	0.40 (2)**	0.05 – 0.74
Palliative treatment only	0.23 (4)	-0.15 – 0.61	0.29 (4)	-0.38 – 0.96
Mixed protocols	0.61 (6)**	0.29 – 0.93	0.40 (5)	-0.14 – 0.93
Q statistic <i>p</i>	0.301		0.505	

Untreated control data. Relevant notes as per Table 7-1.

Medical treatment stage

Is any particular stage in the cancer journey especially fruitful for intervention? Before analysing data for which stages of medical treatment might prove most opportune for therapy, the spread of n (untreated control data) across the 11 coded categories was examined in order to see where categories could be meaningfully combined so as to muster sufficient n for analysis. On the basis of the following frequencies, these eight groupings were made: Newly diagnosed (n = 10); Pre- and/or during treatment (11)(comprising pre-treatment, 6; pre- and during, 5); During treatment (24)(note that this categorisation could refer to one specific treatment such as a course of chemo- or radio-therapy, or to a time period spanning more than one component of medical treatment such as surgery and chemo-therapy); During and/or post treatment (16)(comprising during and post-, 7; post / reintegration with normal life, 9); Recurrence / palliative (9)(recurrence / progression, 1; palliative, 9); Survivor (3); Mixed stages (4); Not reported (4). The latter two groupings were left out of this analysis. Q statistic *p*'s between screened and unscreened groups of the one category were tabulated for those that showed n of three or more in both cells. This was done because it was expected that screening may interact across the trajectory of medical treatment stages, since, according to Andersen's (1992) rankings, stages associated with more serious illness ought to suffer higher psychological morbidity, i.e. higher baseline distress regardless of screening, and therefore similar effect sizes could be expected at

those stages between screened and unscreened samples. Results are set out in Table 7-20.

Moderately strong and statistically significant results appear in screened cells for patients who were newly diagnosed, during, or during and/or post medical treatment (0.67, 0.50, and 0.59 respectively, all $p < 0.05$). These effect sizes fall away for unscreened patients in the former two categories, with the difference reaching statistical significance in relation to the newly diagnosed (Q statistic $p = 0.042$) and approaching it for patients during treatment (0.108). However, for the during and/or post medical treatment category effect size is maintained across the divide such that screened and unscreened outcomes are almost perfectly homogeneous (respectively: 0.59 and 0.56, both $p < 0.05$, Q statistic $p = 0.929$). This suggests that there may be benefit in the blanket administration of therapy during the late treatment / life reintegration stage. Freidenbergs et al. (1981-1982) had highlighted this as a most vulnerable period since it is when patients broaden their psychosocial focus beyond mere survival of the diagnosis and initial treatment to take in the wider implications of the disease long-term.

The differences between the various stages do not reach statistical significance, however, for either screened or unscreened patients, meaning that no one stage stands out as yielding results that are better than any other. The finding of Cwikel, Behar, and Rabson-Hare (2000) that the medical treatment stage is the most effective for therapy is not replicated, however, it appears that that stage is opportune for screened patients (0.58, $p < 0.05$) though not unscreened (0.16, $p < 0.10$, and note that the Q statistic p for the screened / unscreened comparison almost reaches 0.10 significance).

Table 7-20. Medical treatment stage

	Screened		Unscreened		Q statistic p across cells
	ES (n)	95% CI's	ES (n)	95% CI's	
Newly diagnosed	0.67 (3)**	0.13 – 1.21	0.08 (6)	-0.09 – 0.25	0.042**
Pre and / or during	0.26 (2)*	-0.05 – 0.57	0.44 (5)	-0.20 – 1.08	/
During	0.50 (8)**	0.12 – 0.88	0.16 (8)*	-0.03 – 0.34	0.108
During and / or post	0.59 (8)**	0.22 – 0.96	0.56 (5)**	0.13 – 0.99	0.929
Recurrence / palliative	0.23 (4)	-0.15 – 0.61	0.43 (4)	-0.26 – 1.12	0.614
Survivor	0.48 (1)*	-0.09 – 1.05	-0.14 (1)	-0.86 – 0.57	/
Q statistic p	0.577		0.287		

Untreated control data. Medical treatment stage groupings are described in the text. Mixed studies and those that did not report on this variable were excluded from the analysis. Other notes as per Table 7-1.

Specific issues

Some specific questions were to be asked of medical treatment stage data. First, do patients in medical treatment benefit most from coping skills? In Table 7-21 results are set out for the four inclusive therapy categorisations, using untreated control studies, combining studies in the pre and / or during, during, and during and / or post categories to form a grouping for 'in medical treatment' and presenting screened and unscreened outcomes.

For screened patients, all therapy types are effective except education, with scores statistically significant ($p < 0.05$) and in the moderate range from reasonable n (relaxation, 0.41; CBT, 0.46; expressive-supportive, 0.46). For unscreened patients, a striking result appears for education at 0.48 ($p < 0.05$) from an n of seven. However, this result will be inflated by the two third-world outliers mentioned in the therapy types results chapter (Ali & Khalil, 1989; Corchado, 2006). As could be expected, relaxation is only of small benefit to unscreened patients (0.17, $p < 0.05$), and the CBT result (0.09, n.s.) does not even reach statistical significance. Expressive-support comes through quite strongly, however, at 0.39 ($p < 0.10$), which is close to its screened outcome (0.46, $p < 0.05$).

The answer to whether patients in medical treatment benefit most from coping skills is clearly 'no'. Screened patients can benefit from relaxation, CBT, or expressive-support, and unscreened, from expressive-support. A question-mark hangs over education, the effectiveness of which may be a matter of openness regarding cancer information in the particular society such that in closed societies a far bigger effect could be expected, but in most western societies, far smaller. The expressive-support result for unscreened patients at this treatment stage is the most notable feature of this result set.

Table 7-21. Medical treatment stage, therapy types

	Education	Relaxation	CBT	Exp-sup
	ES (n)	ES (n)	ES (n)	ES (n)
Screened	-0.01 (3)	0.41 (13)**	0.46 (11)**	0.46 (4)**
Unscreened	0.48 (7)**	0.17 (10)**	0.09 (6)	0.39 (5)*

Data from patients 'in medical treatment' as defined in the text, untreated controls, inclusively defined therapy types. Exp-sup = expressive support. Other notes as per Table 7-1.

The second question was, what form of therapy most benefits patients in recovery? During and / or post treatment and survivor data were combined to address this, but unfortunately the unscreened results were still very sparse for frequency (Table 7-22). It is interesting to see that relaxation is effective at a moderate to high magnitude of 0.59 ($p < 0.05$), while the value of CBT and expressive-supportive therapies is a little lower compared with the medical treatment stage, with the CBT result losing statistical significance (0.35) and expressive-supportive falling outside the 0.05 level (0.34, $p < 0.10$). Education has a negligible effect at 0.11 (n.s.). This screened data must be

viewed with caution because of the possible uneven distribution of higher scoring 'screened in' studies, but it is safe to say that as patients complete medical treatment and enter the recovery phase after medical treatment, relaxation is a most useful therapy to offer screened patients (of either type – c.f. therapy type results in the previous chapter). See also the next result, which shows that all of the studies contributing to this relaxation effect size were in fact from the during and / or post treatment stage, not from survivor stage.

Table 7-22. Recovery stage, therapy types

	Education	Relaxation	CBT	Exp-sup
	ES (n)	ES (n)	ES (n)	ES (n)
Screened	0.11 (4)	0.59 (6)**	0.35 (4)	0.34 (3)*
Unscreened	0.89 (1)**	-0.14 (1)	-0.14 (1)	0.99 (2)**

Data from patients 'in recovery' as defined in the text, untreated controls. Exp-sup = expressive support. Other notes as per Table 7-1.

The final question asked whether the most effective stage for relaxation therapy is during medical treatment. Results for relaxation therapies by medical treatment stage are shown in Table 7-23.

The heterogeneity indicated by the Q statistic p for the screened comparison (0.056) was investigated as between those stages with an n of more than one and found to be located between the newly diagnosed (0.92, $p < 0.05$, $n = 2$) and during treatment (0.32, $p < 0.05$, $n = 6$) stages (Q statistic $p = 0.045$). However, the newly diagnosed group has a frequency of only two so cannot be taken as a firm indicator of the superiority of this stage, and the only other stage with sufficient n (during and/or post, 0.59, $p < 0.05$, $n = 6$) did not stand out sufficiently to cause statistical heterogeneity.

To the extent that the data does allow the question to be addressed, it appears that the valuable application of relaxation therapies is broader than during treatment, illustrated by a moderate to high effect size for screened patients in the during and/or post- stage (0.59, $p < 0.05$), the high effect size, albeit based on small n, for screened newly diagnosed patients, and, most interestingly, the moderate to high magnitude result for unscreened advanced patients (recurrence / palliative, 0.67, $p < 0.10$, $n = 3$). This latter result, though supported by only three studies, is sufficiently high to encourage consideration of offering relaxation therapies as part of usual care for patients undergoing the stresses of recurrence or late stage disease. The other interesting result in the unscreened column is the non-significant small effect shown during medical treatment (0.25, $n = 6$) which shows that by no means all patients benefit from relaxation therapies at that time, likely because many are not sufficiently distressed to need it.

Though it may not be true that the most effective stage for relaxation therapy is during treatment, it is a stage when the therapy can benefit both screened and unscreened

patients, at least to a small degree. Stages of most interest are, for screened patients, shortly after diagnosis and during and post treatment, and for unscreened, at recurrence or palliative stage.

Table 7-23. Medical treatment stage, relaxation

Medical treatment stage	Untreated control	
	Screened	Unscreened
	ES (n)	ES (n)
Newly diagnosed	0.92 (2)**	0.08 (2)
Pre and/or during	0.38 (1)	0.08 (4)
During	0.32 (6)**	0.25 (6)
During and/or post	0.59 (6)**	-
Recurrence / palliative	-0.12 (1)	0.67 (3)*
Survivor	-	-0.14 (1)
Q statistic <i>p</i>	0.056*	0.456

Relaxation therapy type, untreated control data only. Other notes as per Table 7-1.

Medical variables conclusion

The most dramatic finding of this section is the moderating effect of cancer site, with predominantly breast site samples producing effects significantly lower - less than half the magnitude - of samples predominantly from other individual cancer sites, and no benefit at all over treatment as usual or placebo. Although generally the other analyses were challenged for *n*, the general theme from those results and the cancer site finding is consistent with the hypothesis that patients who are suffering more difficult or socially isolating types or stages of cancer are the ones who produce the higher effect sizes from therapy. However, in the past there has been a disproportionate focus on early stage patients and breast cancer sufferers.

One exception is the result for medical protocol which produces low results for patients in palliative care. The results from palliative stage are not strong either, but better than this. However, relaxation therapies did well for unscreened patients at recurrence / palliative stages. The explanation for this inconsistency is unknown but it may be that this group is particularly difficult to treat, or it could be that the medical protocol category was too blunt a categorisation, or that the low *n*'s involved were open to influence from unknown factors. A further exception - which could be considered rule-proving - is the effectiveness of therapy for newly diagnosed patients. Although patients at this stage may or may not be facing poor prognosis and arduous medical treatment, this group has much to adjust to and can benefit greatly from therapy.

Other points of interest that emerge from this series of results are: medical treatment protocol specificity may not be important for therapy effectiveness; CBT may be as useful to distant spread patients as expressive-support, including that of an existential nature; for unscreened patients, an effective time for therapy is during and/or post medical treatment; screened patients in medical treatment can benefit from the full range of therapies, but unscreened patients at this stage can benefit (moderate magnitude) from expressive-support; unscreened palliative patients can benefit considerably from relaxation therapies.

Studies of more vulnerable patient groups

The results in this chapter convey the message that research and, perhaps clinical practice, needs to refocus on populations of greater need rather than of ready availability for research. For example, there is very little data available on guarded or dismal prognosis cancer types, or regional or distant spread stages, yet it has been seen that patients with advanced disease can greatly benefit from therapy regardless of screening. Also, there are very few studies on cancer sites that could be expected to produce particularly high levels of distress because of the disabling or disfiguring effects of treatment, such as head and neck or oral cancers – only one of each in this dataset. This suggestion gives rise to substantial logistical and ethical difficulties, but there are studies that have found ways around these problems. To assist any researcher who may wish to learn how this was done, lists of all studies in the dataset, whether included in the analyses above or not, which treated patient populations in the less studied vulnerable medical groups are listed in **Appendix N**.

8. THEORETICAL MECHANISMS

Main effects

Effect size data for three broad constructs that may have had moderating or mediating roles in cancer patient distress and, therefore, therapy, were coded from those primary studies that assessed them. These theoretical mechanisms were coping style, self esteem (or self concept) and self efficacy (including perceived personal control, dispositional optimism and the converse of helplessness). An analysis was run to see how the different therapy types impacted each of these mechanisms. Effect size point estimates were calculated using untreated control data for each of these broad constructs at early times against each of the four main inclusively defined therapy types (Table 8-1, below). Q statistic p 's between therapy types are not given due to the statistical dependency of these categories. Those cells with an n of less than three are considered of little or no interest and are shaded.

Table 8-1. Theoretical mechanisms

Theoretical Mechanism	Screening Level	Education	Relaxation	CBT	Expressive-supportive
		ES (n)	ES (n)	ES (n)	ES (n)
Perceived control	Screened	0.03 (2)	0.00 (1)	0.00 (1)	-0.01 (2)
	Unscreened	0.30 (1)	-	0.34 (2)**	0.30 (2)
Self esteem	Screened	-	0.04 (1)	0.04 (1)	0.15 (2)
	Unscreened	0.76 (3)	0.03 (2)	0.31 (2)	0.30 (3)
Self efficacy	Screened	0.26 (1)	0.52 (5)**	0.57 (7)**	0.45 (4)**
	Unscreened	0.23 (3)	0.04 (2)	0.28 (1)	0.22 (3)

Untreated control data. ES = Hedges g effect size point estimate; n = number of studies in subset; * statistically significant at $p < 0.10$; ** statistically significant at $p < 0.05$; two tailed. Cells with $n < 3$ are shaded. Theoretical mechanisms are defined in text above and in more detail in the Methods chapter.

The need to account for the two design confounds has reduced n to levels that are not sufficient for meta-analysis in most cells. However, it can be seen that the self efficacy of screened patients is lifted moderately and statistically significantly by each of relaxation, CBT and expressive-supportive therapies (0.52, $n = 5$; 0.57, $n = 7$; and 0.45, $n = 4$, respectively, each $p < 0.05$). This is not surprising for the cognitive and behavioural therapies, but perhaps it is in relation to expressive-supportive therapies which do not teach skills overtly. Another interesting result is the strong effect that education can have on self esteem with unscreened patients (0.76, $n = 3$). However, the effect is non-significant and based on low n.

Therapy components

The possibility of asking more questions of this subset of the data is all but curtailed by lack of *n*. However, using untreated controls, frequencies were checked to see whether it would be possible to ask to what extent the specific therapy components of coping strategies, problem identification, or problem solving improved perceived control or self-efficacy. Selecting these specific components reduced frequencies to such levels that only three combinations produced *n* of three or more. Results for those are:

- **Coping strategies**, impact on **self efficacy**: screened, 0.21, *n* = 4, *p* = 0.10; unscreened, 0.11, *n* = 3, n.s..
- **Problem solving**, impact on **self efficacy**: screened, 0.64, *n* = 5, *p* = 0.016.

The magnitude of these results corresponds fairly well with the impact that these therapy components had on the three main outcomes (refer to the Therapy characteristics results) hinting at the importance of self efficacy as a moderator or mediator of therapy effectiveness.

9. GENERAL DISCUSSION

This meta-analysis sought to explain some of the conflicting findings in psycho-oncology research by locating effect size moderators relating to study design, patient characteristics, and therapy factors. It has confirmed the one previous finding that was considered to be fairly certain, namely that patients who are distressed at baseline produce significantly higher effect sizes. For that group of patients, this research places beyond doubt the assertion that psycho-oncological therapies are generally effective. More than this, it has found evidence of a pervasiveness of the effect of baseline distress that reaches into other patient characteristics. It has also discovered research design factors that structure the dataset and should be helpful for future reviewers by resolving much unexplained variance and conflict in study results. However, there was little evidence to sustain the expectation that greater intensity of therapy would produce greater results.

The detail of particular results has already been discussed as they were presented. This chapter will only briefly recap each, as its main aim is to discuss broader implications. First, the slightly unusual method will be discussed and then the study design moderators that were identified as a result of using this approach to primary study quality. Then the substantive results are taken in turn, followed by a section setting out recommendations for the application of findings to practice and research contexts. The chapter concludes with a statement of the study's perceived limitations and strengths.

Preliminary analyses

The approach to vetting primary study quality used in this meta-analysis made few assumptions about the quality of particular study designs, choosing, on theoretical grounds, to test the impact of various design features through a series of preliminary sensitivity analyses. The result was that the dataset was permitted to speak to its own dynamics, showing where design features were important to effects and where they were not.

In accordance with Glass's (1976; 1978) observation, randomisation transpired not to be an important feature of design. Indeed, the greatest heterogeneity in design types defined by the method of allocating participants to conditions was between the randomised designs that were stratified and those that were simple – which theoretically should have shown the greatest homogeneity. This outcome shows the nonsense of the assumption and has far reaching implications since so commonly meta-analyses and reviews limit their databases to randomised trials or make other sweeping exclusions without sufficient justification in the context of this domain. The Newell study was one of these (Newell, Simon-Fisher, & Savolainen, 2002) and the vote-count method it used was another example of insensitivity to the nature of the dataset. The result is not only

the waste of much valuable data, but the production of misleading conclusions from high level evidence, as that study showed.

A (probably, the) major cause of the contrast in the effect sizes produced by random stratified and random simple design types was the co-variation of stratified randomised design type with particular sampling characteristics which transpired to predict poor effect size: female gender, (early stage) breast cancer, and lack of screening for baseline distress. These sampling characteristics presumably derived from a striving for scientific homogeneity and sample size, features essential to the bio-medical model of quantitative research, as is randomisation. Thus the distortion in the dataset caused by these studies, and which has resulted in conflict in the literature and the impression being conveyed, at times, that psycho-oncological therapies are ineffective, can be sheeted home to the misplacement of bio-medical norms in this context. No matter how carefully the randomisation procedure is performed or how clever the stratification of participants, such techniques are only 'the icing on the cake' and cannot remedy poor sample selection. It is important that psychologists remember that they are psychologists, not physicians, that their samples bear different characteristics from medical ones and are typically smaller, and that they modify their research approach accordingly. (Recommendations are made in the application section of this chapter.)

In order to tease out the factors that are structuring a dataset, guarding conclusions from such confounding, it is vital that reviewers and meta-analysts do not exclude data on the basis of unsubstantiated presumptions about the 'gold standard' quality of particular design features. Each meta-analytic dataset will have different dynamics resulting from idiosyncrasies of its domain – the participants, the research practices, the conditions under which interventions proceed, and so on – and should be run through a set of design quality criteria and analyses designed especially to fit. It is sobering to think how many syntheses in the present field, let alone in other clinical psychological fields, may have produced distorted conclusions because of the co-variation of study design with substantive moderators.

By testing a range of design features, the present study was able to dismiss many as unimportant, confirm the importance of screening for baseline distress, and discover two other influential features – one which is possibly unique to psycho-oncology and one which has implications for clinical psychology reviews generally. These latter discoveries were, respectively, screening out for psychological history and the nature of the control comparison. Substantive analyses were able to proceed on a firmer foundation once these factors were recognised.

This approach to primary study quality required a willingness to 'float' on a certain amount of imperfection, however. A string of interpretative decisions and priority choices were required in order to whittle down the blips in preliminary analysis results to those important and pervasive enough to be worthy of taking into account in substantive analyses. The usual simplistic default in favour of $p < 0.05$ did not provide all the answers. This method requires judgment and an element of risk taking. The

most difficult of these ‘messy’ decisions was that taken to group studies that screened in for distress with those that screened out for psychological history. However, careful thinking about the source of the discomfort that this empirically indicated decision caused suggested how to proceed – as Glass promised it would (Glass, 1978). It will be recalled that for the purpose of analysing patient characteristics the two were grouped together because empirical homogeneity was the important thing in that context, but for the purpose of exploring therapy type effects the two types of screening were broken out – in fact, so was a third type that screened both ways – so that interaction could be observed. Despite its messiness, it was discovered that the greater tolerance of imperfection that this approach to study design requires conveys a more secure feeling in the end, since imperfection is a more realistic state relative to the human subject matter and to efforts at psychological research and evidence seeking generally. The kind of security provided by reliance upon a claim of randomisation, publication, or arbitrary design criteria such as those measured by ‘the Jadad’ is shown to be both false and dangerous.

Baseline distress

The fact that some researchers screened for distress and others did not reflects the debate in the Introduction as to whether or not all patients should be offered therapy because of the stressfulness of the cancer experience and the reality that distress may not be detected by medical professionals. In other fields of clinical psychology the issue does not arise. As mentioned, researching unscreened samples has also been driven by the desire to achieve large sample size.

In the present study it was shown that baseline distress - simple ‘screening in’ for distress - moderates the effectiveness of therapy when compared with unscreened studies, using the omnibus outcome (anxiety, depression and general distress combined) and untreated control data. Non-significant trends were also found in relation to anxiety and depression outcomes. This comparison provides confirmation of the moderation that Sheard and Maguire (1999) felt confident would be shown with greater sample size than their meta-analysis could muster, and noted in Sherman et al.’s (2004) observation that ‘the rich do not get richer’ from psycho-oncological intervention.

This is an important finding, coming down on the side of the debate that advocates therapy only for patients with proven distress, given that others gain little if anything and given the scarcity of health resources. However, those practitioners who have warned that distressed patients are often overlooked and therefore therapy should be made broadly available also find support in results from this study: Other results showed that there are a number of subpopulations that generally produce greater effect sizes simply on the basis of their defining socio-demographic or medical characteristic, without any proven baseline distress. The unifying explanation may be that those groups are subject to greater pressure – or are less supported to cope with it – due to circumstances associated with their defining characteristic, and therefore they are generally more distressed than other subpopulations with cancer diagnosis. In analysing

the issue this way, Andersen's (1992) conceptualisation of cancer distress is relied upon, that is, that cancer distress is a function of accumulated loss plus insufficient social support. This discussion has obvious but complex implications for distress screening and the design of service delivery models, which will be discussed below.

Screening out and simultaneous screening

Another peculiarity of psycho-oncology is the high proportion of patients receiving intervention who do not have a history of psychological distress. Being able to make some comparison of the difference that such a non-complex background made to results was intriguing and has important potential for future research. On its own, 'screening out' for such history produced effect sizes sufficiently higher than those of unscreened samples to produce formal heterogeneity under anxiety and general distress outcomes when computed using the whole dataset as part of preliminary analyses. A non-significant strong trend remained when the dataset was split for substantive analyses according to the nature of the control condition. Combined with 'screening in' the jump in effect size that resulted in substantive analyses produced statistically significant contrasts from all other forms of screening except screening in alone, and that contrast also approached significance.

It is speculated that while distress motivates patients for change and provides the 'floor space' for improvement to show on psychological measures, patients without psychological complexity in their backgrounds are likely to have the internal and external structures in place to enable them to fully exploit therapy. It may be that these structures also resist the erosion of therapy benefits over time.

However, nearly all of the data for simultaneous screening related to CBT, for which they produced particularly strong and significantly heterogeneous effects compared with screening only one way or the other. Where patients were merely distress screened ('screened in'), effect sizes from CBT were small, but they leapt to very strong magnitudes for samples that were screened both ways. It was suggested that this may be due to a match between the structured and demanding nature of this therapy type and the motivation and resources that samples screened simultaneously both ways can bring to it. It would be important for future research to test whether other therapy types produce the same dynamic. If this hypothesis is correct, some moderation will be predicted by the structure and demands of the therapy type.

There are other implications for both research and practice. By research comparing the response of non-complex with complex distressed patients, including over time, much could be learned about the mechanisms that impact each sample type differentially, which should provide insights to inform more effective and efficient therapies. If patients without distress in their histories prove to be more readily treatable generally or in relation to particular therapies, and if they prove to retain benefits longer, then these factors can be applied to designing service delivery and follow-up monitoring.

Nature of control condition

The finding that studies that used control groups that contained some small element of treatment (treatment as usual, treatment element placebo, attention placebo) produced significantly smaller effect sizes can hardly be surprising. Moderation by this study design feature must surely be widespread in psychological literature, since it is really a function of the inherent nature of effect size indices in that they are a measure of contrast. The surprising thing is that it is not commonly recognised – certainly not in psycho-oncology. Failure to take this variable into account in literature syntheses exposes their conclusions to distortion, so it is important that its ability to explain much variance is exploited in future reviews and meta-analyses.

Substantive analyses

In the discussion of substantive results which follows, main effects are addressed first, and then moderator effects, which have been the primary focus of this study. The latter are headed by allusion to Paul's 'ultimate question' referred to in the Introduction (Paul, 1969), with therapy related variables discussed before patient characteristics. Bear in mind that most results were based on 'early times' data (i.e. for most studies, the first assessment point within six months after intervention).

Main effects – Are psycho-oncological therapies successful?

Main effects over the three outcome variables – anxiety, depression, and general distress - were all small at early times when design feature confounding (screening and nature of the control condition) was not taken into account. Anxiety distinguished itself as the only outcome of the three that produced statistically significant main effects for all four professionally delivered therapy types at both early and late times, suggesting that anxiety may be more reliably and durably treatable. This confirmed the finding by Sheard and Maguire (1999) that anxiety appeared more treatable. Main effects ranged from moderate to very strong for samples that were distressed at baseline.

Similarity of main effect findings to those of Sheard and Maguire (1999), which is the meta-analysis most similar in design to the present study, gives grounds for confidence in them. Replication of their finding of a distinction between studies that screened recruits for distress and those that did not also points to the inappropriateness of continuing to treat psycho-oncological trials as a homogeneous body of work which can be reduced to one or two handy main effect sizes. The field is hugely varied and to discuss therapy effectiveness, while failing to recognise at least a few basic distinctions within it, is grossly misleading.

Referencing the present findings against those of other spheres of clinical psychotherapy is rendered problematic by the uniqueness of this field, which is distinguished by, for example, its frequent use of nurses and social workers as therapists, the frequency with which education is administered as an intervention, the ongoing pressure of chronic and life-threatening illness on patient well-being, the distress that pain brings to many advanced patients, not to mention the frequent lack of

proven baseline distress in research samples. Further, a cursory glance through the abstracts of clinical psychology meta-analyses reveals a great range in results, no doubt associated with a similar great range in methods used and therapies synthesised. However, the magnitude of the Hedges' g 's arrived at in the present study for patients who were distressed at baseline are very encouraging. Also, if compared with the estimate arrived at by Smith and Glass (1977) in their classic meta-analysis of psychotherapy outcome studies, which is at least similar to the present study in the respect that it also admitted the full range of controlled designs, then psycho-oncology does well. Their standardised mean difference (d) reported from 375 studies on mixed outcomes was 0.68. In the present analyses, the comparable effect sizes are presumably those from distressed samples, which were g 's of 0.73 for anxiety, 0.51 for depression, and 0.60 for distress (these figures are against untreated controls as there is very little for distressed patients against treated control comparisons). So psycho-oncological interventions can be regarded as 'successful' for appropriate patients, but the point is reasserted that the nature of this field is so different from others, and so various within itself, that comparisons derived from data that is not cut to recognise some basic distinctions are misleading.

Therapy type – What does what?

Setting aside education, which is distinctive in nature as a 'therapy', the magnitude of effects produced by the different professional therapy types provided few dramatic differences, in accordance with the finding of homogeneity between types found by Devine and Westlake (1995). However, each type has its own characteristic dynamics. This section attempts to provide the reader with a feel for the dynamics of each therapy type, incorporating relevant findings from the patient characteristics results. Results cited are generally from untreated control data, and recall that a detailed summary table showing the impact of different screening levels in relation to the four main therapy types is provided in the Therapy characteristics results chapter.

Education

This therapy type generally performed poorly, yielding results that were null or negligible at first, lifting to negligible to small after six months. Confidence intervals often fell below zero, meaning that some patients were made more distressed by intervention. Best results were against anxiety, where, against untreated controls and without taking into account screening, a strong effect size was obtained. Medium and small results, respectively, were produced against anxiety by treated and untreated control comparisons of patients without the complication of psychological history. These generally poor results fail to replicate the strong findings by Sheard and Maguire (1999: 'group psychoeducation' $g = 1.59$ against anxiety and 0.94 against depression) or Rehse and Pukrop (2003: 'education treatments' against 'emotional adjustment' $d = 0.96$) but differences in categorisation are likely to be responsible, at least in the former case. It is important to note that no study sampled patients who had been screened in

for baseline distress, therefore the effectiveness of educational intervention with this subpopulation could not be analysed.

Characteristics of the primary research are likely to explain these generally and comparatively poor results: Most studies were from Anglo-western countries where cancer information is broadly available in society; most study results were against treated controls; and, as mentioned, none recruited patients with proven baseline distress. Flickers of evidence from studies sampling patients who were likely distressed at baseline, did not have ready access to other sources of information, and were compared with untreated controls show spectacular positive effects against anxiety or general distress, with an increase over time that may prove to be a hallmark characteristic of this therapy type (Ali & Khalil, 1989; Corchado, 2006).

In developed Anglo-western countries most patients have ample access to cancer and treatment information through usual medical treatment and generally. In addition, many research samples are, on average, well educated, and are therefore capable of accessing and understanding information independently. When educational interventions were broken out against income level, although n was sparse, results were generally consistent with the pattern of poorer samples benefitting more. It may be that poorer patients find knowledge about the disease and its treatment relatively more empowering given their more complex and difficult circumstances. (There was an insufficient amount of data to tell whether patients with less education benefitted more or less from educational interventions.) The suggestion here espoused is that educational interventions will be most effective where information is scarce, or a patient's ability to take advantage of it is restricted, consistent with a deficit hypothesis. A small and inconsistent body of evidence to this effect was cited in the Introduction chapter.

It was fascinating to see that statistical heterogeneity in favour of 'older' patients (study mean ages of 53-66 years compared with 41-52 years) was approached in the comparison of unscreened age groups on the omnibus outcome for educational intervention despite the limited number of studies contributing. The deficit hypothesis suggests the explanation that perhaps the older generation in developed countries remains somewhat subject to the restrictive culture surrounding medical information of earlier times, and/or is less confident and proficient with information technology, and so is more in need of education.

Analysed as specific components of therapy packages (e.g. education re cancer or cancer treatments, education re emotion and cancer), a number of quite high effect sizes were obtained for impact of education on the anxiety and general distress of unscreened patients (there was insufficient data against depression). However, these effects were usually not statistically significant. This again suggests that education can form a worthwhile component in therapy packages, but that some individuals will not benefit from it – indeed, some may become more distressed. As this dynamic is against the context of the considerable access to information that is available in the west, it does not

indicate that education *per se* is bad for some people, but it could be that too much education is bad for some - a matter of balance in respecting individual differences.

Anxiety outcome results for 'screened out' samples (patients without psychological history) were interesting in that the lower confidence interval barely dropped below zero for both treated and untreated control comparisons. This means that very few patients with non-complex psychological backgrounds were adversely affected by educational interventions, and most of these patients received a small to moderate benefit. Perhaps the poorer response from unscreened samples is the result of patients with pre-existing psychological complications finding that more information often adds to their burdens, rather than assists. A thread of thought to this effect may be drawn together from the literature: There is contradictory evidence as to the benefit of providing tape recordings of consultations to patients, which might be explained by the patient's prognosis (McHugh et al., 1995; North, Cornbleet, Knowles, & Leonard, 1992). As patients accumulate physical diagnoses, their psychological morbidity also tends to increase (Clarke, 1998). So it may be that ignorance is bliss for patients without the internal or external supports to cope with complex and less than cheerful information. On the other hand, a study that used a video of patient-specialist consultations to prepare newly diagnosed patients for their first clinic found that mental health patients were particularly helped by this educational material (Walker & Podbilewicz Schuller, 2005). Self-efficacy theory (Bandura, 1997) and coping theory (Lazarus & Folkman, 1984) both imply that knowledge of relevant information is important for adjustment, so from those perspectives, the latter result is logical. However, it may be important how, and in what personal circumstances relevant to the patient, information is conveyed e.g. self-efficacy theory requires a non-emotional context for the gaining of mastery, yet cancer news is often highly emotionally charged. In the light of such uncertain ground, research into factors predictive of adverse reactions to educational interventions in cancer patients would be valuable.

Some of the 'indirect' intervention studies are relevant to considering education interventions as they show how improved communication of medical information by doctors can be very effective (Rutter, Iconomou, & Quine, 1996; Stewart et al., 2007). As a means of delivering intervention efficiently to every patient, including those who are immobile, illiterate, or for whom 'psychological therapy' would be unacceptable, indirect interventions that have the immediate target of improving the communication skills of medical staff could prove to be a very rich prospect for further research and development. These interventions not only exploit the access that doctors have to all patients, but also their unique position of authority, expertise, intimate knowledge and trust. Improved communication also produces improved patient satisfaction with doctors' performance and care, which in turn is likely to improve medical treatment compliance and subsequent physical health, coming full circle to lift psychological well-being.

There may also be scope for efficiencies as indicated by the homogeneity of effect sizes yielded by the analysis of different technological modes of therapy delivery to

unscreened patients. Using modern media, high quality educational material can be delivered with great efficiency, but the form of delivery chosen must be designed to reach the old and the poor, who rely more on simple means of communication. Such interventions should be professionally designed to take account of both the content needed to allay anxiety (e.g. the detailed sensory information provided in a pre-surgical pamphlet used by Corchado, 2006) and the form needed to communicate it well (e.g. the many photographs used by Corchado).

Finally, the long term trajectory of educational interventions deserves study. There was evidence of a small effect at late times for unscreened patients (in comparison with treated control groups) and, as mentioned, some individual study evidence that the effects of education may build against anxiety in patients who were likely suffering elevated distress at baseline. If education indeed has the capacity to build adjustment as time goes by, then trajectory is a very important feature of this intervention type.

In sum, findings leave tensions to be resolved around what factors predict effective educational intervention, and much scope for the development of efficient intervention, including by indirect delivery through doctors, efficient technologies, and targeting older and poorer patients. Study with patients suffering baseline distress is needed, together with long term data. More use could be made of theory relating to the cognitive processing of information, self-efficacy and coping.

Relaxation

Relaxation demonstrated medium magnitude early times effects for non-complex and unscreened patients against anxiety, fading to small or less by late times. A less well substantiated but similar magnitude outcome was produced against the general distress outcome, but nothing to speak of against depression. Against treated controls the results faded to small but data were sparse and effects may have been attenuated by packaging with other therapy types. The strongest results were achieved by patients who were distressed at baseline, especially if they also lacked psychological history. This was so for both anxiety and, surprisingly, depression. However, there were only one or two studies contributing to these results. Also on very light study frequency, there was a thread of evidence that suggested that relaxation effects may hold or even increase into medium term (one to six months) for non-complex patients. Of the specific relaxation therapy components that had sufficient data to be tabled, progressive muscle relaxation and diaphragmatic breathing performed best, with (guided) imagery disappointing. These components showed benefits for screened patients (i.e. distressed at baseline and or non-complex) but not for unscreened patients to any worthwhile degree.

Relaxation results are more readily understood than those for education. This therapy is behavioural and as such can be expected to benefit those with proven need, and for the fairly short term, which it does. Compared with factors that can be incorporated into placebos or treatment as usual protocols, its effects are minimal for patients without proven baseline distress. The impact against depression for patients with baseline

distress may be explained by the role that anxiety plays in that disorder, and deserves further investigation.

An interesting non-significant trend emerged from the comparison of older and younger screened patients, showing older patients received twice the effect (moderate magnitude, and note that the type of screening likely influenced the magnitude of this difference). For unscreened patients the trend also favoured older patients, though both effect sizes were much smaller. There was also evidence that unscreened patients suffering distant spread disease and those at recurrence / palliative stages gained moderate strength effect from relaxation. Read together, these results lend support to the deficit hypothesis espoused by this study: The explanation may be that older patients face more loss with less support (elaborated on in the discussion of socio-demographics), including more complex and serious medical conditions, notably pain and dismal prognosis, and therefore suffer more distress, and that relaxation therapies provide them with a useful tool for managing this. Its mechanism of action may be the alleviation of some degree of physical symptom manifestation and/or through restoring some sense of personal control, thereby providing a coping resource or enhancing self-efficacy.

It would be worthwhile researching whether relaxation should be provided to patients who are older and/or who suffer disease recurrence or distant spread as part of usual treatment – especially as it can be delivered relatively inexpensively. It is important that therapies for this subpopulation are as little burden to the patient as possible. Clinical hypnosis is one such relaxation intervention, and proved very effective with terminally ill patients delivered in four sessions by Lioffi and White (2001: $g = 0.92$ against anxiety). It is surprising that more research of this nature has not been carried out, given its particular suitability and apparent effectiveness, but this may be part of the general neglect of late and palliative medical stages in the research literature.

For screened patients in recovery stage (during / post treatment and survivor), relaxation similarly produced the highest results. Once again this may be because it provides a tool for regaining some sense of control during a transition time that may be particularly stressful. Freidenbergs et al. (1981-1982) have observed that it may be only once the shock of the diagnosis stage is behind patients and medical treatment is near completion that patients are able to turn attention from mere survival to absorbing the wider implications of the disease on their lives. Patients in recovery could be provided with relaxation materials quite cheaply (e.g. audio or written instructions for progressive muscle relaxation and deep breathing) and could be taught to use them in a brief session or two by less highly paid staff (e.g. nurses or social workers) so it may be appropriate for this to become part of standard care.

There was also a trend for men to benefit more from relaxation therapies, which was supported by statistical heterogeneity in relation to screened but not unscreened patients, and it appeared that men obtained at least a small benefit from relaxation

therapies regardless of screening, in contrast to the overall result. An explanation for this gender difference will be suggested in the discussion of socio-demographic results.

As a therapy type, relaxation lends itself to manualisation and delivery by lower salaried staff backed up by non-interactive technologies. They are therefore a therapy type that deserves research directed towards refining their efficient use. Efficiencies appear to be available for use with unscreened patients through the use of cheap technologies. More research is needed on patients with initial distress, at recovery, recurrence and late stages, and on patients without psychological history, and with men, including assessment over the long term in order to establish effect trajectory. There also appears to be a complete lack of indirectly delivered relaxation therapy research (the direct target being a spouse, for example) which may also prove to be a worthwhile means of assisting patients who themselves may not be able or willing to undertake therapy.

CBT

The evidence showed clearly that CBT was not effective for unscreened patients. Although the frequencies of studies contributing to effect sizes were considerably reduced by the need to account for the study design feature confounds, there was enough consistency in the untreated control results to conclude that CBT produces its strongest results for patients who enter therapy with heightened distress. When samples are screened in for this status effects are small to moderate in magnitude, but much stronger when they are simultaneously screened out for psychological history. One way screening to produce non-complex samples produced a small effect against each of the outcomes, but this was supported by sparse study frequency. When worked on omnibus data, simultaneous screening produced statistically significant heterogeneity against both screening in and screening out alone, suggesting that CBT is particularly sensitive to an interaction of baseline distress and psychological non-complexity. Effects tend not to lower against treated controls, presumably because the effective components of the therapy are quite distinctive from what is provided as part of usual treatment or placebo conditions. Limited data suggest the lingering of small or medium effects until late times, but this may be explained by a hint in the trajectory analysis that samples simply screened out for psychological complexity do well over time.

The specific CBT components that performed best over all three outcomes (anxiety, depression and general distress) were cognitive restructuring, challenging negative thoughts, self monitoring, and, especially, problem identification and solving. Pleasant activity scheduling had some effect but assertiveness / communication training yielded generally disappointing results. It could be that the more effective interventions are skills that are less well known by the public generally, and since the cancer experience is laden with challenging problems, patients with both the motivation and the strength to learn them can exploit them very profitably.

The combination of CBT with relaxation showed moderate magnitude effects against depression and general distress and strong impact against anxiety, but there was no data available at late times. The combination of CBT with relaxation and expressive-support

proved particularly poor and, as noted earlier, this raises a question about the theoretical consistency of that combination, but it could well be that patient characteristics confounded results since large expensive trials administering comprehensive therapy packages tended to use unscreened early stage breast samples.

Regular homework or skill practice, compared with none, proved beneficial for screened patients doing CBT and/or relaxation therapies, but not for unscreened patients. When all therapy types were combined, there was no significant gain for screened patients, and unscreened patients were actually disadvantaged by homework (except in a one-off dose). It was noted that the finding relating to CBT / relaxation was consistent with the meta-analytic finding of Kazantzis, Deane, and Ronan (2000) that homework for these therapies produced a moderate strength advantage. However, homework can be demanding and it may be that only screened patients have the motivation (baseline distress) or the strength (internal psychological structure and external resources associated with a lack of psychological history) to exploit the opportunity it provides. Further, a note of caution was sounded by Given et al. (2004) who felt that their CBT intervention may have overburdened patients who were already hard pressed by (baseline) distress and undergoing chemotherapy. No doubt in the practice context homework would be better tailored to the needs of individuals. Future research into homework could provide useful insights into the different ways that non-complex and other distressed populations process highly structured CBT and relaxation therapies.

CBT is also a therapy that can readily be simplified to just one or two effective components, such as problem identification and solving, and manualised for very effective administration by therapists on lower salaries. Strong results were obtained by social workers working with patients who were screened both ways. This provides a basis for efficiency as the services of psychologists can be reserved for complex cases.

It was found that expressive-supportive therapies made much stronger short term gains than CBT with samples that were (simply) screened in for baseline distress, although no formal comparison could be made because of the statistical dependence of these inclusively defined therapy types, and the expressive-support sample was very small. The suggestion was also made that on the evidence available at present, an effective strategy for treating distressed patients may be to start with expressive-supportive therapies, later working to consolidate and fix gains with CBT. It would be valuable to compare the long term trajectories of the effects of these therapies in factorial research that also investigated the impact of the combination of both, administered sequentially. It may be that expressive-support, administered first, would lower distress enabling patients to take full advantage of CBT, offered next. On the other hand, the lowering of distress might remove the motivation needed to meet the demands of CBT, or some inherent conflict in their theoretical structures might attenuate results, as mentioned above. Effects may be further moderated by the psychological complexity in patients' backgrounds.

In terms of particular subpopulations, it was found that small to moderate results accrued to distant spread / advanced stage unscreened patients from the specific CBT components of challenging negative thoughts, problem solving and goal setting, and that these were as helpful as existentially oriented expressive-supportive therapy. This is very interesting given the physical burdens under which these patients labour and their realistically negative assessment of the future given their dismal prognoses. Note that smaller benefits accrued to screened patients and that both of these observations rest on very small sample sizes. It is important that more research be conducted with these vulnerable patients.

CBT was shown to produce stronger effects for lower income patients. Although n was sparse, the pattern over all quadrants of the structural confound matrix was consistent and the comparison in the most sensitive quadrant reached statistical heterogeneity. The same could not be said in relation to the education level of patients, but with more contributing studies a similar pattern may be revealed. It could be that lower income groups find that the skills taught by CBT fill a particular need, given the intensity and scope of problems that cancer catalyses for them.

In sum, CBT is for distressed patients and does particularly well for those without psychological distress in their backgrounds. CBT research on unscreened samples generally produces negligible and non-significant effects and it may be time to abandon it, except in relation to those particular subpopulations who have shown benefit from it and very possibly suffer heightened distress as a result of pressures associated with their defining characteristic. Data are needed for distressed patients against treated controls in order to measure the incremental value of this therapy. More data on how non-complex patients - especially those with baseline distress – respond to CBT and to relaxation therapies over time would be interesting and could have important clinical implications for the treatment of both non-complex and complex populations.

Expressive-support

In contrast to CBT, expressive-supportive therapies produced small to moderate benefits for unscreened patients, with evidence that the effect against anxiety may maintain or even lift a little by six months. However, comparison with treated controls saw this effect evaporate, except, interestingly, the small pick-up against anxiety over time. In relation to psychologically non-complex samples the evidence is a little confused, with a few large results popping up against treated controls, and better results tending to present against anxiety, but overall, providing little if any benefit. The real value of the therapy shows when administered to patients who are distressed at baseline. Effect sizes greater than $g = 1.00$ were produced against each outcome, but in every case there were only two studies to support these results and there were no late times data or comparisons with treated controls. There were no studies that screened both ways, and it would be very interesting to learn whether a lack of psychological complexity in the backgrounds of distressed patients makes a difference to the effect size delivered by this therapy and its trajectory over time.

As specific components contributing to this therapy type, expressive-support relating to existential, spiritual, grief or death issues produced consistently stronger results against all outcomes than expressive-support relating to the cancer experience, physical and psychosocial issues. However, contributing sample frequency was light for the existential form of this therapy, and formal heterogeneity between this and the more general component could not be tested for because of their statistical dependence. Given that this therapy appears to perform best for distressed patients, it may be that a focus on existential issues is what is most needed, especially in western cultures where there may be some considerable reticence to talk about such issues socially. Kristeller, Rhodes, Cripe, and Sheets (2005) have been noted for providing a most efficient and effective existential intervention via oncologists, and an existential approach has been demonstrated in studies by C. Chan, Ho, Fu, and Chow (2006), C. L. Chan et al. (2006) Y. M. Chan et al. (2005) and Ho (2007).

One study (Goldberg & Wool, 1985) administered expressive-supportive therapy to the significant others of lung cancer patients yielding a strong but non-significant indirect impact on anxiety for patients. This study shows the potential of the therapy to reach populations that may not be able or willing to present for therapy, and as such, deserves replication. Written emotional expression may provide another way to reach such patients, or provide efficiencies, but it again seems necessary to select patients who are distressed at baseline and/or socially constrained at home (Zakowski, Ramati, Morton, Johnson, & Flanagan, 2004) as only one of three studies using that therapy type showed as much as a small benefit for unscreened patients.

Subpopulations that may be considered to all have in common a greater degree of social isolation produced stronger results from expressive-support: A non-significant trend favoured older unscreened patients; a trend favouring men was posited and found in relation to both screened and unscreened patients, with sufficient sample size to reach statistical significance if these subsets were combined (untreated controls) i.e. men received about three times the magnitude of effect, compared with two times when all therapies were combined; there was evidence that suggested a trend in favour of single patients; and a moderate magnitude effect was delivered to unscreened distant spread patients. These results are consistent with the literature adduced in the Introduction to the effect that therapies that aim to provide a form of social support can be particularly effective for patients who lack quality social support at home (e.g. are socially constrained) – consistent again with the deficit hypothesis presently proposed. However, expressive-support also delivered moderate magnitude effects to both screened and unscreened patients during medical treatment and during recovery stages (in the latter case the unscreened result was strong but supported by only two studies). More will be said about medical stages below, but the fact that these times can be considered especially stressful may explain the response to therapy by all patients – not just those with proven baseline distress – during them.

In conclusion, relative to treated controls the data show that expressive-support is not worthwhile for patients who are not distressed at baseline, but it produces very strong

results for those who are and moderate strength results for unscreened patients from certain subpopulations. It is posited that relative social isolation is the common thread linking many of these samples (older patients, men, singles, late stage patients), and further research of these groups - particularly seeking specific mechanisms of action such as social constraint at home – would be valuable. The therapy also proved worthwhile during medical treatment and at recovery stages amongst unscreened patients, possibly because of heightened distress or feelings of social isolation at these times.

Non-professional

Non-professionally run therapies produced null results generally, and there were few primary studies available. The Institute of Medicine of the National Academies (2008) points out the importance of not confusing a lack of evidence with ineffectiveness:

“...the frequent provision of many [voluntary] services to patients and families indicates that these services likely help patients and their caregivers meet health-related psychosocial needs. The provision of ... logistical and material support when needed also can logically be assumed to decrease patient distress and increase the ability of both patients and caregivers to manage illness and its consequences...” (p.103).

Part of the reason for a lack of evidence from non-professional interventions is the difficulty in controlling research into the effectiveness of complex services that are already up and running. Even when this is possible, the types of intervention that are provided frequently target distress only indirectly through the provision of practical assistance and advice. The aim of voluntary services is not so focussed on reducing distress, but often on preventing it and on providing practical support. They differ also in that patients presenting to them do so at their own initiative, which presumably means that they feel comfortable and motivated to do so. This is likely to produce samples that bias effect sizes.

There were, however, two studies that produced a strong impact on depression. They are not typical non-professional interventions and are instructive. These studies are a pair led by Weber using long term survivors to meet with radical prostatectomy patients (Weber, 2001; Weber et al., 2004; Weber et al., 2007). The first of these studies (the 2001 and 2004 references) was a thesis pilot for the second and achieved stronger results ($g = 0.96$, $n = 30$ compared with $g = 0.59$, $n = 72$). In both, the peer survivors were given two hours training and each was assigned a single patient to meet with for coffee and discussion of physical and emotional issues associated with the diagnosis and its treatment and side-effects, once a week for eight weeks. The focus and quality of the meetings were professionally monitored. The intervention was designed especially as a means to provide men with social support that was acceptable to them on the premise that for most men that level of intervention would be necessary and sufficient yet they would not be likely to attend a support group.

Conclusions drawn in the socio-demographic and medical results section of this report suggest that Weber et al. were quite correct in targeting the needs of men and of prostate cancer sufferers. Perhaps men are more vulnerable to feeling a stigma from 'weakness' caused by cancer treatments, and therefore they may be particularly assisted by therapies that normalise the experience, enabling them to incorporate it within a masculine identity, meaning that non-professional survivors may play a vitally influential role if overseen or complimented by professional work.

Although non-professional interventions generally perform poorly on the outcomes measured by this meta-analysis, it would be foolish indeed to disrespect and dismiss their wider value to the patients who use them. Indeed, it will be noted below that their particular accessibility to breast cancer patients may be part of the reason for the lower effect sizes produced by that subpopulation. Further, if used strategically to target vulnerable populations and with professional backing, they have shown potential to be powerfully effective.

Indirect

Indirect therapies have already been discussed in the context of education and expressive-support as providing exciting potential as a way to access patients with efficiency, and to access subpopulations which may be difficult to reach due to their reticence to receive psychological help or their physical condition.

Other

The very few studies that could not be coded under one of the above types included three written emotional expression interventions which were not greatly successful, but, as noted, the key may be in moderation by social constraint at home (Zakowski et al., 2004).

Existential therapies have been mentioned as part of the discussion of expressive-support, but deserve fuller discussion. Existential issues are characteristic of and central to the cancer experience (Cunningham, 1988), as are loss and grief (Andersen, 1992). Knight (2004, p.243) sums up the thoughts of others saying, "Understanding a person's purpose in life, including religious or spiritual beliefs, provides a context from which to understand their ability to withstand the disruption and losses associated with cancer." Yet given the sensitivities that can exist around spiritual matters in western cultures – and particularly in western psychology – these issues can easily become 'the elephant in the room' that is strenuously avoided rather than addressed. Psychologists need to find ways within the culture of their clients - rather than their own professional culture - to address these issues because of their central importance to cancer patients. The Kristeller et al. study (2005) mentioned earlier, which delivered a very brief spiritual intervention through medical oncologists, demonstrates one particularly efficient approach. It also appropriately exploited the confidence and intimacy of the doctor / patient relationship to safely deliver a sensitively constructed intervention into this very

personal and essential aspect of the patient's being. Replication research would be valuable.

Therapy delivery and theoretical mechanisms – How?

Delivery variables

A good deal of homogeneity resulted from comparisons around therapy delivery variables (mode-, dose- and therapist-related) suggesting that the ‘hows’ of therapy delivery are not the most important factors in producing effect size – or, at least, not factors that require much improvement in the way that they are already managed. It also suggests some scope for efficiencies where therapy types allow. The point was made, however, that intervention research is conducted in rarefied conditions, where it is common to exclude patients with multiple health problems and to use samples with favourable prognoses, in contrast to conditions experienced by practicing psycho-oncologists. Subject to that qualification, which is elaborated further below, there was evidence that brief therapies (the categorisation was one to four hours or sessions, although this derived merely from statistical convenience) can be effective in the short term for screened patients, and non-interactive therapies can do as well as personally delivered ones for unscreened patients. As noted above, it was also found that homework or skill practice provided no incremental benefit to screened patients, and was detrimental for unscreened patients except in a one-off dose, and should therefore be dispensed with. This was said subject to qualification regarding relaxation and cognitive behavioural therapies, and also regarding a one-off introductory dose.

Psycho-oncologist role

Again, it has already been mentioned that less highly paid professionals (social workers) were found capable of very effectively delivering manualisable therapies to screened patients. Where more individualised therapy is needed, as for patients with psychologically complex backgrounds, a specialist in the therapeutic professions would seem appropriate (psychologist, psychiatrist, or trained counsellor). It is suggested that another important role for psychologists is designing and testing modifications to standard medical practices in order to lower patient distress. These systems-oriented ‘interventions’ may comprise therapeutic modification to ‘treatment as usual’ procedures or the training of medical staff in more beneficial means of communication, such as has been suggested in discussion of indirect therapies. Psychologists may also find ways to make inexpensive and simple therapies, such as relaxation, readily available to distressed patients and those in particular subpopulations. A key role would be the development of systems that detect clinically significant distress, and the training of medical staff in sensitivity to its manifestation. The critical importance of this will be discussed further below.

Delivery and psychological background

There are two points that arise from findings concerning the psychological complexity of patients' backgrounds that can only be described as 'hints' in the data, but which, if substantiated by future research, could have important implications for service delivery.

More general clinical literature conveys suggestions psychological background may make a considerable difference to what therapy delivery resources a patient may require and how effective therapy will be. In their review of general psychotherapeutic literature on the influence of client variables on therapy effectiveness, Clarkin and Levy (2004) make the point that clients with the least severe psychological symptoms demonstrate more ready responses. Note that these authors are talking about presenting symptom severity rather than psychological history, but consider that finding against that of the present study that strong early times therapy effects can be produced by patients who are both initially distressed and without psychological complexity. Clarkin and Levy also cite a large study (Shapiro et al., 1995) in which it was shown that those with more severe depression responded substantially more with 16 rather than eight sessions of either CBT or IPT. Beutler et al. (2004) cite the same study in their review, and note that increasing dose intensity did not help those with less impairment. They cite Beutler, Harwood, Alimohamed, and Malik (2002) for a finding of only modest support for increasing the frequency or length of therapy for highly impaired (general psychological) patients. These authorities suggest that ready response can be expected from less severely affected patients and that more therapy may be necessary for the more impaired, although improvement proportional with the amount of therapy invested is not guaranteed.

In the present research, the only account that could be taken of symptom severity and complexity was screening in for distress and out for psychological history performed at participant study recruitment. CBT was the only therapy type that had a sufficiently even distribution of data to allow comparison of patients from all four screening possibilities – known distress, known lack of psychological history, both, and neither. For the omnibus outcome, unscreened samples performed most poorly, followed by those without psychological history, then those who were distressed, with a jump up to those that were screened both ways. For depression, the category of studies that simply screened out produced the lowest effect sizes. It was only the category that screened both ways that produced significantly higher results against both of these outcomes (omnibus and depression), suggesting an interaction between baseline distress and non-complexity.

This result suggests that a patient's psychological background may be important to the therapy resources required for effective treatment. However, CBT may be particularly sensitive to a patient's history of distress because of its structured nature. But it may be that not only the readiness of response to therapy is predicted by a lack of psychological history, but also the relative fixing of therapy effects over time. There was a hint to this effect in the trajectory data for non-complex patients that needs substantiation but

makes sense. The impact of psychological background on both readiness of response to therapy and the fixing of effects both carry obvious and important clinical implications for the design of services. Furthermore, the role that psychological history may play has been invisible in past literature reviews. If this factor does make such a big difference to therapy response, then administrators need to become aware that research showing that standardised brief therapies delivered by lower paid professionals who deliver high impact and perhaps lasting effects for non-complex distressed patients cannot be expected to generalise to psychologically complex patients. It is important that the distinction be brought into view since most of the patients referred to psycho-oncologists are not psychologically simple.

Secondly, there was a hint in the results that delivery features that convey a sense of personal security may be an important feature for distressed patients. From data combining all therapies, there was (non-significant) evidence that more weeks of therapy – rather than more sessions or more hours – may benefit screened patients, as may individual rather than group administration, and therapy delivered by a professional rather than a student. One explanation for these trends may be that distressed patients find the assurance, more than the intensity, of ongoing one-to-one professional contact therapeutic because uncertainty is so central and pervasive in the cancer experience. Providing a sense of security is not discussed in psycho-oncology intervention literature as a possible mechanism of action but may be worth investigation in future. However, it may be that more weeks of therapy is beneficial simply because it provides an opportunity for monitoring and coaching in learned strategies and skills over a longer period of time (clinician, D. M. Baken, pers. com. 7 July 2009).

Theoretical mechanisms

Although foci of the least interest in the present study, the opportunity presented by the collection of studies for the present research was taken to collect outcome data on a few theoretical mechanisms of therapeutic action, namely self-efficacy, coping and self-esteem. (Although social support was a mechanism of considerable interest, it is discussed under the variable used to operationalise it, marital status, in the following section.)

The strongest pieces of evidence that emerged from the limited theoretical mechanism data that survived the necessity of matricising according to study feature confounds were the moderate magnitude significant effects from relaxation, CBT and expressive-supportive therapies on the self-efficacy of screened patients. Problem solving, as a specific CBT component, produced evidence of a moderate to large significant effect on self-efficacy for screened patients, while the effect of coping strategies on the same subset of patients was small and reached only 0.10 level significance. Against the three main outcomes (anxiety, depression and general distress) problem solving also produced moderate magnitude significant effect sizes. From these data alone a causation direction between distress and self-efficacy cannot be implied, but it can be said that problem solving specifically, CBT as a general therapy, and also relaxation and

expressive-supportive therapies, were effective in lifting self-efficacy in screened patients. It is also not possible to compare self-efficacy with other possible mechanisms from the data available, but only to confirm that self-efficacy is associated with effect.

It is easy to see how skills-based therapies such as CBT or relaxation can lift self-efficacy, since they teach skills aimed directly at better equipping patients to cope with the stresses and challenges that they face. How expressive-support can build mastery is not so obvious. In the group setting, however, patients could have their confidence raised from vicarious experience / modelling and persuasion by their peers, which are important means according to self-efficacy theory. The data here confirm that, in accordance with self-efficacy theory, it is possible to learn from both structured and unstructured therapies.

Patient characteristics - And for whom?

Study sampling transpired to be the most important predictor of effect size in the present research. Importantly, it has also suggested a common theoretical thread linking most of the findings (discussed fully below): Patients who are more likely to be distressed at baseline due to vulnerability characterising their socio-demographic or medical status show the greatest gains from therapy. In short, “It’s the sick who need a doctor” (Matthew 9:12) – and by implication, not everyone is 'sick'. The first patient characteristics discussed below are socio-demographic, then baseline distress is returned to in its capacity as a psychological variable, and finally medical variables are discussed.

Socio-demographics

Age.

Age was found to formally moderate effects in favour of older patients (mean sample ages of 53-66 years compared with 41-52 years) in unscreened but not screened samples. The unscreened effect was traced back to trends in education and expressive-supportive therapies. The strong homogeneity of the screened result suggests that distress or psychological non-complexity placed both age groups on the level. When considered in conjunction with the heterogeneity of the unscreened result, this suggested that distress and/or non-complexity were more influential than older age, but that perhaps older age was characterised by more distress.

If differential baseline distress is the invisible moderator at work in this age comparison then this finding, *prima facie*, seems at odds with the consistent finding referred to in the Introduction that younger patients have higher prevalence of anxiety and depression. However, the nature of the data contributing to this meta-analysis help explain this anomaly. Cancer is generally a disease of older age and the great majority of adult research interventions sample from older patients. The age group parameters of the present study reflected this and drew a defining line between young and old well above that which is commonly used to define younger patients (40 years). The means of the

present samples indicate that younger patients, by the usual definition, are represented only in light numbers in the present dataset.

If meeting vulnerability is key for therapy effectiveness then, of the two age subsets defined for the present study, the older one is likely to be the most vulnerable. Patients are likely to lose income and social support as they age, and also to accumulate more and serious physical diagnoses and to have distant spread disease. In terms of Andersen's conceptualisation of cancer distress (Andersen, 1992) older patients have accumulated more loss and have lost more social support. They sit at the convergence of other factors that were found by the present research to predict therapy effectiveness, which themselves may have such action because they heighten baseline distress, namely lower income and singleness (which is likely to be a proxy for reduced social support in Anglo-western societies). Older people are therefore in a particularly vulnerable position, and well placed to gain from therapy – if they can access it.

It is interesting that education and expressive-support are the therapies that the age moderation particularly traces back to. These therapies may fill particular needs. It was speculated that older people may have relatively less free access to information and that therefore the provision of information is particularly valuable. The extent of their need for social support, as a group, may be indicated by the fact that an analysis of the impact of marital status on interventions administered during terminal phase could not be conducted because there were no studies at terminal phase with enough married or partnered patients to qualify for the subset above the median split on the marital status variable. A critical need for social support may drive the expressive-supportive result.

Perhaps the most important point to take from the age finding is simply that older patients can and do produce worthwhile effect sizes *and should be treated*. It should not be assumed that misery is an expected and acceptable part of being old and ill. This may be the expectation, however, not only of some professionals working in the field, but also of many older patients themselves. As mentioned in the Introduction, there is research to suggest that depression itself predicted feeling ill-informed about service availability (Mehnert & Kock, 2008). If patients see their depression as realistic, they may not understand that therapy can still help. Furthermore, the moderating effect was found with unscreened patients. This suggests that either therapy should be offered to *all* patients of senior age regardless of proven distress, or that they should be carefully and routinely monitored for distress and encouraged into therapy accordingly. Being old and ill is difficult. Support can often be needed - and can be effective.

Race.

Previous literature has found little evidence of the differentiation of psychological distress or intervention effect based on race (refer Introduction) but there is not a great deal of it. The fact that, in the present research, analyses could not be conducted on the basis of a 'white v. non-white' split because of lack of non-white data is confronting to the research community. The data relating to service use referred to in the Introduction imply that clinicians in practice face a similar challenge. Campbell et al. (2007) point

out that racial minorities fare worse in cancer incidence and mortality rates, and therefore should be the especial target of intervention and research. The fact that many minorities suffer lower average income is another good reason to design research and services especially to reach them, since income was also found to be associated with effect size.

Gender.

Gender comparisons were made in several ways, including by therapy type and with sex-specific cancers removed. Almost perfect consistency showed a tendency for men to gain more effect than women. An overall comparison showed that the trend reached 0.10 level significance for screened samples, and approached it for unscreened, but once sex-specific cancers were filtered out, the disparity between men's and women's screened scores narrowed a good deal, leaving a trend in the same direction but without a statistically significant difference. However, from small *n*, significance appeared or was approached in the disparity for screened samples administered relaxation therapies and for screened and unscreened samples administered expressive-supportive therapies.

Considering the evidence as a whole, there is enough consistency here to say that men generally benefit more from therapy. Results are consistent with the earlier meta-analytic finding of Rehse and Pukrop (2003) that men produce effect sizes about twice the magnitude of those of women. Removing studies of sex-specific cancers probably made a difference because of the loss of breast cancer studies, which may form a unique population dynamic (refer to the discussion of cancer site, below).

Why might relaxation and expressive-support be especially impacting for men?

Generally, relaxation therapies teach a skill that can be used at the patient's discretion, and so may be considered a means to regaining some personal control. A qualitative study of men (not cancer patients) in a drug trial for depression concluded that the issue of personal control is of particular importance to them (Watts, 2002). Relaxation would therefore meet a point of particular vulnerability and need. Also, relaxation therapies provide a way of dealing with anxiety without talking about the stressor. This may fit with a typically male style for dealing with distress. Recall the recent German study described in the Introduction where male patients were found more likely to manifest addictive disorders than anxiety disorders (Krauss, Ernst, Kuchenbecker, Hinz, & Schwartz, 2007). Perhaps men are more comfortable dealing with anxiety in a way that is physical and under their own control.

Although it would take many out of their typical coping style, expressive-support may also fill a particular need for men. A tendency to avoid talking about emotions may reflect culturally imposed expectations about 'being strong'. If confidence can be conveyed that makes talking safe, expressive-supportive therapy may provide an emotional outlet and processing opportunity that is rare in the social world of men. In the report of their group psychotherapy intervention for cancer patients undergoing radiotherapy, Forester, Kornfeld, Fleiss, and Thompson (1993) commented that the opportunity for catharsis (to face emotional issues and vent feelings in non-judgmental

safety) seemed to be particularly appreciated by men, who were more distressed than women at baseline and benefited more from their therapy. They also noted that men had an initial tendency to inhibit communication of their feelings and concerns. The real difficulty that arises is that this discomfort about talking may mean that men decline an opportunity for therapy.

Men bring a particular style to group therapies that challenge the ingenuity of researchers and practitioners. Edgar, Rosberger, and Collet (2001) ran a psycho-education intervention for newly diagnosed breast and colon patients. Their coping skills therapists concurred that male patients paid lip service to coping skills but did not practice them, in either individually or group administered modes. Men “seemed to build a protective barrier around them[selves] to prevent having to engage in learning coping skills” (p.300). The therapists also felt that having a mix of genders in their colon cancer groups inhibited disclosure and participation. Perhaps the obedient learning of coping skills felt unacceptably like an admission or acceptance of weakness to the male group members. Helgeson, Cohen, Shulz, and Yasko (2000) wondered whether men may have greater desire for informational support than emotional support. It may be that men perceive it to be more culturally acceptable to do so.

Ironically it seems that men shy away from emotional disclosure in therapy and yet benefit the most from it. This aligns with the vulnerability hypothesis posited by the present study but requires researchers and practitioners to take their shyness into account in designing intervention. Safety around the whole issue of ‘weakness’ may be crucial. The success achieved by prostatectomy survivors who met one-on-one with patients who were adjusting to their diagnosis and treatment effects in the Weber studies (Weber, 2001; Weber et al., 2004; Weber et al., 2007) may have turned on the safety that the design of the study provided. Another means of bringing safety may be in how expressive-support is framed. The term ‘support’ itself suggests weakness and ‘discussion’ may also be threatening, so perhaps these terms should be avoided in the labelling of therapy for men, and replaced with an acceptable suggestion such as ‘education’. Another way that suggests strength and makes intervention acceptable to men is describing it as for the purpose of benefiting the patient's wife or family (clinician, C. Adams, pers. com. 22 June 2009).

These results do not give pointers only for men, but also for women, with implications on the reverse side of the coin. Contrary to the assumption made by a great deal of intervention literature, it seems that both screened and unscreened women do not get more than a small mean effect out of expressive-supportive therapies. Indeed, all the therapy type scores were in the null to small range for women, but the qualification must be noted that they would have been dragged down by effects from studies with samples that were predominantly breast cancer patients. With both female-specific cancer site study groups removed from analyses (breast and gynaecological), screened female patients produced an overall effect size which, though smaller than that for men, was nonetheless high and statistically significant, though the unscreened result remained small and non-significant. The implication of the relatively poorer result for women is

either that they tend to have natural strengths that provide much of what many need in the cancer experience, or, that therapies are not designed for them and are therefore ineffective. The latter scenario seems unlikely given the over-abundance of focus on therapies for breast cancer samples in the literature. The former aligns with the vulnerability hypothesis if it can be assumed that women tend naturally to build and use their own networks for emotional support and do not consider it a sign of weakness to ask for help or express their feelings. More on this in the discussion of cancer site findings.

In sum, the data confirm the finding of Rehse and Pukrop (2003) that men generally benefit more than women from psycho-oncological intervention, and it may be that culturally influenced responses to distress are behind this difference, meaning that intervention provides for needs not otherwise so effectively dealt with by men, as a subpopulation.

Education and income.

Recall that in spite of low study frequency, the distribution and effect sizes tabulated displayed a perfect inverse pattern of association between low income / occupational status and effect size, with statistical significance ($p < 0.10$) showing in the most sensitive of the four confound matrix quadrants. On sparse data, the same pattern was seen for CBT ($p < 0.05$) and it was speculated that, with more data, may have been seen for education also.

It is a truism that with any trauma, the poor are worst affected because they are least able to marshal the material and political resources to buffer themselves from the fullness of consequential losses and difficulties. Poor people therefore have the most to gain from intervention. If cancer diagnosis and treatment is conceptualised as a trauma, the high effect sizes accruing to poorer people in the present study can be seen as a further example of the deficit hypothesis – of vulnerability being met in therapy. The income related findings of the present study are consistent with the finding by Lepore, Helgeson, Eton, and Schulz (2003) that less educated men benefitted more in terms of physical functioning from their educational intervention, and with their observation that since the poor and uneducated suffer more from the disease in terms of mortality and quality of life, they stand more to gain. However, it may be that educational status does not prove to be such a clear predictor of effect generally, as income may be the critical factor and the two do not perfectly covary: It is quite common to be educated but poor, due to singleness or retirement for example, and therefore still exposed to many of the extra complexities that being ill and poor imply.

It can only be hoped that the clarity of the present findings will re-orient future research and clinical efforts towards poorer subpopulations. Coupled with persuasive recent evidence that cancer incidence and mortality are disproportionately higher among the poor (Dalton et al., 2008), there is a strong argument that poorer patients should be the especial target of service provision and research.

Marital status and social support.

The only regularly reported data that could provide a window into the mechanism of social support in relation to therapy effectiveness were study averages on participant marital status. The effect of the construct of interest was muted by a number of technical mediators described previously, but nonetheless (non-significant) trends were found whereby the studies with a lower percentage of married or partnered patients were associated with about double the effect size for both screened and unscreened patients.

Other indicators of social support - which suffer even more inadequacy as proxies for the construct of interest - may be therapy delivery variables such as group or individual mode, whether a significant other was present in therapy, and the intensity of the therapist / patient relationship. Results showed a non-significant trend in favour of therapies delivered one-on-one rather than in group mode or with a significant other, where the patients were screened (results were homogeneous where the patients were unscreened). This is consistent with the importance of quality or intensity of social support over quantity. Specific therapy components relating to the teaching of communication or relationship skills or social network establishment or utilisation may also be thought relevant to social support. However, these components yielded poor effect sizes though mostly from very low study frequencies.

Overall, factors relevant to social support produced limited but interesting evidence of expected tendencies, consistent with the bluntness of the data available. Picking up on Andersen's (1992) conceptualisation that accumulated losses and a lack of social support define cancer related distress, the relevance of this evidence is that it suggests that many single people – or others who lack quality social support – are likely to benefit more from therapy because they are likely to enter it with heightened distress. This factor should therefore become a greater focus of research and clinician awareness.

Baseline distress

The impact of baseline distress on therapy effectiveness was investigated as a substantive variable as well as a preliminary one, as discussed near the beginning of this chapter. Implications of the finding that elevated baseline distress does predict therapy effectiveness are important for research sampling and the targeting of services in practice settings.

Distress screening.

Various authorities were cited in the Introduction to the effect that although a large proportion of cancer patients suffer clinical levels of distress, an equally large or larger proportion does not. Now it is found that distress level before intervention is a moderator of therapy effectiveness. The argument for distress screening is simple and obvious: Not all cancer patients are distressed, and treating cancer patients without proven distress is likely to result in null, negligible or small effect sizes, so precious resources should be directed towards patients exhibiting distress only. The issue then

becomes how to effectively and efficiently screen for distress and reach the individuals and subpopulations in most need. Data noted in the Introduction showed that services are not always effective at reaching needy groups, although many do admit only patients who screen elevated distress. For researchers, as discussed early in this chapter, the challenge becomes to resist the pressure to accumulate a large sample by recruiting unscreened patients, and instead focus on the problem of recruiting subpopulations who are likely to produce higher effect sizes but are more difficult to access for both ethical and practical reasons.

Two qualifications to the general position in favour of distress screening are made: First, it is not known how educational interventions respond to baseline distress, because there were no educational studies that screened in for it. Only a hint of greater effect can be taken from some individual studies where it seems that patients probably had elevated distress at baseline. Further, education serves a functional need that surely applies to most patients, and so should be provided regardless. Cunningham (2000) suggests intervention modules of varying content grades, exposing all patients to some psycho-education and inviting or referring others with particular need on to further more advanced therapy. Second, there are socio-demographic and medical variables that have shown, in the present study, that they moderate effect size despite a lack of baseline screening: to lesser or greater degrees, groupings that show this tendency include the poor, the old, men, non-breast cancer patients, and perhaps distant spread patients. This means that there are certain groups of people, *per se*, who are likely to produce higher effect sizes, which may be associated with higher average baseline distress due to pressures associated with their distinguishing characteristic such as social or economic marginalisation, more serious disease symptoms or prognosis, or more disabling or disfiguring treatment side-effects.

It is suggested that both research and practice contexts use approaches based on risk screening or awareness *and* distress screening. Risk is important because not all patient groups are created equal in their exposure to cancer and its accompanying stress, as has been mentioned. To target cancer treatments at patients equally is therefore "stupid" - to use the term of Danish epidemiologist, Dr. Chris Johansen (co-author of Dalton, 2008, pers. com 23 June 2009). It is also unjust. If it is considered credible that the socio-demographic and medical groups identified by this study as producing higher effect sizes do so because they suffer higher average distress, then risk screening ought to include items that identify them, and should be standard following diagnosis. The recently published Australian clinical practice guideline for the psychosocial care of cancer patients recommends both vulnerability and distress screening (Breast Cancer Centre and National Cancer Control Initiative, 2003). Their model uses a vulnerability checklist and refers vulnerable patients directly for further assessment. Depending on criteria for 'vulnerability', such an approach could prove very expensive. However, at the very least, medical staff should be trained in awareness of the special vulnerabilities that some subpopulations carry and taught to keep them under especial surveillance – especially patients who fall into more than one such subpopulation.

Repeated distress screening also needs to be scheduled for all patients at stages in the cancer journey when effect sizes are highest (refer to the discussion of medical variables, below). Cunningham (1988) advocates regular distress screening for all patients because of the up and down trajectory of cancer-related distress and the reality that distress will not necessarily come to the attention of staff, and this is slowly becoming the best practice recommendation in various countries around the world. There remain challenging issues around reaching and assessing some subpopulations, because of their particular styles of self-reporting distress (men, older patients) or because of social marginalisation or physical incapacitation (low income patients, patients with advanced disease or marked disability). However, these subpopulations are the very ones that associate with heightened cancer-related distress, so research on systems and tools for monitoring distress effectively and efficiently deserve priority. It is pleasing that at world conference level this topic is receiving increasing attention (International Psychosocial Oncology Society, 2009). It may be, for example, that modern touch screen technology will efficiently facilitate distress screening prior to every medical consultation and treatment appointment in developed countries in the foreseeable future, but it will remain important that medical staff are trained to be sensitive to manifestations of distress and make their own enquiries, particularly of patients in at risk subpopulations.

Medical characteristics

There were medical variables that were fairly regularly reported on, which allowed some analysis of them. However, differentiation on these did not serve to predict intervention effectiveness as generally as might have been expected, possibly because of constricted variation in the sense that early stage and breast cancer patients were sampled with disproportionate frequency. Only one medical moderator could be discerned, but there were other results that showed the effectiveness of therapy for particular subpopulations.

Cancer site.

The most spectacular result from the medical variables investigated was moderation by cancer site. After removing the potentially confounding influence of an imbalanced distribution of distress-screening favouring non-breast studies, consolidating screened out with unscreened data to increase power, and removing studies with patients from a mixture of sites in order to sharpen contrast, it was found that studies comprised predominantly of breast cancer patients produced significantly lower effect sizes compared with a grouping of all other single sites combined ($p < 0.05$ for untreated control comparison data and < 0.10 for treated). Breast studies produced only a small effect against untreated controls which disappeared completely when compared with placebo or treatment as usual controls. It was noted that breast cancer studies are so predominant in the literature that they contributed more than twice the number of studies to these analyses than all other single site studies put together. It should be noted, however, that approaching three quarters of these breast studies were for local

spread cancer, whereas only half from other sites were, which may have slightly biased effects against the breast studies. More importantly though, the lack of later stage studies shows another gross imbalance in research focus.

This moderation result aligns with the finding by Edgar, Rosberger, and Nowlis (1992) in their CBT study which compared two intervention points with newly diagnosed patients. They found that breast cancer patients improved over the year regardless of whether their intervention was timed as soon as possible after diagnosis or after four months, and that it was patients with other cancers who benefitted most from their intervention. Their finding suggests that cancer-related distress can be managed without professional intervention more readily by breast cancer patients than by others.

It was speculated that this moderation may be due to the relatively good prognosis and level of functioning associated with breast cancer together with a lack of social stigma and ready access to peer and voluntary support – less loss plus more support, in terms of Andersen’s conceptualisation of cancer distress (Andersen, 1992). Early stage breast cancer carries quite a good prognosis these days (refer Introduction). Also breast cancer prevalence is the greatest of any cancer type and sufficiently high to provide a pool of like individuals for informal and organised mutual support, and women with a common cause can be particularly good at supporting each other. For example, the writer herself received valuable well targeted practical support and advice from a peer survivor who, connected only as the friend of a friend, took the initiative to contact and visit her before she began chemotherapy. The author has since taken the initiative to provide similar support to newly diagnosed women. Although many breast patients may be as distressed as other cancer patients initially, if they are better able to support each other over time, then effect sizes will be ‘floored’ by a lack of baseline distress in many intervention studies. Furthermore, the distress of control groups will drop over time without formal intervention, thus reducing their contrast with treated groups and lowering effect size.

The obvious indication of these findings is a refocusing of research attention away from breast cancer patients – but reserving attention for later stage sufferers and those with established baseline distress. Other cancer populations do pose significant ethical and logistical challenges for empirical investigation, but there are researchers who have overcome these (a list was provided in the Patient characteristics results chapter).

Prognosis, cancer stage, medical protocol.

The skew of study frequencies towards cancer sites with more favourable prognoses, disease stages that are earlier, and curative rather than palliative treatment protocols, was marked. In the case of the comparison cancer sites with different prognoses, the skew was so severe as to completely thwart moderator analysis and the gathering of any other useful data. The cancer stage analysis produced a notable moderate strength result from unscreened distant spread patients suggesting that this subpopulation should be monitored for distress or offered therapy as standard procedure. Relaxation therapies had greatest study frequency and performed well for distant spread patients. The

medical protocol analysis showed low and non-significant effect sizes for both screened and unscreened palliative protocol patients. This was somewhat surprising given the distant spread result and the literature showing that 50% of patients with advanced cancer have psychiatric disorder (Miovic & Block, 2007). It may be that special attention is needed to design effective therapies for this population.

For all of these variables, the imbalanced research focus towards relatively ‘easy’ cases needs to be addressed and, again, the ethics and logistics of doing so are challenging.

Medical treatment stage.

Statistical heterogeneity was not found between any of the medical treatment stages, however, very similar moderate magnitude and statistically significant effect sizes were found for both screened and unscreened patients who were administered therapies during the period late in medical treatment or upon resumption of normal life. While this finding could benefit from replication with more studies, it is consistent with the suggestion put forward by Freidenbergs et al. (1981-1982) mentioned earlier in relation to relaxation therapies, and may be a pointer towards an important window in which to reach patients. With more studies it may also be possible to establish a strong effect for screened patients who are newly diagnosed. Importantly, a moderate to high magnitude effect was produced for unscreened patients at recurrence / palliative stage (but from only three studies).

Read together, these data may suggest that times that demand considerable adjustment are particularly opportune for therapy – probably because they are, again, associated with heightened distress. Such times particularly demand acquiring altered identity, goals and skills, and so require acceptance, learning and the locating of needed resources. These types of adjustment fit more comfortably with the self-efficacy and coping theory concepts (Bandura, 1997; Lazarus & Folkman, 1984) than with the loss / support conceptualisation of cancer-related distress (Andersen, 1992) espoused throughout most of the present paper, since newly diagnosed patients and those returning to normal life after treatment are not at particularly high risk of loss or necessarily particularly lacking social support.

Conclusion

The only firm hypothesis of this research was that the presence of clinically significant baseline distress would predict higher effect size, and that implications from that dynamic would run through other moderation analyses. Confirmatory evidence of the moderating effect of baseline distress was duly found. There was also evidence that a number of subpopulations – mostly defined by socio-demographic characteristics – produce higher effect sizes, and the common thread of probably elevated distress provides an explanation that runs through them all. For the most part, this elevated distress has been attributed to greater loss and less social support (Andersen, 1992) – that is, greater vulnerability in dealing with the cancer experience - but in relation to medical stage (not found to be a formal moderator) an explanation of distressing

transition was attributed. The socio-demographic characteristics associated with higher effect size are greater age, lower income, single marital status (as a proxy for low / poor social support) and male gender. The one medical characteristic that predicted higher effect size was diagnosis with a cancer other than (early stage) breast cancer. It is to be noted that many patients fall into more than one of these groups, and that a high proportion of older patients in particular will do so as they tend to be poorer and single, most will have a cancer other than breast cancer and about half of them will also be male. Likewise many poorer patients will be multiply exposed.

This list appears to cover most cancer patients, and should therefore be of little value for predicting therapy effectiveness. In a world where resources were evenly allocated, this statement would be true. However, as study frequencies in the present research have shown repeatedly, current research practice drastically over-samples the one group of cancer patients that does not fit into these categories: middle class women with early stage breast cancer and without proven baseline distress. In short, the focus of research attention is upside down.

The contradiction between where higher effect sizes are to be found and where the research focus actually is explains much of the conflict and angst in the literature – compounded by the tendency for the big stratified RCTs in particular to sample from the breast cancer group. The narrow sampling focus, especially by these big ‘quality’ studies, has produced apparent evidence that psycho-oncological therapies do not work. The conclusions of reviews and meta-analyses – particularly those that have excluded all but RCTs - have often been distorted accordingly, giving psycho-oncology a dubious name. At the root of this distortion has been the simple neglect of selecting for baseline distress. Yet some data and the experience of many clinicians have all along told a different story.

The likely reason for this unfortunate sampling practice is simply that unscreened early stage breast cancer patients are available for research in the numbers sought – they are effectively a ‘convenience sample’. Given the logistical and ethical difficulties associated with researching clinical populations, the availability of a large homogeneous sample is important. Other considerations of convenience will also have taken a toll. The common practice of requiring dominant language proficiency will have excluded some minority patients, for example. Recall that no analysis of the moderation of race could be performed because the data for non-white samples were so sparse. Similarly, the analysis of income may have been enabled only because retired people were counted in the low income bracket. Gaston Johansson et al. (2000), who utilised a sample of highly educated, married, white women with high incomes for their research, explained that there were barriers to the participation of other groups, including lack of transportation, inability to take time off work, lack of insurance cover, and general financial constraints. In other words, the poor, comprised disproportionately by racial minorities, could not afford the luxury of therapy in the research context. Are they any more able to afford it in the practice context? Obviously, late stage patients pose particular ethical and mobility problems. However, with one in two cancer patients still

succumbing to the disease, meaning that half the cancer population will experience late stage, it is essential to find a way around the obstacles to studying and serving these people. Some researchers have done so. Actually, some researchers have been studying distressed individuals and more vulnerable subpopulations all along. The rest of the field must now follow their lead.

Study limitations and strengths

A meta-analysis is exposed to weakness both from its own design and from the characteristics of the primary studies that contribute to its database. Although steps can be taken to detect and account for some biases associated with primary study design, nothing can be done to compensate for data that were never collected. The limitations of the present study therefore arise from both externally imposed constraints and choices made. However, it is considered that overall meta-analytic method has been applied diligently, and within limitations described below, the main conclusions reached should be both sound and valuable.

Constraints imposed by the data

Rarefication of sample.

Meta-analysis results are necessarily tied to research conditions, which will only approximate to clinical conditions to a lesser or greater degree. There is a concern in the case of the present study that much of the data that contributed to analyses were rarefied in the sense that many of the more physically and psychologically complex cases that are seen by clinicians daily were excluded by recruit selection criteria. It was common to read primary study inclusion / exclusion criteria like these from Antoni et al. (2001, p.21):

Criteria for inclusion were (a) breast cancer diagnosed at Stage II or below and (b) surgery within the last 8 weeks. Potential participants were excluded if they reported a prior cancer, prior psychiatric treatment for serious disorder (hospitalisation or formal diagnosis or psychosis, major depressive episode, panic attacks, suicidal, or substance dependence) major concurrent disease, or lack of fluency in English.

Or these from Walker, Nail, and Croyle (1999, p.1026):

Women who were 18 years of age or older; had a Karnofsky performance status of 70% or higher; were independent in self-care; were oriented to time, place and person; spoke and read English; and were completing RT for stage I or II breast cancer were eligible. The Karnofsky performance cut-off was chosen to avoid burdening patients who had multiple illnesses....The sample was limited to women with localised breast cancer to control for variation in extent of disease and type of RT that could influence both the content of the emotional expression essays and outcomes.

Whilst some of the particular exclusion criteria in such examples were taken into account in the design of the present study by coding ‘screened out for history of psychological distress’, criteria that related to other cancers or serious disease, advanced stage cancer, less than a certain level of physical performance, current psychosis, suicidality or cognitive impairment were not. They were very common exclusions, presumably designed for patient safety, scientific homogeneity, and logistical practicality. However, if such exclusions are extreme to the point of rarefying the sample relative to clinical reality, then study results cannot speak to clinical conditions.

The patients that are referred to psycho-oncologists in the course of their daily work are often both physically and psychologically complex. Indeed, it is often precisely because patients have suffered the ‘double whammy’ of accumulated problems – perhaps multiple serious illness, or the recent loss of a loved one combined with news of the seriousness of their own cancer - that they find they cannot cope (D. M. Baken, pers. com. April 2008). Psycho-oncologists in practice do treat ‘the sick’ - not relatively simple cases. Because of this discrepancy between the research and the clinical context, *all of the conclusions of the present study should be tempered by clinical advice.*

Choice of variables.

This study is also blinkered by the positivist paradigm of which the meta-analytic method is an extreme example. This means that conclusions can be drawn only in areas where attention was focussed in accordance with the *a priori* plan of analysis, to the exclusion of other areas that might in fact be more important. There was a small amount of flex enabled by the iterative approach taken to coding studies, but certainly this study was constrained to those variables for which data were regularly reported. The risk is that there is a group of important variables that are not recognised – or are under-recognised - by the present study. In the Introduction it was noted that Bardwell et al. (2006) had found the influence of psychosocial mechanisms (notably self-efficacy / optimism and social support) to be much more influential in predicting depression prevalence amongst a large sample of breast cancer patients than socio-demographic or medical variables, even including physical functioning / symptom variables. These mechanisms were ones that the present study struggled to tap due to limited data. It needs to be understood that because the present work could not highlight these features does not mean that they are of little importance. Quite the contrary might be true.

Another choice made in relation to variables was that to analyse a great number of them. This was made in order to get the roundest fullest feel of the dataset possible, and proved worthwhile to that end. It may raise a reservation as to the risk of chance significance. In response it is noted that the analyses, though large in number, were substantially planned *a priori*. They could not be entirely pre-planned because how substantive analyses were approached depended on judgments made on the basis of preliminary analyses. A further response is – in keeping with the approach taken throughout this study – to encourage the reader to consider the whole of the evidence

rather than any particular result or alpha level. Conclusions were frequently drawn from more than one analysis and the results of this study fall into a theoretically sensible pattern. It is this overall weight of the evidence that is important.

Choices made relating to method

English language bias.

Because only English language studies were admitted to the present research, a strong Anglo-western bias will be present in results also. There were studies admitted that originated in non-English speaking countries, but there were others that could not be used because of the language requirement, and the impact that their exclusion may have had on results is unknown. What is known is that Anglo-western societies were the best represented by far (more than 75% of studies were from the USA, UK, Canada or Australasia), that they are OECD countries, and that poorer economic status predicts higher cancer morbidity and mortality – and, apparently, higher effect size. Anglo-western societies also have their own cultural and spiritual perspectives on illness, suffering and death, which differ considerably from many eastern or minority cultures. It must be taken into account that the present study is largely a product of this one Anglo-western perspective and accordingly has limited generalisability in culturally sensitive areas of practice, perhaps particularly regarding existential issues and therapy by expressive-support.

Illness specific measures.

As explained in the Introduction, some symptoms of cancer disease and treatment can mimic distress symptoms, and it is therefore best to use illness-specific measures in assessing distress. For the present study, the original intention was to gather data drawn from both illness-specific and general measures and then run a sensitivity analysis to see if they produced significantly different effect sizes. It was later decided to strengthen the dataset by excluding data derived from a non-specific measure where data from an illness-specific measure was available for the same outcome construct in the same study. With hindsight, the writer feels that the original plan was better because it would have made any effect explicit and allowed informed consideration of how to address the matter, including the possibility of an across-the-board exclusion of data derived from non-specific measures.

Having noted the problem, consolation is taken in that the rule applied did reduce the incidence of data derived from non-specific measures in the dataset, and there were only three instances where a non-illness-specific and professionally-rated measure was used – the type of measure most exposed to bias from somatic items. One of those studies delivered an expressive-supportive group therapy by a psychiatrist and distress effect sizes are very large. Another was a problem solving intervention study. While use of the Hamilton Rating Scale for Depression did produce very high effect sizes, this study used non-specific self-reports to measure distress that did also. The other study produced modest scores using the Hamilton Rating Scale for Anxiety. It is hoped that

this weakness in the present study is more theoretical than manifest, but that future reviewers and meta-analysts will do better.

Method approach and execution

It is considered that both the decision to run preliminary analyses of study design features, and the considerable efforts made to collect unpublished data and categorise them in valid and reliable ways, constitute strengths that the present meta-analysis has above many others. The use of preliminary analyses rather than presumptive exclusion criteria not only provided an empirical justification for how design features were then handled, but also gave the opportunity to explore the research dynamics at work in the dataset, leading to findings that should be of considerable future value to the field. This process formed a very firm foundation upon which substantive analyses could be made.

The effort and expense put into obtaining unpublished data, both from database searches and from individual researchers, is rare in psycho-oncology and yet this emphasis on the fundamentals of meta-analytic method proved most worthwhile. Whilst the considerable number of variables analysed also placed strain on the researcher, the effort was rewarding since it produced a theoretically coherent set of conclusions and an important explanation of the conflicting evidence that has dogged the field in the past, together with the evidential basis for some far-reaching recommendations for future research and practice.

Application

Following are two lists – one in relation to practice, the other, research – of recommendations arising out of the present study. It will be noted that some items are substantiated by formal moderation findings and others by evidence of lesser substance or just by ‘hints’ in the findings. Less emphasis is placed on the latter.

Practice recommendations

1. **Risk and distress screening.** Patient characteristics that have predicted or trended towards higher effect sizes (older (and, probably, under 40), poorer, male, single, cancer site other than breast) should be included in a risk or 'vulnerability' screen as described above (under distress screening) together with psychological history and present distress symptoms. Both initial risk and regular distress screening should be built into national practice guidelines and standard medical procedures. Particularly sensitive points for distress screening are shortly after diagnosis, towards the end of medical treatment, at recurrence and at distant disease spread. It is important that patients at risk not be relied upon to identify their own distress and seek help, nor that medical staff be relied upon to detect distress simply in the course of their duties.
2. **Service access through distress.** Elevated distress should be the entry criteria for most therapies. Interventions that fulfil general functional needs (e.g. education) are an exception to be provided to all. However, there are efficiencies that are possible in delivery of those therapies (refer below).

3. **Effective and efficient delivery mechanisms.** Less well paid professionals such as social workers can very effectively deliver manualisable therapies (relaxation and CBT) to distressed patients who do not have psychologically complex backgrounds and whose circumstances are otherwise relatively straightforward. The services of more highly paid professionals (psychologists and psychiatrists) can be reserved for more complex cases and also to give attention to tuning systems of general medical care to reduce and detect distress. Where it is not an integral part of the therapy (as it often is for relaxation and CBT) homework, beyond a one-off introductory dose, it may be best to do without it. Brief therapies can be highly effectively delivered to vulnerable subpopulations. Longer therapies will no doubt be required when cases are complex. It seems likely that much educational material and perhaps some relaxation therapies may be as effectively provided to unscreened patient populations by non-interactive means (e.g. DVD or pamphlet) as by interactive means. Psychological quality in both content and form should be ensured.

Research recommendations

1. **Sample selection.** Subject to allowance for providing for the functional needs of all patients with appropriate educational intervention, research samples should always be either pre-screened for distress or selected on the basis of reason to believe that a particular subpopulation generally suffers heightened distress / produces higher effect sizes. Present research findings emphasise moderation by sample characteristics rather than by therapy or delivery characteristics, and rather than by details of study design. Researchers therefore need to attend to the issue of reaching the right people as a matter of first priority. In some cases payment for time and/or transport will be all that is required to enable poorer or minority people to participate. Another option may be to take the therapy to the patient when he or she is already at hospital for medical treatment or consultation. It has been mentioned that a simple mechanism such as the label used to refer to a therapy type may make a difference to the participation of men, and this may apply to some cultural groups also. Although some sensible method of control is necessary to head off maturation and selection effects, appropriate sampling should not be sacrificed in favour of a particular method of allocating participants to groups (randomisation). It is also more important to reach more vulnerable subpopulations (e.g. those suffering less common cancers) than to attain large sample size. Lower sample size reduces the power to find statistically significant results, but this should be compensated for by higher effect sizes. Furthermore, effect sizes can and should be published regardless of their statistical significance and these data can be combined with those from other small studies by meta-analysis in due course (Kline, 2004). Preliminary research directed solely at how to effectively recruit an appropriate sample may well be required in relation to vulnerable subpopulations.

2. **Patients without history of psychological distress.** It may be that such non-complex patients can take better advantage of some therapies – particularly more demanding cognitive or behavioural therapies – or therapy generally, and retain effects longer. Data distinguishing them should be broken out and analysed for differences in effect size and durability.
3. **Data needed.** Longer term data (six months and one year) are needed in every area. For the sake of enabling studies to be incorporated in future meta-analyses, effect sizes or means and standard deviations at post test, at least, should be provided. Illness-specific measures should be used. Except where a particular combination of therapy types is indicated theoretically, therapies should be delivered individually to enable the effects of particular therapies to be understood. (Refer to Table 6-2 for data gaps relative to specific therapy types.)
4. **Relaxation issues.** Relaxation is an anxiety therapy and yet there were results that indicated that it could be effective against depression where patients were distressed at baseline, inviting research. Care would be required, so as not to overburden distressed patients (Given et al., 2004), but this could be a fruitful area. Relaxation also showed particular potential in relation to older patients, patients at late stages of the disease, patients coming out of treatment, and men, and each of these areas could be usefully developed by further research.
5. **Theoretical mechanisms.** Mechanisms of effectiveness should drive therapy design, including its fit with sample characteristics, and be explicated, measured, and built into analysis where possible. Self-efficacy has demonstrated its relevance in the present study, and the quality of patients' social support should be much more often assessed and factored into result analyses. Delivery modes that convey a particular sense of security (sessions stretched over more weeks, one-on-one, professional therapist) may be interesting to investigate.
6. **Indirect and professionally supervised non-professional interventions.** The considerable potential of these types of intervention to efficiently and effectively reach patients in large numbers or who would otherwise not be accessible is exciting and deserves attention.
7. **Existential issues.** This is a relatively poorly researched area in Anglo-western countries but expressive-supportive therapies relating to existential, spiritual, grief or death issues produced worthwhile results, as did the oncologist delivered intervention by Kristeller et al. (2005). This is another area that deserves more attention.

8. **Middle and lower income countries.** The need for research is great in these countries and effect sizes from non-OECD nations (who mostly used distress screening and untreated control comparisons) were strong.
9. **Reviews and meta-analyses.** To avoid the distortion of conclusions, future syntheses of this body of literature need to take into account the two study variables found to structure the dataset, namely recruit screening (at least for distress, and preferably for psychological history) and the nature of the control comparison. All studies with comparable controls should be admitted, not just RCTs. The biasing of the dataset by the preponderance of low scoring studies of unscreened middle class early stage breast cancer patients should be taken into account in both designing syntheses and drawing conclusions.
10. **Foreign language studies.** There is a need to incorporate the body of trial data reported in foreign languages into meta-analysis of moderators. The opportunity exists for the present dataset to be used as a platform to which data in other languages could be added and analyses re-run. Some allowances for cultural difference would be necessary, but outcomes would be powerfully persuasive.

Closing reflection

When the proposal for funding this research was originally submitted, it was not imagined that such substantial and far reaching recommendations for practice and research would result. However, the approach that was taken has revealed that the confusion besetting the field goes to a deep level, and therefore has broad implications. The poor effects shown by many studies – both primary and review studies - were a product of blind adherence to bio-medical ideals in research i.e. the notion that 'the large randomised double-blind trial' is the model of quality for - despite Cochran's advice that there is no 'gold standard' study design (Cochrane Collaboration, 2006, para. 6.11) and of how ill this design fits the field. It has been seen that, for psycho-oncological interventions, even single blinded trials are not usually possible, and that the patient group most commonly sampled because of their availability in larger numbers – unscreened women with early stage breast cancer – produce the lowest effect sizes. There are also stringent ethical constraints making randomisation impossible for much research on this doubly clinical population (Bredart, Cayrou, & Dolbeault, 2002). One way and another, there is not much left of this bio-medical ideal that works in psycho-oncology, and to hold the field to it – as many reviewers and meta-analysts have done in their primary study inclusion or grading criteria – is to pervert conclusions and consequently mislead administrators and policy makers to the cost of patients. It is hoped that, in future, both researchers and reviewers will feel justified and emboldened by the evidence provided by the present study to think in more creative and effective scientific ways about how best to approach work with those of this special population who need intervention.

References

- Ali, N. S., & Khalil, H. Z. (1989). Effect of psychoeducational intervention on anxiety among Egyptian bladder cancer patients. *Cancer Nursing, 12*(4), 236-242.
- Allard, N. (2006). Day surgery and recovery in women with a suspicious breast lesion: Evaluation of a psychoeducational nursing intervention. *Canadian Oncology Nursing Journal, 16*(3) 137-144.
- Ambler, N., Rumsey, N., Harcourt, D., Khan, F., Cawthorn, S., & Barker, J. (1999). Specialist nurse counsellor interventions at the time of diagnosis of breast cancer: comparing 'advocacy' with a conventional approach. *Journal of Advanced Nursing, 29*(2), 445-453.
- American Cancer Society (2007). *Cancer statistics presentation*. Retrieved 4 December, 2007, from <http://www.cancer.org/docroot/PRO/content/PRO>.
- Andersen, B. L. (1992). Psychological interventions for cancer patients to enhance quality of life. *Journal of Consulting and Clinical Psychology, 60*(4), 552-568.
- Andersen, B. L., Farrar, W. B., Golden-Kreutz, D. M., Glaser, R., Emery, C. F., Crespin, T. R., Shapiro, C. L., & Carson, W.E. (2004). Psychological, behavioral, and immune changes after a psychological intervention: A clinical trial. *Journal of Clinical Oncology, 22*(17), 3570-3580.
- Andrykowski, M. A., & Manne, S. L. (2006). Are psychological interventions effective and accepted by cancer patients? I. Standards and levels of evidence. *Annals of Behavioral Medicine, 32*(2), 93-97.
- Antoni, M., Lehman, J., Kilbourn, K., Boyers, A., Culver, J., Alferi, S., Yount, S., McGregor, B., Arena, P., Harris, S., Price, A., & Carver, C. (2001). Cognitive-behavioral stress management intervention decreases the prevalence of depression and enhances benefit finding among women under treatment for early-stage breast cancer. *Health Psychology, 20*(1), 20-32.
- Arden-Close, E., Gidron, Y., & Moss-Morris, R. (2008). Psychological distress and its correlates in ovarian cancer: A systematic review. *Psycho-Oncology, 17*(11), 1061-1072.
- Badger, T., Segrin, C., Dorros, S. M., Meek, P., & Lopez, A. M. (2007). Depression and anxiety in women with breast cancer and their partners. *Nursing Research, 56*(1), 44-53.
- Badger, T., Segrin, C., Meek, P., Lopez, A. M., Bonham, E., & Sieger, A. (2005). Telephone interpersonal counseling with women with breast cancer: symptom management and quality of life. *Oncology Nursing Forum, 32*(2), 273-279.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bardwell, W. A., Natarajan, L., Dimsdale, J. E., Rock, C. L., Mortimer, J. E., Hollenbach, K., & Pierce, J. P. (2006). Objective cancer-related variables are not associated with depressive symptoms in women treated for early-stage breast cancer. *Journal of Clinical Oncology, 24*(16), 2420-2427.
- Barsevick, A. M., Sweeney, C., & Haney, E. (2002). A systematic qualitative analysis of psychoeducational interventions for depression in patients with cancer. *Oncology Nursing Forum, 29*(1), 73-86.
- Berglund, G., Petersson, L. M., Eriksson, K. C., Wallenius, I., Roshanai, A., Nordin, K. M., Sjöden, P. O., & Häggman, M. (2007). "Between Men": a psychosocial rehabilitation programme for men with prostate cancer. *Acta Oncologica, 46*(1), 83-89.

- Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., and Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York: John Wiley & Sons.
- Bloom, J. R., Ross, R. D., & Burnell, G. (1978). The effect of social support on patient adjustment after breast surgery. *Patient Counselling and Health Education, 1*, 50-59.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2007). Comprehensive Meta-Analysis (Version 2.2.046):<http://www.Meta-Analysis.com>.
- Boyes, A., Newell, S., Girgis, A., McElduff, P., & Sanson-Fisher, R. (2006). Does routine assessment and real-time feedback improve cancer patients' psychosocial well-being? *European Journal of Cancer Care, 15*(2), 163-171.
- Breast Cancer Centre and National Cancer Control Initiative (2003). *Clinical practice guidelines for the psychosocial care of adults with cancer*. Camperdown, N.S.W., Australia: National Breast Cancer Centre.
- Bredart, A., Cayrou, S., & Dolbeault, S. (2002). Re: Systematic review of psychological therapies for cancer patients: Overview and recommendations for future research. *Journal of the National Cancer Institute, 94*(23), 1810-1812.
- Brennan, J. (2001). Adjustment to cancer coping or personal transition? *Psycho-Oncology, 10*, 1-18.
- Bultz, B., Specia, M., Brasher, P., & Page, S. (2000). A randomized controlled trial of brief psychoeducational support group for partners of early stage breast cancer patients. *Psycho-Oncology, 9*, 303-313.
- Campbell, H., Phaneuf, M. R., & Deane, K. (2004). Cancer peer support programmes - do they work? *Patient Education and Counseling, 55*(1), 3-15.
- Campbell, L. C., Keefe, F. J., Scipio, C., McKee, D. C., Edwards, C. L., Herman, S. H., Johnson, L. E., Colvin, O. M., McBride, C. M., and Donatucci, C. (2007). Facilitating research participation and improving quality of life for African American prostate cancer survivors and their intimate partners: A pilot study of telephone-based coping skills training. *Cancer, 109*(2 Suppl), 414-424.
- Carey, M. P., & Burish, T. G. (1999). Etiology and treatment of the psychological side effects associated with cancer chemotherapy: A critical review and discussion. In R. M. Suinn & G. R. VandenBoss (Eds.), *Cancer patients and their families: Readings on disease course, coping, and psychological interventions*. Washington D.C.: American Psychological Association.
- Carlson, L. E., & Bultz, B. D. (2004). Efficacy and medical cost offset of psychosocial interventions in cancer care: Making the case for economic analyses. *Psycho-Oncology, 13*(12), 837-849.
- Chan, C., Ho, R., Fu, W., & Chow, A. (2006). Turning curses into blessings: an Eastern approach to psychosocial oncology. *Journal of Psychosocial Oncology, 24*(4), 15-32.
- Chan, C. L., Ho, R. T., Lee, P. W., Cheng, J. Y., Leung, P. P., Foo, W., Chow, L. W., Sham, J. S., and Spiegel, D. (2006). A randomized controlled trial of psychosocial interventions using the psychophysiological framework for Chinese breast cancer patients. *Journal of Psychosocial Oncology, 24*(1), 3-26.
- Chan, Y. M., Lee, P. W. H., Fong, D. Y. T., Fung, A. S. M., Wu, L. Y. F., Choi, A. Y. Y., Ng, T. Y., Ngan, H. Y. S., and Wong, L. C. (2005). Effect of individual psychological intervention in Chinese women with gynecologic malignancy: A randomized controlled trial. *Journal of Clinical Oncology, 23*(22), 4913-4924.

- Chekryn, J. (1984). Cancer recurrence: Personal meaning, communication, and marital adjustment. *Cancer Nursing*, 7(6), 491-498.
- Chow, E., May, M. N., & Harth, T. (2004). Does psychosocial intervention improve survival in cancer? A meta-analysis. *Palliative Medicine*, 18(1), 25-31.
- Clarke, D. M. (1998). Psychological factors in illness and recovery. *New Zealand Medical Journal*, 111, 410-412.
- Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York: John Wiley & Sons.
- Cochrane Collaboration (2006). *Cochrane handbook for systematic reviews of interventions, version 4.2.6*. Retrieved 30 August 2007, from <http://www.cochrane.org/resources/handbook/hbook.htm>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2 ed.): Lawrence Erlbaum Associates, London.
- Corchado, J. L. (2006). *The effects of preparatory sensory information on breast conservative and mastectomy cancer patients: Mood states and self-concept*. Unpublished thesis 3244765, Widener University School of Nursing, Pennsylvania, USA.
- Coyne, C. A., Xu, R., Raich, P., Plomer, K., Dignan, M., Wenzel, L. B., Fairclough, D., Habermann, T., Schnell, L., Quella, S., & Cella, D. (2003). Randomized, controlled trial of an easy-to-read informed consent statement for clinical trial participation: A study of the Eastern Cooperative Oncology Group. *Journal of Clinical Oncology*, 21(5), 836-842.
- Coyne, J. C., & DeLongis, A. (1986). Going beyond social support: The role of social relationships in adaptation. *Journal of Consulting and Clinical Psychology*, 54(4), 454-460.
- Coyne, J. C., Stephen, S. J., & Palmer, S. C. (2006). Efficacy of psychosocial interventions in cancer care: Evidence is weaker than it first looks. *Annals of Behavioural Medicine*, 32(2), 104-110.
- Cull, A., Gould, A., House, A., Smith, A., Strong, V., Velikova, G., Wright, P., & Selby, P. (2001). Validating automated screening for psychological distress by means of computer touch screens for use in routine oncology practice. *British Journal of Cancer*, 85(12), 1842-1849.
- Cunningham, A. J. (1987). *An evaluation of cognitive behaviour therapy for the relief of psychological distress in cancer patients*. Unpublished thesis NL35787, York University, Toronto, Canada.
- Cunningham, A. J. (1988). From neglect to support to coping: The evolution of psychosocial interventions for cancer patients. In C. L. Cooper (Ed.), *Stress and breast cancer* (pp. 135-154). New York: Wiley.
- Cunningham, A. J. (2000). Adjuvant psychological therapy for cancer patients: Putting it on the same footing as adjunctive medical therapies. *Psycho-Oncology*, 9, 367-371.
- Cunningham, A. J., & Tocco, E. K. (1989). A randomized trial of group psychoeducational therapy for cancer patients. *Patient Education and Counseling*, 14(2), 101-114.
- Cwikel, J., & Behar, L. (1999). Social work with adult cancer patients: A vote count review of intervention research. *Social Work in Health Care*, 29(2), 39-67.

- Cwikel, J., Behar, L., & Rabson-Hare, J. (2000). A comparison of a vote count and a meta-analysis review of intervention research with adult cancer patients. *Research on Social Work Practice, 10*, 139-158.
- Dalton, S. O., Schuz, J., Engholm, G., Johansen, C., Kjoer, S. K., Steding-Jessen, M., Storm, H. H., and Olsen, J. H. (2008). Social inequality in incidence of and survival from cancer in a population-based study in Denmark, 1994–2003: Summary of findings. *European Journal of Cancer, 44*, 2074-2085.
- De Lorenzo, F., Ballatori, E., Di Costanzo, F., Giacalone, A., Ruggeri, B., & Tirelli, U. (2004). Improving information to Italian cancer patients: Results of a randomized study. *Annals of Oncology, 15*(5), 721-725.
- Derogatis, L. R., Morrow, G. R., Fetting, J. H., Penman, D., Piasetsky, S., Schmale, A. M., Henrichs, M., & Carniche, C. L. (1983). The prevalence of psychiatric disorders among cancer patients. *Journal of the American Medical Association, 249*, 751-757.
- Devine, E. C. (2003). Meta-analysis of the effect of psychoeducational interventions on pain in adults with cancer. *Oncology Nursing Forum, 30*(1), 75-89.
- Divine, E. C., & Cook, T. D. (1983). A meta-analytic analysis of effects of psychoeducational interventions on length of post-surgical hospital stay. *Nursing Research, 32*, 267-274.
- Devine, E. C., & Westlake, S. K. (1995). The effects of psychoeducational care provided to adults with cancer: A meta-analysis of 116 studies. *Oncology Nursing Forum, 22*, 1369-1381.
- Dunn, J. S., Steginga, N., Rosoman, N., & Millichap, D. (2003). A review of peer support in the context of cancer. *Journal of Psychosocial Oncology, 21*(2), 55-67.
- Edelman, S., Bell, D. R., & Kidman, A. D. (1999). A group cognitive behaviour therapy programme with metastatic breast cancer patients. *Psycho Oncology, 8*(4), 295-305.
- Edgar, L., Rosberger, Z., & Collet, J. P. (2001). Lessons learned: Outcomes and methodology of a coping skills intervention trial comparing individual and group formats for patients with cancer. *International Journal of Psychiatry in Medicine, 31*(3), 289-304.
- Edgar, L., Rosberger, Z., & Nowlis, D. (1992). Coping with cancer during the first year after diagnosis. Assessment and intervention. *Cancer, 69*(3):817-828.
- Edwards, A. G., Hailey, S., & Maxwell, M. (2004). Psychological interventions for women with metastatic breast cancer (Cochrane Review). *Cochrane Database of Systematic Reviews, 2*.
- Egger, M. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629-634.
- Ell, K., Sanchez, K., Vourlekis, B., Lee, P.-J., Dwight-Johnson, M., Lagomasino, I., Muderspach, L., & Russell, C. (2005). Correlates of depression and receipt of depression care among low-income women with breast or gynecologic cancer. *Journal of Clinical Oncology, 23*(13), 3052-3060.
- Epping-Jordan, J. E., Compas, B. E., Osowiecki, D. M., Oppedisano, G., Gerhardt, C., Primo, K., & Krag, D. N. (1999). Psychological adjustment in breast cancer: Processes of emotional distress. *Health Psychology, 18*(4), 315-326.
- Eton, D., Lepore, S., & Helgeson, V. S. (2001). Early quality of life in patients with localized prostate carcinoma: An examination of treatment-related, demographic, and psychosocial factors. *Cancer, 92*, 1451-1459.

- Fawzy, F. I. (1999). Psychosocial interventions for patients with cancer. What works and what doesn't. *European Journal of Cancer*, 35(11), 1559-1564.
- Fawzy, F. I., Fawzy, N. W., Arndt, L. A., & Pasnau, R. O. (1995). Critical review of psychosocial interventions in cancer care. *Archives of General Psychiatry*, 52(2), 100-113.
- Fawzy, N. W. (1995). A psychoeducational nursing intervention to enhance coping and affective state in newly diagnosed malignant melanoma patients. *Cancer Nursing*, 18(6), 427-438.
- Forester, B., Kornfeld, D. S., & Fleiss, J. L. (1985). Psychotherapy during radiotherapy: Effects on emotional and physical distress. *American Journal of Psychiatry*, 142(1), 22-27.
- Forester, B., Kornfeld, D. S., Fleiss, J. L., & Thompson, S. (1993). Group psychotherapy during radiotherapy: effects on emotional and physical distress. *The American Journal of Psychiatry*, 150(11), 1700-1706.
- Freidenbergs, I., Gordon, W., Hibbard, W., Levine, L., Wolf, C., & Diller, L. (1981-1982). Psychosocial aspects of living with cancer: A review of the literature. *International Journal of Psychiatry in Medicine*, 11(4), 303-329.
- Friedman, L. C., Kalidas, M., Elledge, R., Chang, J., Romero, C., Husain, I., & Dulay, M., Liscum, K. R. (2006). Optimism, social support and psychosocial functioning among women with breast cancer. *Psycho-Oncology*, 15(7), 295-603.
- Galinsky, M. J., & Schopler, J. H. (1994). Negative experiences in support groups. *Social Work in Ambulatory Care*, 20, 77-95.
- Gaston Johansson, F., Fall Dickson, J. M., Nanda, J., Ohly, K. V., Stillman, S., Krumm, S., and Kennedy, M. J. (2000). The effectiveness of the comprehensive coping strategy program on clinical outcomes in breast cancer autologous bone marrow transplantation. *Cancer Nursing*, 23(4), 277-285.
- Given, C., Given, B., Rahbar, M., Jeon, S., McCorkle, R., Cimprich, B., Galecki, A, Kozachik, S., Devoss, D., Brady, A., Fisher-Malloy, M. J., Courtney, K., & Bowie, E. (2004). Does a symptom management intervention affect depression among cancer patients: Results from a clinical trial. *Psycho-Oncology*, 13(11), 818-830.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis. *Educational Researcher*, 5, 3-8.
- Glass, G. V. (1978). Reply to Mansfield and Busse. *Educational Researcher*, 7, 3.
- Goldberg, R. J., & Wool, M. S. (1985). Psychotherapy for the spouses of lung cancer patients: Assessment of an intervention. *Psychotherapy and Psychosomatics*, 43(3), 141-150.
- Goodwin, P. J. (2005). Support groups in advanced breast cancer: Living better if not longer. *Cancer*, 104(Suppl 11), 2596-2601.
- Gotay, C. C., Moinpour, C. M., Unger, J. M., Jiang, C. S., Coleman, D., Martino, S., Parker, B. J., Bearden, J. D., Dakhil, S., Gross, H. M., Lippman, S., Albain, K. S. (2007). Impact of a peer-delivered telephone intervention for women experiencing a breast cancer recurrence. *Journal of Clinical Oncology*, 25(15), 2093-2099.
- Gotay, C. C., & Stern, J. D. (1995). Assessment of psychological functioning in cancer patients. Special issue: Psychosocial resource variables in cancer studies:

- Conceptual and measurement issues. *Journal of Psychosocial Oncology*, 13(1-2), 123-160.
- Gottlieb, B. H., & Wachala, E. D. (2007). Cancer support groups: A critical review of empirical studies. *Psycho-Oncology*, 6(5), 379-400.
- Graves, K. D. (2003). Social cognitive theory and cancer patients' quality of life: A meta-analysis of psychosocial intervention components. *Health Psychology*, 22(2), 210-219.
- Greenhouse, J. B., & Iyengar, S. (1994). Stensitivity analysis and diagnostics. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis*. (pp. 383-398). New York: Russell Sage Foundation.
- Greer, S. (1987). Psychotherapy for the cancer patient. *Psychiatric Medicine*, 5, 267-279.
- Hagedoorn, M., Van Yperen, N. W., Coyne, J. C., van Jaarsveld, C. H., Ranchor, A. V., van Sonderen, E., & Sanderman, R. (2006). Does marriage protect older people from distress? The role of equity and recency of bereavement. *Psychology and Aging*, 21(3), 611-620.
- Hall, J. A., Tickle-Degnen, L., Rosenthal, R., & Mosteller, F. (1994). Hypotheses and problems in research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 17-28). New York: Russell Sage Foundation.
- Hedges, L. V. (1994). Statistical considerations. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of research synthesis*. (pp. 29-38). New York: Russell Sage Foundation.
- Helgeson, V. S., Cohen, S., Schulz, R., & Yasko, J. (2000). Group support interventions for women with breast cancer: Who benefits from what? *Health Psychology*, 19(2), 107-114.
- Helgeson, V. S., Cohen, S., Schulz, R., & Yasko, J. (2001). Long-term effects of educational and peer discussion group interventions on adjustment to breast cancer. *Health Psychology*, 20(5), 387-392.
- Hepworth, S. I. (2004). *An intervention implemented by medical staff to address anxiety related to fears of cancer recurrence inpatients who have been treated for cancer of the head and neck*. Unpublished thesis, University of Manchester, Manchester, UK.
- Herrman, H., Patrick, D., Diehr, P., Martin, M., Fleck, M., Simon, G., & Buesching, D. (2002). Longitudinal investigation of depression outcomes in primary care in six countries: The LIDO study. Functional status, health service use and treatment of people with depressive symptoms. *Psychological Medicine*, 32(5), 889-902.
- Hewitt, M., & Rowland, J. H. (2002). Mental health service use among adult cancer survivors: Analyses of the National Health Interview Survey. *Journal of Clinical Oncology*, 20(23), 4581-4590.
- Ho, T. H. R. (2007). Psychophysiological effects of psychosocial interventions: An example of breast cancer patients in Hong Kong (China). *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 67(10-A).
- Holland, J. C. (1990a). Clinical course of cancer. In J. C. Holland & J. H. Rowland (Eds.), *Handbook of psycho-oncology* (pp. 75-100). New York: Oxford University Press.
- Holland, J. C. (1990b). Historical overview. In J. C. Holland & J. H. Rowland (Eds.), *Handbook of psycho-oncology* (pp. 3-12). New York: Oxford University Press.

- Holland, J. C. (2002). History of psycho-oncology: Overcoming attitudinal and conceptual barriers. *Psychosomatic Medicine*, 64(2), 2006-2221.
- Holland, J. C., & Lesko, L. (1990). Chemotherapy, endocrine therapy, and immunotherapy. In J. C. Holland & J. H. Rowland (Eds.), *Handbook of psycho-oncology*. New York: Oxford University Press.
- Honda, K., & Goodwin, R. D. (2004). Cancer and mental disorders in a national community sample: Findings from the National Comorbidity Survey. *Psychotherapy and Psychosomatics*, 73(4), 235-242.
- Houby, I. (1987). *A short-term intervention for cancer patients at risk for enduring distress*. Unpublished thesis 8821987, California School of Professional Psychology, San Diego, California, USA.
- Hughes, K. N. (2006). *Expressive writing and breast cancer: Outcomes and linguistic analyses*. Unpublished thesis 3252267, Utah State University, USA.
- Institute of Medicine of the National Academies, Committee on psychosocial services to cancer patients/families in a community setting. (2008). Psychosocial health services. In N. E. Adler & A. E. K. Page (Eds.), *Cancer care for the whole patient: Meeting psychosocial health needs*. (pp. 81-151). Washington D.C.: The National Academies Press.
- International Psychosocial Oncology Society (2009, June). *Proceedings of the 11th World Congress, Vienna, Austria*. To be published online at <http://www.ipos-society.org>.
- Jacobsen, P. (1990). Psychological reactions to cancer surgery. In J. C. Holland & J. H. Rowland (Eds.), *Handbook of psycho-oncology* (pp. 117-133). New York: Oxford University Press.
- Jacobsen, P., Donovan, K., Swaine, Z., & Watson, I. (2006). Management of anxiety and depression in adult cancer patients: Toward an evidence-based approach. In A. Chang, P. Ganz, D. Hayes, T. Kinsella, H. Pass, J. Schiller, R. Stone & V. Strecher (Eds.), *Oncology: An evidence-based approach*. (pp. 1552-1579). New York: Springer-Verlag.
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17, 1-12.
- Katz, M., Irish, J., & Devins, G. (2004). Development and pilot testing of a psychoeducational intervention for oral cancer patients. *Psycho-Oncology*, 13(9), 642-653.
- Katz, M. R., Rodin, G., & Devins, G. M. (1995). Self-esteem and cancer: Theory and research. *Canadian Journal of Psychiatry*, 40(10), 608-615.
- Kazantzis, N., Deane, F. P., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice*, 7(2), 189-202.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioural sciences*. (3 ed.). Pacific Grove, C.A: Brooks/Cole.
- Kissane, D. W., Bloch, S., Smith, G. C., Miach, P., Clarke, D. M., Ikin, J., Love, A., Ranieri, N., & McKenzie, D. (2003). Cognitive-existential group psychotherapy for women with primary breast cancer: a randomized controlled trial. *Psycho-Oncology*, 12(6), 532-546.
- Kissane, D. W., Grabsch, B., Clarke, D. M., Christie, G., Clifton, D., Gold, S., Hill, C., Morgan, A., McDermott, F., & Smith, G. C. (2004). Supportive-Expressive Group Therapy: The Transformation of Existential Ambivalence into Creative

- Living While Enhancing Adherence to Anti-Cancer Therapies. *Psycho-Oncology*, 13(11), 755-768.
- Kline, R. B. (2004). What's wrong with statistical tests - and where we go from here. In *Beyond significance testing* (pp. 61-91). Washington D.C.: American Psychological Association.
- Knight, S. J. (2004). Oncology and hematology. In P. M. Camic & S. J. Knight (Eds.), *Clinical handbook of health psychology: A practical guide to effective interventions*. (2nd ed., pp. 233-261). Ashland O.H., U.S.: Hogrefe & Huber Publishers.
- Krauss, O., Ernst, J., Kuchenbecker, D., Hinz, A., & Schwartz, R. (2007). Predictors of mental disorders in patients with malignant diseases: Empirical results [German]. *Psychotherapie Psychosomatik Medizinische Psychologie*, 57(7), 273-280.
- Krischer, M. M., Xu, P., Meade, C. D., & Jacobsen, P. B. (2007). Self-administered stress management training in patients undergoing radiotherapy. *Journal of Clinical Oncology*, 25(29), 4657-4662.
- Kristeller, J. L., Rhodes, M., Cripe, L. D., & Sheets, V. (2005). Oncologist Assisted Spiritual Intervention Study (OASIS): Patient acceptability and initial evidence of effects. *International Journal of Psychiatry in Medicine*, 35(4), 329-347.
- Kruijver, I. P. M., Garssen, B., Visser, A. P., & Kuiper, A. J. (2006). Signalling psychosocial problems in cancer care: The structural use of a short psychosocial checklist during medical or nursing visits. *Patient Education and Counseling*, 62, 163-177.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Lepore, S., Helgeson, V., Eton, D., & Schulz, R. (2003). Improving quality of life in men with prostate cancer: A randomized controlled trial of group education interventions. *Health Psychology*, 22(5), 443-452.
- Lepore, S. J., & Coyne, J. C. (2006). Psychological interventions for distress in cancer patients: A review of reviews. *Annals of Behavioural Medicine*, 32(2), 85-92.
- Lerman, C., Schwartz, M. D., Miller, S. M., Daly, M., Sands, C., & Rimer, B. K. (1996). A randomized trial of breast cancer risk counseling: Interacting effects of counseling, educational level, and coping style. *Health Psychology*, 15(2), 75-83.
- Leszcz, M., & Goodwin, P. J. (1998). The rationale and foundations of group psychotherapy for women with metastatic breast cancer. *International Journal of Group Psychotherapy*, 48(2), 245-273.
- Lev, E. L. (1997). Bandura's theory of self-efficacy: Applications to oncology. *Scholarly Inquiry for Nursing Practice*, 11, 21-37.
- Levenson, J. L., Hamer, R. M., & Rossiter, L. F. (1990). Relation of psychopathology in general medical inpatients to use and cost of services. *American Journal of Psychiatry*, 147, 1498-1503.
- Levenson, R. W., Carstensen, L. L., & Gottman, J. M. (1993). Long-term marriage: Age, gender, and satisfaction. *Psychology and Aging*, 8(2), 301-313.
- Lin, M. L., Tsang, Y. M., & Hwang, S. L. (1998). Efficacy of a stress management program for patients with hepatocellular carcinoma receiving transcatheter arterial embolization. *Journal of the Formosan Medical Association ('Taiwan yi zhi')*, 97(2), 113-117.

- Lioffi, C., & White, P. (2001). Efficacy of clinical hypnosis in the enhancement of quality of life of terminally ill cancer patients. *Contemporary Hypnosis, 18*(3), 145-160.
- Love, A. W., Grabsch, B., Clarke, D. M., Bloch, S., & Kissane, D. W. (2004). Screening for depression in women with metastatic breast cancer: A comparison of the Beck Depression Inventory Short Form and the Hospital Anxiety and Depression Scale. *Australian and New Zealand Journal of Psychiatry, 38*(7), 526-531.
- Lovejoy, N. C., & Matteis, M. (1997). Cognitive-behavioral interventions to manage depression in patients with cancer: Research and theoretical initiatives. *Cancer Nursing, 20*(3), 155-167.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications.
- Luebbert, K., Dahme, B., & Hasenbring, M. (2001). The effectiveness of relaxation training in reducing treatment-related symptoms and improving emotional adjustment in acute non-surgical cancer treatment: A meta-analytical review. *Psycho-Oncology, 10*(6), 490-502.
- Maguire, P., Hopwood, P., Tarriner, N., & Howell, T. (1985). Treatment of depression in cancer patients. *Acta psychiatrica Scandinavica* (Suppl), 320, 81-84.
- Maguire, P., Tait, A., Brooke, M., Thomas, C., & Sellwood, R. (1980). Effect of counselling on the psychiatric morbidity associated with mastectomy. *British Medical Journal, 281*(6253), 1454-1456.
- Manne, S. L., Ostroff, J. S., Winkel, G., Fox, K., Grana, G., Miller, E., Ross, S., & Frazier, T. (2005). Couple-focused group intervention for women with early stage breast cancer. *Journal of Consulting and Clinical Psychology, 73*(4), 634-646.
- Manne, S., Ostroff, J. S., & Winkel, G. (2007b). Social-cognitive processes as moderators of a couple-focused group intervention for women with early stage breast cancer. *Health Psychology, 26*(6), 735-744.
- Manne, S. L., Rubin, S., Edelson, M., Rosenblum, N., Bergman, C., Hernandez, E., Carlson, J., Rocereto, T., & Winkel, G. (2007). Coping and communication-enhancing intervention versus supportive counseling for women diagnosed with gynecological cancers. *Journal of Consulting and Clinical Psychology, 75*(4), 615-628.
- Massie, M. J. (2004). Prevalence of depression in patients with cancer. *Journal of the National Cancer Institute. Monographs, 32*, 57-71.
- McHugh, P., Lewis, S., Ford, S., Newlands, E., Rustin, G., Coombes, C., Smith, D., O'Reilly, S., & Fallowfield, L. (1995). The efficacy of audiotapes in promoting psychological well-being in cancer patients: A randomised, controlled trial. *British Journal of Cancer, 71*(2), 388-392.
- McLachlan, S. A., Allenby, A., Matthews, J., Wirth, A., Kissane, D., Bishop, M., Beresford, J. & Zalberg, J. (2001). Randomized trial of coordinated psychosocial interventions based on patient self-assessments versus standard care to improve the psychosocial functioning of patients with cancer. *Journal of Clinical Oncology, 19*(21), 4117-4125.
- McPherson, C. J., Higginson, I. J., & Hearn, J. (2001). Effective methods of giving information in cancer: A systematic literature review of randomized controlled trials. *Journal of Public Health Medicine, 23*(3), 227-234.

- Mehnert, A., & Kock, U. (2008). Psychological comorbidity and health-related quality of life and its association with awareness, utilization, and need for psychosocial support in a cancer register-based sample of long-term breast cancer survivors. *Journal of Psychosomatic Research*, *64*, 383-391.
- Meyer, T. J., & Mark, M. M. (1995). Effects of psychosocial interventions with adult cancer patients: A meta-analysis of randomized experiments. *Health Psychology*, *14*(2), 101-108.
- Miovic, M., & Block, S. (2007). Psychiatric disorders in advanced cancer. *Cancer*, *110*(8), 1665-1676.
- Mitchell, A. J., Karr, S., Coggan, C., & Herdman, J. (2008). Acceptability of common screening methods used to detect distress and related mood disorders - preferences of cancer specialists and non-specialists. *Psycho-Oncology*, *17*, 226-236.
- Moher, D., Cook, T. D., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *The Lancet*, *354*, 1896-1900.
- Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel group randomized trials. *BioMedcentral Medical Research Methodology*, *1*(2) available online only: <http://www.biomedcentral.com/147-2288/2281/2282>.
- Mosher, C. E., & Danoff-Burg, S. (2005). A review of age differences in psychological adjustment to breast cancer. *Journal of Psychosocial Oncology*, *23*(2-3), 101-114.
- Moynihan, C., Bliss, J. M., Davidson, J., Burchell, L., & Horwich, A. (1998). Evaluation of adjuvant psychological therapy in patients with testicular cancer: randomised controlled trial. *BMJ (Clinical research ed.)*, *316*(7129), 429-435.
- Mystakidou, K., Tsilika, E., Parpa, E., Katsouda, E., Galanos, A., & Vlahos, L. (2006). Psychological distress of patients with advanced cancer: Influence and contribution of pain severity and pain interference. *Cancer Nursing*, *29*, 400-405.
- National Cancer Institute. (2009). *Surveillance epidemiology and end results: Breast cancer statistics fact sheet*. Retrieved 26 May, 2009, from <http://seer.cancer.gov/statfacts/html/breast.html>.
- National Comprehensive Cancer Network (2005). *Clinical Practice Guideline in Oncology, Distress Management*. Retrieved 2007, from http://www.nccn.org/professionals/physician_gls/f_guidelines.asp.
- Newell, S. A., Sanson-Fisher, R. W., Girgis, A., & Bonaventura, A. (1998). How well do medical oncologists' perceptions reflect their patients' reported physical and psychosocial problems? Data from five oncologists. *Cancer*, *83*(8), 1640-1651.
- Newell, S. A., Simon-Fisher, R. W., & Savolainen, N. J. (2002). Systematic review of psychological therapies for cancer patients: Overview and recommendations for future research. *Journal of the National Cancer Institute*, *94*(8), 558-584.
- New Zealand Department of Statistics (2009). *2006 Census statistics*. Retrieved 26 March 2009, from http://www.stats.govt.nz/census-outputs/quickstats/snaps_hotplace2.htm?id=1000008&type=region&ParentID=.
- New Zealand Health Information Service (2007). *Cancer: New registrations and deaths 2003*. Retrieved 15 September, 2007, from <http://www.nzhis.govt.nz/publications/cancer.html>.

- New Zealand Health Information Service. (2009). *Major causes of death, numbers and percentages*. Retrieved 26 March 2009, from <http://www.nzhis.govt.nz/moh.nsf/pagesns/32#01>.
- Nezu, A. M., Nezu, C. M., Felgoise, S. H., McClure, K. S., & Houts, P. S. (2003). Project Genesis: Assessing the efficacy of problem-solving therapy for distressed adult cancer patients. *Journal of Consulting and Clinical Psychology, 71*(6), 1036-1048.
- North, N., Cornbleet, M. A., Knowles, G., & Leonard, R. C. (1992). Information giving in oncology: A preliminary study of tape-recorder use. *British Journal of Clinical Psychology, 31*(3), 357-359.
- Northouse, L. L. (1988). Social support in patients' and husbands' adjustment to breast cancer. *Nursing Research, 37*, 91-95.
- Norton, T. R., Manne, S. L., Rubin, S., Hernandez, E., Carlson, J., Bergman, C., & Rosenblum, N. (2005). Ovarian cancer patients' psychological distress: The role of physical impairment, perceived unsupportive family and friend behaviors, perceived control, and self-esteem. *Health Psychology, 24*(2), 143-152.
- Osborn, R. L., Demoncada, A. C., & Feuerstein, M. (2006). Psychosocial interventions for depression, anxiety, and quality of life in cancer survivors: Meta-analyses. *International Journal of Psychiatry in Medicine, 36*(1), 13-34.
- Owen, J. E., Klapow, J. C., Hicken, B., & Tucker, D. C. (2001). Psychosocial interventions for cancer: Review and analysis using a three-tiered outcomes model. *Psycho-Oncology, 10*, 218-230.
- Pai, A., Drotar, D., Zebracki, K., Moore, M., & Youngstrom, E. (2006). A meta-analysis of the effects of psychological interventions in pediatric oncology outcomes of psychological distress and adjustment. *Journal of Pediatric Psychology, 31*(9), 978-988.
- Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). *Global Cancer Statistics, 2002*. Retrieved 10 October 2007, from <http://intl-caonline.amcancersoc.org/cgi/content/abstract/55/2/74>.
- Pascoe, S., Edelman, S., & Kidman, A. (2000). Prevalence of psychological distress and use of support services by cancer patients at Sydney hospitals. *Australian and New Zealand Journal of Psychiatry, 34*, 785-791.
- Paul, G. L. (1969). Behavior modification research: Design and tactics. In C. M. Franks (Ed.), *Behavior Therapy: Appraisal and status* (pp. 29-62). New York: McGraw-Hill.
- Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design and analysis*. Hillsdale, N.J.: Lawrence Erlbaum.
- Redd, W. H. (1995). Behavioural research in cancer as a model for health psychology. *Health Psychology, 14*, 99-100.
- Redd, W. H., Montgomery, G. H., & DuHamel, K. N. (2001). Behavioural intervention for cancer treatment side effects. *Journal of the National Cancer Institute, 93*(11), 810-823.
- Rehse, B., & Pukrop, R. (2003). Effects of psychosocial interventions on quality of life in adult cancer patients: Meta analysis of 37 published controlled outcome studies. *Patient Education and Counseling, 50*, 179-189.
- Reis, V. A. (2007). Young women with breast cancer: A meta-analysis of age differences across quality of life domains. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67*(10-B), 6074.

- Rosenbaum, D. I. (2006). *What's good for the gander is good for the goose: Helping cancer patients to cope by treating their spouses*. Unpublished thesis 3241620, State University of New York, Binghamton, USA.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183-192.
- Ross, L., Boesen, E. H., Dalton, S. O., & Johansen, C. (2002). Mind and cancer: Does psychosocial intervention improve survival and psychological well-being? *European Journal of Cancer*, *38*(11), 1447-1457.
- Iconomou, G., & Quine, L. (1996). Doctor-patient communication and outcome in cancer patients: An intervention. *Psychology & Health*, *12*(1), 57-71.
- Savard, J., Simard, S., Giguère, I., Ivers, H., Morin, C. M., Maunsell, E., Gagnon, P., Robert, J., & Marceau, D. (2006). Randomized clinical trial on cognitive therapy for depression in women with metastatic breast cancer: Psychological and immunological effects. *Palliative & Supportive Care*, *4*(3), 219-237.
- Scheier, M. F., Helgeson, V. S., Schulz, R., Colvin, S., Berga, S., Bridges, M. W., Knapp, J., Gerszten, K., & Pappert, W. S. (2005). Interventions to enhance physical and psychological functioning among younger women who are ending nonhormonal adjuvant treatment for early-stage breast cancer. *Journal of Clinical Oncology*, *23*(19), 4298-4311.
- Scheier, M. F., Helgeson, V. S., Schulz, R., Colvin, S., Berga, S. L., Knapp, J., & Gerszten, K. (2007). Moderators of interventions designed to enhance physical and psychological functioning among younger women with early-stage breast cancer. *Journal of Clinical Oncology*, *25*(36), 5710-5714.
- Scholz, C. E. (1999). Psychological interventions for children with cancer undergoing invasive medical procedures: A meta-analysis of the research. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *59*(9-B), 5110.
- Schroevers, M., Ranchor, A. V., & Sanderman, R. (2003). The role of social support and self-esteem in the presence and course of depressive symptoms: A comparison of cancer patients and individuals from the general population. *Social Science & Medicine*, *57*, 375-385.
- Sellick, D. M., & Crooks, D. L. (1999). Depression and cancer: An appraisal of the literature for prevalence, detection, and practice guideline development for psychological interventions. *Psycho-Oncology*, *8*, 315-333.
- Sheard, T., & Maguire, P. (1999). The effect of psychological interventions on anxiety and depression in cancer patients: Results of two meta-analyses. *British Journal of Cancer*, *80*(11), 1770-1780.
- Sherman, A. C., Mosier, J., Leszcz, M., Burlingame, G. M., Ulman, K. H., Cleary, T., Simonton, S., Latif, U., Hazelton, L., & Strauss, B. (2004). Group Interventions for patients with cancer and HIV disease: Part III. Moderating variables and mechanisms of action. *International Journal of Group Psychotherapy*, *54*(3), 347-387.
- Siegel, K. (1990). Psychosocial oncology research. *Social Work in Health Care*, *15*, 21-43.
- Silberfarb, P. M. (1982). Research in adaptation to illness and psychosocial intervention. *Cancer*, *50*(Suppl), 1921-1925.
- Smedslund, G., & Ringdal, G. I. (2004). Meta-analysis of the effects of psychosocial interventions on survival time in cancer patients. *Journal of Psychosomatic Research*, *57*(2), 123-131.

- Smith, J. E., Richardson, J., Hoffman, C., & Pilkington, K. (2005). Mindfulness-based stress reduction as supportive therapy in cancer care: Systematic review. *JAN Journal of Advanced Nursing*, 52(3), 315-327.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752-760.
- Snowdon, J. (2001). Is depression more prevalent in old age? *Australian and New Zealand Journal of Psychiatry*, 35(6), 782-787.
- Spiegel, D. (2002). Effects of psychotherapy on cancer survival. *Nature Reviews / Cancer*, 2, 383-389.
- Spiegel, D., Sephton, S. E., Terr, A. I., & Stites, D. P. (1998). Effects of psychosocial treatment in prolonging cancer survival may be mediated by neuroimmune pathways. *Annals of the New York Academy of Science*, 840, 674-683.
- Stanton, A. L. (2006). Psychosocial concerns and interventions for cancer survivors. *Journal of Clinical Oncology*, 24(32), 5132-5137.
- Steel, J. L., Nadeau, K., Olek, M., & Carr, B. I. (2007). Preliminary results of an individually tailored psychosocial intervention for patients with advanced hepatobiliary carcinoma. *Journal of Psychosocial Oncology*, 25(3), 19-42.
- Stewart, M., Brown, J. B., Hammerton, J., Donner, A., Gavin, A., Holliday, R. L., Whelan, T., Leslie, K., Cohen, I., Weston, W., and Freeman, T. (2007). Improving communication between doctors and breast cancer patients. *Annals of Family Medicine*, 5(5), 387-394.
- Tatrow, K., & Montgomery, G. H. (2006). Cognitive behavioural therapy techniques for distress and pain in breast cancer patients: A meta-analysis. *Journal of Behavioral Medicine*, 29(1), 17-27.
- Telch, C. F., & Telch, M. J. (1986). Group coping skills instruction and supportive group therapy for cancer patients: A comparison of strategies. *Journal of Consulting and Clinical Psychology*, 54, 802-808.
- Thomas, S. P., Groer, M., Davis, M., Droppleman, P., Mozingo, J., & Pierce, M. (2000). Anger and cancer: An analysis of linkages. *Cancer Nursing*, 23(5), 344-348.
- Timms, M. W. H. (1990). Aspects of psychological intervention with breast cancer patients. *Irish Journal of Psychological Medicine*, 7(2), 168-176.
- Trask, P. C., Paterson, A. G., Griffith, K. A., Riba, M. B., & Schwartz, J. L. (2003). Cognitive-behavioral intervention for distress in patients with melanoma: Comparison with standard medical care and impact on quality of life. *Cancer*, 98(4), 854-864.
- Trijsburg, R. W., vanKippenberg, F. C. E., & Rijpma, S. E. (1992). Effects of psychological treatment on cancer patients: A critical review. *Psychosomatic Medicine*, 54, 489-517.
- van't Spijker, A., Trijsburg, R. W., & Duivenvoorden, H. J. (1997). Psychological sequelae of cancer diagnosis: A meta-analytical review of 58 studies after 1980. *Psychosomatic Medicine*, 59(3), 280-293.
- Walker, B., Nail, L. M., & Croyle, R. (1999). Does emotional expression make a difference in reactions to breast cancer? *Oncology Nursing Forum*, 26, 1025-1032.
- Walker, M. S., & Podbilewicz-Schuller, Y. (2005). Video preparation for breast cancer treatment planning: Results of a randomized clinical trial. *Psycho Oncology*, 14(5), 408-420.
- Watts, P. M. (2002). *A grounded theory study of depression in a sample of men participating in a clinical trial examining the effect of a dietary supplement on*

- depression*. Unpublished thesis, Massey University, Palmerston North, New Zealand.
- Weber, B. A. (2001). *The effect of dyadic intervention on self-efficacy, social support, and depression for men with prostate cancer*. Unpublished thesis 3034641, Case Western Reserve University (Health Sciences), Ohio, USA.
- Weber, B. A., Roberts, B. L., Resnick, M., Deimling, G., Zauszniewski, J. A., Musil, C., Yarandi, H. N. (2004). The effect of dyadic intervention on self-efficacy, social support, and depression for men with prostate cancer. *Psycho Oncology*, 13(1), 47-60.
- Weber, B. A., Roberts, B. L., Yarandi, H., Mills, T. L., Chumbler, N. R., & Wajzman, Z. (2007). The impact of dyadic social support on self-efficacy and depression after radical prostatectomy. *Journal of Aging and Health*, 19(4), 630-645.
- Williams, K. (2003). Has the Future of Marriage Arrived? A Contemporary Examination of Gender, Marriage, and Psychological Well-Being. *Journal of Health and Social Behavior*, 44(4), 470-487.
- Wong-Kim, E. C. (2000). Predicting depressive symptoms in breast cancer patients: Racial comparisons. *Dissertation Abstracts International: Section A: Humanities and Social Sciences*, 61(3-A), 1166.
- Worden, J. W., & Weisman, A. D. (1984). Preventive psychosocial intervention with newly diagnosed cancer patients. *General Hospital Psychiatry*, 6(4), 243-249.
- Wortman, P. M. (1994). Judging research quality. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Yancik, R. (1997). Cancer burden in the aged: An epidemiologic and demographic overview. *Cancer*, 80, 1273-1283.
- Zabalegui, A., Sanchez, S., Sanchez, P. D., & Juando, C. (2005). Nursing and cancer support groups. *Journal of Advanced Nursing*, 51(4), 369-381.
- Zabora, J., Brintzenhofesoc, K., Curbow, B., Hooker, C., & Piantadosi, S. (2001). The prevalence of psychological distress by cancer site. *Psycho-Oncology*, 10, 19-28.
- Zakowski, S. G., Ramati, A., Morton, C., Johnson, P., & Flanigan, R. (2004). Written emotional disclosure buffers the effects of social constraints on distress among cancer patients. *Health Psychology*, 23(6), 555-563.