http://researchspace.auckland.ac.nz

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

• Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
• Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
• You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.
The Effectiveness and Cost-Effectiveness of the Green Prescription Physical Activity Intervention: a Cluster Randomised Controlled Trial in Primary Health Care

By Carolyn (Raina) Elley

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in General Practice,

The University of Auckland, 2003
Abstract

This thesis assesses the effectiveness and cost-effectiveness of the ‘Green Prescription’ physical activity intervention amongst less-active adults in primary care, using the Auckland Heart Study (AHS) questionnaire to assess change in activity.

The validity of two physical-activity questionnaires, the AHS and the GSS questionnaires, was assessed initially, comparing their performance with 7-day activity diaries and pedometers amongst less-active adults in primary care. A cluster randomised controlled trial was then conducted in the Waikato region of New Zealand, with 42 general practices randomised to give the Green Prescription or ‘usual care’. A systematic screening process identified less-active 40-79 year-old patients. Main outcome measures included change in physical activity, quality of life (SF-36), coronary risk, and blood pressure, over a twelve-month period. Costs of the programme and offset costs of primary and secondary health care utilisation, productivity, and exercising, were collected prospectively for cost-effectiveness analysis from a societal perspective.

The AHS questionnaire was found to have adequate reliability and validity, and to be the most appropriate measurement tool for use in primary care research. In the Green Prescription trial, 74% of general practitioners (n=117) and 66% of screened eligible patients (n=878) participated. Follow-up rate, at one year, was 85% (n=750). Mean total energy expenditure increased by 9.4 kcal/kg/week (p=0.001) and leisure exercise by 2.7 kcal/kg/week (p=0.02) or 34 minutes/week more in the intervention group than the control group (p=0.04). The proportion of the intervention group undertaking 2½ hours/week of leisure exercise increased by 9.72% (p=0.003) more than in the control group (NNT=10.3). SF-36 measures of ‘general health’, ‘role-physical’, ‘vitality’, and ‘bodily pain’ improved significantly more in the intervention group (p<0.05), as a result of the Green Prescription intervention. There was a trend towards decreasing blood pressure, no increase in adverse events, and no statistically significant difference in four-year risk of coronary heart disease. The cost of delivering the Green Prescription was $170.45 per participant from a funder’s perspective and $37.16 (95%CI:-$945.21,$1019.53) from a societal perspective. To increase leisure-time exercise by one
hour per week cost $25.36 per month from a programme funder’s perspective and $5.47 (95%CI:-$138.90,$149.84) per month from a societal perspective.

The AHS questionnaire was considered reliable and valid. The Green Prescription intervention is effective in increasing physical activity and improving quality of life over 12 months without evidence of adverse effects. The intervention is more cost-effective than other physical activity interventions reported in the literature and may be cost-saving in terms of long-term health gains.
Acknowledgements

I would like to acknowledge my supervisors, Dr Ngaire Kerse and Associate Professor Bruce Arroll, for their continued support and advice, without which this thesis would not have been possible. I would also like to thank Professor Boyd Swinburn for his supervision during the first year of the PhD and for his sound advice about the design of this study.

I would also like to acknowledge the financial support of this study from the New Zealand Heart Foundation, the Hillary Commission (Sport and Recreation, New Zealand), the Waikato Medical Research Foundation, the Royal New Zealand College of General Practitioners Research and Education Charitable Trust, and the University of Auckland Department of General Practice and Primary Health Care Research Fund. Thank you also to the Auckland Medical Research Foundation, and the Maurice and Phyllis Paykell Trust for their travel grants for the dissemination of the results of this study at national and international conferences.

Many people have helped make this study possible, not least of all the general practitioners, practice nurses, receptionists, practice managers, and patients of the participating general practices in the Waikato. In addition, I would like to thank the Pinnacle Independent Practitioners' Association for assisting with the administration and organisation of this study. The assistance of Dr Brett Anderson in the development of the ACCESS software for the data entry and calculation of the physical activity and cardiovascular variables was invaluable. Thank you, also, to the research assistants, Jan Gaskin, Moira Johnson, Sharon Matangi-Nixon, Hayley Gaddes, Helen Dunn, Chris Drent and Ruth Boyce for their enthusiasm and dedication to the project.

The statistical advice and support provided by Elizabeth Robinson was much appreciated, as was the critical appraisal of the design of the study, by the reference group. The reference group included Professor Rod Jackson, Associate Professor Toni Ashton, and Professor David Thomas, amongst others, from the University of Auckland and Diana O'Neill from the, then, Hillary Commission. I am also indebted to Associate Professor Richard Milne for the inspiration and advice about the design of the cost-
effectiveness study. Thank you also to the TADs team of the University of Auckland Goodfellow Unit, in particular, Dr Richard Fox who provided the motivational interviewing training to the intervention clinicians. I am also indebted to Drs Shanti Ameratanga and Tim Kenealy for their feedback and support in the PhD writing group.

I am very grateful to my mother, Val Elley, for her encouragement and practical support. I owe a huge debt to my father, Warwick Elley, for the many hours he spent discussing the ideas for this thesis, and his pain-staking proof reading of this thesis. Most importantly, thank you to my partner, David, and our children, Ben and Leo, for their unfailing tolerance, support, and encouragement.
Table of Contents

ABSTRACT ..........................................................................................................................II

ACKNOWLEDGEMENTS .................................................................................................... IV

TABLE OF CONTENTS ......................................................................................................... VI

LIST OF TABLES .................................................................................................................. X

LIST OF FIGURES ................................................................................................................ XIII

LIST OF ABBREVIATIONS .................................................................................................. XV

1 INTRODUCTION .............................................................................................................. 1

1.1 Rationale for the Thesis ............................................................................................... 1

1.2 Objectives of the Thesis .............................................................................................. 6

1.3 Structure of the Thesis ................................................................................................. 7

2 LITERATURE REVIEW .................................................................................................... 8

2.1 Introduction .................................................................................................................. 8

2.2 Background .................................................................................................................. 8

2.2.1 Physical Inactivity and Cardiovascular Disease ..................................................... 8

2.2.2 Prevalence of Physical Inactivity in New Zealand ................................................. 18

2.2.3 Measurement of Cardiovascular Risk ................................................................. 21

2.2.4 Physical Inactivity and Quality of Life ................................................................. 23

2.2.5 Development and Implementation of the Green Prescription Intervention ............. 24

2.3 Systematic Review of Physical Activity Questionnaire Reliability and Validity Studies 27

2.3.1 Introduction ........................................................................................................... 27

2.3.2 Measurement of Physical Activity using Self-report Instruments ....................... 28

2.3.3 Previous Reviews of Physical Activity Questionnaires ....................................... 32

2.3.4 Methods ............................................................................................................... 34

2.3.5 Studies Identified and Critical Appraisal ............................................................. 39

2.3.6 Discussion of the Review ....................................................................................... 39

2.4 Systematic Review of Physical Activity Intervention Trials in Primary Health Care 50

2.4.1 Introduction ........................................................................................................... 50

2.4.2 Methods ............................................................................................................... 51

2.4.3 Systematic Reviews Identified and Critical Appraisal ......................................... 53

2.4.4 Identification and Critical Appraisal of Additional Physical Activity Intervention Trials in Primary Health Care ................................................................. 63

2.4.5 Discussion of the Review ....................................................................................... 68

2.5 Systematic Review of Cost-Effectiveness Studies of Physical Activity Interventions in Primary Health Care ................................................................. 70

2.5.1 Introduction ........................................................................................................... 70

2.5.2 Methods ............................................................................................................... 71

2.5.3 Studies Identified and Critical Appraisal ............................................................. 73
4.3.4 Participation Rates ........................................ 144
4.3.5 Participant Characteristics ................................. 149
4.3.6 The Relationship between Cardiovascular Risk and Usual Level of Physical Activity .......................... 156
4.3.7 Quality of Life Characteristics ......................... 160
4.3.8 Adequacy of Randomisation .............................. 163
4.3.9 Positive Predictive Value of Screening Question ... 168
4.3.10 Rates of Intervention Delivery .......................... 168
4.3.11 Final Outcomes ........................................... 168
4.3.12 Adequacy of Randomisation .............................. 169
4.3.13 Intracluster Correlation Coefficients .................. 178
4.3.14 Summary ................................................ 181

4.4 Results of the Green Prescription Cost-Effectiveness Study .................................................. 181
4.4.1 Introduction ................................................ 181
4.4.2 Programme Costs ......................................... 182
4.4.3 Offset Costs ............................................... 187
4.4.4 Incremental Cost-effectiveness Ratios and Cost-effectiveness ........................................ 195
4.4.5 Summary ................................................ 197

5 DISCUSSION .................................................. 199
5.1 Introduction .................................................. 199

5.2 Discussion of the Physical Activity Questionnaire Reliability and Validity Study ....................... 199
5.2.1 Main Findings ............................................. 199
5.2.2 How the Findings relate to Previous Research ..................... 199
5.2.3 Strengths and Limitations of the Study .................. 203
5.2.4 Implications and Recommendations for Future Research .................................................. 205

5.3 Discussion of the Green Prescription Cluster Randomised Controlled Trial .................................. 205
5.3.1 Pilot Study Findings ..................................... 205
5.3.2 Baseline Analysis of the Randomised Controlled Trial .......................................................... 206
5.3.3 Main Findings of the Cluster Randomised Controlled Trial .................................................. 209
5.3.4 Methodological Strengths and Limitations of the Study ..................................................... 210
5.3.5 How the Findings relate to Previous Research ..................... 218
5.3.6 Implications of the Findings .............................. 223
5.3.7 Recommendations for Future Research and Potential Implementation ................................. 226

5.4 Discussion of the Cost-Effectiveness Study .......................................................... 228
5.4.1 Main Findings ............................................. 228
5.4.2 Strengths and Limitations of the Study .................. 230
5.4.3 How the Findings relate to Previous Research ..................... 232
5.4.4 Implications and Recommendations for Future Research .................................................. 235

6 CONCLUSION ............................................... 240

REFERENCES .................................................. 245

APPENDICES .................................................. 269

1 Auckland Heart Study (AHS) Physical Activity Questionnaire
2 Green Script Study (GSS) Physical Activity Questionnaire
3 Physical Activity Questionnaire Reliability and Validity Studies Evidence Tables
4 Seven-Day Physical Activity Diary
5 Study Participation Screening Form
6 Study Participant Information Sheet
7 Study Participant Consent Form
8 Current Physical Activity Status (Stage of Change)
9 The Green Prescription (developed by the Hillary Commission)
10 Form Faxed to Exercise Specialists
11 Baseline Research Protocol
12 Baseline Recording Sheet
13 Notice For Practice Waiting Rooms
14 Medication Codes
15 Medical Condition Codes
16 Practice Information Form
17 Form for Recording Excluded Patients
18 Clinician Consent Form
19 Clinician Questionnaire
20 Follow-up Recording Sheet
21 Follow-up Research Protocol
22 Patient Feedback Form
23 Patient Letter of Results
24 General Practitioner Letter of Results
25 Record of Visits to the General Practitioner
# List of Tables

Table 2-1 Relative Risk of Cardiovascular Mortality for Inactive compared with Active Individuals, and the estimated Cardiovascular Mortality attributable (Attributable Fraction) to Inactivity for Middle-aged and Older Adults in New Zealand.......................................................... 21

Table 2-2 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Test-Retest Reliability and Validity Studies of Physical Activity Questionnaires from 1990 to 2002......................................................... 36

Table 2-3 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Trials and Systematic Reviews of Physical Activity Counselling in Primary Health Care................................. 51

Table 2-4 Summary of Systematic Reviews of Counselling for Physical Activity in Health-Care Settings ........................................................................................................................................ 57

Table 2-5 Physical Activity Counselling Trials based in Primary Care that have not previously been Reviewed by Systematic Reviews.................................................................................................................. 65

Table 2-6 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Cost-Effectiveness Studies of Physical Activity Interventions in Primary Health Care.................. 72

Table 2-7 Evidence Table of Cost-Effectiveness Studies undertaken of Interventions that included Physical Activity Promotion within Primary Health Care or the Community ............................................... 74

Table 3-1 Sample Size Calculations for Different Variables for the Green Prescription Cluster Randomised Controlled Trial ........................................................................................................... 103

Table 4-1 Estimated Means for Total Energy Expenditure (kcal/kg/wk), Total Moderate and Leisure Moderate Activity (hours/week) from the AHS Questionnaire, 7-day Diary and GSS Questionnaire ........................................................................................................................ 132

Table 4-2 Test-Retest Reliability of the AHS questionnaire: Intraclass Correlation Coefficients (ICC) and Spearman’s Correlation Coefficients (SCC) for Total Energy Expenditure, Total Moderate and Vigorous Activity and Leisure-time Moderate Activity........................................................................... 133

Table 4-3 Validity of the AHS Questionnaire and the GSS Questionnaire compared with the 7-day Diary for Total Energy Expenditure, Total Moderate Activity and Leisure Moderate and Vigorous Activity using Spearman’s Correlation Coefficients................................................................................. 134

Table 4-4 Validity of the AHS Questionnaire, and the GSS Questionnaire, compared with the Pedometer recording for Total Moderate Activity and Total Energy Expenditure using Spearman’s Correlation Coefficients.................................................................................. 137

Table 4-5 Pilot Study Mean Baseline Characteristics of Intervention and Control Patients ........................................................................................................................................................... 139

Table 4-6 Patients found to have Medical Reasons for Exclusion Post-Enrolment .................................................................................................................................................................. 148

Table 4-7 Characteristics of General Practitioners ........................................................................................................................................................................ 149

Table 4-8 Physical Activity and Cardiovascular Measures of the Study Population at Baseline .................................................................................................................................................. 154

Table 4-9 National Non-Māori Population Estimates* compared with Non-Māori Study Participant Means for Cardiovascular Risk Factors (with 95% confidence intervals)...................................................................................... 155
Table 4-10 Relationship of Four-year Cardiovascular Risk of Study Participants to Health, Activity and Demographic Variables (Multiple linear regression analysis, grouped by medical practice) 157

Table 4-11 Relationships of Variables to Total Energy Expenditure (Natural Log) of Study Participants. (Univariate linear regression analysis, grouped by practice) ........................................ 158

Table 4-12 Relationships of Variables to Total Energy Expenditure (Natural Log) of Study Participants (Multiple linear regression analysis, grouped by practice) ........................................ 159

Table 4-13 Baseline Demographic and Clinical Characteristics of Less-active 40-79 year-old Patients in General Practice, by Intervention and Control Group. Mean values and (standard deviations) are presented, unless otherwise specified ............................................................................. 166

Table 4-14 Baseline Physical Activity Characteristics of the Intervention and Control Groups. Mean values and (standard deviations) are presented, unless otherwise specified ........................................ 167

Table 4-15 Mean Baseline SF-36 Quality of Life Scores (and standard deviations) for Intervention and Control Patients ........................................................................................................... 167

Table 4-16 Mean Changes in Physical Activity, Cardiovascular and Quality of Life Outcomes in the Control and Intervention Groups, at 12 Months ............................................................................. 169

Table 4-17 Changes in Antihypertensive Medication Use from Baseline to Follow-up, amongst the Intervention and the Control groups ........................................ 171

Table 4-18 Change in Weight, Body Mass Index and Serum Cholesterol Concentration of Intervention and Control Groups over 12-months ........................................ 172

Table 4-19 Multi-level, Time-series Analysis of Adverse Events for Intervention and Control Patients (comparing the year prior to study enrolment with the year following study enrollment) .... 172

Table 4-20 Average Change in Leisure-time Physical Activity and Total Energy Expenditure by Baseline Stage of Change amongst Intervention Participants ........................................ 174

Table 4-21 Per-protocol Sensitivity Analyses of Primary Outcome Measures ........................................................................................................ 175

Table 4-22 Sensitivity Analysis using Multiple Regression of Change in Physical Activity, Clinical, and Significant Quality of Life Outcomes for the Intervention Group compared with the Control Group, after adjusting for Age, Gender, Socio-economic Status, and Clustering ............................................................................. 176

Table 4-23 Sensitivity Analysis using Multiple Regression of Change in Physical Activity, Clinical, and Significant Quality of Life Outcomes for the Intervention Group compared with the Control Group, after adjusting for Baseline Outcome Measures and Clustering ............................................................................. 177

Table 4-24 Intra-class Correlation Coefficients (ICCs) for Demographic and Clinical Characteristics in a Consecutively Screened Sample of Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices ............................................................................. 179

Table 4-25 Intra-class Correlation Coefficients (ICCs) for Physical Activity Characteristics in a Consecutively Screened Sample of Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices ............................................................................. 180

Table 4-26 Intra-class Correlation Coefficients (ICCs) for Change in Variables from Baseline to 12-month Follow-up amongst Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices ............................................................................. 180

Table 4-27 Costs of Set-up and Co-ordination of the Green Prescription Programme Nationally .......... 183

Table 4-28 Costs of Sports Trust attributable to the Green Prescription Study Participants for 12 months (2001/2002) ........................................................................................................... 185
Table 4-29 General Practice Staff Costs of Delivering the Green Prescription Intervention .......... 186
Table 4-30 Average Cost of Follow-up Advice from General Practitioner or Practice Nurse during the 12 months following Intervention Delivery .......................................................... 186
Table 4-31 Offset Patient Costs Associated with Exercising for the Year between Baseline and Follow-up
.............................................................................................................................. 187
Table 4-32 Change in Primary Health Care Use by Study Participants for the Year-Before and the Year-
After Baseline (adjusted for inflation and discounted at 5%)............................................. 189
Table 4-33 Change in Costs of Public Hospital Admissions and Outpatient Use for the Year-before and
the Year-after Study Enrolment (costs adjusted for inflation)............................................. 190
Table 4-34 Change in the Cost of Days of Work due to Sickness or Accident during the Year-before
compared with the Year-after Study Enrolment for Intervention and Control Study
Participants .................................................................................................................. 191
Table 4-35 Offset Costs per Patient for the Intervention Group compared with the Control Group
(Intention-to-treat Analysis) ......................................................................................... 192
Table 4-36 Incremental Cost per Patient of the Green Prescription Programme, including Programme and
Offset Costs and Savings, by Intention-to-treat ............................................................. 193
Table 4-37 Incremental Monthly Cost-Effectiveness Ratios (CER) for the Green Prescription vs. Usual
Care from Programme-Funder and Societal Perspectives with Sensitivity Analysis ........... 196
Table 4-38 Cost Effectiveness Ratios for the Green Prescription Compared with Project Active
'Lifestyle' and 'Structured' Physical Activity Promotion Programmes a.............................. 197
List of Figures

Figure 3-1 Waikato Region of New Zealand and Location of the Green Prescription Study.............................. 100

Figure 3-2 Flow Diagram of the Design of the Cluster Randomised Controlled Trial of the Green Prescription Intervention ................................................................. 107

Figure 4-1 Participation Rates in the Physical Activity Questionnaire Reliability and Validity Study... 131

Figure 4-2 Bland-Altman Graphs of the Auckland Heart Study (AHS) and Green Script Study (GSS) Questionnaires compared with the 7-day Diary Estimations of Total Moderate Activity and Total Energy Expenditure (TEE) ............................................................. 135

Figure 4-3 Pilot Study Mean Changes in Physical Activity Variables, Leisure-time Moderate and Vigorous-Intensity Activity (minutes per week) and Total Energy Expenditure (kcal/kg/week), for Intervention and Control Patients ........................................ 140

Figure 4-4 Pilot Study Mean Change in Systolic and Diastolic Blood Pressure for the Intervention and Control Patients, from Baseline to 2-month Follow-up ........................................ 141

Figure 4-5 Pilot Study Mean Changes in Lipid Concentrations (mmol/l) and 5-year Cardiovascular Relative Risk Reduction (proportion) for Intervention and Control Patients from Baseline to 2-Month Follow-up .................................................. 142

Figure 4-6 Pilot Study Mean Changes in SF-36 Quality of Life Scores for Physical Function, Role Physical and Bodily Pain, for the Intervention and Control Patients from Baseline to 2-month Follow-up .................................................. 142

Figure 4-7 Pilot Study Mean Changes in SF-36 Quality of Life Scores for General Health, Vitality, Social Functioning, Role Emotional and Mental Health, for the Intervention and Control Patients from Baseline to 2-month Follow-up .................................................. 143

Figure 4-8 Process of Recruitment of General Practitioners and Randomisation of Practices in the Green Prescription Cluster Randomised Controlled Trial ................................................................. 145

Figure 4-9 Progress of Patient Screening, Recruitment and Follow-up Stages of the Trial following Pre-randomisation of Practices ................................................................. 147

Figure 4-10 Prevalence of Hypertension in Study Participants compared with the New Zealand Population Prevalence* (non-Māori only) ................................................................. 151

Figure 4-11 Prevalence of known Diabetes amongst Study Participants compared with New Zealand Population Estimates* ............................................................................. 152

Figure 4-12 Prevalence of known Diabetes amongst Study Participants compared with New Zealand estimates by Age and Ethnic Group ................................................................. 152

Figure 4-13 Baseline SF-36 Quality of Life Scores, ‘Physical Function’, ‘General Health’, and ‘Mental Health’, of Study Participants by Age Group ................................................................. 160

Figure 4-14 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Male Study Participants 40-64 years of age compared with New Zealand Gender-matched and Age-similar Norms .................................................................................. 161
Figure 4-15 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Male Study Participants 65-79 years of age compared with New Zealand Gender-matched and Age-similar Norms .................................................................................................................. 161

Figure 4-16 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Female Study Participants 40-64 years of age compared with New Zealand Gender-matched and Age-similar Norms .................................................................................................................. 162

Figure 4-17 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Female Study Participants 65-79 years of age compared with New Zealand Gender-matched and Age-similar Norms .................................................................................................................. 162

Figure 4-18 Practice Sizes of Intervention and Control Practices .................................................................................................................. 163

Figure 4-19 Number of Rural and Urban Practices in the Intervention and Control Groups .................................................................................................................. 164

Figure 4-20 Gender of General Practitioners in the Intervention and Control Groups .................................................................................................................. 165

Figure 4-21 Number of Full-time and Part-time General Practitioners in the Intervention and Control Groups .................................................................................................................. 165

Figure 4-22 Proportions of Participants achieving 2 ½ Hours of Moderate or Vigorous Leisure Activity per Week at Baseline and 12-month Follow-up .................................................................................................................. 170

Figure 4-23 Distribution of Change in Programme and Offset Costs (Offset) for Control Patients (0) and Intervention Patients (1) in New Zealand Dollars for the Year before Baseline compared with the Year following Baseline .................................................................................................................. 194

Figure 4-24 Distribution of Change in Programme and Offset Costs (Offset) for Control Patients (0) and Intervention Patients (1) in New Zealand Dollars for the Year before Baseline compared with the Year following Baseline with Outliers* Removed .................................................................................................................. 195
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHS</td>
<td>Auckland Heart Study physical activity questionnaire (Appendix 1) (Jackson 1989)</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>BP</td>
<td>Blood Pressure</td>
</tr>
<tr>
<td>CER</td>
<td>Cost Effectiveness Ratio</td>
</tr>
<tr>
<td>Chol</td>
<td>Cholesterol</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CSC</td>
<td>Community Services Card</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular Disease</td>
</tr>
<tr>
<td>DALY</td>
<td>Disability Adjusted Life Years gained</td>
</tr>
<tr>
<td>D.eff.</td>
<td>Design Effect (Inflation factor used to increase estimated sample size when using a cluster design)</td>
</tr>
<tr>
<td>GP</td>
<td>General Practitioner</td>
</tr>
<tr>
<td>GRx</td>
<td>Green Prescription</td>
</tr>
<tr>
<td>GSS</td>
<td>Green Script Study physical activity questionnaire (Appendix 2) (Swinburn, Walter et al. 1998)</td>
</tr>
<tr>
<td>HDL</td>
<td>High Density Lipoprotein</td>
</tr>
<tr>
<td>ICC</td>
<td>Intraclass Correlation Coefficient</td>
</tr>
<tr>
<td>IPA</td>
<td>Independent Practitioners’ Association</td>
</tr>
<tr>
<td>ITT</td>
<td>Intention-to-treat</td>
</tr>
<tr>
<td>LDL</td>
<td>Low Density Lipoprotein</td>
</tr>
<tr>
<td>LYG</td>
<td>Life Years Gained</td>
</tr>
<tr>
<td>Km²</td>
<td>Square kilometres</td>
</tr>
<tr>
<td>Kcal</td>
<td>Kilocalories</td>
</tr>
<tr>
<td>Kj</td>
<td>Kilojoules</td>
</tr>
</tbody>
</table>
Kg Kilograms
MET Ratio of the associated metabolic rate for a specific activity divided by the resting metabolic rate.
Mm Hg Millimetres of mercury
Mmol/L Millimoles per Litre
N Number in Sample
NNT Numbers Needed to Treat
QALY Quality Adjusted Life Year
RCT Randomised Controlled Trial
RR Relative Risk
RRR Relative Risk Reduction
sd Standard deviation
SES Socio-Economic Status
SOC Stage of Change according to the Transtheoretical model of readiness to change
SF-36 Short-form 36 Quality of Life Questionnaire
SF-6D Short-form 6 Dimension preference measure scale developed by Brazier
SPARC Sport and Recreation New Zealand (formerly Hillary Commission)
TADS Tobacco, Alcohol and Other Drug Early Intervention Project, Goodfellow Unit, University of Auckland
TC Total Cholesterol
TEE Total Energy Expenditure
TG Triglycerides
UK United Kingdom
US United States
VO2max Maximum oxygen consumption (with maximum perceived effort)
1 Introduction

1.1 Rationale for the Thesis

This thesis reports the results of a comprehensive evaluation of the Green Prescription physical activity intervention in New Zealand general practice.

There is now substantial epidemiological evidence to implicate a sedentary lifestyle as a risk factor for obesity, diabetes, cardiovascular disease, osteoporosis and various other disease states (Paffenbarger, Hyde et al. 1986; Blair, Kohl et al. 1995; Lee, Hsieh et al. 1995). This research has led to major position statements, such as the U.S. Surgeon General’s report on “Physical Activity and Health”, which reviewed the evidence for health gains from regular physical activity (Centers for Disease Control and Prevention 1996). This report concluded that people of all ages benefit from regular physical activity and that significant benefits can be obtained from 30 minutes of moderate-intensity physical activity on most days of the week (Centers for Disease Control and Prevention 1996). The report noted that physical activity reduces the risk of premature mortality in general, and of coronary heart disease, hypertension, colon cancer and diabetes mellitus, in particular, as well as improving mental health. Physical activity has also been associated with improved quality of life and hence improved health outcomes particularly amongst the elderly (Spirduso and Cronin 2001). Preventative health strategies have developed accordingly (Thompson, McAfee et al. 1995).

The Ministry of Health in New Zealand has identified eight population-attributed risk factors, which have a significant impact on mortality and health expenditure in New Zealand (Ministry of Health 1999). These factors include smoking, alcohol consumption, inadequate fruit and vegetable intake, physical inactivity, diabetes (type 2), obesity, high blood pressure and total cholesterol.

Within this political context and because of the evidence, physical activity interventions have been developed within primary health care to improve health outcomes. Primary health care seems an ideal context in which to offer a physical
activity intervention. Primary health care services have personal contact with a large proportion of the population, since at least 80% of the population visit the general practitioner annually (Ministry of Health 1998). General practice is also a recognised source of health advice. Many preventative services based in general practice have been shown to be more acceptable to the public, and more effective at recruiting participation, than community-based initiatives (Clover, Redman et al. 1996; Giorgi, Giordano et al. 2000). However, evidence of the effectiveness of these interventions, both in terms of increases in physical activity and health, is sparse, particularly long-term effectiveness. Consequently, the research on which this thesis is based aimed to determine whether an intervention developed for primary health care can be effective in increasing physical activity levels and improving health in the long-term.

In New Zealand, at least 32% of adults do not achieve the recommended 2 ½ hours of moderate physical activity per week, which is estimated as the minimum required for health-benefit (Sport and Recreation New Zealand 2003). Several national initiatives promoting physical activity have been undertaken in New Zealand including the "Green Prescription", promoted by the Hillary Commission. Recommended levels of physical activity are outlined on this prescription, which is given to patients by the general practitioner. In conjunction with other promotions the Green Prescription aims to achieve greater numbers of physically active New Zealanders. As such, it is designed to reduce morbidity, premature mortality, and health costs for individuals and the New Zealand government. It is assumed that increasing the physical activity of New Zealanders would reduce morbidity and mortality, particularly that associated with cardiovascular disease (Russell, Worsley et al. 1987). Another expected outcome of the Green Prescription initiative is that the increased physical activity will result in an improvement in the quality of life of many New Zealanders. Therefore, the effectiveness of a physical activity intervention in primary health care needs to assess health outcomes such as cardiovascular risk and quality of life as well as physical activity. This thesis reports such outcomes.

The intervention was designed using a behavioural change approach based on the transtheoretical model, which is discussed more fully in Chapter 2 (Prochaska and DiClemente 1983). Patients need to be receptive to the messages, appreciate that physical activity will improve their health, and have few barriers to increasing
physical activity in practice. If general practitioners are aware of these factors and tailor their counselling accordingly, the intervention is more likely to succeed. The initiative also involves on-going support to the patient receiving a Green Prescription in the form of guidance from community-based exercise specialists, the option of community-run activity programs, and on-going encouragement from the surgery. This support may be important for the maintenance of change in patients' levels of physical activity. Clinicians and exercise specialists were taught motivational interviewing techniques and assessment of the 'stage of change', based on the transtheoretical approach, to maximise patient physical activity behaviour change (Marcus and Simkin 1994).

For an initiative like the Green Prescription programme to succeed, general practitioners must be willing to accept and implement the strategy. They must also demonstrate the skills to motivate a change in patient behaviour, as has been demonstrated in evaluations of similar interventions overseas (Riddoch, Puig-Ribera et al. 1998). The efficacy of the primary care physician or general practitioner in increasing physical activity has been demonstrated (Anonymous 1994; Cupples and McKnight 1994; Calfas, Long et al. 1996). However, the gains in physical activity in these studies were either short-term, involved only those with pre-existing cardiovascular disease or were in the context of a programme addressing multiple risk factors. In additions, the gains were modest. Furthermore, the evaluation of health outcomes from physical activity counselling in primary care has rarely been attempted. Where it has, outcomes have been disappointing (Halbert, Silagy et al. 2000; Dubbert, Cooper et al. 2002). Further study of the effectiveness of such primary care interventions in the 'real world' in New Zealand is necessary to understand the true impact of the brief primary care intervention.

New health initiatives and quality assurance components need to be evaluated in practice to see if they are achieving their specified aims (Marwick, Grol et al. 1992). Comprehensive evaluations of interventions, such as the Green Prescription, provide evidence for health funders, to address health outcome targets (Ministry of Health 1999). Health outcomes are defined as measurable changes in health status.

One possible barrier to the success of health initiatives such as the Green Prescription programme and other quality assurance initiatives, is that there is often a variable
uptake by general practitioners. A committee of the World Organisation of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA) examined the influences involved in motivating general practitioners to change. Part of the quality assurance cycle, and an important motivator for general practitioners, was found to be the provision of feedback to the individual general practitioner, on the effect of their interventions (Marwick, Grol et al. 1992). In the present study, we encouraged general practitioner involvement in the study, provided feedback as to how their patient participants fared, and disseminated outcome-based evidence of the effect of their intervention.

For the purposes of this study, the Green Prescription intervention was developed one-step further to include systematic screening of patients at the practice reception. This screening identified those who were not achieving the recommended level of physical activity and prompting of general practitioners was done to deliver the intervention to those eligible and willing. As such, the Green Prescription became a systematic practice-based and patient-prompted intervention rather than an opportunistic clinician-initiated intervention. The rationale for this approach was to allow all those eligible to receive the intervention and to maximise the potential population benefit. This strategy also streamlined the process in that the general practitioner was not responsible for ascertaining present level of activity or determining who would be willing to receive such an intervention.

Interventions may sometimes be effective but lack practical application in real life circumstances. Therefore, it was important to test an intervention that was both sustainable and acceptable to all participants. The Green Prescription intervention has had a good level of acceptability from clinicians and patients in New Zealand over several years. The history of the development and uptake of this intervention is presented in Chapter 2. Briefly, the Green Prescription programme was already being used in some regions of New Zealand at the time of this study. A short-term randomised controlled trial of the Green Prescription had already been undertaken which demonstrated increases in physical activity over a six-week period (Swinburn, Walter et al. 1998). Indeed, the intervention was introduced into the Waikato region mid 1999 at the same time as the randomised controlled trial for this thesis was being designed. This timing allowed for a preliminary assessment of effectiveness before
widespread adoption of the intervention, but also ensured that an intervention with demonstrated short-term effectiveness, good face validity and sustainability was being assessed.

It has not previously been shown that Green Prescriptions facilitate a long-term increase in physical activity. Nor has it been shown that any physical activity intervention in primary care improves the health of the patients involved. The previous randomised controlled trial of exercise advice from a general practitioner supplemented with a Green Prescription showed a short-term (6-week) improvement in physical activity when compared with exercise advice alone (Swinburn, Walter et al. 1998). The control group in this study received advice, thus potentially underestimating the true impact of the initiative. In addition, potential selection bias caused by having the enrolled doctors select patients for inclusion in the study may limit the generalisability of the results (Swinburn, Walter et al. 1998). Thus, there was a need for a randomised controlled study evaluating the effect of a prescription for exercise in terms of long-term exercise levels and health outcomes within the current setting of general practice in New Zealand.

In order to carry out the present study, a cluster randomised controlled trial design was chosen based within general practice, to ensure a rigorous evaluation of effectiveness and reduce the risk of contamination of intervention. A discussion of this methodology is set out in Chapters 2 and 3. The population with the most to gain from a physical activity intervention includes those who are less active and those at high risk of cardiovascular disease, the rationale for which is presented in Chapter 2. Accordingly, sedentary patients aged between 40 and 80 years of age and visiting their usual general practitioner were chosen as the study population. The trial was also defined as 'pragmatic' in that it was based in every-day general practice to ensure generalisability of findings to real-life practice.

An important criterion, when justifying differential allocation of 'scarce' resources, is that of cost effectiveness. Therefore, a cost-effectiveness analysis of the intervention was also carried out. With the increasing push for cost-effective health care, a study addressing this aspect has the potential to make a major contribution to current understanding of the utility of the health dollar. This study may inform future
decisions concerning the allocation of funds to preventative health initiatives like the Green Prescriptions.

The cost-effectiveness analysis includes both costs and offset costs to the funder to allow funding decisions to be made, but also includes costs and offset costs to the participants and to the community at large. This approach is called a 'societal perspective' and outlines the real costs and benefits to society of a health intervention. Furthermore, while a cost-effectiveness analysis provides the cost per unit health outcome (cost-effectiveness ratio), a comparison with another intervention is necessary to produce a comparative cost-effectiveness analysis (Drummond, O'Brien et al. 1997). Hence, the cost-effectiveness of the Green Prescription intervention is compared with that of two other physical activity interventions from the United States and the United Kingdom. The methods and results for the cost-effectiveness evaluation are presented in Chapters 3 and 4, respectively.

Issues of physical activity assessment had to be addressed prior to the evaluation of the effectiveness of a physical activity intervention. Appropriate measurement of physical activity within large epidemiological trials is an area for debate. A summary of the issues is presented below. This thesis uses physical activity questionnaires to assess physical activity. Although the use of questionnaires is a recognised method for assessing physical activity, no questionnaires were found that had been validated within primary health care amongst less-active middle-aged or older adults, the target population of the intervention. Yet, the validity of an instrument should be context and population-specific (Rennie and Wareham 1998). To achieve this purpose, a physical activity questionnaire was adapted for use in the trial, and a reliability and validity study undertaken within primary health care amongst less-active adults prior to the commencement of the principal study. The methods used in the reliability and validity study are presented in Chapter 3 and the results in Chapter 4.

1.2 Objectives of the Thesis

In the light of the foregoing considerations, this thesis addresses the three topics of adapting and validating an appropriate physical activity questionnaire, evaluating the effectiveness of the 'Green Prescription' intervention, and assessing the cost-effectiveness of this intervention. Specifically the objectives of the thesis are:
1. To assess the reliability and validity of a physical activity questionnaire adapted for use amongst less-active adults within a primary health care setting.

2. To assess the effectiveness of the Green Prescription physical activity intervention within primary health care over a one-year period in terms of physical activity, cardiovascular risk, and quality of life, amongst less-active 40 to 79 year-old adults.

3. To assess the cost-effectiveness of the Green Prescription physical activity intervention.

1.3 Structure of the Thesis

The thesis is divided into six chapters, including introduction, literature review, methods, results, discussion, and conclusions. The introduction has presented the rationale for the study and its objectives. The literature review sets the thesis within the wider context of what is known in the literature about physical activity and cardiovascular disease and quality of life. A history of the development of the Green Prescription intervention is also presented. Three systematic reviews are included, which assess existing background knowledge for the three objectives of this thesis. A justification for the methods used in this thesis is also presented as the final section of the literature review.

The methods, results and discussion chapters are divided into the three objectives dealt with in the thesis. The first objective involves a study of reliability and validity of two adapted physical activity questionnaires. The second includes a pilot study and the cluster randomised controlled trial of the Green Prescription intervention. The third comprises a cost effectiveness evaluation of the Green Prescription intervention. Results and implications for widespread implementation and future research from this work are discussed in the conclusion in Chapter 6.
2 Literature Review

2.1 Introduction

This literature review provides the context for the thesis. The first section presents a background summary of evidence for the relationship between inactivity and cardiovascular disease and the link between physical activity and quality of life. This section also includes a short history of the evolution of the Green Prescription programme. The next three sections deal with the three thesis objectives, outlining what is known from the literature already and identifying gaps that this thesis addresses. Lastly, the literature review presents the rationale for using a randomised controlled trial with cluster randomisation to demonstrate the effectiveness of the Green Prescription programme in a pragmatic setting.

2.2 Background

2.2.1 Physical Inactivity and Cardiovascular Disease

2.2.1.1 Relative Risk

Based on longitudinal epidemiological studies, physical inactivity has been recognised as a risk factor for cardiovascular disease for more than half a century. As early as 1953, Morris and colleagues identified an increased incidence of coronary heart disease and mortality amongst sedentary civil servants compared with individuals in active jobs such as postal workers (Morris, Heady et al. 1953). They were also able to demonstrate the same increased incidence amongst ‘sedentary’ bus drivers compared with ‘active’ bus conductors (Morris, Kagan et al. 1966).

Paffenbarger and others found a dose-response relationship between physical activity (miles walked, stairs climbed, sports played and energy expenditure per week) and all-cause mortality after following up 16,936 Harvard alumni for 16 years (Paffenbarger, Hyde et al. 1986). This study may have had limited generalisability because only men participated and they were predominantly of a higher socio-economic and educated background. However, the evidence has remained consistent
in subsequent studies and systematic reviews (Kampert, Blair et al. 1996; Lee and Skerrett 2001). There is a particularly strong relationship between physical inactivity and risk of coronary heart disease as demonstrated by meta-analyses and systematic reviews (Powell, Thompson et al. 1987; Kohl 2001). A large meta-analysis by Berlin and Colditz in 1990 showed a relative risk of coronary heart disease of 1.3 -1.9 for sedentary, compared with active, adults with respect to occupational and non-occupational physical activity (Berlin and Colditz 1990). This finding has been confirmed by many subsequent studies (Lindsted, Tonstad et al. 1991; Blair, Kohl et al. 1995; Morgan and Clarke 1997; Bijnen, Caspersen et al. 1998). Further evidence to support the dose-response relationship between activity and cardiovascular risk comes from the Nurses’ Health Study where there was a graded reduction in coronary artery disease risk with increased walking (Manson, Hu et al. 1999).

There is also evidence for a graded reduction in risk of all-cause mortality with increasing cardiovascular fitness. Blair and colleagues followed a cohort of 9,777 men for 5 years measuring physical fitness and all-cause mortality. There was a drop in all-cause mortality of 44% between the most and the least fit. There was also evidence that the least fit benefited the most by an increase in cardiovascular fitness (Blair, Kohl et al. 1995).

However, because all these epidemiological studies are non-experimental, there is always the risk of unrecognised confounding factors. Therefore, although there appears to be a relationship between physical activity and primary prevention of cardiovascular disease and reduced mortality, it is not completely certain. Although there are randomised controlled trials for secondary prevention using lifestyle interventions such as physical activity that have shown improved cardiovascular outcomes or reduced mortality, there are no such trials for primary prevention (O’Connor, Buring et al. 1989; Miller 1997; Ebrahim and Davey-Smith 1999; Garcia-Palmieri 2000; Ketola, Sipila et al. 2000; Coats 2001; Jolliffe, Rees et al. 2003). Primary prevention intervention trials have shown improvements in cardiovascular risk with increased exercise, such as blood pressure, lipid profiles and risk of diabetes, but not cardiovascular outcomes or mortality. Randomised controlled trials with cardiovascular and mortality outcomes would be needed to more convincingly demonstrate the relationship between physical activity and primary prevention of
cardiovascular disease and the magnitude of that relationship. However, these trials would have to be very large with a long follow-up to demonstrate change. The existing evidence for benefit from physical activity in improving cardiovascular risk factors means that such a trial would not only be expensive and logistically difficult, but may face opposition due to the ethical issues surrounding withholding physical activity advice.

Even so, previous reviews of observational studies have found that the better the quality of the study, the stronger the relationship demonstrated between physical activity and cardiovascular disease outcomes (Powell, Thompson et al. 1987; Berlin and Colditz 1990). This finding increases the likelihood that there is a causal relationship between physical inactivity and cardiovascular disease.

In New Zealand research, the relative risk of cardiovascular mortality associated with physical inactivity has been estimated to be between 1.3 and 1.8 (Ministry of Health 1999). These figures are consistent with a review of primarily longitudinal studies, by Stephenson and colleagues (Stephenson, Bauman et al. 1999).

2.2.1.2 Is there a Dose-Response Relationship?

There appears to be a dose-response relationship between physical activity and cardiovascular mortality, based on observational studies (Kohl 2001). The greater the usual energy expenditure, the lower the risk of all-cause mortality after controlling for several confounding risk factors (Kesaniemi, Danforth et al. 2001). However, the physical activity threshold to reduce cardiovascular risk in terms of intensity, duration and frequency, is inconclusive (Kohl 2001). Measurements of physical activity in studies have often been dichotomised to produce a comparison of risk of high activity to moderate activity or sedentary lifestyles. This makes a comprehensive understanding of the relationship between energy expenditure and risk, difficult, particularly when the measurement, definitions, and categories of intensity and quantity of physical activity vary from one study to the next.

A recent systematic review of observational studies examining physical activity and all-cause mortality also found clear evidence of a linear dose-response relationship between total energy expenditure and all-cause mortality in men and women (Lee and Skerrett 2001). Many studies measured only energy expenditure associated with
leisure time or occupations, rather than total energy expenditure. The measurement of only one type of activity limits interpretation of change in energy expenditure, because reductions or increases in other areas, such as domestic activity, are ignored and not measured. For example, while leisure exercise may increase, total energy expenditure may actually decrease due to compensatory reductions in domestic, transport, or occupational physical activity.

Despite these limitations, current evidence shows that adherence to the recommended 30 minutes of moderate-intensity physical activity on most days of the week, was associated with a reduction in all-cause mortality risk of 20-30% (Lee and Skerrett 2001). A 20-30% reduction in risk of ischaemic stroke has also been observed with similar amounts of physical activity or walking in a large study of 72,500 participants, conducted over a period of eight years (Hu, Stampfer et al. 2000). This amount of physical activity represents approximately 1000 kcal/week (4,200 kilojoules) energy expenditure in addition to daily light activity and basal metabolic rate. Further risk reductions were associated with greater energy expenditure. There was also some indication that there was at least some benefit from lower amounts of light activity, possibly as low as 500 kcal/week (Lee and Skerrett 2001). There was not enough data collected by the studies reviewed to ascertain if different frequencies, duration or intensities of activity were significant, assuming constant energy expenditure.

These figures are based on observational studies only, and no randomised controlled trials of interventions were found that demonstrated reduced mortality in primary prevention. However, there is some evidence of a weakly graded response from randomised controlled trials of physical activity and short-term health outcomes (Oja 2001). Outcomes include cardiovascular fitness, blood pressure, lipid profiles and glucose tolerance, which are indicators of cardiovascular and other morbidity and mortality.

2.2.1.3 Cardiovascular Fitness

Cardiovascular fitness, more than physical activity, is closely correlated with health outcomes and mortality. The relationship is clearly curvilinear with an asymptote at higher fitness levels (Blair, Kohl et al. 1995; Blair, Cheng et al. 2001). The less clear relationship with physical activity may be partly due to the difficulty entailed in
measuring physical activity accurately in everyday life (Kesaniemi, Danforth et al. 2001).

A review of the literature showed that three ten-minute sessions of moderate activity throughout the day were as effective at improving cardiovascular fitness as one 30-minute session (ie the energy expenditure was the same) (Hardman 2001). However, it was also found that cardiovascular fitness increased more with high intensity activities than low or moderate intensity activities undertaken for a longer duration, assuming constant energy expenditure (Hardman 2001).

2.2.1.4 Blood Pressure:

2.2.1.4.1 Acute Effects:

The acute post-exercise hypotensive effect can be a reduction as great as 18-20 mm Hg systolic and 7-9 mm Hg diastolic for up to 12-16 hours (Thompson, Crouse et al. 2001). The best response is seen in mildly hypertensive people, where a daily exercise regime can lower blood pressure for most of each day. Average reductions in blood pressure post exercise (12-16 hours), including normotensive individuals, were more modest (2.1/0.3 mm Hg) using an exercise threshold of about 40% of VO2 max (ibid).

2.2.1.4.2 Chronic Effects

Two good quality meta-analyses report that the blood pressure response to aerobic physical activity is similar for frequencies between three and five times per week and for sessions between 30 and 60 minutes (Halbert, Silagy et al. 1997; Fagard 2001). Exercise trials included in the meta-analyses varied from four to 52 weeks. Exercising at 30-85% of VO2max reduces blood pressure by 3/2 mm Hg in normotensive participants and 7/6 mm Hg in the hypertensive group, according to one meta-analysis (Fagard 2001). An average reduction of 4.7/3.1 mm hg was found in the other meta-analysis (Halbert, Silagy et al. 1997) and 5.4/1.9 mm hg in another review (Thompson, Crouse et al. 2001). All three reviews found very little difference in blood pressure reduction with moderate versus vigorous intensity exercise (eg: 40 - 85% VO2 max). There was some evidence that sustained effects on blood pressure are achieved after three sessions of exercise (Thompson, Crouse et al. 2001).
beneficial effect can be lost in 1-2 weeks of not exercising (Thompson, Crouse et al. 2001).

The most recent high quality meta-analysis of physical activity interventions and blood pressure found that aerobic exercise (30–60 minute sessions, 3-5 times per week) was associated with a significant reduction in blood pressure (Whelton, Chin et al. 2002). Mean systolic and diastolic blood pressure reductions of -3.84 mm Hg and -2.58 mm Hg, respectively, were found. Concurring with previous reviews, there was no obvious advantage of vigorous over moderate intensity activity. The modest reduction in blood pressure may reflect the fact that more long-term trials, and more community-based trials with less exercise ‘supervision’, were included in this meta-analysis, compared with previous reviews. This may mean lower compliance with exercise and an apparent lower reduction in blood pressure when compared with previous estimations.

2.2.1.5 Lipids

It has been estimated that for every 0.026-mmol/L increase in serum high-density lipoprotein (HDL) concentration, the risk for a coronary heart disease event is reduced by two percent in men and at least three percent in women (Leon and Sanchez 2001). The same review stated that one percent reduction in serum low density lipoprotein concentration (LDL) is associated with a two to three percent lower risk of coronary heart disease (Leon and Sanchez 2001).

2.2.1.5.1 Acute Effects

Physical activity acutely reduces triglycerides (3-15%) and increases HDL (4-43%) 18-72 hrs after exercise (Thompson, Crouse et al. 2001). There is no obvious threshold for the effect but it seems to be proportional to energy expenditure. The effects are probably mediated by increased enzyme activity, causing the delay in response. Lowered LDL (4-38%) is partly due to haemodilution and is seen only after prolonged high intensity activity (eg marathon running) (ibid).
2.2.1.5.2 Chronic Effects

Meta-analyses and reviews have noted variable relationships between change in lipid profile and increased physical activity. The most consistent finding is an increase in HDL of 4-5% with regular moderate or vigorous activity (Halbert, Silagy et al. 1999; Leon and Sanchez 2001). Average reductions of 0.08–0.10 mmol/L have been reported for triglycerides, LDL and total cholesterol in one meta-analysis (Halbert, Silagy et al. 1999). These estimates are similar to findings of other reviews based on both randomised controlled trials and observational studies (Kesaniemi, Danforth et al. 2001; Leon and Sanchez 2001). There is inadequate data to evaluate a dose-response relationship, but most studies involved exercise energy expenditures of over 500 kcal/week (Leon and Sanchez 2001).

The lipid profile response to exercise is equivalent for young and old. One review of the experimental literature suggested HDL in men might respond more to exercise (Wilmore 2001). However, there was no difference found between the response of men and women in terms of their total cholesterol/HDL ratios or other lipid levels in another review of randomised controlled trials and observational studies (Leon and Sanchez 2001; Wilmore 2001). Study participants with previous cardiovascular disease or worse baseline lipid profiles tended to improve their lipid profiles with exercise more than “healthy” participants were able to achieve (Wilmore 2001). Even so, there appears to be considerable individual variation in physiological responses to exercise (Kesaniemi, Danforth et al. 2001).

Furthermore, HDL changes may be partly explained by weight-loss from exercise. A review of studies about health effects and weight-loss from exercise showed that while observational studies suggested an independent effect of exercise on lipids, many intervention studies did not address the issue of body weight as a confounding factor (Williams 2001). When body weight was controlled for in the intervention studies, the increase in HDL was no longer statistically significant. He concluded, also, that the reduced coronary heart disease and cardiovascular disease observed with increased fitness might also be primarily due to weight control (ibid).
2.2.1.6 Weight Reduction

Intervention trials of weight-loss from physical-activity interventions with short-term follow-up (up to 16 weeks) tend to involve twice the energy expenditure (2200 vs 1100 kcal/week) of long-term studies (more than 26 weeks). This was reported in a recent review of the literature (Ross and Janssen 2001). These short-term follow-up studies resulted in three times the weight reduction (0.18 vs 0.06 kg/week) and fat loss (0.21 vs 0.06 kg/week) compared with long-term follow-up studies. There was also a dose-response relationship between exercise-induced weight loss and body fat reduction in short-term studies but not in long-term studies. Physical activity tended to be associated with reductions in abdominal and visceral adiposity (Ross and Janssen 2001).

Recently, the American College of Sports Medicine made a position statement about appropriate interventions for weight loss. This statement suggests that although the recommended 30 minutes of moderate activity five days per week (equivalent to approximately 1000 kcal/week) is of benefit for cardiovascular risk, twice that amount may be needed for weight-loss (>2000 kcal/week) (Jakicic, Clark et al. 2001).

2.2.1.7 Diabetes Prevention

Diabetes is a significant risk factor for cardiovascular disease and mortality. Regular physical activity is associated with a decreased risk of developing type 2 diabetes. A review of cohort studies showed that a 6% decrease in risk of developing diabetes was associated with each 500 kcal expended on leisure-time physical activity per week, after adjustment for age (Kesaniemi, Danforth et al. 2001).

There is also evidence from intervention trials that physical activity and dietary interventions can impair progression to type 2 diabetes from impaired glucose tolerance, by 50-60%. This evidence was derived from the Malmo intervention trial in Sweden (Eriksson and Lindgarde 1991), the Diabetes Prevention study in the United States (Charatan 2001) and the Da Qing study in China (Pan, Li et al. 1997). The Da Qing study included some groups with exercise alone, in which reductions in the progression to diabetes were reported (Kelley and Goodpaster 2001). There is not enough data to work out dose-response relationships, but the United States intervention study was based on the recommendation of 30 minutes of physical
activity on most days of the week (Anonymous 1999). The benefit of the exercise and the dietary intervention was also effective in significantly reducing mortality rates to the level of the normoglycaemic population at 12 years follow-up of the Malmo trial (Eriksson and Lindgarde 1998).

2.2.1.8 Glucose Metabolism and Diabetes Control

Physiological studies have demonstrated that exercise decreases insulin resistance in peripheral tissues and improves insulin-stimulated glycogen metabolism in skeletal muscle. However, this may be an acute effect rather than a training effect. Exercise also improves post-prandial glucose (not fasting levels) and may increase insulin secretion post-prandially. If this last effect were proven, it would be unique as neither diet nor pharmacological agents can increase early phases of post-prandial insulin secretion. There is also an indication that exercise may acutely decrease the hepatic production of glucose in type 2 diabetes. Overall, metabolism of glucose in the peripheral tissues of individuals with diabetes becomes closer to ‘normal’ with exercise (Kelley and Goodpaster 2001) (Thompson, Crouse et al. 2001).

A review of randomised controlled trials showed that even a single session of exercise can improve glucose control in people with type 2 diabetes and ameliorate insulin resistance in other subjects (Thompson, Crouse et al. 2001). Hepatic production of glucose the next day is reduced. However, the response lasts for only a few days. A session of moderate activity for 45-60 minutes has been shown to reduce blood glucose levels in type 2 diabetics by 1-2 mmol/L, while blood glucose levels of people without diabetes remain constant, even during vigorous exercise (Thompson, Crouse et al. 2001).

One third of studies in one review found no long-term decrease in HBA1c from exercise interventions for individuals with type 2 diabetes, despite increases in fitness (Kelley and Goodpaster 2001). Where a decrease in HBA1c was found, this was modest (0.5-1.0%). (The minimum therapeutic criteria indicated by the United States Federal Drug Administration for approving pharmacological glucose lowering medications is about 1% reduction in HbA1c) (Kelley and Goodpaster 2001). Because the effect is modest, no dose-response relationships have been revealed.
It is difficult to separate the effect of exercise and diet on diabetes control, as they are often advised together in intervention studies. Furthermore, medications may be altered with improving and worsening diabetic control, potentially masking the effect of exercise (Kelley and Goodpaster 2001).

2.2.1.9 Coagulation, Haemostatic, and Other Cardiovascular Effects

Changes in the balance of factors promoting and inhibiting coagulation probably play a role in both the potentially harmful and protective effects of exercise. For example, the formation of thrombosis is an important process in cardiovascular events such as myocardial infarction and stroke. Strenuous exercise can increase platelet activation and adhesion, particularly amongst previously sedentary individuals who have a higher catecholamine response to exercise (Rauramaa, Li et al. 2001). This process may contribute to cardiovascular events or sudden death occasionally seen in strenuous exercise. However, platelet activity is inhibited during moderate exercise (eg: 50% of VO2 max) and platelet adhesion decreases with regular exercise.

Raised fibrinogen levels are an independent risk factor for cardiovascular disease (Kesaniemi, Danforth et al. 2001). Observational studies show an inverse relationship between fibrinogen levels and physical activity levels. However, randomised controlled trials have been unable to demonstrate a consistent effect on fibrinogen (ibid). Tissue plasminogen activator is increased acutely with exercise and plasminogen activator inhibitor (PAI-1) is decreased, but only in strenuous exercise. Regular physical activity seems to have little long-term effect on these factors (ibid). Adipose tissue plays an active role in thrombus formation and may be involved in exercise induced changes in PAI-1 production. The role of adipose release of PAI-1, insulin and other cytokines as a response to exercise is not completely understood (Rauramaa, Li et al. 2001).

Regular exercise also improves myocardial contractility and coronary perfusion. In fact, exercise improves arterial compliance and endothelial function, in general. It is thought that greater sheer stress with exercise enhances synthesis of nitric oxide from endothelial cells. Nitric oxide may slow the development of atherosclerosis. Nitric oxide may also reduce risk of acute coronary events by relaxing smooth muscle and
inhibiting the proliferation of smooth muscle, platelet aggregation and leukocyte adhesion to vessel walls (Clarkson, Montgomery et al. 1999) (Stewart 2002).

2.2.1.10 Summary

In summary, there is substantial observational evidence that increasing physical activity level is associated with improved cardiovascular outcomes. However, experimental evidence for an independent effect of physical activity on primary cardiovascular events and mortality is lacking, due to the large study population and time that would be required to demonstrate such an effect. Ethical and compliance issues would also be encountered in such a trial.

However, there is strong experimental evidence to show a relationship between increasing levels of physical activity and intermediate outcomes predictive of cardiovascular outcomes. In particular, it has been shown that increasing level of physical activity amongst 'less active' and 'at risk' individuals can increase cardiovascular fitness. Increasing physical activity can also lower blood pressure, improve serum lipid profiles, reduce weight, lower risk of diabetes development and improve glycaemic control. In addition, some of the mechanisms are now understood by which physical activity may reduce the risk of coagulation defects and clot-formation, important in the onset of coronary and cerebrovascular events.

2.2.2 Prevalence of Physical Inactivity in New Zealand

The prevalence of 'physical inactivity' in a country depends on the definition used or the question asked. Unfortunately, there has not been agreement over a standard survey question or measurement tool for asking about physical activity. This limits the ability to compare prevalence estimates of inactivity between countries (Stephens and Caspersen 1994).

Some instruments used to measure national prevalence of inactivity have not been validated, such as those used in the United States (Sallis and Saelens 2000). Furthermore, relying on self-report as a measure of physical activity levels is fraught with inaccuracies (Sallis and Saelens 2000; Lamonte and Ainsworth 2001). Throughout the developed world, such as the United Kingdom, United States, Australia and New Zealand, reported proportions of the 'inactive' population vary
considerably. Prevalence of inactivity is estimated to be 32% (Sport and Recreation New Zealand 2003) to 42% in New Zealand (Ministry of Health 1999). In Australia, the prevalence has been estimated to be 44% (Stephenson, Bauman et al. 1999) compared with 60% in the United States (Berlin and Colditz 1990) and 70% in the United Kingdom and Finland (Haapanen-Niemi, Vuori et al. 1999). If there were a close causal relationship between cardiovascular disease and physical inactivity, then one would expect large differences in cardiovascular disease incidence between these countries, which is not the case. It is more likely that the definition of inactivity or the question used in each country is slightly different.

Evidence-based recommendations were made in the United States by the Centres for Disease Control and Prevention, and the American College of Sports Medicine, and disseminated in the Surgeon General’s report in 1996 (Pate, Pratt et al. 1995; Centers for Disease Control and Prevention 1996). These reports concluded that every adult should accumulate 30 minutes or more of moderate-intensity physical activity on most, preferably all, days of the week (Pate, Pratt et al. 1995). The American Heart Association has made similar recommendations (Stewart 2002). Since that time, most definitions of physical inactivity have been based on this recommendation. However, New Zealand surveys tend to ask about the accumulation of 2 ½ hours of moderate activity per week, while researchers from the United States ask about 30 minutes of moderate activity on most days of the week. This discrepancy may partly explain the difference in prevalence figures.

2.2.2.1 Incidence of Cardiovascular Disease in New Zealand

The incidence of death from coronary heart disease, hypertension or cerebrovascular accidents, which make up most of the cardiovascular disease in New Zealand, was 109/100,000 (0.0011) for female and 195/100,000 (0.0020) for male populations in 1996. These rates accounted for 9,581 deaths during that year (Hay 1999).

2.2.2.2 Population Attributable Risk

An estimated 2143 deaths (29,137 years life lost) can be attributed to physical inactivity per year in New Zealand (Ministry of Health 1999). Using the concept of population attributable risk (PAR) and assuming the population was 3.8 million in
1996, approximately 922 cardiovascular deaths amongst men and 507 cardiovascular deaths amongst women could be attributed to physical inactivity in New Zealand.  

2.2.2.3 Fraction of Cardiovascular Mortality Attributable to Physical Inactivity

A theoretical “attributable fraction” \((AF)\) or amount of a cardiovascular mortality attributable to a risk factor can be calculated.\(^2\) Therefore, approximately 24% of all cardiovascular mortality could be attributed to physical inactivity in New Zealand, if the relative risk for cardiovascular mortality is 1.8 for inactive compared with active individuals and the prevalence of inactivity is 0.4. In other countries, the attributable fraction of cardiovascular mortality due to physical inactivity has also been estimated. This figure was estimated to be 35% in the United States (Kesaniemi, Danforth et al. 2001) and between 22% and 39% amongst middle-aged Finnish men after adjustment for other risk factors (Haapanen-Niemi, Vuori et al. 1999). The higher attributable fractions are due to the estimation of higher prevalence \((p)\) of inactivity in some countries (e.g. 60% in the United States and 70% in Finland).

In New Zealand, the impact of physical inactivity on cardiovascular mortality will depend on the relative risk and the prevalence of physical inactivity in each gender and age group. Relative risks and prevalence of inactivity for each subgroup have

\[\begin{align*}
\text{PAR} &= \frac{I_{t} - I_{u}}{I_{t}} \\
&= p*I_e + (1 - p)*I_u \\
&= p(RR - 1)*I_u \\
I_{t} &= \text{incidence of cardiovascular mortality (only including CHD, hypertension and CVA deaths) of the total population and is estimated to be 0.002 (195/100,000) for males and 0.0011 (109/100,000) for females based on mortality statistics (Hay, 1999).} \\
p &= \text{the proportion of the population exposed to the risk (prevalence) which is estimated to be 0.4 (Ministry of Health, 1999).} \\
RR &= \text{relative risk of cardiovascular mortality and is estimated to be 1.8 - 1.3 (Berlin and Colditz, 1990; Ministry of Health, 1999).} \\
\text{And: } RR &= I_e/I_u \\
I_e &= \text{incidence of cardiovascular mortality of the exposed population (inactive). } I_e \text{ is therefore estimated to be 0.002727 (1/367) for males and 0.0015 (1/667) for females.} \\
I_u &= \text{incidence of cardiovascular mortality of the unexposed population (active). } I_u \text{ is therefore estimated to be 0.001515 (1/660) for males and 0.000833 (1/1200) for females.} \\
\text{Assuming the population was approximately 3.8 million in 1997-9 and } PAR &= \frac{I_{t} - I_{u}}{I_{t}}. \\
\text{For males: } PAR &= 0.002 - 0.001515 = 0.000485 \\
&= 0.000485 * 1,900,000 \text{ population} = 922 \text{ deaths} \\
\text{For females: } PAR &= 0.0011 - 0.000833 = 0.000267 \\
&= 0.000267 * 1,900,000 \text{ population} = 507 \text{ deaths} \\
\text{AF} &= p (RR - 1) * I_u / I_t \\
\text{And, } AF &= \frac{(p * (RR - 1))}{(p * (RR - 1) + 1)} \\
&= (0.4 * (1.8 - 1)) / (0.4* (1.8 - 1) + 1) \\
&= 0.24
\end{align*}\]
been obtained from the literature and the attributable fractions (AF) are presented in
Table 2-1 (Ministry of Health 1999).

Table 2-1 Relative Risk of Cardiovascular Mortality for Inactive compared with
Active Individuals, and the estimated Cardiovascular Mortality attributable
(Attributable Fraction) to Inactivity for Middle-aged and Older Adults in New
Zealand.

<table>
<thead>
<tr>
<th>Age</th>
<th>RR (female)</th>
<th>AF (%)</th>
<th>RR (male)</th>
<th>AF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45-64</td>
<td>1.8</td>
<td>24</td>
<td>1.8</td>
<td>24</td>
</tr>
<tr>
<td>65-74</td>
<td>1.4</td>
<td>11</td>
<td>1.8</td>
<td>19</td>
</tr>
<tr>
<td>75+</td>
<td>1.3</td>
<td>13</td>
<td>1.3</td>
<td>11</td>
</tr>
</tbody>
</table>

RR refers to relative risk; AF refers to attributable fraction

However, these calculations use univariate analysis techniques and assume that risk
factors are randomly distributed and do not interact with other risk factors. These
assumptions are unlikely to be the case. Therefore, these results are crude estimates
only (Macera and Powell 2001).

2.2.2.4 Summary

In summary, levels of physical inactivity are high in New Zealand and are responsible
for a significant portion of the cardiovascular morbidity and mortality. If interventions
are developed to increase levels of physical activity, substantial health gains may
result. To assess the effect of these interventions on cardiovascular outcomes,
intermediate outcomes that empirically reflect cardiovascular risk can be measured,
such as blood pressure.

2.2.3 Measurement of Cardiovascular Risk

In considering evaluation of physical activity as impacting on cardiovascular events,
estimated overall 'cardiovascular risk' can be used as a surrogate outcome. An
individual's risk of having a cardiovascular event in the next 4-12 years can be
estimated using an equation that was originally developed from a study which took
place in Framingham in the United States (Anderson, Odell et al. 1990). A
cardiovascular event includes angina pectoris, myocardial infarction, peripheral
vascular disease, congestive heart failure, transient ischaemic attack and stroke. Since the initial development of the equation, equations to predict coronary heart disease development, cerebrovascular disease event, any cardiovascular event, myocardial infarction, and mortality from cardiovascular event for any period from 4 to 12 years, have been developed (ibid).

The Framingham equations were developed to predict risk for those who had not had cardiovascular events in the past. Recently, data from several generations from the Framingham study have been analysed and an equation for predicting coronary heart events in people with previous cardiovascular disease has been developed (D'Agostino, Russell et al. 2000). This equation may be less accurate for predicting events in older women (personal communication with D'Agostino), however, it offers the best prediction for secondary events to date.

There are other limitations with these equations. Firstly, they were developed using an American population. There is some evidence to suggest that the risk of some diseases varies with ethnic group, particularly cardiovascular disease. In this case, the equations may not be valid for the New Zealand population with its unique ethnic mix. Secondly, only a certain number of variables are taken into account in the equations, including blood pressure, serum lipid concentration, diabetes status, gender, age, left ventricular hypertrophy, and smoking. There may be other factors independent of the traditional cardiovascular variables that could independently predict the outcome, such as energy expenditure, intensity of exercise undertaken, stress, or serum C-reactive protein (Ridker, Rifai et al. 2002).

The Framingham 5-year cardiovascular risk equation has been shown to adequately predict proportions of the New Zealand population who will have a cardiovascular event in the next 5 years (unpublished data, Richard Milne). However, the equation was also shown to predict an event in individual men, only marginally better than using 'age' as a predictor, based on data from a cohort of over 6,000 New Zealand men and women (unpublished data, Richard Milne). The equation was no better than 'age' at predicting the first cardiovascular event in individual women.
2.2.4 Physical Inactivity and Quality of Life

Other positive outcomes associated with physical activity include improved mental health (Dunn and Dishman 1991; Glenister 1996; Yeung 1996; Stephenson, Bauman et al. 1999; Miser 2000; Kesaniemi, Danforth et al. 2001). Although much of the evidence is based on observational studies, one review of randomised controlled trials was able to show a causal relationship between increasing physical activity and improved mental health (Glenister 1996). There is also experimental evidence of the beneficial effects on mood acutely, even after only one episode of exercise (Yeung 1996).

Health-related quality of life refers to the cognitive, social, physical and emotional functioning of a person (National Health Committee 1998). Improved quality of life has also been associated with increasing levels of physical activity (McAuley, Courneya et al. 1991; Steptoe, Rink et al. 2000; Shephard 2001). Quality of life is also an important predictor of physical functioning, particularly amongst older age groups (Spirduso and Cronin 2001), although more randomised controlled trials are needed to establish a causal relationship. A linear relationship between quality of life and physical activity amongst older adults has been reported in one review (Kesaniemi, Danforth et al. 2001).

Consequently, this thesis used quality of life as an outcome measure to assess one aspect of effectiveness of the physical activity intervention in primary health care. The use of the SF-36 questionnaire to assess health-related quality of life is described in Chapter 3.

The preceding discussion justifies physical activity as an important area to attempt intervention. In addition, relevant outcomes from a physical activity intervention include amount of activity, cardiovascular risk, and health-related quality of life. These three outcomes were chosen as measures of effectiveness in the cluster randomised controlled trial of a physical activity intervention, described in Chapters 3 and 4. Consideration of a viable physical activity intervention for evaluation follows, with a description of the development of the Green Prescription intervention.
2.2.5 Development and Implementation of the Green Prescription Intervention

2.2.5.1 Introduction

In response to the high prevalence of inactivity and its health cost, initiatives to promote physical activity have been developed throughout the world, including initiatives in primary health care. In New Zealand, the Green Prescription physical activity counselling initiative was developed for use in general practice. This section describes how the intervention was developed and implemented. The theoretical basis upon which the intervention was developed is also summarised.

2.2.5.2 Primary Care as the Setting for an Intervention

The choice of primary health care as a setting for a physical activity intervention was made for various reasons. Firstly, general practice offers access to a majority of the population. Over 80% of the population visit their general practitioner at least yearly (Ministry of Health 1998). Secondly, the population visiting the general practice expects to receive health advice. Receiving advice about physical activity from a health professional may then be seen in the context of their continued medical and preventative care. Thirdly, the general practitioner or practice nurse knows the risk factors for each individual, and appropriate targeting of risk factors can be carried out. In addition, the general practitioner or practice nurse often has knowledge about each individual’s psychosocial context and potential barriers to behaviour change, so these factors can be addressed. Other brief interventions, such as smoking cessation, have been shown to be effective within a primary health care context (Silagy and Stead 2003).

2.2.5.3 Development and Implementation

The Green Prescription intervention was developed in the mid 1990s by the Hillary Commission (now known as Sport and Recreation New Zealand or SPARC), a government and publicly funded organisation that promotes physical activity and sport in New Zealand. The Green Prescription intervention was based on exercise prescription schemes developed in the United Kingdom. Thirty-seven general practitioners from Auckland and Dunedin tested the Green Prescription on sedentary patients in a randomised controlled trial undertaken during 1995 and 1996. A short-
term increase in self-reported physical activity was demonstrated (Swinburn, Walter et al. 1998). A qualitative study, investigating attitudes and perceptions of general practitioners about the Green Prescription intervention and its implementation, assisted in the development of resources for general practitioners and patients (Swinburn, Walter et al. 1997). Pacific Island Heartbeat adapted the resources for Pacific Island languages (Cook Island Maori, Tongan and Samoan).

The Green Prescription intervention was officially launched in the northern region of New Zealand by the then Minister of Sport, Fitness, and Leisure (Murray McCully) in May 1997. Exercise specialists at regional sports foundations were trained about their support role in the Green Prescription intervention, and its implementation. Training of general practitioners was achieved through local Independent Practitioners Associations (IPAs). A public media campaign was held in Northland and Auckland using radio, bus posters and mail, to encourage patients to ask their general practitioner for a Green Prescription.

By May 1998, a survey of 317 Northland and Auckland general practitioners showed that 65% had prescribed a Green Prescription and 97% were giving verbal advice about increasing physical activity (Communication with the Hillary Commission). In July 1998, a nation-wide rollout of the intervention was started through IPAs, Māori Health Providers, individual general practitioners, and practice nurses. Computer prescribing of the intervention was also introduced into general practice. Green Prescription area managers were employed to liaise between the Hillary Commission, regional sports foundations, general practitioners, practice nurses and other health professionals throughout New Zealand. Patients who had received a Green Prescription and who either contacted exercise specialists or were referred by their clinician to an exercise specialist at a regional sports foundation, were sent out quarterly newsletters. Direct referral from health clinician to exercise specialist, by fax, established telephone-based exercise specialist support of patients. Referrals continued to increase yearly, with 4,000 referrals made from mid 2000 to mid 2001, and 4,925 made from 2001 to 2002 in the following year (Figures supplied by the Hillary Commission).

Surveys of random samples of between 400 and 1000 general practitioners were conducted annually from 1999-2003 to ascertain uptake of the initiative. The
proportion of general practitioners, nationally, that had given a Green Prescription increased steadily from 47% in 1999-2000 to 73% in November 2002. However, the average number prescribed, remained low (ranging from 2.7 to 5 per general practitioner per month over that period). The 2002 survey indicated that practice nurses also initiated the intervention in 35% of the practices surveyed. The reasons recorded for giving a Green Prescription were most frequently noted as ‘weight’ (93%), ‘diabetes’ (61%), and ‘high blood pressure’ (60%) (Information provided by the Hillary Commission). ‘Time’ appeared to be the major impediment to more frequent use of the Green Prescription.

Funding for the Green Prescription initiative was provided by the Ministry of Health during 1999/2000. Pharmac took over the funding from 2000 until 2003. A project to investigate how the intervention could better meet the needs of Māori was commenced in November 2000 (Information provided by the Hillary Commission).

2.2.5.4 Transtheoretical Behaviour Change Model

Following the initial studies of the Green Prescription intervention (Swinburn, Walter et al. 1997; Swinburn, Walter et al. 1998), motivational interviewing techniques were introduced to increase the effectiveness of the intervention. These techniques were based on the transtheoretical approach, which has been suggested as an effective way to deliver physical activity advice (Marcus and Simkin 1994; Marcus, King et al. 1996). Motivational interviewing tailors the counselling approach to the ‘readiness to change’ or ‘stage of change’ of the patient. In the context of physical activity, the stages of change refer to the following. The ‘precontemplator’ stage means the patient is not doing adequate physical activity, but also does not want to increase their activity. The ‘contemplator’ stage means they are considering increasing activity. The ‘preparation’ stage means they are preparing to increase activity. The ‘active’ stage means that they are already increasing their physical activity. The ‘maintenance’ stage refers to individuals who are maintaining adequate activity levels.

The original transtheoretical theory was developed to change smoking behaviour (Prochaska and DiClemente 1983). Subsequently, the model has been applied to other behaviour change interventions for alcohol and substance abuse, smoking cessation, dietary change, HIV risk behaviour and exercise (Dunn, Deroo et al. 2001). Many of
the studies of primary care physical activity intervention have used this model (Graham-Clarke and Oldenburg 1994; King, Sallis et al. 1998; Goldstein, Pinto et al. 1999; Harland, White et al. 1999; Dunn, Deroo et al. 2001; Kirk, Higgins et al. 2001; Sarkin, Johnson et al. 2001; Pinto, Friedman et al. 2002; Woods, Mutrie et al. 2002).

The evidence for a greater effectiveness of the transtheoretical approach, compared with other approaches, is not certain. A systematic review of 29 randomised controlled trials of behaviour change found that motivational interviewing was effective at changing substance abuse behaviour (Dunn, Deroo et al. 2001). However, there was inadequate evidence to demonstrate its effectiveness in other lifestyle behaviour change, including physical activity. A meta-analysis of 91 studies involving physical activity found that there was evidence of a relationship between stage and actual level of activity and most changes in activity were in the direction predicted by the transtheoretical theory (Marshall and Biddle 2001). Even so, there was not clear evidence that physical-activity behaviour change occurs in a series of qualitatively different stages. More work on standardising the stages and improving reliability is needed (ibid).

Using stage of change as an outcome is of limited significance, if not backed up by an actual change in physical activity. It is therefore inappropriate to use ‘stage of change’ as a principal outcome without also measuring actual change in physical activity, which has been done in several physical-activity intervention studies (Pinto, Lynn et al. 2001; Woods, Mutrie et al. 2002).

Once the importance of physical activity to health outcomes was established, and a reasonable intervention to evaluate was identified, information about measurement of the main outcomes and evidence for the effect of other interventions was sought.

2.3 Systematic Review of Physical Activity Questionnaire Reliability and Validity Studies

2.3.1 Introduction

This section presents a background to physical activity assessment in epidemiological study. The advantages and limitations of using physical activity questionnaires for assessment are discussed, with particular reference to their potential use in primary
health care research. The requirements of a 'good quality' physical activity questionnaire and the importance of validating questionnaires amongst appropriate populations and in relevant settings are justified.

Previous reviews of physical activity validity and reliability studies are presented and discussed. In addition, a systematic review of the physical activity validity and reliability studies published in the literature from 1990 to 2002 was conducted for this thesis. Evidence tables are included in Appendix 3. The section concludes with the rationale for conducting the physical activity questionnaire reliability and validity study for this thesis.

2.3.2 Measurement of Physical Activity using Self-report Instruments

The issue of what intensity of activity is necessary to lower disease risk remains contentious (Wilson, Paffenbarger et al. 1986). Part of the reason for the lack of agreement is the difficulty in measuring everyday physical activity accurately, amongst individuals and populations (Pols, Peeters et al. 1995; Pols, Peeters et al. 1997; Wareham and Rennie 1998). The use of the physical activity questionnaire for assessment of physical activity levels is the most practical method for large-scale epidemiological study (Washburn 2000). More objective methods, such as direct and indirect calorimetry and doubly-labelled water techniques, are either impractical in an epidemiological setting or their expense prohibits large-scale use (Ainsworth, Leon et al. 1993). Measures of fitness, such as VO2 max (maximum oxygen consumption), are often reasonably correlated with the amount of vigorous-intensity activity that an individual undertakes, but are often not correlated with the amount of moderate- and lower-intensity activity (Hopkins, Wilson et al. 1991).

Physical activity questionnaires are frequently used in epidemiological study. They are low-cost and practical, can be designed or adapted for particular study populations, are well tolerated by respondents, and have been shown to be reasonably accurate (Pereira, FitzGerald et al. 1997). However, physical activity self-report instruments do have limitations. Firstly, self-report bias has been demonstrated (social desirability bias) (Sallis and Saelens 2000). Self-report has sometimes been found to overestimate physical activity by as much as 100-300%, particularly when estimating vigorous activity when compared with estimates from objective measures, such as
doubly-labelled water or accelerometry (Klesges, Eck et al. 1990; Sallis, Strikmiller et al. 1996). However, moderate activity tends to be more accurately reported or even underestimated in physical activity questionnaires (Hayden, Sallis et al. 1998; Buchowski, Townsend et al. 1999). Categorical data, such as the number of days of at least 60 minutes of physical activity, seem to be more accurately recalled and reported, than continuous data, such as duration and intensity of certain activities. (Prochaska, Sallis et al. 1999) However, categorical data lacks the precision of continuous data necessary to detect small amounts of activity change and to investigate subtle dose-response relationships between physical activity and health outcomes.

Cognitive skills are required to interpret and fill out physical activity questionnaires, which can be challenging, particularly for very young, very old, or less literate respondents (Sallis and Saelens 2000). This phenomenon may explain why one review found that interviewer-administered questionnaires often have higher validity than self-administered questionnaires (ibid). Interviewer administration allows explanation and assistance by the interviewer. However, another study found higher reliability correlations for questionnaires mailed and self-administered questionnaires, compared with telephone interview-administered questionnaires (Washburn, Smith et al. 1993). Recall bias usually implies an overestimate of exercise participation. Yet, some studies have shown that self-administered questionnaires have produced lower estimates of leisure activity than interviewer-administered questionnaires (Vuillemin, Oppert et al. 2000).

In a non-blinded trial situation, interviewer assistance may increase the chance of interviewer bias compared with self-administered questionnaires. In addition, interviewer-administered questionnaires may not be practical for screening patients in a busy primary care setting, where self-administered questionnaires would be more efficient. Therefore, a self-administered questionnaire was chosen for use in the Green Prescription trial.

When conducting physical activity questionnaires there needs to be agreement about definitions of terms such as ‘vigorous’ or ‘moderate’ intensity activity between respondents, interviewers, and researchers. This agreement is difficult to achieve. However, the use of empirically based physical activity compendia can help
standardise the interpretation of certain physical activities as light, moderate, or vigorous, according to their energy expenditure classification (Ainsworth, Haskell et al. 2000). Therefore, such compendia were used in the Green Prescription trial, as described in Chapter 3.

Self-report physical activity instruments also lose accuracy if the information reported is not complete (Sallis and Saelens 2000). Many physical activity recall instruments do not ask about all categories of activities or appropriate modes of activity for the respondent population, so large portions of daily activity may not be recorded. For example, one study found that physical activity levels of African American women were underestimated because the questionnaire asked only about leisure-time physical activity (ibid). Much of their activity was undertaken while working at their occupation and while undertaking household tasks. Therefore, it is important for self-report instruments to ask about activities appropriate to the study population, and measure frequency, intensity, type, and duration of activity, making sure that leisure-time, domestic, occupational, and transport activities are included (Bauman, Armstrong et al. 1999). Furthermore, estimates of light activity, inactivity, and rest, need to be made to produce a realistic estimate of total energy expenditure. Again, such information is necessary to detect subtle dose-response relationships between physical activity levels and health outcomes (Pereira, Fitzerald et al. 1997; Kesaniemi, Danforth et al. 2001). These issues were considered when selecting and adapting a questionnaire for use in the Green Prescription trial.

The fact that a variety of physical activity self-report measures have been developed and are being used in epidemiological study, makes it difficult to compare results between studies (Stephens and Caspersen 1994). However, if we are to feel confident about the results of a survey or study of physical activity, the reliability and the validity of the instrument used to estimate physical activity should be demonstrated. In many situations, this has not been done. For example, the measures used to estimate national levels of physical activity in the United States have not been validated, (Sallis and Saelens 2000) although, those used to estimate New Zealand national levels, have been validated (Hopkins, Wilson et al. 1991). Furthermore, the reliability and the validity of the instrument used to measure physical activity are important for other reasons.
When assessing the reliability of self-report physical activity instruments, test-retest reliability results will depend not only on the quality of the instrument, but also on several other factors. Besides the cognitive and memory abilities of the participants, the period asked about in the questionnaire, and the length of time between administrations of the questionnaire, will also influence results. For example, if the two administrations of a questionnaire ask about different periods of activity, discrepancies in answers may reflect real change in activity levels rather than low reliability of the tool.

It is also important to assess the criterion-related validity of questionnaires used. At present, there is no ideal objective gold standard that measures all the aspects and categories of physical activity. Objective direct measures of physical activity, such as doubly-labelled water, heart rate monitoring, and motion sensing, have been favoured over objective indirect measures, such as cardio-respiratory fitness (eg. VO2max) and body composition, as they are more closely related to actual activity (Sallis and Saelens 2000). For example, cardio-respiratory fitness has a variable relationship with actual physical activity, particularly in younger people (Sallis and Saelens 2000) and older adults (Tager, Hollenberg et al. 1998). Body composition is influenced by many other variables (such as energy intake).

The use of subjective measures, such as activity diaries as comparators for validation, risks the same recall bias as questionnaires. In one study, activity diaries were found to over-estimate physical activity by 40-60% amongst women (Jakicic, Polley et al. 1998). However, direct measures, such as heart rate monitoring, have also been shown to over-estimate and under-estimate energy expenditure. For example, one study compared heart-rate monitoring and activity diaries for estimating energy expenditure, using energy intake adjusted for changes in body energy stores as a reference (Kalkwarf, Haas et al. 1989). This study found group energy expenditure was over-estimated by heart rate monitoring by 2-9%, but that errors in estimating individual energy expenditure ranged from -53% to 67%. Activity diaries, on the other hand, underestimated group energy expenditure by 2-6%, with individual errors from -39% to 56%. This finding would suggest that while both methods are adequate for group activity estimation, neither is particularly good at estimating individual activity (Kalkwarf, Haas et al. 1989).
While objective measures do not risk 'recall bias', none of the objective measures discriminate type, frequency, duration, and intensity of component activities, so subjective measures, such as activity diaries, have been widely used to supplement assessment of validity (Ainsworth, Montoye et al. 1994; Montoye, Kemper HCG et al. 1996; Wareham and Rennie KL 1998).

Activity diaries are particularly useful if they are structured to measure the same classifications of activities, over a similar time span, as the questionnaire. Types, durations and subjective intensities can be recorded which are equivalent to those measured by the questionnaire. Activity diaries have been used to validate most well-known physical-activity questionnaires (Arroll, Jackson et al. 1991; Jacobs, Ainsworth et al. 1993; Richardson, Leon AS et al. 1994; Wareham and Rennie KL 1998). Interestingly, the levels of validity comparing questionnaires with activity diaries, are often similar to or greater than those achieved against doubly-labelled water for the same questionnaires (Philippaerts, Westerterp et al. 1999; Bonnefoy, Normand et al. 2001; Conway, Irwin et al. 2002).

Despite the limitations of 'self-report', physical activity questionnaires do often demonstrate good correlation with more objective measures. The use of self-report instruments is justified, particularly if they are used to measure change in group physical activity, rather than absolute energy expenditure of individuals. Even so, it is essential that such instruments demonstrate good reliability, content validity, and criterion-related validity.

It is therefore important to review the literature on physical-activity questionnaire validity and reliability prior to choosing, adapting and validating an appropriate physical activity questionnaire for use in a study.

2.3.3 Previous Reviews of Physical Activity Questionnaires

The first comprehensive review of physical activity assessment in epidemiological research was carried out in 1985 (LaPorte, Montoye et al. 1985). This review divided self-report techniques into several categories, including 'recall' (detailed records of activities for periods of one to seven days), or 'quantitative histories', (asking about periods lasting more than a week) or 'general surveys'. Validity and reliability levels of the questionnaires were not included.
Another review was published in 1992, although it was not a systematic review (Ainsworth, Montoye et al. 1994). This review concluded that no "gold standard" measurement techniques exist. They stressed the need for accurate and valid questionnaires that measure the various types of physical activity, and can be generalised to populations of different socio-economic, gender, age, and ethnic groups (ibid).

Montoye et al published a review of physical activity questionnaires in 1996 (Montoye, Kemper et al. 1996). The authors present a summary of the six physical activity questionnaires most commonly used amongst adults in epidemiological studies, as a practical guide to choosing an appropriate instrument for estimating physical activity. Correlation coefficients for validity and reliability when compared with other recall methods, doubly labelled water, activity monitoring, and physical fitness measures, were presented in table format. In addition, physical activity questionnaires used in national surveys amongst the elderly, and amongst children, were discussed. Questionnaire reliability and validity results were also summarised.

The authors concluded that questionnaires are useful in large population surveys but have limitations, particularly for estimating activity of young children and the elderly. The development and validation of further questionnaires suitable for particular age groups and settings was recommended. In addition, the authors note that assessment of validity is limited by the lack of a 'gold standard' comparator, and to date, the validity and test-retest reproducibility of questionnaires have not been adequately studied. However, Montoye does suggest that doubly labelled water may be an effective criterion 'gold standard', useful in validating instruments that estimate total energy expenditure (Montoye, Kemper et al. 1996).

Another useful resource for physical activity questionnaires was published in 1997 (Pereira, FitzerGerald et al. 1997). This article presented reproductions, instructions about use, and data on reliability and validity, for most physical activity questionnaires used in epidemiological study. The questionnaires were grouped as suitable for use in the general population, suitable for older adults, or as major population-based surveys. The reliability and validity of each questionnaire were reported from the literature. The methods used and sample populations tested in the validation studies, were described in detail.
The next review found that four physical-activity questionnaires commonly used amongst older adults, the Modified Baecke, Zuphthen, Yale and PASE questionnaires were as reliable and valid as questionnaires for broader age groups. The measurement of physical activity in the older age group is particularly difficult because most of their activity is in the light to moderate categories, which are the most difficult to estimate accurately (Washburn 2000).

The most recent review was undertaken by Sallis and Saelens. This review examined self-report physical activity instruments that had been validated against objective measures (doubly-labelled water, heart rate monitoring or motion sensors) (Sallis and Saelens 2000). The review also assessed the content validity, by assessing whether the following dimensions were measured by each questionnaire: Type, frequency, intensity, and duration of activity, and sedentary, moderate, vigorous, leisure, work, household and transport activities (ibid).

All the reviews of physical activity questionnaires described above, focussed on the questionnaires themselves, rather than systematically reviewing validity and reliability studies. Some of the reviews presented the results of reliability and validity studies for each questionnaire (Pereira, FitzerGerald et al. 1997; Sallis and Saelens 2000). However, the latest review did not include physical activity questionnaires validated against activity diaries (Sallis and Saelens 2000). In addition, many studies have been published since these reviews. Therefore, an updated systematic review of reliability and validity studies of physical activity questionnaires was conducted for this thesis.

All published studies of the reliability or validity of physical-activity questionnaires were identified by a search, as described below, and reviewed to assess the level of reliability and validity considered adequate in the literature. Comments on the quality of the studies were also made. The review was also undertaken to locate any physical activity questionnaires that had been validated for use amongst adults in primary health care.

2.3.4 Methods

This sub-section presents the methods used for a systematic review of physical-activity reliability and validity studies published in English from 1990 to 2002. To be
consistent with the most recent systematic review, which focussed on the content validity of currently used questionnaires (Sallis and Saelens 2000), the period of review chosen was 1990 until the present. Validity studies prior to this date were excluded, because prior to the 1990s, most questionnaires did not include all categories of activity and focussed mostly on vigorous activity. It was only during the 1990s when the importance of moderate- as well as vigorous-intensity activity was appreciated, that the questionnaire designs reflect this emphasis and were more likely to include appropriate questions.

The present literature review addresses the following specific questions:

1. What are the methods used and the quality of reliability and validity studies carried out in studies to date?

2. What level of criterion-related validity is considered adequate for use in epidemiological study? ('Criterion-related' validity is the measure of how similar the results of one instrument are, to those of a 'gold standard' or 'reference standard', such as doubly labelled water or activity diary.)

3. Have any physical activity questionnaires been validated for adults within the setting of primary health care?
2.3.4.1 Search Strategy

Several electronic medical databases were searched. Key words were mapped to relevant subject headings for each database. All databases and terms used are listed in Table 2-2.

Table 2-2 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Test-Retest Reliability and Validity Studies of Physical Activity Questionnaires from 1990 to 2002

<table>
<thead>
<tr>
<th>Databases</th>
<th>Physical activity subject headings</th>
<th>Questionnaire subject headings</th>
<th>Validity and reliability subject headings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medline (1966-2002)</td>
<td>Exercise</td>
<td>Questionnaires</td>
<td>Reliability</td>
</tr>
<tr>
<td>SportDiscuss (1949-2002)</td>
<td>Physical fitness</td>
<td>Interviews</td>
<td>Reproducibility of results</td>
</tr>
<tr>
<td></td>
<td>Exertion</td>
<td>Health survey</td>
<td>Specificity</td>
</tr>
<tr>
<td></td>
<td>Leisure</td>
<td>Short survey</td>
<td>Test reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Comparative study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Validity</td>
</tr>
</tbody>
</table>

2.3.4.2 Inclusion Criteria

To be included in the review, the study needed to assess test-retest reliability, or criterion-related validity, of physical activity questionnaires. The questionnaires had to have been used to assess present levels of physical activity amongst adults for epidemiological purposes, and be self- or interviewer-administered. Studies that used physical activity diaries, logs or other records, motion sensors, heart-rate monitoring or doubly-labelled water estimation of energy expenditure, as direct reference measures, were included in the validation review. Studies using indirect criterion measures such as fitness (eg. VO2max or body composition) were not included unless direct criterion measures were also used.

2.3.4.3 Data Abstraction

All titles were assessed and abstracts read, of studies thought to fulfill inclusion criteria. The Medline search produced 428 papers. Titles or abstracts were reviewed to
produce 84 papers that potentially fulfilled inclusion criteria. The Embase search initially produced 93 papers from which 31 were potentially eligible for inclusion. A total of 61 articles fulfilled inclusion criteria, from these two databases, after a more careful examination of abstracts or full papers. Six more studies that fulfilled inclusion criteria were found from the 77 papers identified from the SPORTDiscus database. Lastly, Cinahl database was searched, which produced 142 studies from which no new studies fulfilling inclusion criteria were identified. A further five studies were identified by examining reference lists from previous reviews.

In total, 72 studies fulfilled the inclusion criteria. Most of these studies were read in full, except for a few that were part of theses or in papers that could not be obtained. In these cases, information was obtained from abstracts. Data abstracted for the systematic review included study design, questionnaire used, number, type and recruitment rates of study participants, statistical methods used, results, and authors' interpretation of results.

2.3.4.4 Critical Appraisal

The present author was unable to locate a validated quality assessment system for reliability and validity studies from the literature. The quality of diagnostic-test validation studies is related. Criteria for assessing quality of diagnostic test validation studies are presented below (Lijmer, Mol et al. 1999).

- Spectrum (population)
- Verification (complete, different reference tests or partial)
- Interpretation of test results (blinded or not blinded)
- Patient selection (consecutive or non-consecutive)
- Data collection (prospective, unknown or retrospective)
- Details of test (sufficient or insufficient)
- Details of reference test (sufficient or insufficient)
- Details population (sufficient or insufficient)

Of the eight criteria, an inappropriate or poorly described study population (spectrum), the use of unsuitable reference tests, an insufficient description of the test,
and a lack of blinding, were found to have had the greatest potential for exaggerating odds ratios (Lijmer, Mol et al. 1999).

Although these criteria are not directly transferable to physical activity reliability and validity studies, some aspects have been used for the critical appraisal in this systematic review. One criterion that is of importance is the quality of the 'gold standard' comparator used in the validity studies. Furthermore, the method of recruitment of the study sample (random, consecutive, convenience or volunteers) will affect generalisability, as will the population from which the sample is drawn and the participation rates. For example, a sample drawn from University alumni is likely to have an overrepresentation of higher socio-economic and educated individuals. Both these variables are associated with higher levels of usual physical activity. In addition, the results from such a study may not reflect the ability of less educated groups to fill out a self-administered questionnaire. The size of the sample will also influence the precision of the results, and the generalisability of findings. Therefore, the sample size, method of recruitment, participation rate, and appropriateness of the population from which the sample was drawn, were all used to assess the quality of the studies.

Physical-activity validity studies are difficult to 'blind'. For this reason, blinding was not used as a criterion for quality assessment. The choice of statistical tests used in the analysis of reliability and validity was assessed, because the choice of an inappropriate test, such as unadjusted Pearson's correlation coefficients for skewed data, may inflate the results. Therefore, measures of study quality included:

1. Choice of reference standard, ranging from highest to lowest quality (use of more than one reference comparator improves quality):
   a. Doubly-labelled water
   b. Heart Rate Monitoring
   c. Activity Monitor
   d. Activity Diary or Log
2. Study Population (size, selection, and participation rates, of sample population). Sample sizes were considered large if they were greater than 100, moderate between 30 and 100, and small if they were less than 30.

3. Methods of analysis

As most papers identified in the review included a reliability study, the most detailed quality assessments were carried out on reliability assessment. The implications of these assessments also affect the internal and external validity of the accompanying validity studies.

2.3.5 Studies Identified and Critical Appraisal

Seventy-two articles were identified containing reliability and validity studies that fulfilled the inclusion criteria. Studies are listed in four evidence tables, in order of year of publication (Appendix 3: Tables 7-1 to 7-4). These articles included 49 studies of questionnaire test-retest reliability (Table 7-1). Twenty-seven studies assessing validity of physical activity questionnaires compared with physical activity diaries or logs were found in the 72 articles (Table 7-2). Table 7-3 presents the 25 studies assessing validity of physical activity questionnaires compared with motion-sensors or heart rate monitoring located in the articles. Lastly, Table 7-4 presents nine studies assessing validity of physical activity questionnaires compared with doubly-labelled water.

2.3.6 Discussion of the Review

This systematic review assessed the studies of reliability and validity of physical activity questionnaires, published from 1990 to 2002 (Appendix 3), which met pre-specified inclusion criteria. No studies were identified that assessed the reliability and validity of physical activity questionnaires within a primary care context. The methodologies, and quality of the studies, were variable, which made comparison difficult.

In some studies, total energy expenditure or total activity was reported more reliably than other activity sub-groups (Table 7-1, Appendix 3) (Bharathi, Sandhya et al. 2000; Wareham, Jakes et al. 2002). However, this finding was not always the case and
there was a lot of variability in results, which made it difficult to generalise about the findings. For example, in some studies there were low test-retest reliability correlation coefficients for occupational activities (Wareham, Jakes et al. 2002) or household activities (Chasan-Taber, Erickson et al. 2002) compared with other categories. Other studies found occupational (Pols, Peeters et al. 1995) and household activities (Richardson, Leon et al. 1994) were the most reliably reported.

As has been noted previously, this review found that vigorous or 'heavy' activities tended to be reported more reliably than other activity categories (Jacobs, Ainsworth et al. 1993; Richardson, Leon et al. 1994). Vigorous activities were often more accurately reported in the validity studies, as well (Ainsworth, Leon et al. 1993; Chasan-Taber, Rimm et al. 1996; Ainsworth, Bassett et al. 2000; Singh, Fraser et al. 2001). In fact, reported levels of validity for total activity or vigorous activity were often greater than for other activity categories, and were sometimes the only correlations reported (Arroll, Jackson et al. 1991; Richardson, Ainsworth et al. 2001; Schuler, Richardson et al. 2001; Young, Jee et al. 2001; Wareham, Jakes et al. 2002). Light-intensity activity tended to be the least accurately reported in physical activity questionnaires compared with other reference measures (Richardson, Leon et al. 1994; Suleiman and Nelson 1997).

Motion sensors tended to be more highly correlated with reported walking, sport or vigorous activity, than with total activity or other activity categories (Table 7-3, Appendix 3) (Richardson, Ainsworth et al. 1995; Ainsworth, Bassett et al. 2000; Singh, Fraser et al. 2001). Even so, correlation coefficients of activity questionnaire estimates, compared with motion sensors estimates, tended to be low (e.g. r = 0.2-0.4). This was probably due to the fact that questionnaires and motion sensors are measuring different aspects of physical activity, as well as due to the imprecision of physical activity recall.

Men tended to report physical activity more consistently over time than women, although this was not always the case (Pols, Peeters et al. 1995; Pols, Peeters et al. 1997; Wareham and Rennie 1998). It is hard to know whether this finding meant that reporting was more reliable amongst men, or that actual activity varied less, than with women. In the validity studies (Tables 7-2 to 7-4, Appendix 3), results also showed greater validity when activity was reported by men than when reported by women.
(Pols, Peeters et al. 1995; Richardson, Ainsworth et al. 1995; Pols, Peeters et al. 1997; Pols, Peeters et al. 1997; Richardson, Ainsworth et al. 2001). However, sometimes the reverse was true (Ainsworth, Leon et al. 1993; Suzuki, Kawakami et al. 1998; Singh, Fraser et al. 2001). It may be that some questionnaires were more suited to the usual activities of men. In addition, men tend to participate more in sporting and vigorous activities that are more reliably and more accurately reported, than light and moderate activities.

There did not seem to be an agreed upon standard for accepting reliability or validity as adequate for epidemiological study, although most chose to consider correlation coefficients of at least 0.6, adequate for reliability. The correlation coefficients considered adequate for validity varied depending on the reference measure used for comparison. Coefficients of 0.2 to 0.5 were often considered adequate for motion sensors, while coefficients of at least 0.6 were expected for activity records or diaries. The acceptability of agreement also varied depending on the activity categories compared. For example, greater agreement was expected for vigorous activities than moderate or light activities, due to the difficulty of recalling the latter types, accurately. Assessing reliability and adequacy of the questionnaires was not only complicated by differences in the ability to recall, but also by the fact that some discrepancies are due to real activity change rather than the inadequacies of the questionnaire.

This systematic review was limited by the fact that the included studies represented only those located through the search of the stated databases. Bibliographies of individual studies were not examined for further studies and the grey literature was not searched. It was assumed that the present search was representative rather than exhaustive. The systematic review was also limited by the exclusion of studies using indirect measures of activity, such as VO2max.

The strengths of this systematic review are that it is the most up-to-date, and the first to systematically locate reliability and validity studies primarily, rather than assessing the reliability and validity studies of specific questionnaires. This review examines what is generally considered as adequate reliability and validity in the literature. In addition, it provides a critical appraisal of some of the methods used in the studies.
2.3.6.1 The Quality of Reliability and Validity Studies

The quality of reliability and validity studies was variable. It was difficult to compare the studies because of the use of different definitions, energy expenditure calculations, categorisation of activity, and analytical methodologies. However, almost all of the 49 reliability studies used correlation coefficients to assess test-retest reliability, with a significant proportion using more than one method of statistical analysis to assess agreement, which aids in interpreting the internal validity of results.

Many studies used Pearson correlation coefficients, without supplementing the analysis with non-parametric statistics (Rauh, Hovell et al. 1992; Dipietro, Caspersen et al. 1993; Bonnefoy, Kostka et al. 1996; Pols, Peeters et al. 1996; Allison, Keller et al. 1998; Suzuki, Kawakami et al. 1998; Bharathi, Sandhya et al. 2000). Pearson’s correlation coefficients tend to produce higher values than Spearman rank order correlation coefficients when data are skewed, which is frequently the case with physical activity data. However, some studies appropriately conducted log-transformation of the skewed data prior to calculating Pearson’s correlations (Ainsworth, Jacobs et al. 1993; Ainsworth, Leon et al. 1993; Chasan-Taber, Rimm et al. 1996; Chasan-Taber, Erickson et al. 2002). Other studies conducted square root transformation to attain a normal distribution of data (Wolf, Hunter et al. 1994).

Although Pearson’s product-moment and Spearman’s rank order correlation coefficients have been the most commonly used methods for measurement of reliability, Booth points out that both these methods measure association, not agreement (Booth, Owen et al. 1996). Also, Pearson’s correlations assume a bivariate situation, yet reliability is univariate, and Pearson’s correlations are not able to detect trends over time, which is important when assessing reliability (Patterson 2000). Alternative methods such as intraclass correlation coefficients and limits of agreement (for continuous data) (Bland and Altman 1986; Patterson 2000) and Kappa statistics (for categorical data) have been recommended for assessment of repeatability (Booth, Owen et al. 1996). Presentation of confidence intervals has also been recommended (Patterson 2000), although few studies presented these (Washburn, Smith et al. 1993; Pols, Peeters et al. 1997; Sobngwi, Mbanya et al. 2001).
The use of correlation coefficients alone to assess agreement between two measurement techniques may actually be misleading (Bland and Altman 1986). The scale or absolute values from the two measurement techniques may be quite different but still produce high correlation coefficients. To assess agreement, Bland and Altman developed a graphical depiction whereby the differences between the values obtained by each measurement technique are plotted for each individual against the mean of the two values (Bland and Altman 1986). Bland-Altman plots are rarely presented in physical activity questionnaire reliability and validity studies. When they are presented, in five studies, results were variable, possibly due to the imprecision of assessing physical activity (Lakka and Salonen 1992; Pols, Peeters et al. 1997; Pols, Peeters et al. 1997; Lowther, Mutrie et al. 1999; Liu, Woo et al. 2001). Some studies used the Bland-Altman approach to calculate repeatability coefficients, without presenting the graphs (De Abajo, Larriba et al. 2001; Sobngwi, Mbanya et al. 2001).

Lastly, the absence of statistically significant differences between means is not the same as being the 'same', and is an inappropriate measure of similarity. This was a common method but was usually supplemented by an analysis of correlation (Bonnefoy, Kostka et al. 1996; De Abajo, Larriba et al. 2001; Sobngwi, Mbanya et al. 2001). Whether correlation coefficients were statistically significant was more a reflection of sample size than degree of agreement.

The variability in intervals between administrations of the questionnaire, and variability of the period asked about, makes comparison between studies difficult. The interval between administrations of questionnaires in test-retest reliability studies varied from a few days (Lowther, Mutrie et al. 1999) to five years (Ropponen, Levalahti et al. 2001). While the appropriate interval between the administrations will be influenced by the period asked about in the questionnaire, too short an interval will mean that the participants are likely to remember what they answered on the previous administration. Too long a period will mean that the two administrations of the questionnaire are estimating physical activity from different times, producing different results. These differences are then interpreted incorrectly as measurement error (Patterson 2000).

Some studies have shown that the correlations vary depending on whether the time period asked about, is the same as in the initial administration, or different, (eg.
One study used a wide range of intervals between questionnaire administrations, which is a further source of misinterpretation of results (Roeykens, Rogers et al. 1998). In fact, the length of time between administrations in Roeykens' study was shown to affect the correlation of results. Therefore, it is important to have both administrations asking about the same period, if possible. Patterson suggests questionnaires ask about the previous 7 days and be repeated in the following 7 days to assess reliability (Patterson 2000). However, this approach risks recall of previous answers. It is important, however, to use a test-retest interval that is appropriate for the period asked about in the questionnaire. Therefore, an appropriate interval (11 days), between repeated administrations of the 3-month and the 2-week recall physical-activity questionnaires chosen for this thesis, was used in the reliability study reported in Chapters 3 and 4.

2.3.6.1.1 Participants

It is also difficult to compare results from the studies in the systematic review, when the sample size, sample selection methods, and participation rates are so variable. Again, this variability will affect the internal and external validity of the findings.

Most early studies used a moderate-sized sample (30-99 participants). Many studies used convenience sampling or volunteer participants (Ainsworth, Jacobs et al. 1993; Bonnefoy, Kostka et al. 1996; Sfakianos 1999; De Abajo, Larriba et al. 2001; Harada, Chiu et al. 2001; Sobngwi, Mbanya et al. 2001). Volunteers were also used in all the validity studies involving doubly labelled water.

Volunteer samples, used in validity and reliability studies, were often from highly educated or higher socio-economic backgrounds. Some involved University alumni (Ainsworth, Jacobs DR et al. 1993; Jacobs, Ainsworth et al. 1993; Chasan-Taber, Erickson et al. 2002). One study included members of an aerobics class (Lowther, Mutrie N et al. 1999), while two other studies included physical therapists or physical therapy alumni (Miller, Freedson et al. 1994; Roeykens, Rogers et al. 1998). Two studies included members of a religious denomination (Singh, Tonstad et al. 1996; Singh, Fraser et al. 2001); one was limited to monozygotic twins (Ropponen, Levalahti et al. 2001); while two studies included people with chronic medical conditions (Allison, Keller et al. 1998; Liu, Woo et al. 2001). These methods of
sample selection may limit generalisability unless the questionnaires were intended for use in that particular population. This was the case with several studies involving specific sub-population, particularly ethnic groups, when the questionnaires were to be used in a larger epidemiological study of that population (Kriska, Knowler et al. 1990). A few studies selected their study participants systematically or randomly from a population (Arroll, Jackson et al. 1991; Wolf, Hunter et al. 1994; Booth, Owen et al. 1996; Chasan-Taber, Rimm et al. 1996; Suleiman and Nelson 1997; Wareham, Jakes et al. 2002). The results of these studies have better generalisability to the general population, particularly if they have large sample sizes (Booth, Owen et al. 1996). Participation and study completion rates were often low, particularly when systematic or random selection was used (Suleiman and Nelson 1997), which may also limit generalisability of findings.

Several studies excluded outlying results in their analyses. These exclusions were carried out by removing from the analysis, all participants who claimed that their activity had changed significantly over the period of the study (Bonnefoy, Kostka et al. 1996; Pols, Peeters et al. 1996; Singh, Tonstad et al. 1996; Pols, Peeters et al. 1997). Alternatively, outlying results were excluded in a sensitivity analysis, with the assumption being that such wide variations must be due to real activity change, rather than measurement error (Wolf, Hunter et al. 1994). Although such exclusion may help to reduce the effect of real activity change that lowers agreement between measures, the exclusions are subjective and may have the effect of demonstrating greater reliability of the tool than is the case. The lack of standardisation of criteria for inclusion in analysis makes comparison of correlation coefficients and degrees of agreement difficult when assessing the reliability and validity of the physical activity questionnaires presented in the literature.

The reliability studies of the highest quality tended to use large, randomly selected population-based samples (Washburn, Smith et al. 1991; Washburn, Smith et al. 1993; Chasan-Taber, Rimm et al. 1996; Wareham, Jakes et al. 2002). In addition, appropriate statistical tests were used (Ainsworth, Leon et al. 1993; Booth, Owen et al. 1996; Pols, Peeters et al. 1997) with appropriate intervals between administration or multiple administrations (Ainsworth, Leon AS et al. 1993). The assessment of several questionnaires simultaneously produced particularly useful comparative
results in six studies, although the repeated recalling of activities may influence results (Rauh, Hovell et al. 1992; Ainsworth, Jacobs et al. 1993; Jacobs, Ainsworth et al. 1993; Wilbur, Holm et al. 1993; Pols, Peeters et al. 1996; Philippaerts and Lefevre 1998).

Only nine validity studies used doubly-labelled water as a ‘gold standard’ reference comparator in validity studies of physical activity questionnaires (Schultz, Harper et al. 1994; Schuit, Schouten et al. 1997; Philippaerts, Westerterp et al. 1999; Starling, Matthews et al. 1999; Bonnefoy, Normand et al. 2001; Leenders, Sherman et al. 2001; Staten, Taren et al. 2001; Conway, Irwin et al. 2002; Conway, Seale et al. 2002). Three studies used heart rate monitoring as their reference comparator (Singh, Tonstad et al. 1996; Sobngwi, Mbanya et al. 2001; Wareham, Jakes et al. 2002). Twenty-two studies used motion sensors (Table 7-3) and 27 studies used physical activity logs or diaries to assess the validity of the questionnaires (Table 7-2).

The quality of a few studies was increased by using several reference comparators (direct and indirect estimates of physical activity) (Jacobs, Ainsworth et al. 1993; Leenders, Sherman et al. 2000). Patterson states that although there is no clear gold standard of physical activity measurement, the use of several reference methods aids the assessment of the ‘construct’ validity of an instrument (Patterson 2000). Although the studies that used doubly-labelled water as their gold standard were of high quality due to their choice of reference comparison, the external validity of all these studies was limited by the small samples (13-35) of volunteers. Presumably, the small sample sizes are a reflection of the cost of the doubly-labelled water technique. Many of the doubly-labelled water validations of physical activity questionnaires reported low correlation or regression coefficients (Conway, Irwin et al. 2002; Conway, Seale et al. 2002).

2.3.6.2 Levels of Reliability and Validity considered adequate for Questionnaire use in Epidemiological Study

The levels of validity of questionnaires currently used in epidemiological study are discussed below. Of interest were the authors’ interpretations of what was considered adequate levels of reliability and validity.
Most Pearson’s and Spearman’s correlation coefficients of test-retest reliability ranged from 0.5 to 0.8 for physical activities (Ainsworth, Leon et al. 1993; Pols, Peeters et al. 1996; Pols, Peeters et al. 1997; Richardson, Ainsworth et al. 2001). Intraclass correlation coefficients for test-retest reliability also ranged from 0.5 to 0.8 (Philippaerts and Lefevre 1998; Roeykens, Rogers et al. 1998; Weller and Corey 1998). Although, these levels of reliability seem variable and often low, they were considered adequate by the authors of the studies. Correlation coefficients for physical activity questionnaires compared with physical activity diaries tended to range from 0.5 to 0.9 for total activity as well as other categories of activity such as leisure-time activity (Arroll, Jackson et al. 1991; Ainsworth, Leon et al. 1993; Pols, Peeters et al. 1997). Even lower correlation coefficients were accepted in the literature as adequate for physical activity questionnaires compared with motion-sensors (e.g. 0.2 to 0.6) (LaPorte, Montoye et al. 1985; Richardson, Leon et al. 1994; Lowther, Mutrie et al. 1999; Washburn and Ficker 1999). These ‘accepted’ levels of correlation are low in comparison to previous recommendations. The review of reliability of physical activity self-report instruments by Sallis and Saelens (2000), reports weighted correlation coefficients as ‘acceptable’ if greater than 0.6. Perhaps low correlation coefficients were accepted because of the real variability of physical activity over time as well as the imprecision of physical activity recall.

In summary, based on comments from these studies, reliability correlation coefficients of 0.6, are usually judged as acceptable for epidemiological study, while lower validity coefficients are judged as acceptable (0.5 to 0.6). Correlation coefficients as low as 0.2 or 0.3 may be acceptable when questionnaires are compared with heart rate monitoring, motion sensors, or doubly labelled water. Often, if correlation coefficients for motion sensors were very low or non-significant, they were ignored in the discussions and results from other comparators used to justify validity.

Low correlations of these direct objective techniques may be found because they measure different aspects of activity when compared with ‘recall’ techniques, and may be influenced by other factors, such as emotion when using heart rate monitoring for example. Alternatively, motion sensors may not detect certain activities, such as light activities or upper body activities, or may not discriminate between different intensities of activity. Doubly labelled water techniques are most useful when
comparing estimates of total energy expenditure, not components of activity, as this technique is unable to assess type, duration, or intensities of specific activities.

In addition, reliability and validity studies with adequate sample sizes of greater than 30 (ideally greater than 100), random or systematic selection of participants, and good participation rates have the best external validity. Furthermore, a variety of statistical techniques to measure similarity, which are appropriate for the potentially skewed nature of the data, improve the internal validity of results.

2.3.6.3 Physical Activity Questionnaires Validated in Primary Health Care

No physical activity questionnaires were found, that had been validated in a primary health care setting, amongst adults. There was only one questionnaire found that had been validated as a screening tool for adolescents in primary health care (Prochaska, Sallis et al. 2001). The validity of an instrument is context- and purpose-specific (Patterson 2000). Accordingly, a questionnaire found valid amongst one population, in one context, may not be valid in another.

The findings from this systematic review support the claims of Little and colleagues, who identified a major gap in the literature of physical activity assessment tools validated for use within primary health care (Little and Margetts 1996).

This gap may have contributed to the inability of some studies to detect a change in physical activity amongst sedentary adults in primary care. For example, a recent randomised controlled trial of a physical activity intervention in primary health care was able to demonstrate differences in physical fitness long-term amongst some groups (The Writing Group for the Activity Counseling Trial Research Group 2001). However, the study was unable to show the differences in regular physical activity between groups, using a previously validated questionnaire, although the increased fitness was likely to have resulted from increases in physical activity. This example demonstrates the difficulty in detecting change in usual activity level and the importance of having physical activity assessment tools validated for the appropriate target population and setting.

In other words, the external validity of physical-activity measurement tools is influenced by the choice of appropriate population for validation and the frame of
reference of exposure (Rennie and Wareham 1998). Physical activity questionnaires have often been validated amongst active (Lowther, Mutrie et al. 1999) or educated (Roeykens, Rogers et al. 1998) volunteers (Bonnefoy, Kostka et al. 1996) rather than less active or randomly selected subjects. However, less-active adults visiting their family doctor make up the population of interest for testing the effectiveness of most physical activity interventions in primary health care.

To be suitable for use in primary care, questionnaires need to be self-administered and filled out within 10-15 minutes. To allow comprehensive analysis, all major areas of activities should be asked about in the questionnaires, including leisure-time, occupational, and domestic activities (Wareham and Rennie 1998). The questionnaires must also comply with recommended data standards for physical activity measurement, which require recording type, frequency, intensity and duration of each activity (Bauman, Armstrong et al. 1999).

In addition, most of the questionnaires assessed in the studies reviewed are not suitable for use in the proposed study because they do not estimate total energy expenditure. Several questionnaires do not ask about occupational activity (Minnesota, Paffenbarger’s, 7-Day recall, Cardia, Godin, modified Baeke, Zutphen questionnaires). Nor do many questionnaires ask about domestic (Paffenbarger’s, 7-Day recall, Cardia, Godin, Baeke questionnaires), or transport activities (Minnesota, Paffenbargers, 7-Day recall, Cardia, Godin, Baeke, PASE, Yale questionnaires) (Sallis and Saelens 2000).

2.3.6.4 Conclusion

This thesis attempts to address these gaps in the literature by choosing at least one physical activity questionnaire suitable for use in primary health care and assessing reliability and validity in the primary care context amongst ‘less active’ adults. The physical activity questionnaire found to be most suitable from the studies reviewed was the 3-month recall physical-activity questionnaire developed for the Auckland Heart Study (AHS) (Appendix 1) (Jackson 1989). This questionnaire took aspects from previously validated questionnaires and assessed all types of activities. Therefore, total energy expenditure could be estimated. Although it had not been validated in primary care, it had been validated amongst a large randomly selected
sample (n=113) from a New Zealand population of adults (Arroll, Jackson et al. 1991). The validity study used a variety of statistical tests to assess validity, and levels of validity were good, when compared with an activity diary. Details of this questionnaire are included in Chapter 3.

A second questionnaire was chosen because it had been used in primary care in the original Green Script Study (GSS) (Appendix 2) (Swinburn, Walter et al. 1998). This questionnaire was short and could be adapted to estimate total energy expenditure. Although it had been used in primary care and found to be reliable, it had not been validated in that context. A description of the reliability and validity study of these questionnaires in primary care is described in Chapter 3. The AHS physical-activity questionnaire was subsequently chosen to assess physical activity in the Green Prescription cluster randomised controlled trial conducted for this thesis.

2.4 Systematic Review of Physical Activity Intervention Trials in Primary Health Care

2.4.1 Introduction

Previous systematic reviews, evaluating the effectiveness of community-based lifestyle interventions, including physical activity interventions, for primary disease prevention have identified few well-designed studies (Hillsdon and Thorogood 1996; Ebrahim and Davey-Smith 1999). Where benefit has been shown long-term, this has been minimal compared with pharmacological interventions (Ebrahim and Davey-Smith 1999). The aim of this section of the literature review was to focus on physical activity counselling interventions in primary health care. In particular, the review addresses the following questions:

- What is the evidence for the effectiveness of physical activity counselling in primary health care for primary disease prevention?

- Have previous systematic reviews answered this question adequately? (What was the validity and quality of their research questions, methodology, and evaluation?)
• Are any additional studies available that add to the findings of previous systematic reviews?

2.4.2 Methods

2.4.2.1 Search Strategy

A literature search was carried out using the databases and key words outlined in Table 2-3. Researchers in the area of physical activity promotion were asked about other on-going or completed trials. References of individual studies and systematic reviews were examined to locate any further studies.

Table 2-3 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Trials and Systematic Reviews of Physical Activity Counselling in Primary Health Care

<table>
<thead>
<tr>
<th>Databases</th>
<th>Physical activity subject headings</th>
<th>Health promotion subject headings</th>
<th>Primary health care subject headings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embase</td>
<td>Physical fitness</td>
<td>Health promotion</td>
<td>General practice</td>
</tr>
<tr>
<td>Psychlit (now Psycinfo)</td>
<td>Leisure activities</td>
<td></td>
<td>Family practice</td>
</tr>
<tr>
<td>Sportdiscuss</td>
<td>Exertion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best Evidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cinahl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cochrane database of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>systematic reviews</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.4.2.2 Inclusion Criteria

Systematic reviews of controlled trials of lifestyle interventions including physical activity advice in health care settings directed towards adults were included in the analysis. Inclusion criteria for the search for any additional, or more recent, primary studies were determined separately. These criteria include controlled trials of lifestyle interventions for primary prevention in primary health care, involving adults, with
physical activity as a component of the intervention as well as an outcome measure. Only papers published in English up until the end of 2002 were examined.

2.4.2.3 Data Abstraction

Titles and abstracts were reviewed by the author and the full papers of those likely to fulfill inclusion criteria, retrieved and read. Those fulfilling inclusion criteria were examined and their quality assessed by the author. Systemic reviews were examined for several characteristics. These characteristics included review question(s), inclusion and exclusion criteria, search strategy employed, and data abstraction methods. The reviews were also examined for assessment of heterogeneity and appropriateness of analysis, methods used for quality assessment of each trial and results of those assessments, findings from short-term and long-term trials, overall conclusions, and recommendations. Information about study design, nature of the intervention, outcome measures and results were abstracted from primary studies.

2.4.2.4 Critical Appraisal

Results of previous systematic reviews were critically appraised with respect to the quality of the review. Assessment of quality was based on criteria from the literature (Guyatt and Drummond 2002). In particular, each review was assessed by considering the following questions:

1. Did the review address a sensible clinical question, using appropriate inclusion criteria?

2. Was the search strategy detailed and exhaustive?

3. Was data abstraction appropriate? (e.g. how many reviewers?)

4. Were the studies assessed for quality?

5. Was heterogeneity assessed?

6. Were results interpreted appropriately?

Results from primary studies identified were critically appraised with respect to their quality, based on a validated system used in the literature (Schulz, Chalmers et al.)
1995; Khan, Daya et al. 1996). This system is consistent with the Jadad system of grading the reporting of controlled trials on the basis of whether they were described as randomised and double blind, and whether there was a description of withdrawals and drop-outs (Jadad, Moore et al. 1996). The specific criteria used in this review of the literature included the following:

1. Was randomisation adequate? (concealment of treatment allocation, generation of allocation sequence)

2. Were all enrolled participants included in the analysis according to allocation of intervention? (intention-to-treat analysis)

3. Was blinding adequate or appropriate?

2.4.3 Systematic Reviews Identified and Critical Appraisal

Eight systematic reviews were identified (Ashenden, Silagy et al. 1997; Ebrahim and Smith 1997; Eaton and Menard 1998; Simons-Morton, Calfas et al. 1998; Eakin, Glasgow et al. 2000; Lawlor and Hanratty 2001; Eden, Orleans et al. 2002; Petrella and Lattanzio 2002). The description and critical appraisal of each review are presented in Table 2-4.

The first systematic review of randomised trials of lifestyle advice in primary care settings concluded that both brief and intensive advice were effective short- and long-term (Ashenden, Silagy et al. 1997). However, some of the trials included in this review were of multicomponent interventions so the effective component could not be isolated. The positive health outcomes of some trials could not necessarily be attributed to physical activity advice, as several interventions also included dietary and other advice (Anonymous 1994; Graham-Clarke and Oldenburg 1994). In addition, the review included studies conducted solely with patients with previous cardiovascular disease (Cuppes and McKnight 1994) or at high cardiovascular risk (Graham-Clarke and Oldenburg 1994). Some studies that were included in the review were not randomised trials and did not analyse according to intervention allocation (Logsdon, Lazaro et al. 1989). Allocation was not concealed and assessment was not blinded in the studies. Accommodation, and regression to the mean, of blood pressure was not taken into account in the design of one trial (Anonymous 1994; Anonymous 1995; Khan, Daya et al. 1996). This system is consistent with the Jadad system of grading the reporting of controlled trials on the basis of whether they were described as randomised and double blind, and whether there was a description of withdrawals and drop-outs (Jadad, Moore et al. 1996).
Accordingly, the positive findings of this systematic review must be taken with caution.

The systematic review of Ebrahim and Smith (1997) was not explicit about the review question, data extraction or quality assessment. The authors reviewed only randomised controlled trials, but included workplace trials as well as primary care trials (Ebrahim and Smith 1997). They also included all lifestyle interventions and used health measures as outcomes rather than physical activity, so again, relevance of the findings is limited. Positive health outcomes were found by the meta-analysis, including change in blood pressure (even after exclusion of trials that incorporated pharmaceutical treatment), cholesterol levels, and smoking rates. However, the authors still concluded that the effectiveness of health education to prevent heart disease is in doubt because no difference in mortality could be detected. As the maximum follow-up period was five years, and many studies included patients at low risk in primary prevention trials, this was not surprising. It was evident, however, that the magnitude of change was a lot smaller than with pharmaceutical interventions. Cost-effectiveness studies and longer follow-up of trials may help assess the comparative worth of such interventions more appropriately.

The systematic review by Simons-Morton and colleagues (1997) included studies with physical activity as an outcome measure, but still included multi-component intervention trials in all clinical settings (Simons-Morton, Calfas et al. 1998). Therefore, the conclusion that physical activity can be increased by primary care interventions still does not answer whether physical activity counselling is effective in primary care. The authors also note the many methodological problems with existing trials. They recommend that more randomised trials be conducted that achieve high follow-up rates and include minority groups and both genders. They also stress the need for cost-effectiveness analyses, and research into the dissemination and implementation of the initiatives (ibid).

The systematic review by Eaton and Menard (1998) was very thorough and limited the review to studies in primary care (Eaton and Menard 1998). The authors found many studies were of poor quality and attained low follow-up rates. Therefore, although at least half of the trials had positive results, both short and long-term, the
authors conclude that scientific evidence of efficacy is limited and any effects are modest (ibid).

The systematic review by Eakin and colleagues (2000) focussed on investigating which strategies are most effective in primary health care to improve physical activity levels (Eakin, Glasgow et al. 2000). Although the authors found that particular strategies were effective for short-term gain, the opposite was sometimes the case for long-term gains. Overall, it seemed that interventions that used behavioural approaches were generally as effective as those that did not, and there was no obvious difference in the effectiveness between interventions delivered by clinicians compared with those delivered by non-clinicians. However, the diversity of study designs precluded any comprehensive comparison of strategies. The authors also scored the studies with respect to whether they fulfilled the criteria of the RE-AIM model (reach, efficacy, adoption, implementation and maintenance). However, few studies assessed reach, adoption, implementation or (long-term) maintenance (ibid).

The systematic review by Lawlor and Hanratty (2001) focussed on primary care physical activity interventions within the usual consultation. The review was of good quality and assessed the quality of studies according to validated criteria (Schulz, Chalmers et al. 1995). These criteria included adequacy of concealment of allocation of randomisation (and whether this was stated or not), the randomisation process, exclusions after randomisation, and double blinding. The authors concluded that effectiveness of physical activity promotion in primary care is “modest at best” (Lawlor and Hanratty 2001).

The next systematic review included studies of physical activity counselling by primary or secondary care physicians, that assessed physical activity, fitness, or health outcomes (Petrella and Lattanzio 2002). This review identified 13 randomised and non-randomised controlled trials. Most were short-term and many involved programmes addressing other risk factors in addition to physical inactivity. Many interventions used behavioural theories, such as the transtheoretical model, or social cognitive theory. However, those without a behavioural theory base were just as likely to show positive outcomes. The authors of this review found that while there was evidence of short-term physical-activity behaviour change, evidence for long-term
change and cost-effectiveness was limited. More investigation was recommended to clarify this and to ascertain the generalisability of the findings (ibid).

The most recent systematic review assessed physical activity counselling by clinicians in primary health care (Eden, Orleans et al. 2002). The inclusion criteria differed from previous reviews in that only studies that involved physical activity counselling which was at least in part delivered by the primary care clinician (doctor or nurse) were included. The review also aimed to assess which interventions were most effective and what risk of harm was associated with the interventions. Studies were ranked on the basis of the US Preventive Services Task Force current methods for assessing quality of studies and admissible evidence for making recommendations (Harris, Helfand et al. 2001). Seven randomised controlled trials and one non-randomised controlled trial met the inclusion criteria, representing 9,054 adult study participants. Only one trial assessed harm (The Writing Group for the Activity Counseling Trial Research Group 2001). This study did not have a 'usual care' control comparison, so it is difficult to assess the risk of harm due to physical activity counselling. The review also assessed the efficacy of interventions compared with usual care controls (six trials), separately from interventions that were compared with other interventions (two trials). The results were mixed and often associated with poor adherence, although it seemed that in certain circumstances, physical activity counselling was effective. There were also short-term trials and methodological flaws with many of the studies. Because of these methodological difficulties, the authors of the review were unable to assess the efficacy or effectiveness of the interventions to be able to generalise to primary health care.
<table>
<thead>
<tr>
<th>Systematic review</th>
<th>Search Strategy</th>
<th>Results</th>
<th>Critical appraisal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study questions /inclusion &amp; exclusion criteria</strong></td>
<td><strong>Included studies up until May 1995</strong>&lt;br&gt;Medline, Psychlit, Sociofile, Cinahl, EMBASE and Drug databases searched&lt;br&gt;MeSH** terms to retrieve RCTs according to Cochrane search strategy, terms relating to general practice (eg &quot;family practice&quot;, &quot;primary health care&quot;, &quot;physician&quot;, and terms relating to &quot;exercise&quot;)&lt;br&gt;Six studies identified</td>
<td><strong>Short-term</strong>* significant increases in physical activity in 1/1 trial.&lt;br&gt;Long-term** significant increases in physical activity in 3/3 trials but 1 not ITT**. (the other 2/6 trials had no exercise outcomes measured.)&lt;br&gt;BP, lipid or weight changes significant in 3/4 trials&lt;br&gt;Both brief and intensive effective but not possible to isolate effective component as most interventions multifactorial (eg. Diet and exercise)&lt;br&gt;Quality checked by adequacy of randomisation, analysis of all randomised, and blinding. Two studies were of good quality. The other 4 studies had at least one criterion compromised</td>
<td><strong>Study question</strong> was explicit and consistent with inclusion criteria although included two studies with no exercise outcome measures and 1 trial of angina patients only&lt;br&gt;<strong>Search strategy</strong> was quite extensive and appropriate. No hand search&lt;br&gt;<strong>Data abstraction</strong> by one reviewer and checked by two others&lt;br&gt;<strong>Quality of studies</strong> were assessed appropriately&lt;br&gt;<strong>Heterogeneity</strong> outcomes, duration, multiple intervention and who delivered it acknowledged so results not pooled&lt;br&gt;<strong>Interpretation of results</strong> was appropriate and limited by paucity, heterogeneity, and quality of trials</td>
</tr>
<tr>
<td><strong>(Ashenden, Silagy et al. 1997)</strong>&lt;br&gt;Hypotheses: 1. Provision of advice is more effective in eliciting lifestyle-related behaviour change than providing no advice. 2. That providing intensive advice is more effective than providing brief advice.&lt;br&gt;English language reports of randomised trials, investigating lifestyle advice in a general practice setting. Brief advice = single occasion, all others = intensive. Primary and secondary prevention studies included.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(Ebrahim and Smith 1997)</strong>&lt;br&gt;Aim: To assess the effectiveness of multiple risk factor intervention in reducing cardiovascular risk</td>
<td>Search strategy not reported&lt;br&gt;Fourteen studies identified, four of which included exercise advice and were based in</td>
<td>Risk factor reductions in all but two studies. Meta-analysis on 11 studies (3 removed to reduce heterogeneity).</td>
<td></td>
</tr>
</tbody>
</table>
factors, total mortality and CHD mortality
Randomised controlled trials of lifestyle interventions in workforce and in primary care and followed for at least 6 months.

Random effects analysis took study variation into account. Pooled results showed significant reductions in blood pressure by 4.2/2.7 mmHg (even when studies using pharmacological treatment removed, differences remained significant 2.3/1.1 mmHg), rates of smoking, and cholesterol, but not for mortality.
Concluded effectiveness of health education to prevent heart disease is in doubt.
Quality not reported for individual studies

3 (Simons-Morton, Calfas et al. 1998)
1. Can interventions delivered in health care settings increase the level of physical activity or cardiorespiratory fitness in patients without disease (primary prevention)? 2. In patients with existing CVD (secondary prevention)? 3. What types of intervention are most effective? 4. What additional research is needed? 5. What practice recommendations can be made?
For primary prevention: Included
For primary prevention trials: Index medicus computerised database for past 30 years.
MeSH terms included physical activity and exercise terms as well as “primary care providers”, “physicians”, “nurses”, “physician assistants”
Twelve studies identified for primary prevention of which 10 were initiated in primary care
Short-term results 6/8 studies showed significant increases in PA$^d$. Long-term results in 3/7 studies showed significant increases in PA. Decay of effect over time.
“we conclude that changes in physical activity can be initiated by intervention in a primary care setting”
Quality: 5/12 not randomised, 1 ignored randomisation in analysis, and follow-up rates sometimes not reported or low (40-60%). No standardization of PA definitions or outcome measures, sometimes without reliability/validity data.

Search strategy was not reported
Data abstraction not reported
Quality of studies commented on, generally, in discussion
Heterogeneity of participants and use of pharmacological treatments acknowledged and adjusted for.
Interpretation of results was not consistent with results or discussion, which showed significant risk factor improvement and discussed possibility of latency of effects on CHD mortality.

Study question and inclusion criteria clearly stated and justified
Search strategy limited to one database
Data abstraction described which variables were abstracted but did not state who abstracted or who checked
Quality of studies commented on in discussion and recommendations to improve methods made
Heterogeneity of methodologies extensively discussed in discussion. Appropriately, no pooling of results was done
<table>
<thead>
<tr>
<th>Controlled trials; interventions delivered in clinical practice settings, exercise or fitness outcome measures, reported in English</th>
<th>Analyses by gender not done, racial and socio-economic make-up usually not reported.</th>
<th>Interpretation of results was consistent with results. Recommendations were made to improve quality of studies to answer more specific questions about the effectiveness and increase generalisability</th>
</tr>
</thead>
</table>
| **4 (Eaton and Menard 1998)**  
1. What is the quality of the evidence that physical activity counselling in primary care office practice is efficacious?  
2. If efficacious, how generalisable are these results to normal primary care office practice?  
Studies must include assigning participants to control or intervention status, interventions performed in doctor’s office practice, and exercise had to be assessed at least 4/52 after intervention and interpretable as a dichotomous variable. | Computerized searches of: Medline, Dialog (R) of Dissertation abstracts, Sci Li Reference from 1961 to 1997. MeSH terms included “exercise”, “physical fitness”, “trials”, “met-analysis”, and “outcome assessment”  
Search limited to English language. Searched references of papers and reviews. Content experts contacted about other published and unpublished papers.  
Eight studies were identified | 4/4 studies showed short-term positive effect. Long-term results 2/4 studies showed positive effect, but 1 of these was not ITT  
The scientific evidence for the efficacy of PA promotion in primary care is modest at best  
Quality of studies was mixed: assessed for recruitment rate, selection and confounding bias, measurement error, competing interventions, ITT and generalisability. |
| **5 (Eakin, Glasgow et al. 2000)**  
What strategies are practical and effective to use in family practice to enhance levels of patient physical activity? | Databases searched: Medline (1980-1998), psychological abstracts, ERIC and HealthStar; Website for the Journal of Family Practice, bibliographies of selected studies and previous | Study question was explicit and consistent with inclusion criteria  
Search strategy was extensive and appropriate but no hand search.  
Data abstraction by two reviewers independently and described in detail  
Quality of studies was assessed appropriately. Recommendations about validated PA measures, made  
Heterogeneity of methodologies acknowledged and results not pooled but odds ratios compared  
Interpretation of results was appropriate and limited by paucity, heterogeneity, and quality of trials |
| --- | --- | Study question was explicit and consistent with inclusion criteria  
Search strategy was extensive and appropriate but no hand search.  
Data abstraction by single reviewer |
Inclusion criteria were: controlled trial, intervention initiated in primary care, PA outcome measure. Excluded studies solely on populations with CVD.

<table>
<thead>
<tr>
<th>Key words: physical activity, randomised controlled trials, physical activity counselling, primary care or medical office, exercise interventions, health promotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search limited to English language</td>
</tr>
<tr>
<td>15 studies identified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quality of studies was assessed in terms of internal and external validity using RE-AIM model. (Not clear if rating system has been validated. Did not assess all aspects of RE-AIM. Also study design scores did not add up)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneity of methodologies acknowledged and results not pooled</td>
</tr>
</tbody>
</table>

**Interpretation of results:** As methodologies are so different, conclusions drawn about intervention characteristics most likely to be effective may be confounded by study design.

<table>
<thead>
<tr>
<th>Study question</th>
<th>not explicit, although an aim and inclusion criteria are stated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search strategy</td>
<td>was extensive and appropriate although no hand searching</td>
</tr>
<tr>
<td>Data abstraction</td>
<td>independently abstracted by two reviewers using structured format</td>
</tr>
</tbody>
</table>

**Quality of studies was assessed according to randomisation, concealment, selection bias, blinding and ITT** (Schulz, Chalmers et al. 1995)

<table>
<thead>
<tr>
<th>6 (Lawlor and Hanratty 2001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aims to determine the effect of advice given in routine primary care consultations on levels of physical activity</td>
</tr>
<tr>
<td>Controlled trials that assessed the effectiveness of PA advice within routine consultation in primary care by health professional, with PA as outcome measure. Studies in any language</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Key words: exercise therapy, physical fitness, physical activity, jogging, walking, bicycling, swimming, physicians family, family practice, primary health care, general practitioner, primary care, general practice, health promotion, health behaviour, health education, reverse for longterm. Brief counselling (3-10 mins) is as effective as lengthy counselling (&gt;15 mins). Tailoring advice to readiness more effective than not for short but not long term. Written material seems to help. No clear advantage of follow-up or theory based interventions. Who delivered intervention not important in short term. Interventions targeting &gt;50 year-old patients less effective short-term. Method scores ranged from 4 to 9 out of 10, with implementation rates rarely reported and some analyses not ITT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Potential bias: Both randomised trials clustered but clustering not taken into account in analysis, lack of concealment, and lack of randomisation of rest of trials. Only 1 trial used objective measure (accelerometer) in subgroup 0-8weeks follow-up: effective in 4/6 studies; &gt;4months follow-up: effective in ½ studies (both RCTs' negative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Current evidence suggests that advice in routine primary care consultations is not effective at</td>
</tr>
</tbody>
</table>
7 (Petrella and Lattanzio 2002)

Objectives: 1. To establish studies of physical activity counselling or exercise prescriptions in family practice; 2. To determine whether they demonstrated improved physical activity, fitness or health outcomes; 3. To examine study strengths and weaknesses; 4. To identify intervention strategies, instruments and outcome measures used; 5. To identify areas for future research.

Inclusion criteria: RCTs or controlled trials of interventions delivered by physicians in primary, secondary or tertiary health care settings, that measured physical activity or cardiorespiratory fitness at follow-up.

<table>
<thead>
<tr>
<th>Lifestyle advice, behaviour change, brief intervention</th>
<th>Producing sustained increases in physical activity levels. More trials needed as many studies poor quality.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eight studies identified</td>
<td>Exercise prescriptions are effective in increasing physical activity levels. Most interventions were part of multiple risk factor programs. Long-term physical activity results are not available. Most interventions aimed at changing behaviour. Time and skills required appear to be main barriers to interventions. Cost effectiveness remains unknown. More investigation is needed into long-term effects and whether results are generalisable. 6/13 studies were RCTs. 7/13 were quasi-experimental. 5/13 studies assessed long-term effectiveness of at least 12 months. Only one of these was reported as finding positive PA effects at 12 months, although one study was not yet complete.</td>
</tr>
</tbody>
</table>

**Heterogeneity of methodologies acknowledged so results not pooled**

**Interpretation of results** was appropriate and limited by paucity and poor quality of many trials

**Study question** Objectives were explicit and consistent with inclusion criteria

**Search strategy** used only two electronic databases and reference searching. One unfinished study was included in the review (ACT trial).

**Data abstraction** was very systematic and well presented including: design, frequency, duration, intensity of exercise, intervention type, measurement methods and outcome data. Did not specify who scrutinized

**Quality of studies** Type of trial specified, but did not discuss randomisation and analysis adequacy of studies, nor blinding.

**Heterogeneity** of methodologies was acknowledged making comparison difficult

**Interpretation of results** appropriate
8 (Eden, Orleans et al. 2002)

1. Do adults counselled by primary care clinicians improve or maintain physical activity behaviour?
   
   Included controlled trials of PA counselling in primary care. Some components of intervention delivered by GP or practice nurse and reported behavioural outcomes.

   Included studies from 1994 to March 2002
   
   Cochrane database of systematic review and registry of controlled trials, Medline and HealthStar databases, and searched reference lists.

   1. Physical fitness
   2. Counselling
   3. Patient education
   4. Health education

   Eight studies identified

   Two good quality and six fair quality studies (USPSTF criteria). (Harris, Helfand et al. 2001)

   Short-term results: small significant increases in PA were seen in 3/4 studies reported.

   Long-term results: small significant increases seen in 3/6 studies reported

   Evidence inconclusive due to methodological issues and poor patient adherence. Too few studies to discern any relationship between components of intervention and efficacy

| Study question | was explicit and consistent with inclusion criteria |
| Search strategy | was quite extensive and appropriate but no hand search |
| Data abstraction | by one reviewer; quality assessed by 2 reviewers |
| Quality of studies | were assessed appropriately |
| Heterogeneity | of methodologies, outcomes and duration of follow-up acknowledged so results not pooled |
| Interpretation of results | was appropriate |

---

*Short-term refers to less than 6 months follow-up. # Long-term refers to more than 6 months follow-up.

**MeSH refers to medical subject headings standardised for each database.

a CHD refers to coronary heart disease. b ITT refers to intention-to-treat. c CVD refers to cardiovascular disease. d PA refers to physical activity or fitness

4 RE-AIM refers to Reach (the percentage and representativeness of patients), Efficacy, Adoption (percentage and representativeness of settings that are willing to adopt intervention), Implementation (extent to which intervention is delivered as intended), and Maintenance (individual and system level).

f RCT refers to randomised controlled trial
Identification and Critical Appraisal of Additional Physical Activity Intervention Trials in Primary Health Care

The literature search using the strategy described in Table 2-3 revealed over 900 studies, of which 31 primary studies fulfilled the inclusion criteria. All studies except four (Dubbert, Cooper et al. 2002; Green, McAfee et al. 2002; Hillsdon, Thorogood et al. 2002; Pinto, Friedman et al. 2002) had been reviewed by previous systematic reviews. A summary of the design, findings and critical appraisal of these studies are presented in Table 2-5.

The first study (Hillsdon, Thorogood et al. 2002) found that physical activity advice in primary health care was not effective. However, the Zelen technique of randomising those who completed a physical activity questionnaire, prior to consent, contributed to the poor ‘follow-up’ rate. This was because those who did not consent, as well as the large number who were lost to follow-up, were all included in the final analysis. As such only 32% of the ‘brief negotiation’ intervention participants and 33% of the ‘direct advice’ participants, who had been randomised, finished the study. Any increases in activity were well diluted, using an intention-to-treat analysis. This is demonstrated by the large difference in effect found when only those who completed the trial were included in the analysis. Although intention-to-treat analysis is the most appropriate method, trials that have 32-33% follow-up rates have questionable validity and risk a type two error (missing an effect that exists).

The second study included a computer-based telephone intervention, but recruited from a multi-site medical practice, presumably including family practitioners (Pinto, Friedman et al. 2002). The authors reported positive effects at three months. However, this outcome was the proportion that had met recommended levels of physical activity. There was no difference in energy expenditure using the appropriate intention-to-treat analysis. There was also no significant difference in any physical activity outcome at six months follow-up. Although this study had good follow-up rates (>80%), only 35% of eligible responders participated in the study. This study differs from previous ones in that none of the counselling was delivered in the practice and may therefore have limited relevance to physical activity counselling by clinicians in primary care practices.
The next study aimed to assess the effect of different levels of follow-up support in the STEPS physical activity intervention to increase walking amongst the elderly (Dubbert, Cooper et al. 2002). All participants received brief physical activity advice from the nurse. The control group received no further follow-up. The two intervention arms received either 20 follow-up telephone calls from the nurse or 10 from the nurse and 10 automated motivational calls. All groups increased walking over the 12 months. There was no significant difference between the nurse-only intervention group and the control group. However, the intervention that involved automated motivational calls in addition to nurse calls did produce positive self-reported physical activity outcomes compared with those of the control group. No other significant differences in health-related variables or fitness were detected between the groups. The study was well designed and used a modified intention-to-treat technique for analysis (although those who did not attend follow-up (15%) were not included in the analysis, which may have overestimated the effect) (Dubbert, Cooper et al. 2002). The lack of a true control group (that received no physical activity advice) limited the ability of the study to demonstrate physical activity, fitness, and health effects of the physical activity programme (although all groups improved fitness from baseline). It was interesting to note that the inclusion of 10 ‘automated’ follow-up telephone calls interspersed with 10 personalised follow-up calls achieved a positive outcome, whereas 20 personalised follow-up calls did not.

In the fourth study, 62% of enrolled intervention patients refused the intervention (Green, McAfee et al. 2002). Although the follow-up rate was very good (81% overall, including 80% of patients randomised to intervention), most intervention participants had not received the intervention. Therefore, treatment effect was likely to be diluted, as the increase in treatment effect with analysis of those who received the intervention only, demonstrated. Statistically significant increases in physical activity scores were found at follow-up amongst intervention patients compared with control patients adjusting for baseline (p<0.05). However, the baseline physical-activity levels were not balanced and were quite a bit lower in the control group than the intervention group. There was no significant difference when ‘change in activity score’ was the efficacy variable. (A statistically significant difference was found only when those who did not receive the intervention were excluded.) A strength of this study was the blinding of activity assessment.
Table 2-5 Physical Activity Counselling Trials based in Primary Care that have not previously been Reviewed by Systematic Reviews

<table>
<thead>
<tr>
<th>Author, date</th>
<th>Intervention and Physical Activity Outcome measures</th>
<th>Results</th>
<th>Critical appraisal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> (Hillsdon, Thorogood et al. 2002)</td>
<td>Intervention delivered by health promotion specialist. 30 minutes of direct advice based on ‘health belief’ model (DA); or brief negotiation based on motivational interviewing (BN), both with follow-up telephone calls. Control = no intervention</td>
<td>73% response to original questionnaire, 45% of these “insufficiently active”, 13% ineligible, 1,658 randomised. 55% BN, 52% DA and 71% controls consented to study. Follow-up rates: BN: 32% of randomised, 59% of consented; DA: 33% of randomised, 62% of consented; Control: 56% of randomised, 80% of consented.</td>
<td><strong>Randomisation:</strong> by household (cluster) prior to consent so not all randomised were enrolled (Zelen method). Allocation of intervention type in opaque envelope opened after consent at baseline, controls consented at 11 months after baseline, so potential for selection bias - participation rates varied between control and intervention. However, baseline results were balanced. Method of selection of sub-sample for motion-sensor measurements was not clear. <strong>Analysis:</strong> Sample size calculated. <strong>ITT</strong> using imputed control values for all randomised for int vs control but only of enrolled for BN vs DA. Also did sub-analysis of those who completed trial. Clustering not accounted for in analysis because claimed that clustering had little effect of standard errors (figures not reported). Large difference in follow-up rates between groups. Low follow-up rates of randomised</td>
</tr>
<tr>
<td><strong>Design, study population (n)</strong></td>
<td>Self-reported PA at 12 months from 28-day logbooks based on modified Minnesota leisure-time Activity Questionnaire. (Sub sample of 40 asked to wear Tri-Trac motion sensors to validate self-report) Also, change in blood pressure and body mass index (BMI) in BN vs DA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up 11/12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

65
| 2 | (Pinto, Friedman et al. 2002) | Intervention: telephone-linked physical activity counselling system (TLC-PA) promoting moderate intensity PA. All calls initiated by users and counselling activated by key pads. Based on transtheoretical, social cognitive, and decision-making theories. | Random-selection of contacts, 2884 screened, 1738 completed interview, 609 eligible (35%), 298 enrolled (49% of eligible). 249 (83%) completed 3-month follow-up, 243 (81%) completed 6-month follow-up. ITT analysis at 3/12 showed that the TLC-PA group was more likely to meet recommendations for moderate or vigorous physical activity. EE only significantly higher when only those who completed study were included in analysis. Results not maintained at 6/12. | (32% BN and 33%DA) with ITT diluted treatment effect. |
|   | RCT of sedentary adults >25 years of age with a ‘sub-optimal diet’ from a multi-site medical practice. Excluded people with medical conditions such as Alzheimers, recent heart or renal disease. (Sedentary refers to <5 x 30 mins moderate or <3 x 20 mins vigorous PA per week). Comparison group received an automated intervention promoting health eating (TLC-Eat) (n=298) | Outcome measures: moderate intensity physical activity energy expenditure using 7-Day Physical Activity Record (Blair, Haskell et al. 1985). |  |

<p>| 3 | (Dubbert, Cooper et al. 2002) | STEPS program. All three arms received baseline brief advice from nurse about increasing walking (aim: 3x20 minutes per week). Two intervention arms with different types of follow-up support: 1. 20 nurse-initiated calls; 2. 10 nurse-initiated calls plus 10 motivational calls through automated phone system. | 53% (253/475) of invited patients responded. 83% of those were eligible. 85% (181/212) of randomised patients completed 12 month follow-up. All groups improved walking and fitness from baseline. The group receiving the nurse plus automated calls walked significantly more frequently than those with no phone contacts. No other significant | Randomisation: Individual randomisation after screening for eligibility. |
|   | RCT of 60-80 year-old primary care patients (mostly male). Inclusion: stable health, noninstitutionalized, independent in activities of daily living, willing to increase activity. Excluded: if already doing 20 minutes of walking at least 2 days per week. (n=212) | Outcome measures: self-reported | Analysis: Used ‘modified ITT’ approach (all randomised patients excluding those who withdrew, or became ineligible due to deteriorating health) – this method risks selection bias for final analysis. Self-report PA measures validated by subsample of 51 that |</p>
<table>
<thead>
<tr>
<th>Three arms including control of brief advice with no follow-up. 12 months follow-up but 10-month measures were used as final outcomes.</th>
<th>walking from activity diaries, measured at 6 months (initiation) and 10 months (maintenance). Tranquilisation of results from activity diaries with PA interviews and 2 administrations of 7-day PAR questionnaire at baseline and 6 months. Objective measures of endurance and mobility also used. Quality of life (SF-36 and falls, injuries and healthcare utilisation.</th>
<th>differences between groups. No change in quality of life (QOL). Some correlation between physical QOL measures (but not mental health) with increased walking. No evidence of increased adverse events with STEPS programme.</th>
<th>wore accelerometers. Good follow-up rate (85%)</th>
<th><strong>Blinding:</strong> Data collector was blinded to allocation of intervention. Also, nurse delivering intervention was blinded to adherence to walking and other self-report follow-up data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4</strong> (Green, McAfee et al. 2002)</td>
<td>RCT 18-65 year old patients who exercised &lt;15 minutes daily and wanted to increase their activity recruited from mailed health risk assessment. Excluded patients with pre-existing CVD or diabetes. From one family physicians practice. (n=316) Baseline and 6-month follow-up assessments by telephone.</td>
<td>Intervention delivered by telephone (3 x monthly phone calls each 20-30mins), using motivational interviewing from health educator. Used mean of 5.6 calls. Main outcome measures physical activity score (PACE) administered by telephone. Proportion reaching recommended levels of moderate or vigorous PA; Moderate activity energy expenditure; proportion in each stage of change.</td>
<td>54% response rate to original recruiting questionnaire, of those, 51% eligible, 88% consented and randomised (316), but only 38% of intervention patients received intervention (rest declined), but 80% follow-up of randomised intervention patients and 82% control. Significantly increased self-reported PA amongst intervention patients compared with control patients when baseline levels were controlled for, although difference in change in PACE score did not reach significance (ITT). Even larger gains with post-hoc analysis of those who received intervention. No ‘dose’ effect of counselling.</td>
<td><strong>Randomisation:</strong> appropriate using random-number generator. Baseline variables reasonably balanced between groups, although PACE score slightly higher in intervention group. <strong>Analysis:</strong> ITT analysis still showed significant effect despite intervention delivery rate was 38%. But this was only in follow-up PACE levels adjusting for baseline. ‘Change in PACE’ level was not significant. (Was significant if analysed only those who received intervention) <strong>Blinding:</strong> Assessor was blinded to allocation of intervention.</td>
</tr>
</tbody>
</table>

*RCT refers to randomised controlled trial *ITT refers to ‘Intention-to-treat analysis*
2.4.5 Discussion of the Review

The evidence for the effectiveness of physical activity counselling interventions in primary health care is inconclusive. Although there are indications that such interventions are effective short-term, studies of at least 12 months are needed to establish long-term effectiveness. In addition, most studies had methodological limitations. There is also a lack of evaluation of such interventions amongst lower socio-economic and minority groups, and few studies assessed 'harm' or cost effectiveness.

Overall, most of the systematic reviews were of high quality. The specific conclusions drawn by each systematic review appeared to depend on the inclusion criteria set for the review. The final systematic review was of very good quality and had inclusion criteria most relevant to the trial presented in this thesis (Eden, Orleans et al. 2002). The conclusion of this review was that evidence is inadequate to conclude whether physical activity counselling by health professionals in primary health care is effective. There is no strong evidence to suggest that physical activity interventions delivered by health professionals are any more effective than those delivered by other professionals in primary care (Eakin, Glasgow et al. 2000). While one review suggested that behaviourally-based interventions and continued follow-up may help the effectiveness of interventions, (Simons-Morton, Calfas et al. 1998) another review found no evidence to support this conclusion (Eakin, Glasgow et al. 2000). However, the lack of standardised protocols of the studies makes it difficult to assess the effectiveness of certain aspects of the intervention.

The four additional primary care studies that were reviewed above did not change the conclusions reached by previous systematic reviews, that evidence is inadequate to assess the effectiveness of physical activity counselling in primary care. Primary care clinicians were not the ones to deliver the intervention in three out of the four additional studies. The study that used a nurse and automated system to deliver the intervention demonstrated some increase in a physical activity parameter (Dubbert, Cooper et al. 2002). However, the nurse intervention without the automated system did not. The intervention was delivered at the practice in only one study (Hillsdon, Thorogood et al. 2002). Recruitment rates were sometimes very low (e.g. 35%) (Pinto, Friedman et al. 2002). Follow-up rate was very low in another study (32-33%) (Hillsdon, Thorogood et al. 2002). The rate of intervention patients
receiving the intervention was very low in yet another study (38%) (Green, McAfee et al. 2002). The positive effect found in the latter study can not be relied upon because comparison of follow-up values were used, yet baseline values were not balanced (Green, McAfee et al. 2002). Again, the design, the poor participation rates, and other methodological issues, limited the ability of these studies to detect a positive effect.

The current review of the literature has several potential limitations. Firstly, not all original trials were reviewed because this had been done several times previously. Considering the results of trials published since the last systematic review along with the conclusion reached by previous systematic reviews is not the same as considering all trial results together. In addition, the review only included published trials. However, there is evidence to suggest that the inclusion of unpublished trials may lower the estimate of treatment effect, particularly in meta-analyses (Egger, Juni et al. 2003). Therefore, the inclusion of unpublished trials, is unlikely to have changed the findings from the current review.

Only studies published in English were included in the reviews. This may have resulted in a conservative bias. A review of meta-analyses has shown that studies published in languages other than English were more likely to have been of lower quality and more likely to have shown a positive effect (Egger, Juni et al. 2003).

The US Preventive Services Task Force (USPSTF) made their recommendations on the basis of the most recent systematic review (Eden, Orleans et al. 2002). In particular, “the USPSTF found insufficient evidence to determine whether counselling patients in primary care settings to promote physical activity leads to sustained increases in physical activity among adult patients” (U. S. Preventive Services Task Force 2002). These recommendations have also been published in the American Family Physician (Anonymous 2002). Studies published after the systematic reviews do not change this conclusion.

However, the USPSTF also concluded that “multicomponent interventions combining provider advice with behavioural interventions” appear the most promising (U. S. Preventive Services Task Force 2002). These interventions often include, “patient goal-setting, written exercise prescriptions, and individually tailored follow-up assistance provided by specially trained staff” (U. S. Preventive Services Task Force 2002).
Therefore, it is clear that further randomised controlled trials are needed to help clarify the issue. Ideally, the study should have a high recruitment and follow-up participation rate, involve a diversity of regional, socio-economic and cultural populations, and include assessments of cost-effectiveness and ‘harm’, as well as ‘benefit’.

The randomised controlled trial of the Green Prescription intervention in primary health care conducted for this thesis, was designed to add to this literature. In particular, the present author aimed to conduct a high quality trial, involving a wide range of general practices and patients, from different regional, socio-economic, and ethnic backgrounds. Recruitment and follow-up methods that improved recruitment and follow-up rates were planned to maximise participation rates and generalisability of the findings. The trial was also designed to include assessments of potential adverse effects, as well as physical activity and health benefits. Lastly, a cost-effectiveness evaluation was designed prospectively, alongside the randomised controlled trial. A systematic review of cost-effectiveness studies of physical activity interventions in primary health care is presented in the next section of this chapter to demonstrate existing evidence and methods used by previous studies.

2.5 Systematic Review of Cost-Effectiveness Studies of Physical Activity Interventions in Primary Health Care

2.5.1 Introduction

Significant economic savings could be made if a larger proportion of the population became active. These savings could be achieved in both direct health care costs, and indirect costs in terms of prolonged productivity, as a result of decreased morbidity and mortality, particularly due to cardiovascular disease. In 1987, it was estimated that if an additional 10% of the New Zealand population increased their activity to recommended levels, $24.75 million per year could be saved in direct and indirect costs (Russel, Worsley et al. 1987). This figure was updated to $55 million per year by a report commissioned by the Hillary Commission in 1990s (Hillary Commission 1998). Another study estimated a saving of $48 million if the 31% of the population that were sedentary (being active for less than 3 hours per day) was reduced to 21% (Jensen, Sullivan et al. 1993).
However, the cost of increasing physical activity levels amongst sedentary individuals needs to be considered in order to assess the cost-effectiveness of promoting physical activity. As mentioned in the previous section, there are few assessments of cost-effectiveness of physical activity interventions for adults in primary health care. This systematic review was carried out to answer the following questions:

- What is the cost-effectiveness of physical-activity interventions based in primary health care?
- What is the quality of the studies conducted to date?

The systematic review was also carried out to identify appropriate interventions to use as comparators to the Green Prescription intervention in the cost-effectiveness study of this thesis.

2.5.2 Methods

2.5.2.1 Search Strategy

A literature search was carried out in several databases. The strategy used is described in Table 2-6. Searching article references and hand searching was also used. Nine papers were identified, representing eight economic studies.

2.5.2.2 Inclusion Criteria

Studies were included if they were economic analyses of lifestyle interventions that included physical-activity advice, designed for adults, and were based in primary care. Only cost analyses conducted in association with lifestyle intervention trials were included. Economic analyses based on surveys or non-experimental studies, or theoretical studies, were excluded. Only studies published in English were included.

2.5.2.3 Data Abstraction

Data were abstracted on the study population, intervention, design, outcome measures, results, and quality of the study. Full papers were reviewed.
Table 2-6 Literature Search Strategy of Electronic Databases, and Subject Headings used to locate Cost-Effectiveness Studies of Physical Activity Interventions in Primary Health Care

<table>
<thead>
<tr>
<th>Databases</th>
<th>Physical activity subject headings</th>
<th>Cost-effectiveness headings</th>
<th>Primary health care subject headings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embase</td>
<td>Physical activity</td>
<td>Health economics</td>
<td>General practice</td>
</tr>
<tr>
<td>Psychlit (now Psycinfo)</td>
<td>Fitness</td>
<td>Cost minimization analysis</td>
<td>Family practice</td>
</tr>
<tr>
<td>Sportdiscuss</td>
<td>Physical fitness</td>
<td>Cost benefit analysis</td>
<td>Primary medical care</td>
</tr>
<tr>
<td>Cinahl</td>
<td>Exertion</td>
<td>Cost effectiveness analysis</td>
<td></td>
</tr>
<tr>
<td>Cochrane database of systematic reviews</td>
<td></td>
<td>Cost utility analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Health care costs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Costs and cocts analysis</td>
<td></td>
</tr>
</tbody>
</table>

2.5.2.4 Critical Appraisal

A quality checklist for health economics studies has been published previously (Drummond and Jefferson 1996). The critical appraisal undertaken of the papers in the systematic review of this thesis, is based on Drummond’s checklist and includes the following aspects:

1. Study design:
   - There is a clear research question and the economic importance is discussed,
   - The viewpoint of the analysis is stated and justified,
   - The choice of comparative programmes is explained.

2. Data Collection:
   - The sources of economic information are stated,
   - The details of the design and results of the effectiveness study are available,
   - Primary outcome measures, and the methods used to value health states, are explicit,
   - Productivity changes are reported separately,
• Quantities and unit costs are reported and the method for estimating these is explained,
• Currencies, adjustments for inflation, and currency conversions, are reported

3. Analysis and Interpretation:
• The time horizon of costs is explicit and appropriate discount rates used,
• Effectiveness of the intervention has been demonstrated,
• Statistical tests and confidence intervals, where appropriate, are reported,
• Appropriate sensitivity analyses are carried out,
• Relevant alternative programmes are compared,
• Incremental analyses are reported,
• The study question is answered and conclusions are appropriate

2.5.3 Studies Identified and Critical Appraisal

Table 2-7 includes the nine papers that fulfilled the inclusion criteria. The first five papers include cost analyses of lifestyle interventions that include physical activity advice. The other four papers include cost-analyses of physical activity interventions.
Table 2-7 Evidence Table of Cost-Effectiveness Studies undertaken of Interventions that included Physical Activity Promotion within Primary Health Care or the Community

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants; Intervention</th>
<th>Design; Outcomes</th>
<th>Results</th>
<th>Critical Appraisal</th>
</tr>
</thead>
</table>
| 1     | (Langham, Thorogood et al. 1996)  
Costs and cost effectiveness of health checks conducted by nurses in primary care: the Oxcheck study. | 4,121 35-64 year old patients in primary care; Nurse health cardiovascular screening with health education and follow-up | Cost-outcome analysis of RCT (3-year follow-up); 1. Cost of health check programme 2. Cost per 1% reduction in coronary risk | 1. UK£29.27 per patient. 2. £1.46 - £2.25 per 1% reduction (nearly twice as much for men as women) | Study Design Clear research question and importance. Viewpoint not explicit (presumably costs to practice). No comparative programme was used, so not 'cost-effectiveness'. 
Data Collection Actual costs used, sources clear, and research expenses excluded. Based on RCT. Outcome measures clear. Quantity and unit costs not presented separately. No productivity costs estimated. Inflation adjustments not stated. 
Analysis and Interpretation Time horizon stated but no discounting apparent. No statistical tests or CIs* Incremental costs of healthcare utilisation not calculated (these were not recorded for control patients). Sensitivity analyses appropriate (to allow for assumptions about nurse research time, and to account for possible incremental increase in medication use) Inappropriate comparison with costs of cervical screening. |
| 2     | (Wonderling, McDermott et al. 1996)  
Costs and cost effectiveness of cardiovascular screening and intervention: the British family heart study | 4,185 men (40-59yrs) and 2,827 of their partners, from 13 general practices Nurse-led cardiovascular risk assessment and health education program. | Cost-outcome analysis of RCT; 1. Cost of programme 2. Cost to NHS 3. Cost per 1% reduction in coronary risk at 12 months | 1. UK£63 per person per yr for programme 2. £77 per man and £13 per woman to NHS. 3. £5.08 per man and £5.78 per woman to reduce risk by 1% | Study Design Clear question and importance. Health service viewpoint explicit, although discussed costs to patient (e.g. transport) in discussion. No comparison programme in this paper. 
Data Collection Sources of costs explicit. Design and results of RCT available. Clinical and resource use costs taken from trial and unit costs from external estimates (appropriate). Currency stated, VAT included and 6% discount rate used, based on Treasury rate. Productivity changes not considered as health service perspective. 
Analysis and Interpretation Although intervention was
3 (Wonderling, Langham et al. 1996)

**What can be concluded from the Oxcheck and British family heart studies: commentary on cost effectiveness analyses**

| Compared results from Oxcheck (4,121, 35-64 year-olds) with British family heart study (4,185 men who were 40-59yrs, and 2,827 of their partners) | Cost-effectiveness analysis of 2 RCTs. Also comparison with other lifestyle programmes. |
| Nurse health cardiovascular screening with health education programs and follow-up | 1. Life-years gained (LYG) 2. Cost per life-year gained |
| Cost per LYG: **Oxcheck:** £34,800 (for 1yr duration of program) to £1,500 (for 20 yrs duration) **British family heart study:** £29,300 (for 1yr duration) to £900 (for 20yr duration) |

**Study Design** Clear question and importance. Viewpoint not stated, but was stated in previous paper as health services perspective. Choice of comparison programmes justified

**Data Collection** Based on the two studies reviewed above. Discounting and adjustments explicit. Identified differences in methods used to estimate costs in the two studies.

**Analysis and Interpretation** Difference in time horizons adjusted for and presented estimated cost/LYG if the programmes lasted for 1 year compared with if they lasted for 20 years (valuable sensitivity analysis). Incremental analyses reported where possible and acknowledged risk of accommodation bias and non-returner bias from studies in potentially overestimating effectiveness, and therefore cost-effectiveness. Relevant alternative programmes compared in LYG table showing huge variation, likely to be at least in part due to variability in methods and analyses used in different studies. Appropriate conclusions and possible implications

4 (O’Neill, Normand et al. 1996)

**Cost effectiveness of personal health education in primary care for people with angina in the greater**

| 688 patients aged less than 75, with angina for at least 6 months | 1. Cost of drug and health services for each group |
| Nurse-led brief advice on coping with disease and lifestyle approach | No significant difference in overall health service costs between the groups, yet intervention patients had |

**Study Design** Clear question. Viewpoint not stated (although assumed a health services perspective). No comparative programmes used, and no cost-effectiveness ratios calculated. Therefore, this is not a cost-effectiveness analysis.

**Data Collection** Details and results of RCT summarised. Actual costs associated with secondary care were collected. GP costs estimated from Oxcheck study. Research costs were not excluded from intervention costs and community care costs not
### Analysis and Interpretation

Change in drug costs calculated only for those upon which there were complete data (Not ITT** and possible bias) Incremental change was reported, but this risks bias when baseline costs so different (? adequacy of randomisation). As no overall statistically significant difference in health service costs (including intervention costs) could be detected between intervention and control groups, it was concluded that the benefits of the intervention were achieved with no extra cost. This is a flawed inference, because the statistical significance of difference will depend on the sample size and how many overall service costs the authors chose to include in their analysis. No significant difference is not synonymous with being the ‘same’ cost. CI’s (or P values) not presented. No allowance for discounting or inflation was reported. No comparisons with other programs. Study question inadequately answered.

### Study Design

Clear question and importance. Viewpoint stated (government health care funder). Both comparative programmes explained

### Data Collection

Sources of costs explicit. Data on indirect costs and benefits, and CVD-related quality of life, were estimated from previous study on patients with MI, which may not be representative (This was acknowledged). Primary outcome results available. Unit costs and quantities not reported separately. Discounting of 5%, and currency, reported. Inflation adjustments not reported. Productivity changes not reported separately, although excluded for sensitivity analysis.

<table>
<thead>
<tr>
<th>1. No significant effect on CVD risk factors compared with routine care</th>
<th>2. Cost per QALY for males: $AUD152,000 to $204,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>82 GPs randomised to three arms. 757 patients with at least one CVD risk factor.</td>
<td>Cost-effectiveness analysis alongside a cluster RCT</td>
</tr>
<tr>
<td>Routine care; Lifestyle intervention using video; Lifestyle intervention using video and self-help group</td>
<td>1. BP, BMI, cholesterol, smoking status</td>
</tr>
</tbody>
</table>

---

**Notes:**

- CVD: Cardiovascular Disease
- BP: Blood Pressure
- BMI: Body Mass Index
- MI: Myocardial Infarction
- QALY: Quality Adjusted Life Years
- ITT: Intention to Treat
<p>| <strong>6</strong> (Moffett, Torgerson et al. 1999) | <strong>187</strong> 18-60 year old patients with mechanical back pain of 4 weeks to 6 months duration; Exercise classes for low back pain given by physiotherapist compared with usual primary health care | <strong>Cost description along side RCT</strong> | <strong>Analysis and Interpretation</strong> Data analysis was performed on only those who returned for follow-up (65% or less for each outcome measure) so not ITT (source of non-returner bias). Programmes compared with each other. CT's and P values reported for effectiveness outcomes but not for cost-effectiveness. Cost per LYG and QALY calculations were not incremental analyses and not appropriate (except for post-hoc sub group analysis of high-risk subjects) as effectiveness had not been demonstrated. Conclusions acknowledge study faults. |
| <strong>Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences</strong> | <strong>Cost-outcomes analysis along side RCT</strong> | <strong>Study Design</strong> Not a clear economic question. Viewpoint of NHS and societal perspective was explicit. No comparative programme, so not 'cost-effectiveness' |
| <strong>187</strong> 18-60 year old patients with mechanical back pain of 4 weeks to 6 months duration; Exercise classes for low back pain given by physiotherapist compared with usual primary health care | <strong>Outcomes:</strong> | <strong>Data Collection</strong> Self-report health-care utilisation to estimate costs, although unit costs and quantities not reported separately. Details and results of RCT available. Productivity changes reported separately. |
| <strong>Exercise classes for low back pain given by physiotherapist compared with usual primary health care</strong> | <strong>Cost of moving one person:</strong> | <strong>Analysis and Interpretation</strong> Statistical tests explained appropriately, confidence intervals presented and intention-to-treat analysis undertaken. However, study under-powered due to study resources, which may risk type 2 error, especially with respect to economic analysis. No sensitivity analyses and no comparative programmes compared. Incremental analyses reported appropriately. Authors concluded that the intervention was cost-effective, although evidence presented may not conclusively support this, due to lack of significance. |
| <strong>1. Assessment of disability, back pain and quality of life at 6 and 12 months</strong> | <strong>1. UK£650 to move person out of sedentary group.</strong> | <strong>Study Design</strong> Clear question and importance. Viewpoint not explicit (programme costs only). No comparative programmes identified. |
| <strong>2. Days off work and health care utilisation</strong> | <strong>2. £2,500 to move one person to</strong> | <strong>Data Collection</strong> Summarised as 'top-down' excluding only costs associated with research. Annual equivalent costs of equipment included. Unit costs and quantities not reported separately. Productivity costs not presented. Primary outcomes |
| <strong>7</strong> (Stevens, Hillsdon et al. 1998) | <strong>714</strong> inactive people (45-74 yrs). Leisure centre and home-based exercise programme initiated by | <strong>Cost-outcomes analysis along side RCT</strong> | <strong>Study Design</strong> Clear question and importance. Viewpoint not explicit (programme costs only). No comparative programmes identified. |
| <strong>Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old</strong> | <strong>Outcomes:</strong> | <strong>Data Collection</strong> Summarised as 'top-down' excluding only costs associated with research. Annual equivalent costs of equipment included. Unit costs and quantities not reported separately. Productivity costs not presented. Primary outcomes |</p>
<table>
<thead>
<tr>
<th>Study Design</th>
<th>Clear question and importance. Societal viewpoint explicit. Based on non-randomised controlled trial.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Collection</td>
<td>Sources of costs stated and justified. Programme set-up costs not included in analysis. Results and detail of controlled trial presented. Currency and exchange rates stated. Quantities and unit costs reported separately.</td>
</tr>
</tbody>
</table>
**Prevent falls. 2: Controlled trial in multiple centres.**

and hospitalisation due to fall

(Over 12 months)

**Analysis and Interpretation** (As above) Time horizon appropriate. Incremental analyses carried out. Discounting not stated. Statistical testing described. CI’s not reported but one-way sensitivity analyses were carried out using 125th and 75th percentile, and total costs, for implementation when calculating CERs, which is appropriate. Results presented to be comparable to previously reviewed study.

---

* Confidence intervals *a* Randomised controlled trial ** Intention-to-treat

* Blood pressure *b* Body mass index

* Cardiovascular disease

* Cost effectiveness ratios
2.5.4 Discussion of the Review

The variability of outcome measures, research designs, and time spans, makes comparison of cost-effectiveness of interventions difficult. However, when considering studies with similar outcome measures, the British family heart study intervention was slightly more cost-effective than the Oxcheck study intervention (Wonderling, Langham et al. 1996). A primary-care-based nurse-led exercise intervention was more cost-effective at preventing falls in the elderly than the community-based nurse-led intervention, but less effective at preventing injuries or hospitalisation (Robertson, Devlin et al. 2001; Robertson, Gardner et al. 2001). When considering the interventions that did have an effect, the cost effectiveness ratios for life years gained were comparable to other interventions (Wonderling, Langham et al. 1996). This was particularly so if it was assumed that 'benefit' would continue at similar rates over time. However, these interventions involved multiple components, including increased medication use, so these conclusions may not be relevant for physical activity counselling, alone. It does appear, however, that exercise promotion to prevent falls in the elderly, is relatively cost-effective when compared with other interventions (Robertson, Devlin et al. 2001).

In one study, the costs associated with small increases in physical activity seemed large (UK£650 to move one person out of the sedentary category) (Stevens, Hillsdon et al. 1998). However, the range of health benefits that may accrue from increased physical activity may make such an intervention seem more cost-effective. Unfortunately, the number of potential benefits from exercise, and the time required to observe long-term benefits, make such cost-effectiveness difficult to demonstrate within a randomised controlled trial context.

The quality of the studies in this systematic review was variable. Only three studies were true 'cost-effectiveness' studies (Wonderling, Langham et al. 1996; Salkeld, Phongsavan et al. 1997; Robertson, Devlin et al. 2001). Only two of these were valid because one study conducted a cost-effectiveness analysis despite no statistically significant effect being demonstrated during the accompanying randomised controlled trial (Salkeld, Phongsavan et al. 1997), which should have precluded cost-effectiveness analysis (Drummond, O'Brien et al. 1997).
'Cost-effectiveness' evaluations of programmes involve the comparison of two or more alternative programmes (Drummond, O'Brien et al. 1997). Some would argue that comparison of the costs and effects of a programme compared with 'usual care' is a cost-effectiveness evaluation. However, when estimating the costs (and cost-effectiveness ratios) associated with a programme, including health-care savings, incremental costs are used (costs with the programme minus costs of 'usual care'). If the incremental costs per outcome are described without the comparison of another programme, then this analysis is referred to as 'cost-outcome' study, not a 'cost-effectiveness' study (Drummond, O'Brien et al. 1997).

Four studies included 'cost-outcome' analyses, by calculating cost-effectiveness ratios of intervention compared with control groups (Langham, Thorogood et al. 1996; Wonderling, McDermott et al. 1996; Stevens, Hillsdon et al. 1998; Robertson, Gardner et al. 2001). Two studies were 'cost descriptions' eg. (O'Neill, Normand et al. 1996; Moffett, Torgerson et al. 1999). These studies only presented costs and overall outcomes, but did not present cost per unit outcome (Drummond, O'Brien et al. 1997).

The two studies by Robertson and colleagues, used falls and admissions prevented as their efficacy variables (Robertson, Devlin et al. 2001; Robertson, Gardner et al. 2001). Consequently, the cost per increase in physical activity outcome could not be estimated.

In fact, cost per unit of physical activity gained was reported only in one of the economic analyses included in the present systematic review (Stevens, Hillsdon et al. 1998). Therefore, this was the only study that could be used as a comparative study for the present thesis, if the cost of increasing physical activity is used as a cost-effectiveness outcome measure. Even so, the physical activity outcomes, used by Stevens and colleagues, were categorical. Cost-effectiveness ratios were presented as the cost of moving one person out of a sedentary group and cost of moving someone to the recommended level of activity. Consequently, the cost-effectiveness ratios for the continuous physical activity variables measured in this thesis, could not be compared with those of the Stevens study, although the categorical data could be compared.

Therefore, another study was located to provide a further comparative programme for the continuous physical-activity outcome measures from the Green Prescription study. This study was community-based rather than based in primary health care. The study chosen was
a community-based study of two physical activity interventions (Project Active) that used similar outcome measures of effectiveness and cost-effectiveness to the Green Prescription evaluation presented in this thesis (Sevick, Dunn et al. 2000).

This systematic review found few cost-effectiveness or cost-outcome studies of lifestyle interventions, including physical activity counselling, conducted in primary health care. The cost per unit increase in physical activity has not been estimated. Furthermore, none of the studies assessed physical activity interventions initiated during everyday consultations by usual clinicians. The cost-effectiveness analysis presented in this thesis was designed to address these issues. The methods used in this analysis are described in Chapter 3, section 3.4.

2.6 Rationale for the Cluster Randomised Controlled Trial Design

2.6.1 Introduction

This section describes the rationale behind choosing a cluster randomised controlled trial to assess the effectiveness of the Green Prescription intervention in general practice. Issues associated with conducting such trials in a community setting are discussed.

2.6.2 Randomised Controlled Trials

Randomised controlled trials are justified where there is uncertainty about the treatment effect of an intervention. The US National Institutes of Health describe the clinical trial as “the most definitive tool for evaluation of the applicability of clinical research.” It represents “a key research activity with the potential to improve the quality of health care and control costs through careful comparison of alternative treatments” (Friedman, Furberg et al. 1985). If the only difference between a control and an intervention group is the intervention, outcome differences over time are likely to be due to the intervention.

“Randomisation tends to produce study groups comparable with respect to known as well as unknown risk factors, removes investigator bias in the allocation of subjects, and guarantees that statistical tests will have valid significance levels.” (Friedman, Furberg et al. 1985)

The randomisation process should be carried out independently, preferably by computer generated random sequence. Allocation of randomisation should not be known until after
study participants are enrolled, to avoid selection bias. It has been found that treatment effects and odds ratios can be exaggerated by as much as 41% if randomisation of participants is not carried out correctly (Schulz, Chalmers et al. 1995). Randomised controlled trial designs have been recommended as the most appropriate design to use for the evaluation of physical activity interventions in primary care (Iliffe, Tai et al. 1994).

Consequently, a randomised controlled trial design was selected as the study design to assess the effectiveness of the Green Prescription intervention.

2.6.3 Methodological Issues of Lifestyle Intervention Trials

2.6.3.1 Rationale for Community-based Trials

Lifestyle interventions have been developed to prevent diabetes, cardiovascular and other diseases. Well-designed randomised controlled efficacy trials have demonstrated that lifestyle interventions, such as dietary and physical activity programmes, can reduce risk and progression to disease (Anonymous 1997; Halbert, Silagy et al. 1999; Knowler, Barrett-Connor et al. 2002; Whelton, Chin et al. 2002). On the basis of this evidence, lifestyle interventions have been introduced into general practice, and are assumed to be effective, although this may not be the case (Tai, Gould et al. 1997; Hillary Commission 1998; Moore, Summerbell et al. 2001). The effectiveness of these interventions may be very different in everyday general practice compared with strictly controlled efficacy trials. Thus, there is a need to evaluate effectiveness in real-world circumstances, with usual general practitioners or nurses delivering the intervention in everyday consultations. However, exporting randomised controlled trial methodology to community or every-day primary care consultations raises methodological issues that threaten the validity of results. These issues, and techniques to reduce potential bias, should be considered when designing lifestyle intervention trials in primary health care (Moore, Summerbell et al. 2001). A discussion of these issues and techniques is presented below.

2.6.3.2 Rigour of Lifestyle Intervention Evaluation

The quality of studies, designed to evaluate effectiveness of lifestyle interventions in community or health care settings, has often been low. Yet a randomised controlled trial design is recommended for lifestyle intervention studies, to reduce potential bias (Tai and
Iliffe 2000). However, a minority of lifestyle intervention evaluations in community and primary care settings have been randomised controlled trials (Riddoch, Puig-Ribera et al. 1998). A review of all published articles in the British Journal of Sports Medicine from 1991 to 1995 revealed that only 3% of articles reported studies that had used a randomised controlled trial design (Tai and Iliffe 2000). Exercise intervention studies have tended to use before-after comparison or ‘uncontrolled’ trial designs.

Intention-to-treat analysis refers to the analysis of outcomes of all participants originally enrolled in the study, according to their allocation of randomisation, regardless of compliance (Hollis and Campbell 1999). If some participants fail to complete the study, actual or estimated outcomes are allocated to them. If outcomes are estimated, a conservative assumption should be made, such as assuming ‘no change’ from baseline, or assuming a change equal to the mean change of the control group. Intention-to-treat analyses are employed to reduce the bias associated with analysing only results from participants that are ‘compliant’ with the intervention and study protocol. In addition, those who do not complete the study may differ in some systematic way to those who do complete the study. If those who do not complete the study are excluded from analysis, this may bias the results and limit the generalisability, or external validity of the results.

However, intention-to-treat analyses are often not conducted in physical activity-intervention trials. For example, intention-to-treat analyses were not conducted in any of the eight randomised controlled trials of exercise and osteoporosis reviewed by Block (1997). The intention-to-treat approach is more important for pragmatic trials of effectiveness than trials of efficacy, as such an approach estimates the overall effect within a population (Hollis and Campbell 1999).

2.6.3.3 Contamination of Intervention and the Cluster Design

There is a risk of contamination of the intervention to control groups within non-blinded randomised controlled trials in community or health care settings, when individual randomisation is used. Contamination will dilute the demonstrated effect, which may lead to a type 2 error (finding no effect when one exists). Estimating the extent of contamination and inflating the sample size accordingly has been recommended, to counteract the potential ‘dilution’ of effect (Slymen and Hovell 1997; Torgerson 2001). Although a formula has been
developed to inflate the sample size for an individually-randomised trial, this formula assumes that the proportion of control participants that are ‘contaminated’ by the intervention can be accurately estimated. Levels of contamination are rarely measured, and are difficult to predict. In a review of cancer trials using individual randomisation and the Zelen’s technique, Torgerson found that an average of 18% (range 10%-36%) of patients chose to swap from one treatment arm to the other (Torgerson 2001). He uses this figure to estimate the range of likely contamination in an individually randomised trial. However, volunteering to swap from one treatment arm to another is not the same as control patients being exposed to the intervention.

Contamination may arise from communication between intervention and control patients in the waiting rooms of general practices, for example. Patients cannot usually be blinded to the intervention and yet have a high chance of having contact with control patients, passing on information from the intervention, causing ‘contamination of intervention’ by patients (Edwards, Braunholtz et al. 1999). Concern amongst control patients who believe they are “missing out” could also influence behaviour change within the control group and alter the study outcome.

Furthermore, it is difficult for health-providers to turn on and off the delivery of advice to patients, depending on their individual randomisation to intervention or control. Consequently, there is a high likelihood of ‘contamination of intervention’ to control patients by health providers who are delivering the intervention to others within a trial. Alternatively, asking the provider not to deliver the advice may alter ‘usual care’, as some lifestyle advice may traditionally be part of ‘usual practice’.

The other option, to reduce the risk of contamination, is to use a cluster randomised controlled trial design.

“A cluster randomisation trial is one in which intact social units, or clusters of individuals, rather than individuals themselves, are randomised to different intervention groups” (Donner and Klar 2000)

Thus, randomising at a group or cluster level is also recommended when delivery of an intervention is likely to affect others within the cluster (Campbell and Grimshaw 1998; Edwards, Braunholtz et al. 1999; Eccles, Grimshaw et al. 2000). However, using a cluster
randomised controlled trial design requires sample size inflation to overcome the loss of statistical power, associated with randomising by practice, rather than individual. (This is discussed in more detail in the next sub-section.) If the inflation required due to clustering is less than the inflation required to counter contamination in an individually-randomised trial, then a cluster randomised trial would be more cost-efficient (Slymen and Hovell 1997; Torgerson 2001).

There are other reasons for using a cluster randomised design, besides risk of contamination. If the intervention involves, for example, education of practitioners, but outcomes are measured at the patient level (e.g. (Kerse, Flicker et al. 1999)), then a cluster design is practical. When the intervention is delivered at a group level, a cluster randomised design is also sensible, such as class-based educational interventions. Cluster randomised controlled trials are being used increasingly where delivery of an intervention is at a group level (Raab and Butcher 2001; Yudkin and Moher 2001).

A cluster design may also be used when there is a risk of infection and re-infection due to proximity of patients randomised to control or intervention group (Donner and Klar 2000). For example, if a trial were being carried out for a treatment for headlice, randomisation by class would be sensible to avoid re-infection of intervention patients by control patients.

2.6.3.4 Multilevel Analysis

There are methodological issues to consider when using a cluster randomised design, rather than an individually-randomised design. Taking the effect of the practice into account during analysis allows for the group effect of the intervention (Campbell, Mollison et al. 2000). Patients within one practice, as members of a group, are more likely to share similarities, compared with patients from different practices. This is because members both influence, and are influenced by, group membership (Goldstein, Pinto et al. 1999). Furthermore, although randomisation is at the group level, outcomes are measured at the individual level (Sashegyi, Brown et al. 2000) (Wood and Freemantle 1999). Therefore, two traditional assumptions of randomised controlled trials are violated. These two assumptions include firstly, that all individuals are independent, and secondly, that analysis is at the level of randomisation (Sir Ronald Fisher, 1935) (Donner and Klar 2000). It is therefore necessary to compensate for the
correlated nature of the data (clustering) by inflating sample size and using statistical methods of analysis that account for clustering.

2.6.3.4.1 Sample Size

Responses to interventions, as well as health and demographic variables, tend to be more similar within practices than between practices (Campbell, Mollison et al. 2000; Moore, Summerbell et al. 2001). Accordingly, statistical power is reduced by the lack of independence of individual responses, within clusters. The sample size needs to be inflated to compensate for the dilution of statistical power. A statistical measure of the interdependence within each cluster is the intra-class correlation coefficient (ICC). The ICC can be used to calculate the sample size inflation factor required, referred to as the design effect (Campbell, Mollison et al. 2000).

The design effect of cluster-randomised trials is the ratio of the total number of subjects required using cluster randomisation to the number required using individual randomisation. (Kerry and Bland 1998) The design effect \( d_{eff} \) can be calculated for different outcome variables using the equation \( d_{eff} = 1 + (m-1) \times ICC \), where \( m \) is the cluster size. Sample size is then inflated accordingly (ie. simple random sample size \( \times d_{eff} \) = adjusted sample size).

Appropriate sample size calculation requires estimation of intra-class correlation coefficients obtained from previous research. Publication of intra-class correlation coefficients from cluster-randomised trials has been recommended (Donner, Brown et al. 1990; Kerry and Bland 1998).

2.6.3.4.2 Analysis of Results

In general, the variance within groups randomly drawn from the total population is assumed to be random. If members of a sample are drawn in clusters from the population then this variance is not necessarily random (Campbell, Mollison et al. 2000; Raab and Butcher 2001). Consequently, the effect of clustering should also be taken into account during analysis (Bland and Kerry 1997; Campbell, Mollison et al. 2000). Whereas point estimates of each cluster were estimated in the past to produce values for analysis, complex statistical models now take into account individual values, but allow for the effect of clustering (Campbell, Mollison et al. 2000; Raab and Butcher 2001). Accordingly, a statistical package that allows
for clustering should be used to analyse the results of a cluster randomised controlled trial, as was carried out in the current trial.

Although some studies have used cluster randomisation in health services research, many have not accounted for the cluster design during calculation of sample-size or during analysis, thus producing underpowered studies or spurious results (Campbell and Grimshaw 1998; Campbell, Mollison et al. 2000). Reviews of randomised trials using cluster designs have also highlighted the problem of inappropriate analysis techniques for cluster randomisation and lack of information about levels of intra-class correlation (Donner, Brown et al. 1990; Simpson, Klar et al. 1995). There is usually an appropriate unit of analysis for intervention, depending on who delivers the intervention and what environmental influences are likely to have an effect on the outcome of the intervention. “Errors” in the selected unit of analysis were found in 70% of studies in one review of intervention trials aimed at influencing physician practice (Divine, Brown et al. 1992).

Accordingly, the present author chose a cluster randomised design to assess the effectiveness of the Green Prescription intervention. Consideration of the cluster design was therefore allowed for both within sample size calculations and analysis of results. An attempt to monitor potential contamination of intervention was also made and intraclass correlation coefficients were calculated.

2.6.3.5 Recruitment and Follow-up Rates

Low participation rates of eligible participants will limit the generalisability of study findings. Recruitment rates tend to be low if general practitioner referral is relied upon within lifestyle intervention trials (Bell-Syer and Moffett 2000). If practice registers are used, participation of eligible patients is still often below 50%, as occurred in a recent physical activity trial (The Writing Group for the Activity Counseling Trial Research Group 2001). Furthermore, patients are frequently recruited from only one or two practices, which may not be representative of the quality or variability of intervention delivery and patient response found across medical practices (Halbert, Silagy et al. 2000; Hillsdon, Thorogood et al. 2002). Attrition rates in lifestyle intervention studies are often high. One review of exercise interventions for the treatment of osteoporosis found attrition rates varying from 4 to 50% with most over 20% (Block 1997). This is a common problem in exercise intervention
studies in primary care also, where attrition rates greater than 50% are common, such as in a recent physical activity trial in general practice (Hillsdon, Thorogood et al. 2002).

Accordingly, the present author used techniques to encourage high participation and follow-up rates in order to ensure reasonable generalisability of results. In addition, methods were chosen to maximize the delivery of the intervention to reduce the risk of a type 2 error, from low rates of intervention delivery.

2.6.3.6 Timing of Recruitment

If a cluster design is adopted, practices can be randomised prior to patient recruitment. However, this approach risks selection bias due to lack of concealment of randomisation allocation while recruiting patients, producing imbalance of recruitment rates and baseline characteristics (Moore, Summerbell et al. 2001). This imbalance may affect results. Inadequate allocation concealment has been associated with odds ratios exaggerated by 41%, according to a meta-analysis from the Cochrane pregnancy and childbirth database (Schulz, Chalmers et al. 1995). An alternative approach involves recruitment of all patients prior to randomisation of practices. However, this approach is more time-consuming and increases the time between enrolment and intervention delivery (Moore, Summerbell et al. 2001). It requires patients to return to the practice for the intervention, is more expensive, (Moore, Summerbell et al. 2001) and may be more logistically difficult if many practices are taking part.

2.6.3.7 Lack of Blinding

Participants and assessors in trials involving lifestyle interventions are usually not blinded, as the author of one meta-analysis of exercise and blood pressure commented (Whelton, Chin et al. 2002). Ethics committees often do not allow deception of participants as to the nature of the intervention, so blinding of participants is difficult. Even in clustered trials, informed consent is recommended from all participating members of the control and intervention groups (Edwards, Braunholtz et al. 1999). Lack of blinding of assessors in health services trials is associated with an exaggeration of odds ratios by 17% (Schulz, Chalmers et al. 1995). Some trials have been able to blind their assessors (The Writing Group for the Activity Counseling Trial Research Group 2001). However, this strategy often requires
measurement to be carried out away from usual practices, if randomisation was at the practice level. It also relies on the participants not disclosing their allocation of randomisation to assessors. If the health provider knows the allocation of randomisation, there is also the risk that these providers will offer a co-intervention to intervention patients or control patients differentially, thus biasing results.

Therefore, the present author used methods that minimised the risk of selection and assessor bias, due to pre-randomisation of practices and lack of blinding. This objective was achieved by ensuring patients remained blind to allocation of randomisation until after enrolment. In addition, few exclusion criteria were used, screening and recruitment of patients was largely independent of the ‘un-blinded’ researchers, and objective measures and self-administration of questionnaires were used. These techniques are described in detail in Chapter 3.

2.6.3.8 Summary

In summary, pragmatic trials of lifestyle interventions are justified. A randomised controlled trial is the most appropriate design for assessing effectiveness of such interventions. Randomising by cluster reduces the risk of ‘contamination’ of intervention when the intervention is delivered to individuals. A cluster design is also appropriate where the intervention is delivered at a group level. Sample-size calculations and analysis of results should take into account the cluster design. Intention-to-treat analysis is also important. However, it appears that there is a lack of rigour in many community-based studies of lifestyle interventions. These issues were taken into account by the author during the design of the Green Prescription randomised controlled trial. This design is described in more detail in Chapter 3 (Methods).

2.7 Conclusions of the Literature Review

Physical activity is important to health, especially cardiovascular health. Primary health care is a sensible setting to influence health behaviours of the population. A feasible physical activity intervention, the Green Prescription, has been developed for the New Zealand primary care setting. A long-term evaluation of its effectiveness is needed. The first literature review reported in this chapter revealed the need for adapting and validating an instrument to estimate change in physical activity amongst a ‘less active’ population within the setting of
general practice. The second and third literature reviews identified the need for further trials to demonstrate effectiveness and cost-effectiveness of physical activity interventions in primary care, because evidence is inconclusive and there is a lack of high quality studies. There are important methodological issues to be considered in conducting lifestyle intervention trials in every-day primary health care, including measurement problems and issues surrounding cluster randomisation.

This thesis is intended to contribute to the literature by assessing the following objectives: Firstly, to assess the reliability and validity of a physical activity questionnaire adapted for use amongst less-active adults within a primary health care setting. Secondly, to assess the effectiveness of a practice-based Green Prescription physical activity intervention within primary health care over a one-year period, in terms of physical activity, cardiovascular risk, and quality of life, amongst less-active 40 to 79 year-old adults, and lastly, to assess the cost-effectiveness of the Green Prescription physical activity intervention.
3 Methods

3.1 Introduction

This chapter presents the aims and methods of the three related studies in this thesis. The first section describes the methods used for physical-activity questionnaire selection, adaptation, and assessment of reliability and validity. The second section describes the methods used in the pilot study and the cluster randomised controlled trial to assess the effectiveness of the Green Prescription intervention in primary health care. The third section describes the methods used to assess the cost-effectiveness of this intervention.

3.2 Methods of Physical Activity Questionnaire Validity and Reliability Study

3.2.1 Aims

In order to assess the effectiveness of the Green Prescription intervention within a randomised controlled trial, an appropriate and valid method for estimating physical activity was required. Chapter 2, section 2.3 discussed different methods for physical activity assessment and, although physical activity questionnaires had limitations, they were still considered the most practical method for physical activity assessment in epidemiological study. To ensure that at least one valid instrument could be found, two questionnaires were adapted and tested for validity. Therefore, the aims of this study were:

1. To select two existing physical activities questionnaires and adapt them for use in epidemiological research in primary health care

2. To assess the criterion-related validity of these questionnaires among less-active adults in primary health care, and to assess the test-retest reliability of one of the questionnaires.

3.2.2 Physical Activity Questionnaire Selection and Adaptation

The Auckland Heart Study (AHS) 3-month recall physical-activity questionnaire (Appendix 1) and the Green Script Study (GSS) 2-week recall physical-activity questionnaire (Appendix
2) were selected because they had been used amongst adults in recent New Zealand studies. The use of either of these questionnaires in the subsequent Green Prescription evaluation would potentially allow comparison of the findings with those of previous studies. In addition, the AHS questionnaire had already demonstrated good validity in a high quality validation study (Arroll, Jackson et al. 1991).

Furthermore, both questionnaires could be easily adapted to facilitate calculation of total energy expenditure and to assess all categories of activity, including leisure, household, occupational activities and rest. Both questionnaires already assessed intensity, frequency and duration of activities. The importance of these characteristics has been explained in Chapter 2. The questionnaires were also adapted to allow self-administration within a practice setting. In particular, patient instructions were printed clearly, but briefly, at the beginning of the questionnaires. Several examples of appropriate and representative activities for each category were included with the questions to prompt respondents.

The questionnaires were initially administered to several volunteers in the age group to test the time taken, and to investigate any difficulties with the format or wording for self-administration. Once an acceptable format had been established, the validity of the questionnaires and the test-retest reliability of one questionnaire, were assessed on consecutive ‘less active’ patients recruited within a general practice setting. The origin of the two questionnaires is set out below.

3.2.2.1 Auckland Heart Study (AHS) Questionnaire

The AHS questionnaire was developed by Jackson from components derived from several sources (Jackson 1989). These sources included, the Stanford Five City 7-day physical-activity questionnaire, (Sallis, Haskell WL et al. 1985) and the Harvard Alumni Study physical-activity questionnaire (Paffenbarger, Wing et al. 1978). Elements were also derived from the Health Insurance Project (HIP) Job questionnaire, (Shapiro, Weinblatt et al. 1965) and household activity questions developed by Scragg (Scragg 1983). The AHS questionnaire has been used in the Auckland Heart Study, (Jackson 1989) and the Auckland Blood Pressure Control Study (Arroll 1992). Arroll and colleagues (1991) validated the AHS questionnaire amongst a random sample of 113 adults in New Zealand.
3.2.2.2 **Green Script Study (GSS) Questionnaire**

The GSS questionnaire has been used previously in a randomised controlled trial with sedentary patients, evaluating the Green Prescription physical-activity counselling programme in primary health care over a 6-week period (Swinburn, Walter et al. 1998). The origin of the questionnaire is unclear (personal communication with the authors). The test-retest reliability of the GSS questionnaire had been established previously, (Swinburn, Walter et al. 1998) but a criterion related validity study had not been carried out.

3.2.3 **Study Population**

This validation study was conducted within two medical practices, in different towns, each with two general practitioners and two practice nurses. Recruitment of patients was carried out in the waiting room of each practice over a 5-day period in November of 1999. Inclusion criteria comprised “less active” 40 to 79 year-old patients visiting their general practitioner during the week of recruitment. Patients were defined as “less active” if they answered “No” to the following question. “As a rule, do you do at least half an hour of vigorous or moderate exercise (such as walking or a sport) on five or more days of the week?”

Exclusion criteria included presence of unstable angina, uncontrolled congestive heart failure, progressive and debilitating conditions, serum cholesterol concentration of over 9 mmol/L, systolic blood pressure over 220 mm Hg or diastolic blood pressure over 120 mm Hg. These criteria were based on recommended contraindications to advising exercise to the elderly (Haskell 1997). Patients were also excluded if they did not speak English or were acutely injured or unwell. If the general practitioner or staff considered the candidate unsuitable for participation on medical or other grounds, the patient was not included.

All 40-79 year-old patients visiting the four general practitioners over a five-day period in November 1999 were screened for inactivity as they entered the practice. Those who fulfilled the inclusion criteria were invited to take part in the study.

3.2.4 **Reliability**

Test-retest reliability was assessed for the AHS questionnaire. This questionnaire was self-administered at baseline, within the practice and repeated between one and two weeks later
(mean of 11 days), at home. The completed questionnaires were then posted back to the author.

3.2.5 Validity

Study participants completed both the AHS and the GSS physical-activity questionnaires, initially. Participants were then given a 7-day activity diary and a pedometer to wear for seven days. Diaries and pedometer records were obtained from participants at the end of the week, for analysis.

3.2.6 Physical Activity Diary and Pedometer

The physical activity 7-day diary (Appendix 4) and the pedometer were used as ‘gold standards’ to assess the criterion-related validity of the questionnaires. The 7-day activity diary was developed by Arroll and used in the previous validity study of the AHS questionnaire (Arroll, Jackson et al. 1991). The pedometers used were Omron HJ-003, battery-powered and digitally displayed step counters. Pedometers were calibrated against actual number of steps, in each individual, prior to the pedometer record. All participants were asked to wear the pedometer at their waist for a week, and record the number of steps for each day, on a form. During the same week, respondents were also asked to fill in the 7-day activity diary by listing all moderate and vigorous leisure, domestic, and occupational activities, performed throughout each day. The duration and intensity of each activity was also recorded, as were the number of hours spent sleeping and resting in bed, in order to calculate total energy expenditure.

3.2.7 Analysis

Information about type, duration, frequency and intensity of each activity was recorded in the questionnaires. A metabolic equivalent value (MET) was established for each activity, using a standard compendium (Ainsworth, Haskell et al. 1993). A MET represents the ratio of work metabolic rate to a standard resting metabolic rate of 1.0 kcal (4.184 kJ)/kg/hr (Ainsworth, Haskell et al. 2000). The compendium provides an empirically based coding system of common leisure-time, domestic, and occupational activities and designates each activity with an average MET value. Time spent sleeping or resting in bed was coded as 1.0 MET, as this represents the resting metabolic rate (Ainsworth, Haskell et al. 2000). Time not
accounted for, after asking about sleep, leisure, domestic, and occupational activities was allocated a 1.5 MET value, assuming that time unaccounted for was spent in light activity. Activities of 3.0-4.9 MET were classified as moderate. Activities of greater than 5.0 MET were classified as vigorous (Wilson, Paffenbarger et al. 1986; Lamonte and Ainsworth 2001). This allowed the calculation of energy expenditure in each category, as well as total energy expenditure. Energy expenditure was calculated in kcal/kg/week.

Time spent in light, moderate and vigorous activities, of leisure, occupational or domestic categories were also estimated from the questionnaires and expressed as hours per week. In addition, a computer program was developed to calculate energy expenditure and time spent in different activities from the AHS questionnaire, using ACCESS software (Anderson and Elley CR 2002). The values calculated from the two questionnaires were compared with the validation 7-day diary estimates and pedometer readings.

Outcome measures used for the reliability and validity studies included total energy expenditure, total and leisure moderate-intensity activity, and total and leisure vigorous-intensity activity, as estimated by the questionnaires and the 7-day diary. The outcome measure obtained from the pedometer recordings was steps per day.

Analysis of reliability and validity was performed using SPSS 9.0 statistical software. Spearman’s rather than Pearson’s correlation coefficients were calculated because of the non-parametric nature of the data. Intraclass correlation coefficients were also calculated for reliability data (MacLennan 1993). Bland-Altman graphs of total energy expenditure and total moderate activity were presented as an alternative measure of validity (Bland and Altman 1986).

The validation study was undertaken alongside the pilot study of the subsequent randomised controlled trial. In the ‘control’ practice, data for the validity and test-retest reliability study were collected at the commencement of the pilot intervention study. Within the ‘intervention’ practice, data for the validity and test-retest study were collected after the completion of the 2-month intervention study. This approach was employed to avoid the reliability and validity studies being undertaken while participants were in the process of changing their physical-activity behaviour. The results of the reliability and validity study are presented in Chapter 4, section 4.2.
3.3 Methods of Cluster Randomised Controlled Trial of the Green Prescription Intervention in Primary Health Care

3.3.1 Introduction

This section discusses the main study of this thesis, the cluster randomised controlled trial of the Green Prescription intervention. Firstly, the methods that were used in the pilot study prior to the randomised controlled trial are described in section 3.3.3. (The results of the pilot study are presented in Chapter 4, section 4.3.2.) Secondly, the methods that were used in the cluster randomised controlled trial are described from section 3.3.4. (The results of this trial are presented in Chapter 4, section 4.3.3.)

The effectiveness of the Green Prescription intervention was assessed in the randomised controlled trial by measuring change in physical activity levels, cardiovascular risk status, and quality of life (using the SF-36) over a twelve-month period. This approach was used to determine whether increases in physical activity could be sustained over a one-year period, and whether these increases were enough to achieve clinically, and statistically, significant improvements in health.

3.3.2 Aims

Therefore, the aims of the randomised controlled trial were:

1. To determine whether the Green Prescription physical-activity intervention in general practice is effective in increasing the physical-activity levels of less-active middle aged and older people over a one-year period.

2. To determine whether the Green Prescription physical-activity intervention in general practice is effective in improving health outcomes of less-active middle aged and older people, assessed by change in blood pressure, cardiovascular risk, and quality of life health status.
3.3.3 Pilot Study Design

A pilot study was conducted from November 1999 until February 2000. The purpose of the pilot study was to assess the acceptability of the screening process, intervention delivery, and study protocol, to the practice staff and patients. In addition, the pilot study allowed an estimation of the proportion of patients who would be eligible for such a study, and the likely participation rate. The pilot study was also designed to demonstrate physical activity or health outcome trends in response to the intervention over a short period (two months). Lastly, the pilot study would provide interim data to calculate outcome variable means, standard deviations, and intracluster correlation coefficients that could be used in the sample size calculations for the main cluster randomised controlled trial. This information was particularly useful, because such data are sparse in the literature, particularly on less-active adults in primary health care.

3.3.3.1 Pilot Study Population, Allocation of Randomisation and Protocol

The study population was the same population that participated in the physical-activity reliability and validity study, described above, including the same inclusion and exclusion criteria. Two general practices in the Waikato-Hauraki region took part. Both practices had two general practitioners and two practice nurses working in them. Allocation of each practice to intervention or control was decided by the flip of a coin. Five days were spent recruiting from the waiting room in each practice. The receptionist handed forms to 40-79 year old patients as they arrived, in order to screen for inactivity, as described in the previous section, and to invite eligible patients to participate.

Informed consent was obtained and baseline measures taken from those who were willing to participate in the study, if they fulfilled eligibility criteria. Those in the ‘intervention’ practice were handed a form to establish their readiness to increase activity (‘stage of change’), and asked to give the form to the general practitioner when they entered the consultation. The general practitioner would then give advice about physical activity and a Green Prescription, which was also faxed to exercise specialists at Sport Waikato.

Study participants were co-ordinated two months later at the same practice, to repeat the questionnaires and measures. All measures were undertaken by the author.
3.3.3.2 *Pilot Study Outcome Measures*

Physical activity outcome measures included change in total energy expenditure (kcal/kg/week), leisure moderate-intensity and vigorous-intensity activity (hours per week), and stage of change, over a two-month period. Health outcome measures included change in systolic and diastolic blood pressure, weight, body mass index, serum total cholesterol and high-density lipoprotein concentrations, 5-year cardiovascular risk status, and quality of life, as measured by the SF-36 questionnaire.

Physical activity variables were estimated from the self-administered AHS questionnaire. The blood pressure values were obtained from the average readings of the second and the third recordings from an electronic sphygmomanometer. Weight was obtained from electronic scales and height from a standard tape measure. The serum lipid levels were obtained from finger-prick point of care machines. Demographic, medication, and other health details were obtained by questionnaire.

3.3.3.3 *Pilot Study Analysis*

Data were entered and analysed in SPSS 9.0 database. Baseline demographic and health-related variables were compared to assess whether the two study groups were similar. Mean change in each outcome variable was t-tested by “one sample” and “compared independent samples” methods, using SPSS. Analysis was undertaken on an intention-to-treat basis. Allowance was not made for the effect of clustering in the pilot study. Results of the pilot study are presented in Chapter 4.

3.3.4 *Cluster Randomised Controlled Trial Design*

The cluster randomised controlled trial of the Green Prescription was designed from mid 1999 until data collection commenced in April 2000. The study protocol and sample size calculations were refined following the results of the pilot study. Description of the study setting and population, sample-size calculations, study protocols, outcome measures, intervention, and data analysis, are described below. Wide community and stakeholder consultation was undertaken prior to, and during, the design of the study, and is described at the end of this section.
3.3.5 Study Population

The study catchment area was located within the Central and Eastern Waikato regions of New Zealand (Figure 3.1) with a total population of approximately 200,000, spread over 10,000 km². This area included one city of over 100,000 people, as well as 10 semi-rural and rural towns.

Figure 3-1 Waikato Region of New Zealand and Location of the Green Prescription Study

Source of map: http://communities.co.nz/Home/Waikato.cfm
Figure 3.1 shows the region in which the Green Prescription randomised controlled trial was carried out. The city and towns that participated in the study (in order of size), included: Hamilton, Cambridge, Thames, Huntly, Ngaruawahia, Waihi, Matamata, Morrinsville, Paeroa, Te Aroha, and Ngatea. (Not included on the map are Ngaruawahia, which is a small town between Hamilton and Huntly, and Ngatea, which is a small settlement between Paeroa, Thames and Pokeno.)

3.3.5.1 General Practitioner and Practice Inclusion and Exclusion Criteria

Inclusion criteria for general practitioners were:

1. Practising medicine within a defined geographical area,
2. Involved in primary health care as their central focus and,
3. Working within a general practice setting.

General practitioners could be part-time or full-time. Exclusion criteria included working in large Accident and Emergency clinics that were unlikely to have a stable practice population or a continuing relationship between patient and doctor or practice nurse.

3.3.5.2 Patient Inclusion and Exclusion Criteria

Patient inclusion and exclusion criteria were the same as for the pilot and validity studies. Therefore, all 40-79 year-old patients visiting the general practice to see the doctor or nurse during the study week were screened for ‘inactivity’ by means of the following question. “As a rule, do you do at least half an hour of vigorous or moderate exercise (such as walking or a sport) on five or more days of the week?” This question, used to establish eligibility, was chosen on empirical grounds to reflect those not achieving the recommended 30 minutes of moderate activity on most days of the week (Centers for Disease Control and Prevention 1996). However, this question had not been validated in the past. Patients were also excluded if they did not speak English, were severely injured or too sick to participate, or were planning to leave the region in the next 12 months.

3.3.6 Sample Size and Power Estimation

Sample size calculations were carried out using STATA software. Calculations of sample size were inflated by the design effect of clustering to ensure that the study was powered
adequately to detect clinically relevant changes in outcomes. The formula for the design effect \((D.\text{eff})\) for clustering is explained in Chapter 2. Predicted change in cardiovascular parameters, physical activity, and quality of life, were used as outcome variables to estimate the required sample size. The largest of these estimates was used as the required sample size, to ensure that any differences in all of the primary outcomes could be detected as statistically significant (Friedman, Furberg et al. 1985). The calculations predicted changes in variables that have been achieved in similar lifestyle-intervention studies in the past and that are of clinical significance. Intraclass correlation coefficients, means, standard deviations, and estimated change, for each outcome variable were derived from the literature and from the pilot study results. These values were used in the calculation of sample size and the adjustment for clustering (Table 3-1).

Sample size calculations were based on estimating differences in change of continuous outcome variables, between the intervention and control groups. As a change in outcome variable may be positive or negative compared with the control, two sided tests were used, and taken into account when sample size was calculated (Friedman, Furberg et al. 1985).

For the estimation of sample sizes required to detect statistically significant differences in change of efficacy variables, a number of other assumptions were made. The intervention and control groups were assumed to be of equal size. An average cluster size of 20 was used in the calculations, as this was the number of participants recruited during a five-day period at each practice during the pilot study. A 25% attrition rate was assumed over the 12 months, and the sample size inflated accordingly.

Table 3-1 shows the sample size required to detect statistically significant differences of change in each variable, between the intervention and control groups (alpha = 0.05, power = 90%). Therefore, a sample size of at least 800 participants from at least 40 general practices with an average cluster size of 20 was required.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline Mean</th>
<th>*sd1 (sd2) if different</th>
<th>Difference in Change</th>
<th>Intraclass Correlation Coefficient</th>
<th>Sample size required without clustering</th>
<th>Sample size adjusted for clustering</th>
<th>Sample size adjusted for 25% attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systolic blood pressure</td>
<td>135 mmHg¹</td>
<td>15 mmHg¹³</td>
<td>4.5 mmHg⁴⁺</td>
<td>0.016⁺¹</td>
<td>468</td>
<td>610</td>
<td>800</td>
</tr>
<tr>
<td>5-year cardiovascular risk</td>
<td>9.6%¹</td>
<td>*2.6%¹</td>
<td>1%¹⁺</td>
<td>0.0036¹</td>
<td>286</td>
<td>306</td>
<td>410</td>
</tr>
<tr>
<td>High-density lipoprotein</td>
<td>1.26mmol/L¹</td>
<td>0.36mmol/L¹</td>
<td>0.11mmol/L¹¹</td>
<td>0.0036¹</td>
<td>452</td>
<td>482</td>
<td>640</td>
</tr>
<tr>
<td>Leisure-time physical activity</td>
<td>1 hr/week¹</td>
<td>*3.0 hr/wk¹</td>
<td>1 hour/week¹⁺</td>
<td>0.05¹⁺</td>
<td>276</td>
<td>538</td>
<td>720</td>
</tr>
<tr>
<td>Vitality score from SF-36</td>
<td>65.6¹²</td>
<td>18.5¹²</td>
<td>6</td>
<td>0.01¹</td>
<td>400</td>
<td>476</td>
<td>640</td>
</tr>
</tbody>
</table>

¹ 'sd1’ refers to the standard deviation of the baseline variable, (or standard deviation of the variable amongst the intervention group, if this data were available); 'sd2’ refers to the standard deviation of the variable amongst the control group if this data were available.  
² Standard deviation of change  
³ (Elley 2000); ⁴ (Arroll 1992); ⁵ (Jackson 1989); ⁶ (Arroll and Beaglehole 1992); ⁷ (Halbert, Silagy et al. 1997); ⁸ (Kerse 1998); ⁹ (Campbell, Mollison et al. 2000); ¹⁰ (Anonymous 1994); ¹¹ (Kerry and Bland 1998); ¹² (Halbert, Silagy et al. 1999); ¹³ (Scott, Tobias et al. 1999)
3.3.7 Randomisation

Randomisation and delivery of intervention was at the level of general practice. To preserve independence of randomisation, computer-generated randomisation of practices was carried out by an independent statistician at a distant site. To assist with between-group comparability, randomisation was stratified by practice size (Friedman, Furberg et al. 1985). Solo general practitioners, two and three general practitioner practices, and practices with more than three general practitioners, made up the three strata.

Practices were enrolled in the study before they were randomised to intervention or control groups. However, allocation of randomisation was known by the recruiter at the time of patient recruitment within each practice. To minimise the potential for selection bias, a systematic process for patient selection, independent of the researcher, was used. In addition, patients remained blind to allocation of randomisation until after enrolment. Pre-randomisation of practices was necessary, prior to patient enrolment to ensure delivery of the intervention on the day of patient enrolment in each practice, and to avoid requiring the patient to return to receive the intervention. Therefore, patient participation and receipt of intervention were maximised, and time between enrolment, baseline measurement, and delivery of intervention, was minimised.

In addition, four practices were randomly selected to have no cholesterol testing done (two control and two intervention practices). This step was taken because of limited study resources, and because the sample size required to detect a significant change in cardiovascular risk in the trial, was estimated to be less than that required for other outcomes.

3.3.8 Study Protocol

The research assistant study protocols for patient enrolment and baseline and follow-up assessments are included in Appendices 11 and 21. General practitioners and general practices were recruited initially. Then all the enrolled practices were randomised. The recruitment of patients occurred over the following 12 months. The process of study
recruitment and enrolment are described below and depicted in Figure 3-2. The 12-month follow-up study schedule is described in a subsequent subsection.

3.3.8.1 General Practitioner and Practice Recruitment

Sixty-three practices, comprising 159 general practitioners were located using the telephone book, lists of general practitioners from a local Independent Practitioners' Association and the local promoters of the Green Prescription at the Sports Foundation.

The principal researcher (author) or her research assistant contacted individual general practitioners or practice managers from all these practices by telephone, to invite their participation in the trial. Following this, a letter was sent which explained the aims and design of the trial and what was expected from participant practices. Where requested, the researcher would visit the practice to present the aims and study protocol, before practitioners decided whether they would participate. Participation was voluntary and there was no remuneration offered. The general practitioner or practice manager was then re-contacted a week later to ascertain whether they agreed to participate. If some, but not all, general practitioners at a practice agreed to participate, the practice was still included in the study, but only patients of participating general practitioners were screened and invited to participate. All practices were contacted and enrolled in the study over a six-week period during February and March 2000.

3.3.8.1.1 Patient Enrolment

Rolling recruitment of study participants was undertaken over the following twelve months, alternating control and intervention practices as much as possible. As it was expected that physical activity would vary depending on the season, an even spread of patient enrolment was obtained throughout the year. Furthermore, as the follow-up assessment took place exactly 12 months after the initial assessment, the season was the same at baseline and follow-up for each practice and study patient.

The researcher, or research assistant, normally spent four to five consecutive days in the waiting room of each participating practice, enrolling eligible patients from the waiting room. Occasionally, to fit in with the practice, or the 'half-days' of the general
practitioners, the 4-5 days spanned two weeks. The predicted recruitment rate was between 8 and 10 patients per full-time equivalent general practitioner, as this rate had been achieved during the pilot study.

At each practice, the practice-receptionist handed out a short form to each patient aged between 40 and 80 years, as they arrived at the practice during the study week (Appendix 5). The form screened for ‘inactivity’ and invited those eligible to participate in the study. The form was filled out by each patient and handed back to the receptionist, independent of the researcher. In addition, a notice was put up on the notice board of each practice during the week of recruitment (Appendix 13).

If the patient was interested in taking part and was eligible for participation, the researcher explained the nature of the study and provided one of two information sheets (Appendix 6a or 6b). The information sheets were tailored according to whether serum lipid testing was required. This requirement depended on whether the patient was in a practice that had been randomised to have blood lipid testing done or not. Once informed consent had been obtained (Appendix 7a or 7b), and at least the physical activity status of the patient ascertained, patients in the intervention practices received the intervention from their general practitioner or practice nurse. Patients in the control practices received usual care. All patients were followed up by the research team 12 months later to repeat the study measures.

If patients were known to be ineligible due to exclusion criteria, prior to enrolment, they were not enrolled. However, cholesterol levels, blood pressure readings, and some clinical conditions were not known until after enrolment. Therefore, all patients enrolled, regardless of baseline levels of cholesterol and blood pressure or debilitating/unstable condition, were included in the study follow-up and analysis, despite the fact that they did not meet the criteria. This is a conservative approach and likely to dilute the power of the study, as this group is less likely to have received the intervention from their general practitioner on clinical grounds. Alternatively, their condition may influence the response to the intervention. However, inclusion of all those enrolled is recommended, to reduce the potential for exclusion and recruitment bias when the researcher is not blinded to
allocation of randomisation (Friedman, Furberg et al. 1985; Schulz, Chalmers et al. 1995). Therefore, no patients were excluded from analysis after enrolment.

Figure 3-2 Flow Diagram of the Design of the Cluster Randomised Controlled Trial of the Green Prescription Intervention

All General Practices in the Central and Eastern Waikato invited to take part in the study

Those practices that agree are stratified by size and randomised into two groups

Control
1. Researcher or research assistant in waiting room for 5 days, enrolling 20 patients per practice, on average
2. Receptionist hands every 40-79 year-old patient an activity screen and invitation to participate in study
3. Researcher enrols interested, eligible patients:
   - informed consent
   - questionnaires
   - measurements (before and after GP)
4. Usual care from GP

Intervention
1. Researcher or research assistant in waiting room for 5 days, enrolling 20 patients per practice, on average
2. Receptionist hands every 40-79 year-old patient an activity screen and invitation to participate in study
3. Researcher enrols interested, eligible patients:
   - informed consent
   - questionnaires
   - measurements (before and after GP)
4. Patient takes a prompt to GP or nurse, to indicate to give Green Script and activity counselling
5. Copy of Green Script faxed to exercise specialists

Three telephone-calls from exercise specialists over next 3 months to provide advice and encouragement

GP or practice nurse, provide feedback about physical activity on subsequent visits or by telephone

All participants telephoned 12 months later for face to face visit, for repeat data collection, at original practice. Volunteers in control group are offered intervention after follow-up measures, if appropriate
3.3.9 The Intervention

During 1999, all primary care clinicians in the Waikato had been informed about the use of the Green Prescription and all were offered at least two hours of training in the use of motivational interviewing techniques. After randomisation of practices, the intervention general practitioners and their practice nurses were offered an additional two-hour educational session about how to give physical activity advice and the Green Prescription, using motivational interviewing. The Tobacco, Alcohol and Other Drug Early Intervention Project (TADS) team, from the Goodfellow Unit at the University of Auckland, provided the motivational interviewing training in this study. Those general practitioners and nurses, who were too far from the training venue to attend, were provided with written resources, courtesy of the Goodfellow Unit.

When patients were enrolled in the study, they filled out the ‘stage of change’ form (Appendix 8). At the intervention practices, the patient was also given the ‘stage of change’ form to hand to the general practitioner as the prompt. These forms indicated to the general practitioner or nurse that the patient was enrolled in the study and that they needed to give appropriate physical activity advice and the Green Prescription. The form also indicated to the clinician, the stage of change of the patient, with regard to their present level of physical activity participation, or their level of ‘contemplation’, as discussed in Chapter 2 (Prochaska and DiClemente 1983).

The physical activity goals decided upon by the clinician and patient were usually home-based, walking, or group exercise activities. These goals were written on a standard Green Prescription and given to the patient (Appendix 9). General practitioners sometimes referred the patient to the practice nurse to give the advice and Green Prescription. If patients had booked consultations with the practice nurse rather than the general practitioner, the nurse was responsible for the intervention. The nature of advice given and the type and amount of exercise recommended by the general practitioner, or practice nurse, were left up to the individual health professional.

General practitioners and practice nurses were provided with Green Prescription pads, pamphlets about walking, stretching and the benefits of physical activity for medical
conditions, fridge magnets and record sheets for patients to record which days they achieved their activity goals. These resources were supplied by the Hillary Commission and made up into packs for the general practitioner or nurse to hand to the patient with their Green Prescription, so that each intervention patient would receive the same resources.

A copy of the Green Prescription was faxed to exercise specialists at the local Sports Foundation with the patient's consent. Relevant details such as age, weight and particular health conditions were often included (Appendix 10). The exercise specialists made at least three telephone calls to each enrolled intervention patient (unless after repeated attempts, contact could not be made) and sent out tailored resource materials when appropriate, over the next three months. Intervention patients received quarterly newsletters about the benefits of physical activity and other motivational material. The exercise specialists would also help the person to design appropriate exercise programmes for themselves or direct them to community-based exercise groups suitable for their age, capabilities, and geographical location. All activities in the programme were of 'moderate-intensity' and unlikely to cause harm. The exercise specialists were informed of any relevant medical condition of the patient so that physical activity advice could be tailored accordingly. Motivational interviewing techniques were also used by the exercise specialists.

The staff of each general practice was encouraged to provide feedback and encouragement to participants on subsequent visits to the practice.

3.3.10 Outcome Measures

Change in physical activity levels, systolic and diastolic blood pressure, 4-year coronary heart disease risk-status, and quality of life using SF-36 scores, over a twelve-month period, were the primary efficacy variables for analysis. Self-administered questionnaires and objective measurements using electronic equipment and signed witness statements verifying results were used as much as possible, in order to minimise the potential for bias due to the lack of blinding of researcher and patient.
Baseline data collection was carried out according to a protocol developed by the author (Appendix 11). Data collection required the filling out of several measurement forms (Appendices 12). Due to the patient recruitment method of screening on the day of the practice visit, there was sometimes only a period of ten minutes available for baseline measurements, prior to the patient’s appointment with the general practitioner. Although baseline physical-activity status was determined prior to the commencement of the intervention, other baseline variables were sometimes measured after the consultation with the clinician.

3.3.10.1.1 Physical Activity

Physical-activity outcome measures included change in total energy expenditure, and leisure-time moderate or vigorous-intensity physical activity. All physical activity outcomes were calculated from self-report. Usual levels of physical activity were established using the validated AHS questionnaire (Appendix 1) (Arroll, Jackson et al. 1991; Elley, Kerse et al. 2003a). This questionnaire and the method of its validation are described in detail in the first part of this chapter. The AHS questionnaire, rather than the GSS questionnaire, was chosen to assess physical activity because of the results of the validation study, which are presented in Chapter 4, section 4.2.

3.3.10.1.2 Blood Pressure and Cardiovascular Risk

Gender, age, previous cardiovascular disease, diabetic and smoking status, and cardiovascular medication use, were established using a questionnaire and by referring to patient records where necessary (Appendix 12). Pulse and blood pressure recordings were obtained using a Speidal and Keller OSZ 5 electronic sphygmomanometer, after at least five minutes of sitting quietly. Three recordings were taken and the average of the second and third readings was used in outcome variable calculations. This method has been used in previous trials to reduce the tendency of regression to the mean of baseline blood pressure recordings (Friedman, Furberg et al. 1985). Sphygmomanometers were recalibrated on a regular basis to ensure standardisation of results. The Speidel and Keller OSK electronic sphygmomanometer devices meet the requirements of the European standard for clinical investigation.
Random cholesterol values were obtained from venous blood taken by the researcher. The sample was analysed at a certified laboratory on the day of collection. These variables were required for calculation of the 4-year cardiovascular and coronary heart disease risk scores, using the Framingham and D’Agostino equations (Anderson, Odell et al. 1990; D’Agostino, Russell et al. 2000).

3.3.10.1.3 Quality of Life

Quality of life scores were established by self-administration of the SF-36 questionnaire. The Rand Corporation outcomes short-form 36 health survey (SF-36) is a multidimensional measure of self-perceived general health status or quality of life. This survey has been validated overseas in adult populations in the United States, Britain, and Australia. It has been shown to be suitable for a wide range of age groups and ethnic populations (McHorney, Ware et al. 1993; Medical Outcomes Trust 1993; McHorney, Ware et al. 1994; Hayes, Morris et al. 1995). The SF-36 can also be used to measure changes in health status after medical or lifestyle interventions, and is quite sensitive to subtle changes in health status (Bouchet, Guillemin et al. 1996). It has also been validated for use in primary health care (Brazier, Harper et al. 1992). The SF-36 has been validated, in New Zealand (Wheadon, Kokaua et al. 1994; Scott, Tobias et al. 1999) and has been used amongst older Māori (Durie, Allan et al. 1997). Quality of life was assessed using the eight separate outcomes of the SF-36. These included:

- Physical functioning
- Role physical
- Bodily pain
- General health
- Vitality
- Social functioning
- Role emotional
- Mental health
3.3.10.2 Secondary Outcomes

Changes in the serum lipid concentrations, weight, and body mass index, were also compared between control and intervention groups, for the twelve months between baseline and follow-up. Weight was taken with Digital scales from Wedderburn, which is a certified calibration laboratory.

3.3.10.3 Other Information Collected

3.3.10.3.1 Socio-economic Status

All study participants were asked if they had a community services card (CSC). This card is a health-user card that allows for a government subsidy for primary health care and medications. Eligibility for the card is means-tested on family income. Consequently, the participants were divided into two groups, those with a CSC and those without. Educational achievement level was recorded. Socio-economic status (SES) was also collected at follow-up by means of a validated SES scale that had been derived from the 1996 census in New Zealand. This scale gives each family a score from 0-100 (Davis, McLeod et al. 1997). All study participants were asked their current occupation, their spouse’s occupation, or if retired, their previous occupation and their spouse’s previous occupation. The highest SES of the four categories was allocated as that individual’s SES.

3.3.10.3.2 Ethnicity

Ethnicity status was established by asking the participant to choose as many ethnic groups as they identify with (Appendix 12). This question was based on the ethnicity question asked in the New Zealand 1996 census.

3.3.10.3.3 Medical Conditions and Medications

Changes in medications over the 12 months as well as baseline medications were recorded. Medications and medical conditions were coded according to a system devised by the author (Appendices 14-15). Only those medications or medical conditions that
were likely to have a bearing on cardiovascular or quality of life variables, or were likely to change due to physical activity changes, were included.

3.3.10.3.4 Smoking Status and Alcohol Consumption

Smoking status was recorded as ‘currently smoking’, ‘given up smoking in the previous 12 months’, or ‘non-smoker’. The middle category was included because the Framingham equation classifies those having given up smoking in the previous 12 months as having the same risk as those still currently smoking. Frequency, amount, and type of alcohol consumption, were recorded for each participant at baseline and follow-up.

3.3.10.3.5 Adverse Effects

Self-reported falls in the previous month, injuries in the previous month, and hospitalisations during the previous year, were collected at baseline and 12-month follow-up. Self-reported use of outpatient clinics, physiotherapists, and occupational therapists 12 months prior to and 12 months after baseline were also compared amongst the intervention and control groups. Besides representing potential adverse outcomes of the intervention, these variables were also collected to compare potentially prognostic characteristics of control and intervention subjects at baseline. Hospitalisation rates were also important in the cost-effectiveness evaluation, presented in the next section. Self-report of hospitalisations was verified by comparing hospital records using a patient unique identifier (NHI).

3.3.10.3.6 Practice Characteristics

Information about each practice was recorded, including their fee schedule, funding structure, and personnel. In order to ascertain rates of patient participation, the following were recorded by the researcher or research assistant at each practice (Appendices 16 and 17):

- Number of 40-79 year old patients attending each practice during the study week
- Number screened for the study
Number found to be inactive
Number that agreed to participate in the study
Number excluded (and reasons for exclusion)

Leisure activity levels and demographic information about participating general practitioners and practice nurses was collected after informed consent was obtained (Appendices 18 and 19) at the end of the patient recruitment week.

3.3.11 Follow-up Visit: Description and Schedule

Rolling recruitment of patients for the study had occurred from mid April 2000 until mid April 2001, according to plan. Twelve-month follow-up commenced mid April 2001, with the patients that had been recruited exactly one year previously. Patients were followed up 12 months after baseline assessment, at the original general practices. If the patients had moved within the region, or transport to the surgery was a problem (e.g. unwell, immobile or without transport), follow-up was carried out at the individual’s home or at another practice nearby, although this was rarely necessary.

To facilitate study follow-up, each patient was contacted by telephone or mail and a follow-up appointment was arranged with the researcher or research assistant. Follow-up was spread evenly over the next year, and was finished according to plan in April 2002.

Outcome measures were repeated using similar but extended forms (Appendix 20) according to the follow-up protocol for all intervention and control patients (Appendix 21). These measures allowed the calculation of change in primary and secondary outcomes over the twelve months between baseline and follow-up.

3.3.11.1 Subjective Experience of Intervention

Subjects were also asked whether they had received a Green Prescription in the last 12 months. (This allowed one assessment of potential contamination — control subjects who received the intervention.) All subjects who had received the intervention were asked:

- Whether they had initially increased their activity
• The frequency and duration of any further discussions with the health professional about the intervention
• Whether they felt they had increased their physical activity as a result of the intervention
• Whether they were still doing this activity
• Whether they did their increased activity alone or with others
• How helpful they found the follow-up support from exercise specialists
• What they thought about the intervention (open-ended question)

3.3.11.2 Measures of Compliance, Contamination, and Co-intervention

The numbers of study patients who did not receive the intervention, and the reason why (such as decline by the patient or considered unsuitable by the doctor or nurse after enrolment), were recorded. In addition, the number of patients who received a Green Prescription but declined having the script faxed to Sport Waikato was also recorded. Those who initially increased their activity but discontinued were also noted at follow-up. All subjects, regardless of their participation in the intervention or activity levels, were invited to attend the follow-up visits, and all subjects were included in final analysis, regardless of compliance.

Recording and comparing medications of intervention and control participants, was carried out at baseline and follow-up, to check for medical co-intervention, such as the prescribing of extra antihypertensive medication to intervention patients, by general practitioners. In addition, asking control participants after follow-up if they have received the intervention, and asking general practitioners in control practices if they have been delivering interventions, supplied another check for co-intervention and contamination of intervention to the control group.

To address any ethical concerns or bad feelings over withholding an intervention from control participants, this group was offered the intervention after the completion of the trial. This offer may also have helped to minimise the chances of control patients seeking out a similar intervention during the trial. It is important to note that all study participants...
were aware that they were taking part in a study where half of the participants received a lifestyle intervention from their general practitioner.

3.3.12 Patient and General Practitioner Feedback

At baseline and follow-up, each study patient was given a form with their average blood pressure, pulse, weight and height recorded on it (Appendix 22). Copies of the cholesterol results were also sent to the study patients. All participants that indicated that they would like a copy of the results of the study, were sent a summary of the results after the study was completed (Appendix 23). Approximately 600 of these letters were sent. A letter summarising the results, as well as individual practice results, was also sent to all participating general practitioners after the study was completed (Appendix 24).

3.3.13 Data Management

Data collection was undertaken from April 2000 until April 2002. The development of data collection protocols and training of research assistants is described below. Data coding, double data entry, data cleaning, and data analysis, continued until October 2002. These topics are also described below, as are other procedures put in place to monitor data quality.

The baseline and follow-up protocols were developed by the principal investigator (author of this thesis), prior to the commencement of the baseline and follow-up visits respectively (Appendices 11 and 21). Funding for one 8/10th research assistant was provided for the principal investigator for the two years of data collection. Three research assistants shared this position including a Māori research assistant to collect data from each of the seven Māori Health Clinics. The involvement of a Māori research assistant for Māori Clinics had been recommended during consultation with Māori prior to the commencement of the study. The three research assistants were registered nurses. They had also received training in venepuncture for determination of serum lipid concentration at baseline and follow-up. Half of the data were collected by the principal investigator and the other half by the research assistants.
The principal investigator also trained the research assistants in coding of the data, using a physical activity compendium (Ainsworth, Haskell et al. 1993), and a socio-economic scale (Davis, McLeod et al. 1997). Instruction was given on coding of medical conditions and medications according to the coding system (Appendices 14 and 15). Half of the data were entered into the computer by the principal investigator and the other half was entered by data entry personnel, who were also trained by the principal investigator. All data were re-entered by an independent person for comparison.

3.3.13.1 Data Monitoring and Quality Control

Data collection was standardised by the use of protocols and regular sessions, and by research assistants sitting in and observing data collection with each other at regular intervals, throughout the trial. As much as possible, questionnaires were filled out by the subjects. The research team were versed in ways to clarify or reflect back queries about the questions, rather than make suggested answers, to minimise their influence and risk of potential bias.

Data collection was made as complete as possible. However, with the opportunistic nature of patient recruitment, time was sometimes a factor. Data collection for each subject took about 20-40 minutes. On rare occasions, the patient was not able to stay for the duration of data collection. On those few occasions, the SF-36 questionnaire was given to the subject to complete at home and send back. Some of these were not returned. Other sources of missing data included subjects choosing not to have cholesterol testing done, or subjects declining or unable to attend the follow-up. Every attempt was made to encourage subjects to return for follow-up, if possible. Figures on follow-up rates, and reasons for attrition from the study, are presented in Chapter 4, section 4.3.4.

Calibrated electronic sphygmomanometers and research electronic scales were used in measurements. Re-calibration of sphygmomanometers was done every six to twelve months. Blood pressure, pulse and weight readings were checked by study subjects, who signed witness statements to verify the values recorded. One accredited laboratory was used to analyse all serum samples.
To minimise errors and variability, all forms, questionnaires and procedures were pre- tested on volunteers, and pilot study patients, prior to their use in this study. This pre- testing also ensured the acceptability and feasibility of the study procedures and forms, and clarified any potential areas of misinterpretation.

Coding of each individual's results was checked by two members of the research team to ensure standardisation of coding. The use of compendia, socio-economic scales and look- up coding tables reduced the scope for coder interpretation or inter-rater variability. Regular meetings were also held and samples of coding were discussed amongst the team. Any assumptions that had to be made were written down for future reference by the team and discussed.

Data were entered using an ACCESS program that was specifically developed for the project. The program calculated all physical activity and cardiovascular-risk outcome variables. Double data entry was carried out and coding checked a further time by an independent researcher to validate data entry. Each piece of data was individually compared between the original and re-entered databases, using the ACCESS program, and all discrepancies reviewed by two members of the research team with the original questionnaire to decide on appropriate interpretation.

The data were also checked for any extreme or outlying values (eg: cholesterol or blood pressure values) and verified wherever possible.

3.3.14 Analysis

The positive predictive value of the single screening question for 'inactivity' was assessed by comparing the results with those from the AHS questionnaire, which had been validated. The proportion of those identified by the screening question as 'less active' was compared with the 'less active' proportion identified by the AHS questionnaire. 'Less-active' referred to those achieving less than 2 ½ hours of leisure- time moderate- or vigorous-intensity physical activity per week. The positive predictive value results are presented in Chapter 4, section 4.3.9.
Physical activity variables were calculated in the same way as in the questionnaire validity study and the pilot study described above (section 3.2.7). However, the MET value ascribed to sleep was revised from 1.0 to 0.9, and ‘light activity’ was allocated the MET value of 1.2 instead of 1.5. These revisions were made in the light of subsequent information (Personal communication with an exercise physiologist) (Wilson, Paffenbarger RS et al. 1986; Lamonte and Ainsworth 2001).

Cardiovascular risk scores were estimated for all those less than 75 years of age, as this was the upper age limit for which the equation was valid (Anderson, Odell et al. 1990). Four-year coronary heart disease and cardiovascular risk scores were estimated using the Framingham equation for those without previous cardiovascular disease (Anderson, Odell et al. 1990). For those with previous cardiovascular disease, the D'Agostino equation was used to estimate 4-year coronary heart disease risk (D'Agostino, Russell et al. 2000). The risk of a cerebrovascular event, using the Framingham equation, was added to the D'Agostino estimated risk of coronary event, to estimate overall cardiovascular risk for those with previous cardiovascular disease. Four-year risk was used, as this was the only risk-period predicted by both the Framingham and D'Agostino equations. Data were transferred from ACCESS to Excel and SPSS (10.1) to compare baseline variables and to calculate the eight quality of life variables from the SF-36.

Intracluster correlation coefficients (ICC) were calculated using one-way analysis-of-variance (ANOVA) models within STATA. These models can cope with datasets with large numbers of levels of analysis and inter-group variability. Besides the point estimate of the ICC, the model also estimates 95% confidence intervals and F-statistics. The latter is an indicator of the ratio of the ‘between-group’ variance to the ‘within-group’ variance. The calculation of these statistics is useful for sample size calculations for future study designs. However, these statistics also indicate the demographic and clinical characteristics that are likely to vary more between practices than within, and which characteristics are more likely to be evenly spread.

The STATA statistical program was also used to undertake regression analyses of baseline data. Baseline analysis was carried out using regression models, which allowed
for the correlated data (clustering). General practice was used as the clustering variable and inserted as a random effect. Baseline physical activity data were not normally distributed, so log-transformation of this data was undertaken prior to regression analysis of baseline data. Regression analysis was carried out with all variables hypothesised to be related to physical activity and cardiovascular variables. All variables showing significant relationships (p<0.05) were included in the final multiple regression models. Age and sex were controlled for in each model and allowance was made for clustering by practice.

Final analysis used similar regression models for continuous outcome variables, entering general practice as the clustering variable. Intervention group status was entered as the independent variable in the regression. Changes in outcome measures over the twelve months following baseline readings were the dependent variables. All variables were checked for normal distribution. As change in outcome measures were used as efficacy variables, these were all normally distributed including physical activity data, so log-transformation was not necessary, for final outcomes analysis. Differences between intervention and control groups in change of continuous outcome variables were analysed using random-effects linear regression models in STATA 7.0 (generalised-least-squares). Final analysis of proportions and dichotomous outcomes was conducted using a non-linear mixed model (with random and fixed effects) in SAS 8.2, to allow for clustering by practice. Sensitivity analyses were conducted controlling for baseline and potential confounding factors.

An intention-to-treat analysis of all patients enrolled in the study was performed according to allocation of randomisation, regardless of receipt of intervention or compliance with the intervention. This approach is particularly important for pragmatic trials of effectiveness to account for real-life conditions of intervention delivery and variable compliance (Hollis and Campbell 1999). Baseline results of outcome variables were carried forward for those who did not attend follow-up, for all variables that tended to improve over time. For variables that showed deterioration over time, such as risk of coronary heart disease event, the mean change of the control group was used to estimate change for non-attendees. However, others have suggested that such a conservative approach may risk a type-2 error (Fergusson, Aaron et al. 2002). Therefore, sensitivity
analyses were performed ‘per protocol’ including only those who returned for follow-up in the final analysis, to observe if this changed results substantially.

Analyses of blood pressure were adjusted for medication change, within the regression model, by entering ‘change in antihypertensive and/or lipid-lowering medication’ as a co-variable. In addition, increases and decreases in dose, and change of type, of antihypertensive medications were analysed to look for any significant change or difference between groups.

All statistical analyses were undertaken by the author, except for the analysis of proportions and dichotomous variables, which were carried out by Elizabeth Robinson using the SAS statistical program. The results from all analyses are presented in Chapter 4 section 4.3.

3.3.15 Organisation

3.3.15.1 Research Personnel

The author was the principal investigator, who conducted the trial from her home in Te Aroha, and was funded by the Heart Foundation. Three registered nurses were employed part-time to aid with data-collection. Two of the nurses, and administrative and data-entry support, were based at the Pinnacle Independent Practitioners’ Association in Hamilton. These research support personnel were funded by the Hillary Commission. Two other data-entry personnel and the third nurse were based in Te Aroha. The reference group and supervisors of the study were based at the University of Auckland.

3.3.15.2 Reference Group and Consultation

An advisory committee was formed to comment on the research proposal, as it was being developed. The committee included the two supervisors, Dr Ngaire Kerse, senior lecturer in general practice and primary health care, and Associate Professor Boyd Swinburn (from the Department of Community Health and Director of the Heart Foundation). Sharon Matangi-Nixon was project Māori advisor and local Waikato practice nurse. Professor Rod Jackson, who has significant experience in cardiovascular risk assessment,
as well as physical activity and health research, and Professor David Thomas who advised on qualitative research methodology were both representatives from the Community Health Department at Auckland University. Health economists, also from the department of Community Health, Dr Toni Ashton and Associate Professor Richard Milne, commented on cost-effectiveness design and evaluation. Dr Ross McCormick, Director of the Goodfellow Unit, and Associate Professor Bruce Arroll of the Department of General Practice and Primary Health Care, both of the University of Auckland were also members of the reference group. Bruce Arroll took over as supervisor from Boyd Swinburn in the second year, when Boyd Swinburn took up a new academic position overseas. Hillary Commission representatives included Diana O’Neill, national co-ordinator of the Green Prescription initiative in Wellington and Hayley Gaddes, co-ordinator of the Green Prescriptions initiative in the Waikato. Dr Steven Lillis, a Waikato general practitioner, Waikato Postgraduate Medical Society tutor, regional general practice Training Programme co-ordinator, Medical Council examiner and director of a large Waikato Independent Practitioner’s Association was also a member of the reference committee. The Waikato Sports Trust and local general practitioners were consulted during the development of the research design.

Financial and other support of the project were provided by the National Heart Foundation of New Zealand, the Hillary Commission, the Waikato Medical Research Foundation, the Royal New Zealand College of General Practitioners Research and Education Charitable Fund, the Pinnacle Independent Practitioners’ Association and the University of Auckland. The Waikato Ethics Committee approved the study design in November 1999.

3.3.15.3 Consultation with Māori

A discussion was held with Dr Sue Crengle in March 1999, Department of Māori and Pacific Health, Auckland Medical School, concerning appropriate consultation with Māori. Consequently, prior to the development of the study design, consultation was carried out with the local representatives from Hauraki and Tainui Māori health providers and Iwi. Hauraki Māori Trust Board C.E.O, Te Korowai Haoura O Hauraki general
manager, a Hauraki Kuia and Te Korowai Haora O Hauraki trust board member, and present and past Te Korowai health providers were consulted. A Tainui Kuia and nurse, personnel from Tainui Health Clinics, Whare Haoura, and Ngamiro Health Centre, and a Tainui Trust Board member were also consulted and sent information about the aims and proposed approach of the study. Discussions about physical activity counselling and its effect within the Māori community and the proposed study were carried out with these groups from June 1999 until commencement of the study, April 2000.

Consultation was also undertaken with Te Hotu Manawa Māori (The Māori Heart Foundation) prior to final development of research design. Alterations were made to the research protocol in line with their recommendations.

Following formal and informal consultation, the research design was formulated and final approval from Te Hotu Manawa Māori was obtained. Mr Wayne McLean, CEO of the Tainui Health Providers was contacted and his approval sought before individual Māori health providers were invited. After his approval was obtained, Tainui and Hauraki Māori Health Providers in the Central and Eastern Waikato were invited to take part if they wished. All Tainui health providers invited agreed to participate. Although the one Hauraki Māori Health Provider agreed to participate, due to delayed final approval, it was not possible to include them in the study, as the study had already commenced.

A provisional progress report of baseline findings, especially those pertaining to Māori participants, was provided for Tainui Kuia, Te Aroha Tai-Rakena of Ngamiro Health centre in August of 2001. The principal researcher was also interviewed on Radio Tainui about the project. A final report was sent to Māori Health Providers and Te Hotu Manawa Māori in 2003.

3.4 Methods of Cost Effectiveness Study

3.4.1 Aims

With limited health resources, it is important to assess the cost-effectiveness of interventions to justify their funding. The cost-effectiveness of the Green Prescription
intervention was therefore assessed during the randomised controlled trial. The aims of the cost-effectiveness study were:

1. To calculate incremental cost-effectiveness ratios for the Green Prescription programme compared with usual care in general practice.

2. To compare the cost-effectiveness of the Green Prescription programme with other physical-activity promotion programmes based in primary health care or the community.

3.4.2 Design

The cost-effectiveness analysis of the Green Prescription programme was designed and incorporated prospectively into the trial. The design of the cluster randomised controlled trial is described above. The cost-effectiveness evaluation took a societal perspective, which takes into account changes in indirect costs to the health care provider and the individual, changes in productivity, and direct programme costs. However, component costs are presented to allow analysis from a programme funder’s perspective also, in order to facilitate comparison with other studies.

3.4.3 Outcome Measures

The primary outcomes measured for the cost-effectiveness study were the incremental cost of change, over a one-year period, in self-reported physical activity. Physical activity was measured in terms of total energy expenditure and leisure-time moderate- and vigorous-intensity activity energy expenditure, which allowed comparison of cost-effectiveness with two community-based physical-activity interventions from the literature (Sevick, Dunn et al. 2000). In addition, the cost of moving one sedentary person into the ‘active’ category was also calculated to facilitate comparison with the cost-effectiveness of another primary care-based physical-activity intervention (Stevens, Hillsdon et al. 1998).

3.4.3.1 Cost-Effectiveness Variables

The cost of the programme was calculated by costing all components of the intervention development and delivery. This included general practitioner and practice nurse time,
government subsidies, and cost to the patient. Green Prescription resources, Sport Foundation and Hillary Commission personnel and overhead expenses for the twelve months of intervention, as well as set-up and training costs, were obtained from the respective organisations. Self-reported costs to the individual were also compared between the control and intervention groups. These costs included exercise equipment purchased, sports club or exercise group subscriptions, travel expenses to and from exercise, and any other costs associated with exercise identified by the study participants over the twelve months (Appendix 20).

Primary and secondary health care offset costs were calculated for each individual participant for the 12 months prior to study enrolment and compared with the twelve months after study enrolment. Furthermore, the costs of time off work, accident compensation costs, and costs of allied health therapies, such as physiotherapy, for each individual study participant were compared before and after study enrolment. Again, this was achieved by ascertaining costs for the 12 months prior to, and 12 months immediately after, study enrolment and intervention delivery. Details of cost determination and sources of information for each costing category are presented below.

3.4.3.2 Programme Costs

The Hillary Commission funded the development and implementation of the Green Prescription programme. This included the development, production and distribution of resources, such as exercise pamphlets, fridge magnets, physical activity-scoring cards, and other educational material for general practitioners and patients. These costs also included the installation of free telephone access for patients on the programme. The Commission further funded the salaries of regional programme co-ordinators, a public advertising campaign, and general practitioner training. The national-level costs to set-up and co-ordinate the programme, were obtained from the Hillary Commission, who were responsible for the programme.

All costs incurred in previous years, were adjusted for inflation using the December consumer price index (CPI) from each corresponding year compared with the December 2001 CPI (Statistics New Zealand 2003). A discount rate of 5% was used to calculate
present equivalent values of programme costs from 1996 to 2001 (Drummond, O'Brien et al. 1997). All costs are reported as New Zealand dollars. Where comparisons with programmes from the United States or the United Kingdom were carried out, values were converted to the New Zealand dollar according to the exchange rate of December 2001 (Reserve Bank of New Zealand 2003).

Actual regional Sports Trust personnel and overhead costs, associated with the programme, were obtained from the Trust's accounting department for the year 2001/2002. Average wage costs rather than marginal costs were used, as the exercise specialists were permanent staff of the regional Sports Trust. Delivery costs of the programme within the general practice were estimated using usual consultation charges for participating practices, national award rates for practice nurses, and the time, estimated by general practitioners and practice nurses, for programme delivery.

3.4.3.3 Patient Exercise Costs

Costs to the patient, which were hypothesised to be linked to the intervention, were obtained from participant questionnaires during the trial. These costs included self-reported purchase of exercise or sports shoes, membership fees to exercise groups, costs of exercise equipment, other costs associated with exercise, and travel to and from the location of exercise or physical activity. The costs were calculated for the 12-months between baseline and follow-up. The average costs for the intervention group were compared with the control group to obtain the incremental cost of exercise to the patient. Transport costs to and from exercise-venues were calculated using self-reported distance to activity venue and a running cost of 16.6 cents per km. This running cost was based on calculations carried out by the New Zealand Automobile Association for an average size car of 1601 to 2000 cc engine capacities. (Figures supplied by the New Zealand Automobile Association.) The fixed cost of 50 cents per kilometre, which is calculated from the estimated total daily fixed cost of owning a car, was not included. It was assumed that the study participants would have owned the vehicle regardless of the exercise.
A conservative intention-to-treat analysis was carried out. As intervention participants tend to spend more on exercise-related costs than control participants, it was assumed that those who did not attend follow-up had costs equivalent to their corresponding group.

3.4.3.4 Primary and Secondary Health Care Offset Costs

Costs to the patient and to the health funder were recorded for the year prior to each patient’s enrolment in the study and compared with corresponding costs for the year following enrolment. For each study participant, actual number and type of general practice consultations were obtained from practice records. Numbers of accident-related visits to physiotherapists, chiropractors and osteopaths were obtained from patient questionnaires. Usual consultation charges for each general practice in the region were obtained at baseline, and average charges calculated for each consultation type.

As the consultation charges to patients were taken half way through the two-year period, the same estimations of charges were used in calculations for the year prior and the year after baseline. However, it was also assumed that consultation charges increase from year to year, approximately equal to inflation. Therefore, using the one rate for both years meant that adjustment for inflation was not required. Actual government subsidies for each type of consultation were used. These subsidies did not change over the study period. Health funder costs were therefore adjusted for inflation using the consumer price index ratios of December 2001 compared with December 2000.

Patient charges and subsidies vary according to the type of consultation (e.g. accident or non-accident) and the socio-economic and health-service utilisation status of the patient. During the 2000-2001 period, patient charge-rates for each participating practice were recorded for patients with a low-income ‘community-services-card’ (A1) or a high-user-card (AZ) (those who have consulted with the general practitioner at least 12 times in the previous 12 months). The charges for those without a subsidy card (A3) were also recorded. Government subsidies for each general practitioner consultation were $15.00 for A1 and AZ visits (non-accident-related A3 consultations receive no government funding), and $26.00 for all accident-related visits. The Accident Compensation
Corporation reimbursed ACC-registered physiotherapists, chiropractors and osteopaths at a rate of $19.00 per visit. (Personal communication with ACC.)

Each person in New Zealand has a National Health Identifier (NHI) number, which operates as a unique identifier in primary and secondary health care. The use of this identifier allows tracking of individual primary and secondary health care utilisation. Actual hospital in-patient, out-patient and investigation costs were obtained from the local district health board, which recorded actual costs for each patient from all public regional and base hospitals in the district. These costs were adjusted for inflation using the CPI ratio explained above. Costs for private hospital-use could not be obtained and were excluded from the analysis. However, self-reported private hospital admission-rates were recorded.

The intention-to-treat analysis of health care costs was conducted assuming that there was no change in the rate of health care provision for those in whom data were missing at follow-up.

3.4.3.5 Loss of Productivity Costs

The change in the number of days of illness- and accident-related leave taken from the year before and the year after enrolment in the study, for the intervention and the control groups were obtained from self-report. The average wage for the June quarter from wages, salary and self-employment for those in paid employment was $121.80/day for 2000 and $128.20/day for 2001 (Statistics New Zealand 2003). These figures were used to calculate the cost of loss of productivity due to illness and accident for the year prior to baseline compared with the year after baseline. All costs were adjusted for inflation using the 2001/2000 CPI index ratio to calculate the incremental change in loss of productivity in the intervention and control groups.

The intention-to-treat approach assumed that there was no change in productivity costs over the two years for missing data and for those retired.
3.4.4 Analysis

Actual costs were collected wherever possible for each participant. The differences in change in healthcare and productivity offset costs to the patient and funder for intervention patients compared with control patients, were analysed using a random effects multiple regression model, allowing for clustering by practice, in STATA. Differences in the costs associated with exercise for participants for intervention participants compared with control participants were calculated using the same regression model. Offset costs were included in the final calculation of cost-effectiveness ratios, regardless of whether the difference reached statistical significance. This was because the exclusion of non-significant cost differences may risk a type-2 error because of inadequate sample size. However, 95% confidence intervals were included to demonstrate the range of uncertainty.

Cost-effectiveness ratios were calculated for the Green Prescription programme compared with usual care for incremental changes in physical activity. These outcomes were chosen to allow comparison with other cost-effectiveness analyses in the literature (Stevens, Hillsdon et al. 1998; Sevick, Dunn et al. 2000). All analyses were carried out using an intention-to-treat approach, as stated above. The results from these analyses are presented in Chapter 4, section 4.4.

3.5 Conclusion of Methods

This chapter presented the methods of the three related studies that were necessary to perform a comprehensive evaluation of the Green Prescription intervention in general practice. Before the intervention could be evaluated, a reliable and valid method for assessing physical activity amongst ‘less-active’ adults in primary care had to be established. Next, the methods to test the effectiveness of the intervention were described. Establishing effectiveness is necessary before conducting a cost-effectiveness evaluation, yet the collection of costing data is best achieved prospectively (Drummond, O’Brien et al. 1997). The methods used to conduct the cost-effectiveness study were described in the final part of this chapter. The next chapter presents the results of these three studies.
4 Results

4.1 Introduction

This chapter presents the results from the three inter-related studies that make up this thesis. The first section summarises the results from the test-retest reliability and validity study of the two physical activity questionnaires selected and adapted for use in primary health care research. The second section describes the findings from the pilot study and the cluster randomised controlled trial of the Green Prescription intervention in primary health care. The third section presents the cost-effectiveness evaluation of the Green Prescription that was undertaken alongside the randomised controlled trial. A discussion of these results is included in Chapter 5.

4.2 Results of the Physical Activity Questionnaire Validity and Reliability Study

4.2.1 Introduction

The reliability and validity of the physical activity questionnaires that had been adapted for use amongst less active adults in primary care were assessed to ensure a valid measure of physical activity could be chosen for the subsequent randomised controlled trial. The physical activity questionnaires assessed in this study included the Auckland Heart Study (AHS) and the Green Script Study (GSS) questionnaires Appendices 1 and 2).

4.2.2 Participation Rates

Over the five days of recruiting in the two general practices, 147 patients aged from 40 to 79 years attended. Participation rates are outlined in Figure 4-1. Eighteen of these patients were acutely unwell, were not eligible on medical grounds, or were missed by the receptionist, so were not screened for the study. Of the 129 who were screened at the reception desk for ‘inactivity’, 40% (n = 51) were identified as ‘less active’ and invited to take part in the study. Eighty percent (n = 40) agreed to take part in the reliability and
validity study, although the test-retest study was completed by 71% (n = 36) and the validity study was completed by 67% (n = 34).

**Figure 4-1 Participation Rates in the Physical Activity Questionnaire Reliability and Validity Study**

- 147 patients (40-79 years of age) visited four GPs from 2 practices over 5 days. 18 excluded on medical or other grounds
- 129 screened for inactivity 78 too active
- 51 "less-active" patients eligible 11 declined
- 40 consented & filled out AHS questionnaire at practice (78%).
- 38 completed GSS questionnaire at practice (75%)
- 36 repeated AHS questionnaire after 1-2 weeks, completed 7-day diary, and wore pedometer for 7 days 2 diaries filled out inadequately
- 34 completed 7-day diary adequately (67%)

**4.2.3 Participant Characteristics**

Seventy-five percent (30/40) of participating subjects were female. The average age was 59 years (sd 9.4) and the mean body mass index was 28.9 kg/m² (sd 6.4). There was a wide range of educational levels. While 12.5% (5/40) held tertiary qualifications, 60% (24/40) did not complete secondary school. Ethnic diversity was representative of the region with 10% Māori (4/40) and 90% New Zealand European.
Table 4-1 shows physical activity variables estimated from the 7-day diary, the two administrations of the AHS questionnaire and the GSS questionnaire. The mean number of steps taken per day was 8,238 (sd 3,267).

Table 4-1 Estimated Means for Total Energy Expenditure (kcal/kg/wk), Total Moderate and Leisure Moderate Activity (hours/week) from the AHS Questionnaire, 7-day Diary and GSS Questionnaire

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy AHS, 1st</td>
<td>40</td>
<td>272.66</td>
<td>37.82</td>
</tr>
<tr>
<td>Total energy AHS, 2nd</td>
<td>36</td>
<td>271.71</td>
<td>36.44</td>
</tr>
<tr>
<td>Total energy 7-day</td>
<td>34</td>
<td>272.13</td>
<td>27.46</td>
</tr>
<tr>
<td>Total energy GSS</td>
<td>38</td>
<td>258.34</td>
<td>32.00</td>
</tr>
<tr>
<td>Total moderate AHS, 1st</td>
<td>40</td>
<td>15.60</td>
<td>15.53</td>
</tr>
<tr>
<td>Total moderate AHS, 2nd</td>
<td>36</td>
<td>15.84</td>
<td>14.58</td>
</tr>
<tr>
<td>Total moderate 7-day</td>
<td>34</td>
<td>16.00</td>
<td>11.74</td>
</tr>
<tr>
<td>Total moderate GSS</td>
<td>38</td>
<td>9.49</td>
<td>11.57</td>
</tr>
<tr>
<td>Leisure moderate AHS, 1st</td>
<td>40</td>
<td>3.40</td>
<td>4.53</td>
</tr>
<tr>
<td>Leisure moderate AHS, 2nd</td>
<td>36</td>
<td>3.69</td>
<td>5.23</td>
</tr>
<tr>
<td>Leisure moderate 7-day</td>
<td>34</td>
<td>3.82</td>
<td>5.28</td>
</tr>
<tr>
<td>Leisure moderate GSS</td>
<td>38</td>
<td>1.96</td>
<td>3.52</td>
</tr>
</tbody>
</table>
4.2.4 Reliability of the AHS Questionnaire

There was good test-retest reliability for total energy expenditure, total moderate and leisure moderate activity estimated by the AHS questionnaire. Intraclass correlation coefficients \((r = 0.52\) to \(0.81\)) and Spearman's correlation coefficients \((r = 0.48\) to \(0.71\)) are presented in Table 4-2. Test-retest reliability of the GSS questionnaire has been established previously and was not carried out in this study (Swinburn, Walter et al. 1998). Spearman's rather than Pearson's correlation coefficients were used because of the non-parametric distribution of this data.

Table 4-2 Test-Retest Reliability of the AHS questionnaire: Intraclass Correlation Coefficients (ICC) and Spearman's Correlation Coefficients (SCC) for Total Energy Expenditure, Total Moderate and Vigorous Activity and Leisure-time Moderate Activity

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>I. C. C.</th>
<th>P value</th>
<th>S. C. C.</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy expenditure</td>
<td>36</td>
<td>0.81</td>
<td>&lt;0.01</td>
<td>0.71</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total moderate activity</td>
<td>36</td>
<td>0.74</td>
<td>&lt;0.01</td>
<td>0.59</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total vigorous activity</td>
<td>36</td>
<td>0.52</td>
<td>&lt;0.01</td>
<td>0.67</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Leisure-time moderate activity</td>
<td>36</td>
<td>0.61</td>
<td>&lt;0.01</td>
<td>0.48</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

4.2.5 Validity of the AHS and the GSS Questionnaires

4.2.5.1 Comparison with the 7-day Diary

Spearman's correlation coefficients of the AHS questionnaire and the GSS questionnaire compared with the 7-day diary are presented in Table 4-3. Correlation coefficients were moderate for total energy expenditure \((r = 0.59\) to \(0.74\)), total moderate activity \((r = 0.50\) to \(0.72\)), leisure moderate activity \((r = 0.52\) to \(0.59\)), and leisure vigorous activity \((r = 0.39\) to \(0.99\)). Very few participants took part in vigorous activity. Consequently, vigorous activity correlation coefficients were more variable.
Table 4-3 Validity of the AHS Questionnaire and the GSS Questionnaire compared with the 7-day Diary for Total Energy Expenditure, Total Moderate Activity and Leisure Moderate and Vigorous Activity using Spearman’s Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>1st AHS* P value</th>
<th>2nd AHS** P value</th>
<th>GSS* P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 34</td>
<td>n = 32</td>
<td>n = 33</td>
</tr>
<tr>
<td>Total energy expenditure</td>
<td>0.59 &lt;0.01</td>
<td>0.74 &lt;0.01</td>
<td>0.66 &lt;0.01</td>
</tr>
<tr>
<td>Total moderate activity</td>
<td>0.50 &lt;0.01</td>
<td>0.72 &lt;0.01</td>
<td>0.60 &lt;0.01</td>
</tr>
<tr>
<td>Leisure-time moderate activity</td>
<td>0.52 &lt;0.01</td>
<td>0.59 &lt;0.01</td>
<td>0.55 &lt;0.01</td>
</tr>
<tr>
<td>Leisure-time vigorous activity</td>
<td>0.39 &lt;0.01</td>
<td>0.65 &lt;0.01</td>
<td>0.99 &lt;0.01</td>
</tr>
</tbody>
</table>

*1st AHS: First administration of the Auckland Heart Study questionnaire. **2nd AHS: Second administration of the Auckland Heart Study questionnaire. *GSS: Green Script Study questionnaire

Mean differences for paired estimates from the AHS questionnaire and the 7-day diary, for total energy expenditure (0.29%), total moderate activity (3.96%), and leisure-time moderate activity (1.57%), were very small and not statistically significant. These differences are also depicted in the Bland-Altman graphs of total moderate activity and total energy expenditure of the 7-day diary compared with the AHS questionnaire in Figure 4-2 (a and b). These graphs show mean discrepancies close to zero, but large standard deviations of discrepancies. By contrast, the GSS questionnaire tended to underestimate all activity categories (by 5.93 to 43.08%) by statistically significant amounts (p < 0.01). These underestimates are depicted graphically in the Bland-Altman graphs in Figure 4-2 (c and d).
Figure 4-2 Bland-Altman Graphs of the Auckland Heart Study (AHS) and Green Script Study (GSS) Questionnaires compared with the 7-day Diary Estimations of Total Moderate Activity and Total Energy Expenditure (TEE)

a) AHS vs. 7-day Diary for Total Moderate Activity (hours/week)

b) AHS vs. 7-day Diary for Total Energy Expenditure (Kcal/Kg/week)
c) GSS vs. 7-day Diary for Total Moderate Activity (hours/week)

\begin{center}
\begin{tabular}{|c|c|}
\hline
\textbf{Difference between GSS & 7-day Diary (hrs/wk)} & \\
\hline
+2 SD = 11.9 & \\
\hline
\textbf{Mean} = -5.9 & \\
\hline
\textbf{-2SD} = -23.7 & \\
\hline
\end{tabular}
\end{center}

\begin{center}
\text{Mean total moderate activity (hrs/wk), GSS and 7-day diary}
\end{center}

\begin{center}
\text{\textit{Diff. betw.} Refers to ‘difference between’}
\end{center}

d) GSS vs. 7-day Diary for Total Energy Expenditure (Kcal/Kg/week)

\begin{center}
\begin{tabular}{|c|c|}
\hline
\textbf{Diff. betw. GSS & 7-day diary (Kcal/kg/wk)} & \\
\hline
+2SD = 29.0 & \\
\hline
\textbf{Mean} = -16.1 & \\
\hline
\textbf{-2SD} = -61.2 & \\
\hline
\end{tabular}
\end{center}

\begin{center}
\text{Mean TEE (kcal/kg/wk), GSS and 7-day diary}
\end{center}

\begin{center}
\textit{4.2.5.2 Comparison with the 7-day Pedometer Record}
\end{center}

Table 4-4 shows modest correlation between the physical activity questionnaires and pedometer records for total number of hours of moderate activity (\(r = 0.37\) to 0.61) and total energy expenditure (\(r = 0.22\) to 0.50).
Table 4-4 Validity of the AHS Questionnaire, and the GSS Questionnaire, compared with the Pedometer recording for Total Moderate Activity and Total Energy Expenditure using Spearman’s Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>1st AHS P value</th>
<th>2nd AHS P value</th>
<th>GSS P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Moderate Activity</td>
<td>0.37 0.03</td>
<td>0.51 &lt;0.01</td>
<td>0.61 &lt;0.01</td>
</tr>
<tr>
<td>Total Energy Expenditure</td>
<td>0.22 0.22</td>
<td>0.49 &lt;0.01</td>
<td>0.50 &lt;0.01</td>
</tr>
</tbody>
</table>

*1st AHS: First administration of the Auckland Heart Study questionnaire. **2nd AHS: Second administration of the Auckland Heart Study questionnaire. *GSS: Green Script Study questionnaire

4.2.6 Summary

Test-retest reliability of the AHS physical activity assessed 1-2 weeks after initial administration produced Spearman’s correlation coefficients of 0.48-0.71 and intraclass correlation coefficients of 0.52-0.81 for different physical activity measures (p<0.01). In comparison with the 7-day diary, Spearman’s correlation coefficients were 0.39-0.74 for the AHS questionnaire and 0.55-0.99 for the GSS questionnaire (p<0.01). The Bland-Altman graphs showed that activity levels tended to be underestimated by the GSS but not by the AHS questionnaire, when compared with the 7-day diary. In comparison with pedometer counts, Spearman’s correlation coefficients for leisure time moderate or vigorous intensity activity were 0.37-0.51 for the AHS questionnaire and 0.61 for the GSS questionnaire (p<0.05).

4.3 Results of the Cluster Randomised Controlled Trial of the Green Prescription

4.3.1 Introduction

This section presents the results from the pilot study, followed by the results from the main study, the cluster randomised controlled trial of the Green Prescription intervention in primary health care.
4.3.2 Results of the Pilot Study

4.3.2.1 Pilot Study Participation Rates

The pilot study involved two general practices. A coin was tossed to allocate one practice to intervention (to receive the Green Prescription intervention from their general practitioner) and the other practice was allocated as a control group. Patients were screened over a one-week period and 21 patients were enrolled from each practice.

During the recruitment week, 147 patients in the 40-79 year age group were eligible for physical activity screening at the reception as they entered the practice for their usual consultations. Eighty-eight percent (129/147) of patients in the age group received a screening form. Twelve percent were missed due to the patient being ineligible, or to lack of time. Of the 40% (51/129) of patients identified as 'less active', 82% (42/51) agreed to participate in the intervention pilot study, slightly more than in the reliability and validity study.

4.3.2.2 Pilot Study Participant Characteristics

Thirty-one participants were female, and 11 were male. Four participants identified themselves as Māori and one patient dropped out of the study over the next 3 months. Pilot study baseline cross-sectional cardiovascular risk factors and physical activity characteristics are presented in Table 4-5.

The proportion of screened patients identified as 'less active', was similar to the proportion of 'less active' adults in the New Zealand public (approximately 40%) (Ministry of Health 1998).
Table 4-5 Pilot Study Mean Baseline Characteristics of Intervention and Control Patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intervention Group Mean (SD) n=21</th>
<th>Control Group Mean (SD) n=21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>59.2 (10.1)</td>
<td>57.9 (8.9)</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>136.2 (17.7)</td>
<td>133.6 (19.3)</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>82.9 (9.7)</td>
<td>80.9 (12.6)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>81.1 (17.4)</td>
<td>77.2 (17.0)</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>29.3 (6.3)</td>
<td>28.5 (6.3)</td>
</tr>
<tr>
<td>Cholesterol concentration (mmol/L)</td>
<td>5.5 (0.9)</td>
<td>5.9 (1.2)</td>
</tr>
<tr>
<td>5-yr cardiovascular disease risk (%)</td>
<td>8.7 (6.0)</td>
<td>10.5 (12.4)</td>
</tr>
<tr>
<td>Total energy expenditure (Kcal/kg/week)</td>
<td>243.1 (25.0)</td>
<td>251.6 (39.8)</td>
</tr>
<tr>
<td>Leisure moderate activity (hrs/week)</td>
<td>1.0 (1.0)</td>
<td>1.2 (2.9)</td>
</tr>
</tbody>
</table>

**4.3.2.3 Pilot Study Final Results: Physical Activity Outcomes**

Study patients were co-ordinated at their original practices for follow-up measures, two months after baseline and delivery of the intervention. Measures were taken by the researcher. Change over time in physical activity, cardiovascular risk, blood pressure, serum cholesterol concentrations, and quality of life SF-36 variables, were the efficacy variables. Intervention results were compared with those of the control group and analysed using the SPSS statistical package.

Self-reported leisure-time moderate-intensity activity increased by 2 hours per week in the intervention group and 6 minutes per week in the control group (p < 0.01). Self-reported leisure-time vigorous-intensity activity increased by 19.6 minutes per week in the control group and 14.6 minutes in the intervention group. This difference was not statistically significant. Total energy expenditure increased by 8.8 kcal/kg/wk in the control group and 14.7 kcal/kg/wk in the intervention group. This difference was not statistically significant although there was a significant increase in total energy expenditure from baseline, in the intervention group only (p < 0.01).
4.3.2.4 Pilot Study Final Results: Health Outcomes

Systolic blood pressure decreased by 8.2 mm Hg in the control group and 8.3 mm Hg in the intervention group. There was no statistically significant difference between the groups, but both decreased significantly from baseline (p < 0.01). Diastolic blood pressure decreased by 3.1 mm Hg in the control group and by 7.95 mm Hg in the intervention group. The decrease from baseline was statistically significant in intervention group (p < 0.01) but not in the control group. Difference between the groups did not reach significance (p = 0.08), but a trend was evident. Mean changes in blood pressure in the intervention and control groups, are depicted in Figure 4-4. Of note was the fact that one patient in the intervention group, and two patients in the control group, had blood pressure lowering medications commenced or added, during the study.
As well as a decreasing trend in blood pressure, there was also a trend in weight reduction. Weight decreased in both groups by 0.73 kg, which was not a significant reduction. Body mass index (BMI) decreased by 0.9 kg/m² in the control group and 1.1 kg/m² in the intervention group. These reductions were not statistically significant. Total cholesterol increased in the control group by 0.2 mmol/L and decreased by 0.18 mmol/L in the intervention group, which was a statistically significant difference (p < 0.05). High-density lipoprotein concentration (HDL) increased in the control group by 0.05 mmol/L and decreased by 0.02 mmol/L in the intervention group, which was not a statistically significant difference. Total cholesterol: HDL ratio decreased by 0.06 in the control group and by 0.07 in the intervention group (not significant). Absolute 5-year cardiovascular risk decreased by 1.9% in the control group and by 1.6% in the intervention group. This represented a relative risk reduction of 18% in both groups. These changes were not significantly different from each other or from baseline (Figure 4-5).

**Figure 4-4 Pilot Study Mean Change in Systolic and Diastolic Blood Pressure for the Intervention and Control Patients, from Baseline to 2-month Follow-up**

'SBP' refers to systolic blood pressure; 'DBP' refers to diastolic blood pressure
Figure 4-5 Pilot Study Mean Changes in Lipid Concentrations (mmol/L) and 5-year Cardiovascular Relative Risk Reduction (proportion) for Intervention and Control Patients from Baseline to 2-Month Follow-up

' HDL ' refers to high-density lipoprotein concentration; ' TC ' refers to total cholesterol concentration; ' CVD RRR ' refers to cardiovascular disease relative risk reduction.

Changes in quality of life SF-36 scores showed mixed results, as depicted in Figures 4-6 and 4-7. Only two outcomes demonstrated statistically significant differences in change between the intervention and control groups. These included role physical and social functioning, in which there was a statistically significant improvement amongst the control group compared with the intervention group (p < 0.05).

Figure 4-6 Pilot Study Mean Changes in SF-36 Quality of Life Scores for Physical Function, Role Physical and Bodily Pain, for the Intervention and Control Patients from Baseline to 2-month Follow-up

142
Figure 4-7 Pilot Study Mean Changes in SF-36 Quality of Life Scores for General Health, Vitality, Social Functioning, Role Emotional and Mental Health, for the Intervention and Control Patients from Baseline to 2-month Follow-up

![Graphs showing changes in SF-36 scores](image)

'Ph. Func.' refers to physical function; 'Role Phys.' refers to role physical; 'G. Health' refers to general health; 'Soc. Func.' refers to social functioning; 'R. Emot.' refers to role emotional; 'M. Health' refers to mental health.

4.3.2.5 Acceptability of the Pilot Study Protocol

The recruitment protocol was well tolerated by the staff and patients of the pilot practices and was an efficient way of identifying 'less active' patients. The proportion identified as 'less active' was similar to the prevalence of 'less active' adults in the New Zealand public. Rates of participation and follow-up were high. Baseline and follow-up measurement techniques were also well-tolerated, and fitted in with the practice timetable with minimal disruption.

However, the point-of-care finger prick testing of lipid concentrations showed some variability when repeated tests were done on the same sample. In addition, although there was high correlation between HDL point of care and laboratory estimates (0.99, p<0.01), with no statistically significant difference between the two methods, there was a statistically significant difference between the point of care machine and the laboratory...
estimation of cholesterol concentration (p<0.05). Therefore, it was decided to use laboratory testing of venous samples for lipid concentration determination in the main study cluster randomised controlled trial.

It was also established that the screening procedure was likely to identify a significant number of people with previous cardiovascular disease. Therefore, the Framingham equation, which was used to estimate 5-year cardiovascular risk in the pilot study, would not be suitable for those people. Accordingly, it was decided to use 4-year coronary heart disease risk in the main Green Prescription effectiveness trial (Framingham equation for those without previous cardiovascular disease and D'Agostino equation for those with previous cardiovascular disease). The rationale for this decision is explained more fully in Chapter 2, section 2.2.3.

Apart from these issues, the study protocol underwent minimal change for use in the main cluster randomised controlled trial which commenced two months later, in April 2000.

4.3.3 Results of the Cluster Randomised Controlled Trial of Green Prescription Intervention

The cluster randomised controlled trial was conducted following the completion of the pilot study, which enabled the study protocols to be refined and sample size calculations, checked. The rest of this section presents the results from the main study, the cluster randomised controlled trial. These results represent an assessment of the effectiveness of the Green Prescription in primary health care, with respect to change in physical activity, quality of life, 4-year coronary heart disease risk and blood pressure. The reporting of the randomised controlled trial meets the CONSORT statement requirements.

4.3.4 Participation Rates

One hundred and fifty-nine general practitioners were identified from registers of medical practices in Central and Eastern Waikato and invited to participate in the study. Figure 4-8 depicts the practice recruitment rate. Seventy-four percent (117 general practitioners from 42 practices) agreed and took part. A further four practices had initially agreed to
participate, but had subsequently withdrawn prior to the commencement of patient recruitment at their practices. The reasons for their attrition are discussed below.

Figure 4-8 Process of Recruitment of General Practitioners and Randomisation of Practices in the Green Prescription Cluster Randomised Controlled Trial

Rolling recruitment of patients from the 42 practices occurred over the next 12 months. Researchers spent approximately 5 days at each practice, alternating control and intervention practices. Approximately 8-10 patients, per full-time equivalent general practitioner per week, were enrolled in the study. Figure 4-9 shows the process of patient screening, recruitment and follow-up. Of the 3,433 patients aged 40-79 and attending the study practices over the recruiting period, 33 were too unwell to be screened for inactivity. From the remaining 3,400, 88% (n=2,984) were screened for inactivity at the reception desk. The other 12% were missed or declined screening. Forty-six percent (n = 1,364) of those screened were identified as sedentary. Forty-two patients (4%) were found to be ineligible prior to enrolment, due to medical exclusion criteria, age, or they were planning to leave the area within the next 12 months. Sixty-seven percent (n = 878)
of screened eligible patients were enrolled in the study. Again, the process of screening, recruitment, and enrolment, was well tolerated by practice staff and patients.
Figure 4-9 Progress of Patient Screening, Recruitment and Follow-up Stages of the Trial following Pre-randomisation of Practices
Figure 4-9 also shows that 12-month follow-up measures were completed in 85% of participants (n=750). This follow-up rate was substantially higher than had been allowed for in sample size calculations. Furthermore, attrition rates were very similar in control and intervention groups. While some of the non-attendees declined to attend follow-up, most did not return because they had moved out of the study area, or because they could not be traced by the researchers. There were nine deaths over the period, six in the control group and three in the intervention group.

Post-enrolment, 15 patients (1.7%) were found to meet the exclusion criteria or did not fulfill age criteria. These included nine patients found to meet medical exclusion criteria (four control patients and five intervention patients) (Table 4-6). The general practitioners of the patients with high blood pressure or cholesterol readings were informed of the high readings, with the consent of the patients. Three control patients were found to be over the age of 80 years at baseline and three control patients were found to be under the age of 40 years after they had enrolled. To reduce the risk of recruitment bias, there were no exclusions post enrolments, as discussed in section 3.3.8.1 of Chapter 3. Therefore, results from these 15 patients were included in the analysis.

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>F</td>
<td>60</td>
<td>BP = 199/130 mmHg</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>53</td>
<td>Chol = 9.13 mmol/l</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>49</td>
<td>DBP = 121 mmHg</td>
</tr>
<tr>
<td>I</td>
<td>F</td>
<td>64</td>
<td>Chol = 10.27 mmol/l</td>
</tr>
<tr>
<td>I</td>
<td>M</td>
<td>70</td>
<td>BP = 169/130</td>
</tr>
<tr>
<td>I</td>
<td>M</td>
<td>73</td>
<td>Acutely unwell &amp; progressive illness</td>
</tr>
<tr>
<td>C</td>
<td>M</td>
<td>74</td>
<td>Advanced cancer</td>
</tr>
<tr>
<td>I</td>
<td>F</td>
<td>54</td>
<td>BP = 213/130 mmHg</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>75</td>
<td>BP = 163/122 mmHg</td>
</tr>
</tbody>
</table>

'T' refers to intervention; 'C' refers to control. 'F' refers to female; 'M' refers to male; 'BP' refers to blood pressure in mm Hg. 'DBP' refers to diastolic blood pressure. 'Chol' refers to cholesterol concentration in mmol/L. 'se' refers to serious condition.

Eight participants (0.9%) did not complete major sections of the baseline measures, usually due to time constraints. These included five intervention patients and three control patients. Missing data from these cases were treated in the same way in the analysis as with patients that did not attend follow-up. In other words, no change was assumed over the twelve months where data were missing.
4.3.4.1 Practice Attrition

All practices that commenced patient recruitment completed the entire study. Four practices (7 general practitioners) of the 46 randomised, dropped out before patient recruitment at their practices commenced. One practice did not participate following randomisation because there was a change in practice management. Two other solo practices withdrew just before recruitment started at their respective practices, near the end of rolling recruitment because of unexpected staffing problems. Lastly, a 4-GP practice also withdrew. This practice had already had their week of study participation postponed twice, once by the researcher and once by the practice. Several of the general practitioners then arranged a temporary exchange with practitioners from the United Kingdom, who had not agreed to take part in the study. In addition, time and space became an issue, so the practice withdrew from the study. Again, this was towards the end of the rolling recruitment phase. All four practices that withdrew were control practices.

4.3.5 Participant Characteristics

4.3.5.1 Practice and General Practitioner Characteristics

Practices were stratified according to size, being divided into three categories of solo (n = 18), two or three general practitioners (n = 17), and more than three general practitioner (n = 11) practices. Twenty-one practices were situated in rural or semi-rural towns, and 21 practices were situated in an urban area (Hamilton). Twenty-eight percent of the 117 participating general practitioners were female (n = 33). Characteristics of the participating general practitioners are presented in Table 4-7.

<table>
<thead>
<tr>
<th>General practitioner</th>
<th>Number</th>
<th>Full-time (%)</th>
<th>Part-time (%)</th>
<th>In solo practice (%)</th>
<th>In large practice (&gt;3GPs) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>84 (72%)</td>
<td>81 (96%)</td>
<td>3 (3.6%)</td>
<td>14 (17%)</td>
<td>38 (45%)</td>
</tr>
<tr>
<td>Female</td>
<td>33 (28%)</td>
<td>14 (42%)</td>
<td>19 (58%)</td>
<td>5 (15%)</td>
<td>16 (48%)</td>
</tr>
<tr>
<td>Total</td>
<td>117 (100%)</td>
<td>95 (81%)</td>
<td>22 (19%)</td>
<td>19 (16%)</td>
<td>54 (46%)</td>
</tr>
</tbody>
</table>

*Full time* refers to at least 8 tenths or four full days per week. *Part time* refers to less than this amount.
4.3.5.2 Patient Demographic Characteristics

Baseline characteristics of all patient participants were assessed. Sixty-six percent of study participants were female (n = 582). Educational levels and socio-economic status were similar to the age-matched general population. Fifty-four percent of study participants had a school or other qualification (n = 470) compared with 56% in the age-matched general population (Department of Statistics New Zealand 2002). A community services card is a nation-wide measure of lower income, and was held by 48% of the study population over 45 years of age, compared with 43% of the general population aged over 45 (Department of Statistics New Zealand 2002). Ethnicity of study participants was reflective of the general population with 77% Caucasian and 17% Māori, compared with 80% and 15%, respectively, nation-wide (Department of Statistics New Zealand 2002). The other 6% of study participants were from other ethnic groups.

4.3.5.3 Patient Cardiovascular Characteristics

Within the study population, there were high rates of hypertension (52%), diabetes (10.5%), obesity (43%), and previous cardiovascular disease (19%). Using the Framingham and D’Agostino equations, the average 4-year coronary heart disease risk was 5.6% and the average 4-year cardiovascular disease risk was 7.7% (Anderson, Odell et al. 1991; D’Agostino, Russell et al. 2000). Seventy-eight percent of participants were overweight (BMI > 25) and 41% were obese (BMI > 30). Ninety-three percent of participants had at least one risk factor for cardiovascular disease. These risk factors include blood pressure of greater than 150 mm Hg systolic or 90 mm Hg diastolic, body mass index greater than 25, total cholesterol: HDL ratio greater or equal to 6.5, smoking, or previously diagnosed hypertension, diabetes or cardiovascular disease. These are characteristics identified in the Activity Counselling Trial as cardiovascular risk factors (The Writing Group for the Activity Counseling Trial Research Group 2001). (However, total cholesterol: HDL ratio rather than LDL and a threshold blood pressure of 150/90 mm Hg, rather than 140/90, were used to define hypertension in the current study).

The prevalence of hypertension was high amongst study participants compared with New Zealand average values for similar age groups. The 1996/97 New Zealand
Health survey found the prevalence of diagnosed hypertension from self-report for age groups 45-64 and 65-74 years, was 18.2% and 45%, respectively, compared with 32.9% and 50% of age-matched study participants (Ministry of Health 1998). If those with blood pressure greater than 150/90 mm Hg are added to “diagnosed hypertensive”, 48.8% and 67.5% of study participants for the two age groups were hypertensive. As the New Zealand Health survey figures do not include undiagnosed hypertensive adults, data from the Auckland Heart and Health study were also used for comparison. This Auckland study was a population-based survey of Auckland residents and included data on New Zealand prevalence of previously diagnosed and undiagnosed hypertension (Trye, Jackson et al. 1996; Bullen, Simmons et al. 1998). The Auckland study used similar methodology to that of the present study although it involved a random sample of non-Māori. Therefore, comparisons in Figure 4-10 include results for non-Māori only.

The prevalence of self-reported diabetes was also high compared with the age-matched New Zealand population estimates (Figure 4-11), (Ministry of Health 1998) and the South Auckland population estimates (Figure 4-12) (Simmons, Harry et al. 1999; Kenealy, Scragg et al. 2000).

**Figure 4-10 Prevalence of Hypertension in Study Participants compared with the New Zealand Population Prevalence* (non-Māori only)**

![Graph showing prevalence of hypertension](image)

*(Trye, Jackson et al. 1996; Bullen, Simmons et al. 1998)*
Figure 4-11 Prevalence of known Diabetes amongst Study Participants compared with New Zealand Population Estimates*

'yrs' refers to years of age; 'NZ' refers to New Zealand prevalence; 'Study popn.' refers to study population; *(Ministry of Health 1998)

Figure 4-12 Prevalence of known Diabetes amongst Study Participants compared with New Zealand estimates by Age and Ethnic Group
Mean cardiovascular and physical activity variables are presented in Table 4-8. Cardiovascular characteristics of study participants were compared with the age- and gender-matched population. Table 4-9 compares mean values for this study population with average values obtained from a cross-sectional random sample of Auckland non-Māori adults in 1993-1994 (Jackson, Yee et al. 1995; Bullen, Simmons et al. 1998). Māori within the study were excluded from this analysis, also, to match the ethnic composition of the comparison population. The study population has been categorised by gender and in the age groups of 40-64 years and 65-80 years, in order to be more comparable to data from the 1993-1994 population estimates. The 1993-1994 population estimates used the categories of 35-64 years and 65-84 years.

Smoking rates among study participants were not significantly different from population rates, and lipid profiles were similar or better than those of the comparative population. However, mean systolic and diastolic blood pressure and mean body mass index of the study sample were higher when compared with all age-matched general population categories except for systolic blood pressure for those over 65 years (Jackson, Yee et al. 1995; Bullen, Simmons et al. 1998).
Table 4-8 Physical Activity and Cardiovascular Measures of the Study Population at Baseline

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>58.59 (10.75)</td>
<td>57.52 (11.35)</td>
<td>878</td>
</tr>
<tr>
<td>Total energy expenditure (kcal/kg/day)</td>
<td>35.23 (7.59)</td>
<td>33.08 (5.31)</td>
<td>873</td>
</tr>
<tr>
<td>Total leisure energy expenditure (kcal/kg/day)</td>
<td>1.09 (2.19)</td>
<td>0.79 (1.32)</td>
<td>873</td>
</tr>
<tr>
<td>Leisure moderate or vigorous activity (minutes per day)</td>
<td>13.95 (26.34)</td>
<td>10.49 (17.77)</td>
<td>873</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>89.02 (16.68)</td>
<td>78.84 (18.85)</td>
<td>871</td>
</tr>
<tr>
<td>Body Mass Index (kg/m²)</td>
<td>29.40 (5.10)</td>
<td>30.22 (7.15)</td>
<td>867</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mm Hg)</td>
<td>136.70 (17.89)</td>
<td>134.46 (19.17)</td>
<td>875</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mm Hg)</td>
<td>84.62 (12.55)</td>
<td>80.81 (11.75)</td>
<td>875</td>
</tr>
<tr>
<td>Total cholesterol (mmol/L)</td>
<td>5.48 (1.03)</td>
<td>5.82 (1.00)</td>
<td>728</td>
</tr>
<tr>
<td>High density lipoprotein (mmol/L)</td>
<td>1.16 (0.30)</td>
<td>1.42 (0.38)</td>
<td>728</td>
</tr>
<tr>
<td>TC:HDL ratio</td>
<td>4.96 (1.39)</td>
<td>4.36 (1.32)</td>
<td>728</td>
</tr>
<tr>
<td>4-year risk of coronary heart event*</td>
<td>9.36 (7.46)</td>
<td>3.73 (4.03)</td>
<td>728</td>
</tr>
<tr>
<td>4-year risk of cardiovascular event*</td>
<td>11.68 (8.46)</td>
<td>5.68 (5.29)</td>
<td>728</td>
</tr>
</tbody>
</table>

N = sample size; sd = standard deviation; HDL = high density lipoprotein; TC = total cholesterol;


NB: Five participants failed to complete most components of the baseline requirements, due to lack of time, two patients were unable to be weighed and four patients did not have height measures taken, two due to being wheel-chair bound and two due to lack of time. The sample size was smaller for lipid testing, as four practices were randomly selected not to have cholesterol testing done. This decision was justified because the sample size required to detect a significant change in coronary heart disease risk in the trial was considered to be less than for other outcomes.
Table 4-9 National Non-Māori Population Estimates* compared with Non-Māori Study Participant Means for Cardiovascular Risk
Factors (with 95% confidence intervals).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Men</th>
<th>&lt; 65 years</th>
<th>Women</th>
<th>&lt; 65 years</th>
<th>Men</th>
<th>≥ 65 years</th>
<th>Women</th>
<th>≥ 65 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>(95% confidence interval)</td>
<td>NZ</td>
<td>Study pop</td>
<td>NZ</td>
<td>Study pop</td>
<td>NZ</td>
<td>Study pop</td>
<td>NZ</td>
<td>Study pop</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td>16.9</td>
<td>14.5</td>
<td>14.8</td>
<td>16.9</td>
<td>10.1</td>
<td>17.4</td>
<td>8.4</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>(13.6-20.1)</td>
<td>(8.9-20.1)</td>
<td>(11.7-17.9)</td>
<td>(12.8-21.0)</td>
<td>(7.2-12.9)</td>
<td>(10.1-27.1)</td>
<td>(4.0-8.4)</td>
<td>(3.7-12.7)</td>
</tr>
<tr>
<td>Mean Systolic BP (mmHg)</td>
<td>125.8</td>
<td>132.8</td>
<td>121.2</td>
<td>130.3</td>
<td>143.8</td>
<td>141.8</td>
<td>144.9</td>
<td>142.5</td>
</tr>
<tr>
<td></td>
<td>(124.7-127)*</td>
<td>(131-135)*</td>
<td>(120-122.3)*</td>
<td>(128-132)*</td>
<td>(142-145.7)*</td>
<td>(137-146)</td>
<td>(143.2-146.7)</td>
<td>(139-146)</td>
</tr>
<tr>
<td>Mean Diastolic BP (mmHg)</td>
<td>74.9</td>
<td>85.3</td>
<td>70.7</td>
<td>80.3</td>
<td>73.5</td>
<td>81.9</td>
<td>74.4</td>
<td>79.0</td>
</tr>
<tr>
<td></td>
<td>(74.0-75.8)*</td>
<td>(83.6-87.0)*</td>
<td>(69.8-71.5)*</td>
<td>(79.1-81.4)*</td>
<td>(72.4-74.6)*</td>
<td>(78.8-85.1)*</td>
<td>(73.3-75.4)*</td>
<td>(77.2-80.8)*</td>
</tr>
<tr>
<td>Mean Body Mass Index</td>
<td>26.4</td>
<td>28.8</td>
<td>25.1</td>
<td>29.8</td>
<td>24.9</td>
<td>27.8</td>
<td>24.7</td>
<td>28.9</td>
</tr>
<tr>
<td>(kg/m²)</td>
<td>(26.2-26.7)*</td>
<td>(28.1-29.5)*</td>
<td>(24.8-25.4)*</td>
<td>(29.0-30.5)*</td>
<td>(24.6-25.2)*</td>
<td>(26.9-28.7)*</td>
<td>(24.3-25.1)*</td>
<td>(28.1-29.7)*</td>
</tr>
<tr>
<td></td>
<td>n = 712</td>
<td>n = 124</td>
<td>n = 685</td>
<td>n = 266</td>
<td>n = 480</td>
<td>n = 70</td>
<td>n = 516</td>
<td>n = 130</td>
</tr>
<tr>
<td>Mean Cholesterol (mmol/L)</td>
<td>5.73</td>
<td>5.62</td>
<td>5.60</td>
<td>5.75</td>
<td>5.49</td>
<td>5.18</td>
<td>6.38</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td>(5.66-5.81)</td>
<td>(5.45-5.81)</td>
<td>(5.52-5.68)</td>
<td>(5.63-5.87)</td>
<td>(5.40-5.58)</td>
<td>(4.95-5.41)</td>
<td>(6.29-6.47)</td>
<td>(5.91-6.23)</td>
</tr>
<tr>
<td>Mean HDL (mmol/L)</td>
<td>1.10</td>
<td>1.19</td>
<td>1.40</td>
<td>1.44</td>
<td>1.10</td>
<td>1.17</td>
<td>1.38</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>(1.08-1.13)*</td>
<td>(1.14-1.26)*</td>
<td>(1.37-1.42)</td>
<td>(1.40-1.48)</td>
<td>(1.07-1.13)</td>
<td>(1.09-1.25)</td>
<td>(1.35-1.42)</td>
<td>(1.39-1.53)</td>
</tr>
<tr>
<td>TC:HDL ratio</td>
<td>5.53</td>
<td>4.93</td>
<td>4.28</td>
<td>4.23</td>
<td>5.40</td>
<td>4.66</td>
<td>5.0</td>
<td>4.46</td>
</tr>
<tr>
<td></td>
<td>(5.4-5.67)*</td>
<td>(4.68-5.18)*</td>
<td>(4.17-4.39)</td>
<td>(4.07-4.39)</td>
<td>(5.2-5.5)*</td>
<td>(4.37-4.95)*</td>
<td>(4.90-5.20)*</td>
<td>(4.24-4.68)</td>
</tr>
</tbody>
</table>
(Jackson, Yee et al. 1995; Bullen, Simmons et al. 1998) * Statistically significantly difference (non-overlapping 95% confidence intervals) between Auckland population and study population. *NZ refers to the non-Māori population estimates based on the Auckland Heart and Health Study (Jackson, Yee et al. 1995; Bullen, Simmons et al. 1998). * Study pop = refers to participants of the Waikato Heart, Health and Activity Study, excluding Māori. The study population has been categorised by gender and in the age groups of 40 to 64 years and 65 to 80 years. This was done in order to be more comparable to data from the 1994 data, which were analysed in the categories of 35-64 years and 65 to 84 years.

4.3.6 The Relationship between Cardiovascular Risk and Usual Level of Physical Activity

Four-year cardiovascular risk was found to increase with advancing age, male gender, decreasing total energy expenditure, lower socio-economic status, increasing body mass index and Māori ethnicity. These relationships are demonstrated in Table 4-10, which presents the findings from multivariate linear regression analysis using 4-year cardiovascular risk as the dependent variable, grouped at the level of practice. Other variables were not significantly related to cardiovascular risk in the multiple regression analysis. These variables included rural location, education (which is closely linked to socio-economic status), and self-rated mental health and physical function.

The relationship between levels of physical activity and other variables, such as demographic and cardiovascular variables, was explored using univariate regression analysis, adjusting for clustering by practice (Table 4-11). The physical activity variable used was total energy expenditure in kcal/kg/week. However, as the data were skewed, the natural log was used. Log-transformation produced a more normal distribution to meet the assumptions of linear regression equations. A multiple regression model was constructed using variables that were significant at a p < 0.05 level within the univariate analyses. However, if variables were considered to measure characteristics with a common factor (e.g. blood pressure and cardiovascular risk), then only one variable was chosen for the multiple regression model. Table 4-12 shows the results of this multiple regression analysis. Lower levels of physical activity were associated with those who were female, older, of lower socio-economic status, taking multiple medications, or recently hospitalised.
Table 4-10 Relationship of Four-year Cardiovascular Risk \(^a\) of Study Participants to Health, Activity and Demographic Variables (Multiple linear regression analysis, grouped by medical practice)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Regression Coefficient</th>
<th>Std. Err</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>-7.1965</td>
<td>0.466</td>
<td>0.000</td>
</tr>
<tr>
<td>Increasing age</td>
<td>0.4120</td>
<td>0.021</td>
<td>0.000</td>
</tr>
<tr>
<td>*Increasing total energy expenditure</td>
<td>-4.5036</td>
<td>1.395</td>
<td>0.001</td>
</tr>
<tr>
<td>#Lower economic status</td>
<td>1.6737</td>
<td>0.474</td>
<td>0.000</td>
</tr>
<tr>
<td>$Māori ethnicity</td>
<td>1.6293</td>
<td>0.619</td>
<td>0.008</td>
</tr>
<tr>
<td>Increasing BMI(^b)</td>
<td>0.1205</td>
<td>0.035</td>
<td>0.001</td>
</tr>
</tbody>
</table>

\(^a\) Cardiovascular risk is measured using the Framingham (Anderson, Odell et al. 1990) equation for those without previous cardiovascular disease. For those with previous cardiovascular disease the estimated 4-year risk of a cerebrovascular event, using the Framingham equation, was added to the 4-year risk of a coronary heart disease event using the D'Agostino equation (D'Agostino, Russell et al. 2000) to produce a conservative estimate of overall 4-year cardiovascular risk.

* The natural log of total energy expenditure (kcal/kg/week) was used in the analysis to obtain a normal distribution for regression analyses.

# Economic status was estimated by community services card status, which is allocated according to family income.

$ Ethnicity was analyzed as Māori or non-Māori, with non-Māori as the referent category.

\(^b\) BMI refers to body mass index.
Table 4-11 Relationships of Variables to Total Energy Expenditure (Natural Log) of Study Participants. (Univariate linear regression analysis, grouped by practice)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Regression Coefficient</th>
<th>Std. Err.</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>-0.0550</td>
<td>0.012</td>
<td>0.000</td>
</tr>
<tr>
<td>Increasing age</td>
<td>-0.0043</td>
<td>0.0005</td>
<td>0.000</td>
</tr>
<tr>
<td>Lower economic status</td>
<td>-0.0845</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>Hospitalised in last year</td>
<td>0.0595</td>
<td>0.014</td>
<td>0.000</td>
</tr>
<tr>
<td>CVD Risk</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.000</td>
</tr>
<tr>
<td>Total Medications</td>
<td>-0.0193</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>Attendance at Sec. School</td>
<td>0.0650</td>
<td>0.018</td>
<td>0.000</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>-0.0008</td>
<td>0.0003</td>
<td>0.02</td>
</tr>
<tr>
<td>Diastolic Blood Pressure</td>
<td>0.0007</td>
<td>0.0005</td>
<td>0.2</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>-0.0012</td>
<td>0.001</td>
<td>0.2</td>
</tr>
<tr>
<td>Maori ethnicity</td>
<td>-0.0081</td>
<td>0.016</td>
<td>0.6</td>
</tr>
<tr>
<td>Smoking</td>
<td>0.011</td>
<td>0.011</td>
<td>0.6</td>
</tr>
<tr>
<td>Urban vs. rural</td>
<td>0.0139</td>
<td>0.015</td>
<td>0.4</td>
</tr>
<tr>
<td>Research Assessor</td>
<td>-0.0071</td>
<td>0.006</td>
<td>0.3</td>
</tr>
<tr>
<td>HDL</td>
<td>0.0101</td>
<td>0.016</td>
<td>0.5</td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>-0.004</td>
<td>0.006</td>
<td>0.5</td>
</tr>
</tbody>
</table>

'Sec' refers to secondary
Table 4-12 Relationships of Variables to Total Energy Expenditure (Natural Log) of Study Participants (Multiple linear regression analysis, grouped by practice)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Regression Coefficient</th>
<th>Std. Err.</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>-0.0846</td>
<td>0.014</td>
<td>0.000</td>
</tr>
<tr>
<td>Increasing age</td>
<td>-0.0015</td>
<td>0.0007</td>
<td>0.03</td>
</tr>
<tr>
<td>Lower economic status #</td>
<td>-0.0502</td>
<td>0.013</td>
<td>0.000</td>
</tr>
<tr>
<td>Hospitalised in last year</td>
<td>-0.0028</td>
<td>0.014</td>
<td>0.015</td>
</tr>
<tr>
<td>Increasing CVD Risk</td>
<td>-0.0028</td>
<td>0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>Increasing medications</td>
<td>-0.0079</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>Attendance at high school</td>
<td>-0.0009</td>
<td>0.019</td>
<td>0.96</td>
</tr>
</tbody>
</table>

# Economic status was estimated by community services card status, which is allocated according to family income.
4.3.7 Quality of Life Characteristics

The quality of life scores for each of the eight SF-36 parameter scores were examined at baseline. Several trends with age became apparent. For example, self-rated physical function tended to decrease with age, general health was fairly stable across age-groups, and mental health increased with age, as illustrated in Figure 4-13.

Figure 4-13 Baseline SF-36 Quality of Life Scores, ‘Physical Function’, ‘General Health’, and ‘Mental Health’, of Study Participants by Age Group

Baseline SF-36 quality of life scores for study participants were significantly lower than the New Zealand norms for almost all age and gender categories and for all quality of life variables. Figures 4-13 to 4-16 show the mean difference between study participant SF-36 quality of life scores compared with age- and gender-matched New Zealand norms. The age categories for the New Zealand data were slightly different from those used in this study, but an attempt was made to find categories as closely matching the study sample as possible. Mean scores for 40-64 year-old study patients were compared with New Zealand norms for 45-64 years. Mean scores for 65-79 year-old study patients were compared with New Zealand norms for 65-74 years.
Figure 4-14 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Male Study Participants 40-64 years of age compared with New Zealand Gender-matched and Age-similar Norms

PF = physical functioning; RP = role physical; BP = bodily pain; GH = general health; V = vitality; SF = social functioning; RE = role emotional; MH = mental health

Figure 4-15 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Male Study Participants 65-79 years of age compared with New Zealand Gender-matched and Age-similar Norms
Figure 4-16 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Female Study Participants 40-64 years of age compared with New Zealand Gender-matched and Age-similar Norms

Figure 4-17 Mean Differences (with 95% confidence intervals) in SF-36 Quality of Life Scores of Female Study Participants 65-79 years of age compared with New Zealand Gender-matched and Age-similar Norms

PF = physical functioning; RP = role physical; BP = bodily pain; GH = general health; V = vitality; SF = social functioning; RE = role emotional; MH = mental health
4.3.8 Adequacy of Randomisation

To check the adequacy of randomisation, characteristics of the intervention and control groups were compared to ensure they were similar. Characteristics of general practices and general practitioners from the intervention and control groups are presented below. Figure 4-18 shows practice size of intervention and control practices. Three solo practices and one large (4-GP) practice, all from the control group, withdrew from participation prior to patient recruitment. This explains why there were fewer solo, and very large practices, in the control group compared with the intervention group. In addition, the three solo practices from the control group that withdrew were rural. Despite this, intervention and control groups had similar numbers of rural and urban practices (Figure 4-19). Gender and work characteristics of the intervention and control general practitioners were balanced (Figures 4-20 and 4-21).

The demographic and clinical characteristics of the intervention and control patients were well balanced at baseline, as shown in Table 4-13. Table 4-14 shows various self-reported physical-activity characteristics of the intervention and control groups. These were also balanced, as were the baseline SF-36 quality of life scores, as demonstrated in Table 4-15.

**Figure 4-18 Practice Sizes of Intervention and Control Practices**

![Practice Sizes Graph]
Figure 4-19 Number of Rural and Urban Practices in the Intervention and Control Groups
Figure 4-20 Gender of General Practitioners in the Intervention and Control Groups

Figure 4-21 Number of Full-time and Part-time General Practitioners in the Intervention and Control Groups
Table 4-13 Baseline Demographic and Clinical Characteristics of Less-active 40-79 year-old Patients in General Practice, by Intervention and Control Group. Mean values and (standard deviations) are presented, unless otherwise specified.

<table>
<thead>
<tr>
<th>Patient Characteristic</th>
<th>Intervention (n = 451)</th>
<th>Control (n = 427)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>57.2 (10.8)</td>
<td>58.6 (11.5)</td>
</tr>
<tr>
<td>Systolic BP (mm Hg)</td>
<td>135.1 (19.6)</td>
<td>135.4 (17.9)</td>
</tr>
<tr>
<td>Diastolic BP (mmHg)</td>
<td>82.4 (12.2)</td>
<td>81.8 (12.1)</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>30.0 (6.7)</td>
<td>29.9 (6.4)</td>
</tr>
<tr>
<td>Cholesterol* (mmol/L)</td>
<td>5.78 (1.0)</td>
<td>5.64 (1.0)</td>
</tr>
<tr>
<td>HDL* (mmol/L)</td>
<td>1.33 (0.4)</td>
<td>1.34 (0.4)</td>
</tr>
<tr>
<td>CHD 4-year risk* (% risk)</td>
<td>5.7 (6.2)</td>
<td>5.5 (5.8)</td>
</tr>
<tr>
<td>Number of medications</td>
<td>2.6 (2.5)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td>Female participation: N (%)</td>
<td>301 (67)</td>
<td>281 (66)</td>
</tr>
<tr>
<td>Lower economic status: ** N (%)</td>
<td>205 (45)</td>
<td>211 (49)</td>
</tr>
<tr>
<td>Post high school qualification: N(%)</td>
<td>106 (24)</td>
<td>121 (28)</td>
</tr>
<tr>
<td>European: N (%)</td>
<td>354 (78)</td>
<td>324 (76)</td>
</tr>
<tr>
<td>Smokers: N (%)</td>
<td>78 (17)</td>
<td>76 (18)</td>
</tr>
<tr>
<td>Diabetes: N (%)</td>
<td>46 (10)</td>
<td>46 (11)</td>
</tr>
<tr>
<td>Hypertensive⁵: N (%)</td>
<td>240 (53)</td>
<td>220 (52)</td>
</tr>
<tr>
<td>Previous cardiovascular disease: N(%)</td>
<td>93 (21)</td>
<td>74 (17)</td>
</tr>
<tr>
<td>Obese (BMI&gt;30): N (%)</td>
<td>198 (44)</td>
<td>176 (41)</td>
</tr>
</tbody>
</table>

*Cholesterol testing and risk of 4-year coronary heart disease risk (Anderson, Odell et al. 1990; D'Agostino, Russell et al. 2000) were carried out on a randomly selected sub-sample to contain costs (n=787) and a further 51 participants declined to have cholesterol testing done.

** Economic status was measured at baseline by qualification for a low-income health subsidy card. Forty-three percent of adults over 45 years of age in New Zealand qualify for this card.

⁵Hypertensive refers to a previous diagnosis of hypertension and taking antihypertensive medication or a mean blood pressure of greater than 150 mm Hg systolic or 90 mm Hg diastolic.
Table 4-14 Baseline Physical Activity Characteristics of the Intervention and Control Groups. Mean values and (standard deviations) are presented, unless otherwise specified.

<table>
<thead>
<tr>
<th>Patient Characteristic</th>
<th>Intervention (n = 451)</th>
<th>Control (n = 427)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy expenditure (kcal.kg(^{-1}).week(^{-1}))</td>
<td>237.5 (42.2)</td>
<td>235.7 (45.3)</td>
</tr>
<tr>
<td>Leisure physical activity(a) (kcal.kg(^{-1}).week(^{-1}))</td>
<td>6.0 (12.2)</td>
<td>6.5 (11.1)</td>
</tr>
<tr>
<td>Leisure exercise(\ast) (mins.day(^{-1}))</td>
<td>11.3 (21.7)</td>
<td>12.0 (20.5)</td>
</tr>
<tr>
<td>Leisure exercise &gt;=2.5 hrs/week N(%)</td>
<td>80 (18)</td>
<td>91 (21)</td>
</tr>
</tbody>
</table>

\(a\)Leisure physical activity refers to the energy expenditure of all leisure-time physical activity considered moderate or vigorous by the respondent

\(\ast\)Leisure exercise refers to all moderate (3.0-4.9 MET) and vigorous (>=5.0 MET) leisure-time activities undertaken at least once per two weeks (Lamonte and Ainsworth 2001).

Table 4-15 Mean Baseline SF-36 Quality of Life Scores (and standard deviations) for Intervention and Control Patients

<table>
<thead>
<tr>
<th>SF-36 Quality of Life Category(\ast\ast)</th>
<th>Intervention (n = 451)</th>
<th>Control (n = 427)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>71.3 (23.9)</td>
<td>70.9 (24.6)</td>
</tr>
<tr>
<td>Role physical</td>
<td>57.9 (41.7)</td>
<td>60.4 (41.4)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>61.1 (25.7)</td>
<td>63.9 (26.9)</td>
</tr>
<tr>
<td>General health</td>
<td>62.7 (20.7)</td>
<td>66.1 (20.6)</td>
</tr>
<tr>
<td>Vitality</td>
<td>53.8 (20.6)</td>
<td>56.0 (21.2)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>77.9 (24.6)</td>
<td>77.6 (25.2)</td>
</tr>
<tr>
<td>Role emotional</td>
<td>69.6 (41.3)</td>
<td>68.7 (40.6)</td>
</tr>
<tr>
<td>Mental health</td>
<td>74.5 (17.3)</td>
<td>74.0 (18.2)</td>
</tr>
</tbody>
</table>

\(\ast\ast\)Quality of life scores are derived from the SF-36 questionnaire and represent a score out of 100.
4.3.9 Positive Predictive Value of Screening Question

The screening question used to identify sedentary patients for enrolment in the study was as follows. "As a rule, do you do at least half an hour of vigorous or moderate exercise (such as walking or a sport) on five or more days of the week?" The positive predictive value of the physical activity screening question was 81%. Nineteen percent of those identified as sedentary by the screening question were found to be achieving at least 2 ½ hours per week of moderate or vigorous leisure-time physical activity when compared with results of the AHS questionnaire. Eighty-one percent were achieving less than 2 ½ hours. Therefore, the population identified by the screening question was, on average, less active than the general population, where between 32% and 42% are achieving less than 2 ½ hours per week (Ministry of Health 1998).

4.3.10 Rates of Intervention Delivery

Of the 451 intervention patients, 385 received the intervention from a general practitioner and 66 from a practice nurse. Of the patients who attended follow-up 10/361 (2.8%) control and 370/389 (95%) intervention patients recalled receiving a Green Prescription during the previous year, suggesting that there was minimal contamination between intervention and control groups. The Sports Foundation, 'Sport Waikato', received 410 faxed copies of the study Green Prescriptions, out of a possible 451. Forty-one participants did not want their details and a copy of the Green Prescription faxed to exercise specialists, or they did not receive the intervention from the general practitioner or practice nurse.

Sub-samples of 31 general practitioners and 19 nurses estimated spending an average of 7 minutes and 13 minutes per patient, respectively, delivering the Green Prescription intervention.

4.3.11 Final Outcomes

4.3.11.1 Physical Activity

Physical activity increased in both groups, over the twelve months. However, leisure-time moderate or vigorous physical activity, and total energy expenditure, increased more in the intervention group than the control group (Table 4-16).
## Table 4-16 Mean Changes in Physical Activity, Cardiovascular and Quality of Life Outcomes in the Control and Intervention Groups, at 12 Months.

<table>
<thead>
<tr>
<th>Primary Outcomes Measure</th>
<th>5\textsuperscript{\textdagger} Intervention: (n=451)</th>
<th>5\textsuperscript{\textdagger} Control: (n=427)</th>
<th>*Between-group difference: (n=878)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean change (95% CI)</td>
<td>Mean change (95% CI)</td>
<td>Mean change (95% CI)</td>
<td></td>
</tr>
<tr>
<td><strong>Total energy expenditure (kcal/kg/week)</strong></td>
<td>9.76 (5.85,13.68)</td>
<td>0.37 (-3.39,4.14)</td>
<td>9.38 (3.96,14.81)</td>
<td>0.001**</td>
</tr>
<tr>
<td>*Leisure physical activity (kcal/kg/week)</td>
<td>4.32 (3.26,5.38)</td>
<td>1.29 (0.11,2.47)</td>
<td>2.67 (0.48,4.86)</td>
<td>0.02*</td>
</tr>
<tr>
<td>*Leisure exercise (mins/week\textsuperscript{1})</td>
<td>54.6 (41.4,68.4)</td>
<td>16.8 (6.0,32.4)</td>
<td>33.6 (2.4,64.2)</td>
<td>0.04*</td>
</tr>
<tr>
<td>Systolic BP (mmHg)</td>
<td>-2.58 (-4.02,-1.13)</td>
<td>-1.21(-2.57,0.15)</td>
<td>-1.31(-3.51,0.89)</td>
<td>0.2</td>
</tr>
<tr>
<td>Diastolic BP (mmHg)</td>
<td>-2.62 (-3.62,-1.61)</td>
<td>-0.81(-1.77,0.16)</td>
<td>-1.40(-3.35,0.56)</td>
<td>0.2</td>
</tr>
<tr>
<td>\textsuperscript{c}4-year CHD risk (%)</td>
<td>0.42 (0.23,0.60)</td>
<td>0.52 (0.32,0.72)</td>
<td>-0.10(-0.43,0.23)</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>Quality of life (SF-36):</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning score</td>
<td>3.16 (1.61,4.71)</td>
<td>1.63 (-0.04,3.31)</td>
<td>1.23 (-1.35,3.81)</td>
<td>0.3</td>
</tr>
<tr>
<td>Role physical score</td>
<td>10.53 (6.8,14.3)</td>
<td>4.16 (0.63,7.68)</td>
<td>7.24 (0.16,14.31)</td>
<td>0.045*</td>
</tr>
<tr>
<td>Bodily pain score</td>
<td>6.51 (4.28,8.74)</td>
<td>2.50 (0.15,4.86)</td>
<td>4.01 (0.78,7.24)</td>
<td>0.02*</td>
</tr>
<tr>
<td>General health</td>
<td>5.95 (4.43,7.47)</td>
<td>1.60 (0.22,2.99)</td>
<td>4.51 (2.07,6.95)</td>
<td>0.000**</td>
</tr>
<tr>
<td>Vitality score</td>
<td>5.36 (3.76,6.96)</td>
<td>3.06 (1.44,4.68)</td>
<td>2.30 (0.03,4.57)</td>
<td>0.047*</td>
</tr>
<tr>
<td>Social functioning score</td>
<td>3.02 (0.68,5.36)</td>
<td>2.85 (0.57,5.13)</td>
<td>0.36 (-3.53,4.26)</td>
<td>0.9</td>
</tr>
<tr>
<td>Role emotional score</td>
<td>5.32 (1.43,9.21)</td>
<td>5.70 (2.07,9.32)</td>
<td>-0.38(-5.70,4.94)</td>
<td>0.9</td>
</tr>
<tr>
<td>Mental health score</td>
<td>2.61 (1.17,4.04)</td>
<td>1.63 (0.28,2.98)</td>
<td>0.98 (-0.99,2.95)</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\*Significant at 0.05 level ** Significant at 0.01 level \textsuperscript{5}Unadjusted for clustering \*Adjusted for clustering by medical practice; \textsuperscript{1}CI' refers to confidence interval

\*Leisure physical activity refers to the energy expenditure of all leisure-time physical activity considered moderate or vigorous by the respondent

\*Leisure exercise refers to all moderate (3.0-4.9 MET) and vigorous (>=5.0 MET) leisure-time activities undertaken at least once per two weeks.(Lamonte and Ainsworth 2001)

\textsuperscript{c}4-yr CHD risk refers to 4-year coronary heart disease risk, based on the Framingham and D'Agostino equations. (Anderson, Odell et al. 1990; D'Agostino, Russell et al. 2000) As these equations have only been validated for ages 35 to 75 years, those with ages from 76 to 79 were excluded from the intention-to-treat analysis. Therefore, the sample size for this variable was 802.
Figure 4-15 shows that the proportion of intervention participants who achieved 2½ hours of moderate or vigorous leisure physical activity per week increased by 66/451 (14.6%) compared with 21/427 (4.9%) in the control group (p=0.003). This means that for every 10.3 Green Prescriptions written, one person achieved and maintained 2½ hours of moderate or vigorous leisure-time activity, who otherwise would not have, producing a number needed to treat (NNT) of 10.3. Increases in occupational physical activity contributed substantially to the additional increase in total energy expenditure (p<0.001) although domestic and transport activity did not.

Figure 4-22 Proportions of Participants achieving 2½ Hours of Moderate or Vigorous Leisure Activity per Week at Baseline and 12-month Follow-up

(p=0.003)

4.3.11.2 Quality of Life

Quality of life improved in both groups. However, ‘general health’, ‘role-physical’, ‘vitality’, and ‘bodily pain’ scores on the SF-36 questionnaire improved significantly more in the intervention group compared with the control group. (Table 4-16).

4.3.11.3 Cardiovascular Risk

There was no statistically significant difference in the change of coronary heart disease risk between the two groups. The point estimate difference was equivalent to a 1.7% relative risk
reduction in coronary heart disease risk (0.1%/5.6%), which was substantially lower than that predicted in sample size calculations (Chapter 3, section 3.3.6).

Systolic and diastolic blood pressure improved significantly from baseline in the intervention group but the change did not differ significantly from that achieved in the control group (Table 4-16). Regression analysis of blood pressure change was controlled for antihypertensive medication change.

Interestingly, a greater percentage of control patients had their antihypertensive medications increased, and fewer had their anti-hypertensive medications decreased, with a slightly greater percentage having a change in medication type with equivalent dose, compared with the intervention group (Table 4-17). Although these differences did not reach statistical significance, there was definitely no evidence that intervention patients had increases in their medication compared with the control group.

Table 4-17 Changes in Antihypertensive Medication Use from Baseline to Follow-up, amongst the Intervention and the Control groups

<table>
<thead>
<tr>
<th>Prescribed Antihypertensive</th>
<th>Intervention (n = 389)</th>
<th>Control (n = 361)</th>
<th>Between group % difference (95% CI of difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decrease in dose</td>
<td>18 (4.6)</td>
<td>14 (3.8)</td>
<td>0.75 (-2.1, 3.6)</td>
</tr>
<tr>
<td>Increase in dose</td>
<td>38 (9.8)</td>
<td>37 (10.2)</td>
<td>-0.48 (-4.8, 3.8)</td>
</tr>
<tr>
<td>Change in type, similar</td>
<td>10 (2.6)</td>
<td>15 (4.1)</td>
<td>-1.6 (-4.2, 1.0)</td>
</tr>
<tr>
<td>equivalent dose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (any change)</td>
<td>66 (17.0)</td>
<td>66 (18.3)</td>
<td>-1.3 (-6.8, 4.1)</td>
</tr>
</tbody>
</table>

'CI' refers to confidence interval

Although change in weight, body mass index and cholesterol concentrations were not primary outcomes of the study, these variables are presented in Table 4-18. Again, there was a non-significant trend towards improvement in the intervention, compared with the control group, for these variables except for HDL.
Table 4-18 Change in Weight, Body Mass Index and Serum Cholesterol Concentration of Intervention and Control Groups over 12-months

<table>
<thead>
<tr>
<th>Other Variables Measures</th>
<th>Intervention: (n=451) Mean change (95% CI)</th>
<th>Control: (n=427) Mean change (95% CI)</th>
<th>Between-group difference: (n=878) Mean change (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (kg)</td>
<td>-0.30 (-0.66, 0.06)</td>
<td>-0.14 (-0.49, 0.21)</td>
<td>-0.16 (-0.67, 0.34)</td>
<td>0.5</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>-0.11 (-0.25, 0.02)</td>
<td>-0.05 (-0.18, 0.07)</td>
<td>-0.06 (-0.24, 0.12)</td>
<td>0.5</td>
</tr>
<tr>
<td>Cholesterol (mmol/L)</td>
<td>-0.02 (-0.08, 0.05)</td>
<td>0.01 (-0.05, 0.06)</td>
<td>-0.02 (-0.12, 0.09)</td>
<td>0.7</td>
</tr>
<tr>
<td>HDL (mmol/L)</td>
<td>-0.05 (-0.06, -0.03)</td>
<td>-0.03 (-0.04, -0.02)</td>
<td>-0.017 (-0.04, 0.01)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

'CI' refers to confidence interval

4.3.11.4 Outcomes of Potential Harm

A repeated measures multi-level analysis, using SAS, to show the change in risk of adverse events over time showed no significant difference between the two groups. In particular, outcome measures of potential harm showed no significant difference in self-reported falls (p=0.9) or injuries (p=0.7) in the previous month, or hospitalisations over the previous year (p=0.1) (Table 4-19).

Table 4-19 Multi-level, Time-series Analysis of Adverse Events for Intervention and Control Patients (comparing the year prior to study enrolment with the year following study enrolment)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Yr1 Control (n = 427)</th>
<th>Yr2 Control (n = 427)*</th>
<th>Yr1 Intervention (n = 451)</th>
<th>Yr2 Intervention (n = 451)**</th>
<th>P value**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falls in last month N (%)</td>
<td>47 (11.1)</td>
<td>55 (12.9)</td>
<td>42 (9.3)</td>
<td>49 (10.9)</td>
<td>0.9</td>
</tr>
<tr>
<td>Injury in last month N (%)</td>
<td>60 (14.1)</td>
<td>41 (9.6)</td>
<td>65 (14.4)</td>
<td>39 (8.6)</td>
<td>0.7</td>
</tr>
<tr>
<td>Hospitalisation in last year N(%)</td>
<td>89 (20.8)</td>
<td>90 (21.1)</td>
<td>104 (23.1)</td>
<td>83 (18.4)</td>
<td>0.1</td>
</tr>
</tbody>
</table>

* Intention-to-treat analysis is carried out, bringing baseline findings forward to follow-up for non-attenders.

** P values are calculated using a time-series multilevel analysis that compares the change over time of the intervention group with that of the control group.

As the outcome was the probability of an adverse event, this was more logically presented as odds ratios rather than absolute differences. This analysis also shows the size of the
difference. The odds of having a fall, injury, or being hospitalised, in the year after the intervention compared with the year before were examined to calculate the odds ratio, for the intervention group and the control group. The odds ratios for intervention patients were 1.19 (95% CI 0.82, 1.72) for falls, 0.63 (0.44, 0.896) for injuries, and 0.77 (0.60, 0.996) for hospitalisations. In other words, intervention patients were significantly less likely to have an injury, or be hospitalised, in the year after intervention compared with the year before the intervention. The odds ratios for control patients were 1.22 (0.85, 1.75) for falls, 0.7 (0.48, 1.02) for injuries, and 1.02 (0.796, 1.33) for hospitalisations. In other words, the change in the likelihood of control patients having a fall, injury, or being hospitalised the year before or the year after their enrolment in the study, did not reach statistical significance. However, the odds ratios for the intervention and control groups showed no significant difference from each other.

4.3.11.5 Sub-Group Analyses

Post-hoc analysis by gender showed that leisure-time physical activity increased by 68 minutes (95% CI 29 to 106 minutes) per week in men (n=296), and 20 minutes (95% CI -23 to 63 minutes) per week in women (n=582).

There was no significant difference in the change in leisure-time moderate and vigorous activity (p = 0.8) amongst the intervention group of urban, compared with rural participants. However, there was a difference in overall total energy expenditure. Urban participants increased their total energy expenditure significantly more than rural participants, by 7.55 kcal/kg/week (p = 0.006).

4.3.11.6 Stage of Change

There was a non-significant trend towards greater increases in leisure time activity (3.25 kcal/kg/week, 95% CI 9.26, -2.76) and total energy expenditure (22.3 kcal/kg/week, 95% CI: 50.4, -5.81) amongst the intervention ‘contemplators’ (SOC=2) compared with intervention ‘non-contemplators’ (SOC=1). However, ‘pre-contemplators’ tended to be older than ‘contemplators’ by 5.6 years (95% CI: -1.06, 12.3).
Table 4-20 Average Change in Leisure-time Physical Activity and Total Energy Expenditure by Baseline Stage of Change amongst Intervention Participants

<table>
<thead>
<tr>
<th>Stage of change (SOC) at baseline</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>No response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (%)</td>
<td>13 (3.3)</td>
<td>143 (37)</td>
<td>55 (14)</td>
<td>110 (28)</td>
<td>48 (12)</td>
<td>19 (4.9)</td>
</tr>
<tr>
<td>Average age in years (sd)</td>
<td>63.2 (11.0)</td>
<td>57.6 (11.7)</td>
<td>56.2 (10.3)</td>
<td>56.9 (10.8)</td>
<td>56.7 (8.6)</td>
<td>56.2 (9.9)</td>
</tr>
<tr>
<td>Change in leisure physical activity in kcal/kg/week (sd)</td>
<td>2.77 (4.48)</td>
<td>6.02 (10.86)</td>
<td>6.75 (12.30)</td>
<td>4.30 (14.28)</td>
<td>2.46 (12.57)</td>
<td>5.55 (11.00)</td>
</tr>
<tr>
<td>Change in total energy expenditure in kcal/kg/week (sd)</td>
<td>-10.10 (47.24)</td>
<td>12.18 (49.24)</td>
<td>16.53 (37.38)</td>
<td>12.71 (44.46)</td>
<td>16.90 (42.93)</td>
<td>7.02 (87.03)</td>
</tr>
</tbody>
</table>

Stage of change categories 1-5 refer to: 1. 'Pre-contemplators'; 2. 'Contemplators'; 3. 'Ready to change'; 4. 'Starting to change', and 5. 'Maintaining behaviour change'. The patient's stage of change was determined at baseline using the form in Appendix 10.

4.3.12 Sensitivity Analyses

Several sensitivity analyses were carried out. These included analyses that were 'per-protocol', controlled for potential confounding factors, controlled for baseline variables, conducted using different assumptions about the deterioration of coronary heart disease risk, and without adjustment for medication change.

4.3.12.1 Per-protocol Analysis

In this study, analyses of primary outcomes have been conducted using an intention-to-treat approach. The first sensitivity analysis was conducted 'per protocol' to examine the effects that an intention-to-treat approach may have had on results. A 'per protocol' analysis includes only those who attended follow-up and completed both baseline and follow-up measures in each analysis. Subjects were still analysed according to the group they were randomised to, regardless of compliance with the physical activity advice. The same random-effects regression techniques were used as with the main intention-to-treat analyses, with allowance for clustering by practice. Overall, results were very similar. As expected, for almost every variable, the improvements amongst the intervention group compared with the control group were greater than with an intention-to-treat analysis, except coronary heart disease risk, physical function, and role emotional. These three variables did not improve significantly amongst the intervention group compared with the control group, in either
intention-to-treat or per protocol analysis. Although the effect estimate of most of the outcome measures increased, the precision of the estimates generally decreased. This was probably due to the decrease in sample size, which resulted from excluding those that did not attend follow-up from the analysis.

Table 4-21 Per-protocol Sensitivity Analyses of Primary Outcome Measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Between-group difference: (n=750)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy expenditure (kcal/kg/wk)</td>
<td>10.96 (4.57, 17.34)</td>
<td>0.001</td>
</tr>
<tr>
<td>Leisure exercise (kcal/kg/wk)</td>
<td>2.94 (0.23, 5.66)</td>
<td>0.03</td>
</tr>
<tr>
<td>Duration leisure exercise (mins/wk)</td>
<td>36.8 (-1.8, 75.72)</td>
<td>0.06</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)*</td>
<td>-1.54 (-4.00, 0.91)</td>
<td>0.22</td>
</tr>
<tr>
<td>Adjusted systolic blood pressure (mmHg)#</td>
<td>-1.56 (-4.01, 0.94)</td>
<td>0.22</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)*</td>
<td>-1.59 (-3.85, 0.67)</td>
<td>0.17</td>
</tr>
<tr>
<td>Adjusted diastolic blood pressure (mmHg)#</td>
<td>-1.57 (-3.87, 0.72)</td>
<td>0.18</td>
</tr>
<tr>
<td>4-year Coronary heart disease (%)*</td>
<td>-0.05 (-0.49, 0.40)</td>
<td>0.84</td>
</tr>
<tr>
<td>Adjusted 4-year Coronary heart disease (%)**</td>
<td>-0.05 (-0.49, 0.40)</td>
<td>0.84</td>
</tr>
</tbody>
</table>

SF-36 Quality of Life Variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean change (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical function</td>
<td>0.84 (-3.08, 4.77)</td>
<td>0.67</td>
</tr>
<tr>
<td>Role physical</td>
<td>8.38 (-0.11, 16.87)</td>
<td>0.05</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>4.91 (0.97, 8.84)</td>
<td>0.01</td>
</tr>
<tr>
<td>General health</td>
<td>5.49 (2.25, 8.72)</td>
<td>0.001</td>
</tr>
<tr>
<td>Vitality</td>
<td>2.80 (0.02, 5.58)</td>
<td>0.048</td>
</tr>
<tr>
<td>Social function</td>
<td>0.39 (-5.02, 5.81)</td>
<td>0.89</td>
</tr>
<tr>
<td>Role emotional</td>
<td>-0.58 (-7.24, 6.08)</td>
<td>0.87</td>
</tr>
<tr>
<td>Mental health</td>
<td>1.18 (-1.19, 3.56)</td>
<td>0.33</td>
</tr>
</tbody>
</table>

* Not adjusted for medication changes; # Adjusted for antihypertensive medication change; ** Adjusted for antihypertensive and lipid-lowering medications for coronary heart disease risk; ‘CI’ refers to confidence interval

4.3.12.2 Controlling for Possible Confounding Factors

To ensure randomisation was adequate and baseline values balanced, a series of sensitivity analyses were performed adjusting for baseline and potentially confounding factors. The first analysis examined the effect on physical-activity primary outcomes of controlling for demographic variables. Primary outcomes were also analysed controlling for baseline values to assess the effect of any potential baseline imbalance.
Multiple regression analyses of the physical activity outcomes, controlling for age, gender and socio-economic status are presented in Tables 4-22. This sensitivity analysis made little difference to the physical activity and cardiovascular results. However, difference in change between intervention and control groups, of 'role physical' and 'vitality' SF-36 quality of life scores were no longer statistically significant (p>0.05). 'General health' and 'Bodily pain' maintained a statistically significant difference.

**Table 4-22 Sensitivity Analysis using Multiple Regression of Change in Physical Activity, Clinical, and Significant Quality of Life Outcomes for the Intervention Group compared with the Control Group, after adjusting for Age, Gender, Socio-economic Status, and Clustering**

<table>
<thead>
<tr>
<th>Change in Outcome Variable</th>
<th>Between-group difference: (n=878) Mean change (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy expenditure in kcal/kg/week</td>
<td>9.35 (3.85, 14.84)</td>
<td>0.001</td>
</tr>
<tr>
<td>Leisure energy expenditure in kcal/kg/week</td>
<td>2.60 (0.42, 4.78)</td>
<td>0.02</td>
</tr>
<tr>
<td>Time spent in Leisure exercise in hours per week</td>
<td>0.54 (0.02, 1.06)</td>
<td>0.04</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mmHg)</td>
<td>-1.24 (-3.42, 0.94)</td>
<td>0.26</td>
</tr>
<tr>
<td>Diastolic Blood Pressure in (mmHg)</td>
<td>-1.41 (-3.38, 0.56)</td>
<td>0.16</td>
</tr>
<tr>
<td>4-year CHD risk (%)</td>
<td>-0.10 (-0.43, 0.23)</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>SF-36 Quality of Life Variables:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role physical</td>
<td>6.97 (-0.34, 14.29)</td>
<td>0.06</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>3.67 (0.42, 6.92)</td>
<td>0.03</td>
</tr>
<tr>
<td>General health</td>
<td>4.34 (1.94, 6.74)</td>
<td>0.000</td>
</tr>
<tr>
<td>Vitality</td>
<td>1.98 (-0.28, 4.23)</td>
<td>0.09</td>
</tr>
</tbody>
</table>

‘CHD’ refers to coronary heart disease; ‘CI’ refers to confidence interval

Table 4-23 presents the results from a sensitivity analysis of outcomes after controlling for baseline values of each outcome variable. Physical activity and blood pressure changes in outcomes of intervention compared with control groups became greater. Three of the four significant quality of life variables retained statistical significance, although ‘vitality’ did not.
Table 4-23 Sensitivity Analysis using Multiple Regression of Change in Physical Activity, Clinical, and Significant Quality of Life Outcomes for the Intervention Group compared with the Control Group, after adjusting for Baseline Outcome Measures and Clustering

<table>
<thead>
<tr>
<th>Change in Outcome Variable</th>
<th>Between-group difference: (n=878)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean change (95% CI)</td>
<td></td>
</tr>
<tr>
<td>Total energy expenditure in kcal/kg/week</td>
<td>10.18 (5.14, 15.21)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Leisure energy expenditure in kcal/kg/week</td>
<td>2.68 (0.64, 4.72)</td>
<td>0.01</td>
</tr>
<tr>
<td>Time spent in leisure exercise in hours per week</td>
<td>0.57 (0.13, 1.02)</td>
<td>0.01</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mmHg)</td>
<td>-1.45 (-3.56, 0.66)</td>
<td>0.18</td>
</tr>
<tr>
<td>Diastolic Blood Pressure in (mmHg)</td>
<td>-1.38 (-3.22, 0.47)</td>
<td>0.14</td>
</tr>
<tr>
<td>4-year CHD risk (%)</td>
<td>-0.08 (-0.53, 0.37)</td>
<td>0.7</td>
</tr>
<tr>
<td>SF-36 Quality of Life Variables:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role physical</td>
<td>6.87 (0.20, 13.54)</td>
<td>0.04</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>2.93 (0.02, 5.8))</td>
<td>0.048</td>
</tr>
<tr>
<td>General health</td>
<td>3.73 (1.36, 6.11)</td>
<td>0.002</td>
</tr>
<tr>
<td>Vitality</td>
<td>1.64 (-0.51, 3.79)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

‘CHD’ refers to coronary heart disease; ‘CI’ refers to confidence interval

4.3.12.3 Coronary Heart Disease Risk

The intention-to-treat analysis of primary outcomes was carried out assuming those who did not attend follow-up, had no change from baseline in the variable, except for coronary heart disease risk. This variable tended to deteriorate in both groups over time, with the greatest deterioration in the control group. Therefore, it was assumed that all those who did not attend the follow-up visit had a deterioration in coronary heart disease risk equivalent to the average for the control group, which was equivalent to 0.52%. This assumption produced a difference in absolute risk reduction of -0.10 (95% CI -0.43 to 0.23) between the intervention and control groups. If no change was assumed amongst those who did not attend follow-up, the results do not change significantly. Incremental decrease in absolute coronary heart disease risk is -0.11 % (95% CI -0.39 to 0.16), equivalent to a relative risk reduction of 2%, as baseline absolute risk for intervention patients was 5.6%.
4.3.12.4 Blood Pressure Change and Adjustment for Medication Change

If antihypertensive medications were not controlled for in the regression analysis of systolic and diastolic blood pressures, the results were very similar to those presented above. In particular, there was an incremental decrease in systolic blood pressure of -1.2 mm Hg (95% CI -3.28 to 0.70) and in diastolic blood pressure of -1.39 mm Hg (95% CI -3.04 to 0.25) amongst the intervention group compared with the control group.

4.3.13 Intraclass Correlation Coefficients

Intraclass correlation coefficients (ICC) of demographic and clinical variables measured at baseline, using practice as the clustering variable, are presented in Table 4-24. ICCs of baseline physical activity and quality of life variables are presented in Table 4-25. Table 4-25 shows the ICCs of change in selected variables.

Most ICCs were low (less than 0.1). However, economic status (reflected by the community services card status), educational level and ethnicity had much higher ICCs. Ethnicity had by far the highest ICC (0.384), which was probably because there were seven Māori health-provider clinics out of the 42 participating clinics. Māori health-provider clinics cared for Māori patients predominantly, while other practices cared for New Zealand Europeans, predominantly. Socio-economic status and education are likely to vary by suburb, particularly in urban areas.

The F statistics also reflected a range of 'between-practice' variance to 'within-practice' variance of individual demographic and clinical characteristics. The P value of several variable F statistics (such as for cholesterol), showed that there was very little variance between practices when compared to within-practice variance for these variables.

Intraclass correlation coefficients of change in characteristics were often different from ICCs of baseline variables, such as for cholesterol and physical activity variables, which were considerably higher than for baseline values. On the other hand, ICCs for change in body mass index, and coronary heart disease risk, were lower than for baseline ICCs.
Table 4-24 Intra-class Correlation Coefficients (ICCs) for Demographic and Clinical Characteristics in a Consecutively Screened Sample of Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>ICC (95% confidence interval)</th>
<th>F stat. ANOVA (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community services card status*</td>
<td>870</td>
<td>0.121 (0.052 to 0.189)</td>
<td>3.82 (0.000)</td>
</tr>
<tr>
<td>Gender</td>
<td>878</td>
<td>0.016 (0.000 to 0.044)</td>
<td>1.33 (0.08)</td>
</tr>
<tr>
<td>Age</td>
<td>876</td>
<td>0.050 (0.008 to 0.092)</td>
<td>2.08 (0.000)</td>
</tr>
<tr>
<td>Ethnicity # (Māori/nonMāori)</td>
<td>853</td>
<td>0.384 (0.257 to 0.510)</td>
<td>13.53 (0.000)</td>
</tr>
<tr>
<td>Education level</td>
<td>844</td>
<td>0.123 (0.041 to 0.020)</td>
<td>2.88 (0.000)</td>
</tr>
<tr>
<td><strong>Clinical Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>871</td>
<td>0.043 (0.004 to 0.083)</td>
<td>1.93 (0.0005)</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>875</td>
<td>0.018 (0.000 to 0.047)</td>
<td>1.39 (0.06)</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>875</td>
<td>0.046 (0.006 to 0.087)</td>
<td>2.00 (0.0002)</td>
</tr>
<tr>
<td>Cholesterol concentration</td>
<td>728</td>
<td>0.004 (0.000 to 0.031)</td>
<td>1.08 (0.3)</td>
</tr>
<tr>
<td>HDL** concentration</td>
<td>728</td>
<td>0.039 (0.000 to 0.080)</td>
<td>1.76 (0.004)</td>
</tr>
<tr>
<td>Smoking status</td>
<td>854</td>
<td>0.055 (0.011 to 0.10)</td>
<td>2.18 (0.000)</td>
</tr>
<tr>
<td>Body mass index</td>
<td>867</td>
<td>0.081 (0.027 to 0.14)</td>
<td>2.81 (0.000)</td>
</tr>
<tr>
<td>Coronary heart disease risk</td>
<td>728</td>
<td>0.022 (0.000 to 0.056)</td>
<td>1.43 (0.05)</td>
</tr>
<tr>
<td>Cardiovascular disease risk</td>
<td>728</td>
<td>0.018 (0.000 to 0.051)</td>
<td>1.36 (0.08)</td>
</tr>
</tbody>
</table>

* Community services cards (CSC) allow a greater subsidy for primary health care and are held by approximately 43% of the population over 45 years of age, based on family income. # Ethnicity was dichotomized for analysis purposes to those who identify as ‘Māori’ and those who do not; ** HDL refers to high-density lipoprotein
Table 4-25 Intra-class Correlation Coefficients (ICCs) for Physical Activity Characteristics in a Consecutively Screened Sample of Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices

<table>
<thead>
<tr>
<th>Physical Activity Variables</th>
<th>n</th>
<th>ICC (95% confidence interval)</th>
<th>F stat. ANOVA (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leisure energy expenditure</td>
<td>873</td>
<td>0.014 (0.000 to 0.041)</td>
<td>1.29 (0.1)</td>
</tr>
<tr>
<td>Total energy expenditure</td>
<td>873</td>
<td>0.020 (0.000 to 0.049)</td>
<td>1.42 (0.045)</td>
</tr>
<tr>
<td>Time in leisure exercise</td>
<td>873</td>
<td>0.012 (0.000 to 0.038)</td>
<td>1.25 (0.1)</td>
</tr>
<tr>
<td>Quality of Life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical function</td>
<td>850</td>
<td>0.009 (0.000 to 0.035)</td>
<td>1.19 (0.2)</td>
</tr>
<tr>
<td>Role physical</td>
<td>845</td>
<td>0.018 (0.000 to 0.048)</td>
<td>1.37 (0.06)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>845</td>
<td>0.000 (0.011 to 0.022)</td>
<td>0.95 (0.6)</td>
</tr>
<tr>
<td>General health</td>
<td>840</td>
<td>0.004 (0.000 to 0.028)</td>
<td>1.09 (0.3)</td>
</tr>
<tr>
<td>Vitality</td>
<td>849</td>
<td>0.010 (0.000 to 0.036)</td>
<td>1.20 (0.2)</td>
</tr>
<tr>
<td>Social function</td>
<td>842</td>
<td>0.015 (0.000 to 0.043)</td>
<td>1.30 (0.1)</td>
</tr>
<tr>
<td>Role emotion</td>
<td>832</td>
<td>0.007 (0.000 to 0.031)</td>
<td>1.13 (0.3)</td>
</tr>
<tr>
<td>Mental health</td>
<td>847</td>
<td>0.000 (0.000 to 0.022)</td>
<td>0.76 (0.9)</td>
</tr>
</tbody>
</table>

Table 4-26 Intra-class Correlation Coefficients (ICCs) for Change in Variables from Baseline to 12-month Follow-up amongst Less-active 40-79 year-old Patients selected in Clusters from 42 General Practices

<table>
<thead>
<tr>
<th>ICC of change in variable from baseline to follow-up</th>
<th>N</th>
<th>ICC (95% confidence interval)</th>
<th>F stat. ANOVA (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cholesterol concentration</td>
<td>596</td>
<td>0.041 (0.000 to 0.089)</td>
<td>1.66 (0.009)</td>
</tr>
<tr>
<td>HDL concentration</td>
<td>596</td>
<td>0.024 (0.000 to 0.064)</td>
<td>1.38 (0.07)</td>
</tr>
<tr>
<td>Body mass index</td>
<td>744</td>
<td>0.003 (0.000 to 0.030)</td>
<td>1.05 (0.4)</td>
</tr>
<tr>
<td>Leisure energy expenditure</td>
<td>745</td>
<td>0.049 (0.003 to 0.094)</td>
<td>1.90 (0.0007)</td>
</tr>
<tr>
<td>Time in leisure exercise</td>
<td>745</td>
<td>0.058 (0.009 to 0.108)</td>
<td>2.09 (0.0001)</td>
</tr>
<tr>
<td>Coronary heart disease risk</td>
<td>591</td>
<td>0.000 (0.000 to 0.031)</td>
<td>0.83 (0.7)</td>
</tr>
<tr>
<td>Physical function</td>
<td>727</td>
<td>0.025 (0.000 to 0.061)</td>
<td>1.43 (0.04)</td>
</tr>
<tr>
<td>General health</td>
<td>716</td>
<td>0.044 (0.000 to 0.088)</td>
<td>1.77 (0.003)</td>
</tr>
<tr>
<td>Mental health</td>
<td>726</td>
<td>0.000 (0.000 to 0.026)</td>
<td>0.99 (0.5)</td>
</tr>
</tbody>
</table>

* HDL refers to high-density lipoprotein
4.3.14 Summary

The pilot study had demonstrated that the study protocol was acceptable and efficient, requiring minimal alteration before implementation in the main study, the Green Prescription cluster randomised controlled trial. Increases in leisure-time physical activity were demonstrated in the pilot study, as a result of the Green Prescription intervention.

The main cluster randomised controlled trial of the Green Prescription achieved high levels of participation (74% of general practitioners in the area and 67% of eligible 'less active' adult patients) and follow-up at 12 months (85% of participants). The study population was found to represent a high cardiovascular risk population compared with the age- and gender-matched general population. The trial demonstrated significant improvements in physical activity variables and four of the SF-36 health-related quality of life variables, 'general health', 'bodily pain', 'vitality' and 'role physical' over 12 months amongst the intervention group, compared with the control group. There was also a trend towards improved blood pressure. There was no evidence of increased adverse events with the intervention, and there was some evidence to suggest that adverse events decreased in the intervention group. Sensitivity analyses, conducted subsequently, had little effect on outcomes. Effect size of the intervention tended to increase, although these estimates were sometimes less precise. The results of the pilot and main cluster randomised controlled trial are discussed more fully in Chapter 5, section 5.3.

4.4 Results of the Green Prescription Cost-Effectiveness Study

4.4.1 Introduction

A cost-effectiveness evaluation was designed prospectively and conducted alongside the cluster randomised controlled trial. The results of the cost-effectiveness analysis are presented below. The first section summarises the cost of the Green Prescription programme development, and implementation. These costs were incurred at a national level by the Hillary Commission, at a regional level by the Waikato Sports Foundation, and at a local level by the general practices where the intervention was initiated. The next section presents the offset costs and savings from change in health-care utilisation, productivity, and personal
costs of exercise. The final sections calculate incremental cost-effectiveness ratios and summarise findings.

4.4.2 Programme Costs

4.4.2.1 Green Prescription National Set-up and Co-ordinating Costs

The discounted and annuitized costs incurred by the Hillary Commission at a national level for set-up and co-ordinating the Green Prescription programme are included in Table 4-27. Research and evaluation costs are not included, as they were not part of the programme costs.

The total national set-up and co-ordinating cost from mid 1996 to mid 2002 was $2,861,564. A total of 13,189 patients received Green Prescriptions, and were referred to regional Sports Foundations, during that period. (Figures supplied by the Hillary Commission) Therefore, the programme set-up and co-ordinating costs per patient were $216.97 per Green Prescription recipient that was referred to the Sports Foundation. This is likely to be an overestimate, as approximately 62% of Green Prescription recipients are not referred to the Sports Foundation exercise specialists (personal communication with the Hillary Commission). If set-up and national co-ordinating costs were distributed across all those receiving a Green Prescription (n = 34,708), then the set-up and co-ordinating component of programme cost per patient would be $82.45.
<table>
<thead>
<tr>
<th>Year</th>
<th>Details</th>
<th>Cost (NZ$)</th>
<th>December CPI* (CPI ratio)</th>
<th>Adjusted ** Cost in NZ$</th>
<th>Discounted Cost * in NZ$</th>
<th>Referred GRx recipientsb (all GRx recipients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996/97</td>
<td>Launch, Auckland and Northland, Development of resources, Installation of free telephone line in Auckland and Northland, Grants to regional sports trusts</td>
<td>400,000</td>
<td>989 (1.077)</td>
<td>430,800</td>
<td>549,822</td>
<td>nil</td>
</tr>
<tr>
<td>1997/98</td>
<td>Seminar with GPs and nurses, Re-development of resources, Funding to sports trusts to support programme</td>
<td>180,000</td>
<td>997 (1.068)</td>
<td>192,240</td>
<td>233,669</td>
<td>369</td>
</tr>
<tr>
<td>1998/99</td>
<td>Resources, Co-ordinator positions (6 positions, 3 full-time, 3 part-time), Training, Computer scripting, Advertising,</td>
<td>390,000</td>
<td>1,001 (1.064)</td>
<td>414,960</td>
<td>480,368</td>
<td>1279</td>
</tr>
<tr>
<td>1999/00</td>
<td>Co-ordinator positions (6): Training, computer scripting, promotion, resources, database development, Operations/Administration at a national level</td>
<td>257,000</td>
<td>1,006 (1.059)</td>
<td>476,550</td>
<td>525,396</td>
<td>2616</td>
</tr>
<tr>
<td></td>
<td></td>
<td>93,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total:</td>
<td>450,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000/01</td>
<td>Co-ordinator positions (8): Resources, newsletters, patient support, computer scripting Operations/Administration at a national level</td>
<td>317,000</td>
<td>1,046 (1.018)</td>
<td>498,820</td>
<td>523,761</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>73,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total:</td>
<td>490,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001/02</td>
<td>Co-ordinator positions (8): Resources Computer scripting /database /free telephone line, Promotion, Operations/Administration at a national level</td>
<td>387,000</td>
<td>1,065 (1.0)</td>
<td>548,000</td>
<td>548,000</td>
<td>4925</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$16,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$548,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total:</td>
<td>$2,458,000</td>
<td></td>
<td>$2,861,016</td>
<td>13,189</td>
<td>(34,708)</td>
</tr>
</tbody>
</table>

(re: Figures supplied by the Hillary Commission)

* CPI refers to consumer price index. The CPI ratio refers to the December 2001 CPI: annual December CPI ratio (Statistics New Zealand 2003).
** Adjusted for inflation, using the CPI ratio
* Discounted at a 5% rate (Drummond, O'Brien et al. 1997)

b Green Prescription recipients that were referred to the Sports Trusts for exercise specialists support, which includes 38% of all Green Prescription recipients. ‘GRx’ refers to Green Prescription.

### 4.4.2.2 Regional Sports Foundation Personnel Costs and Overheads

Out of the 451 intervention patients in the present study, 410 (91%) were referred to the Sports Foundation exercise specialists. The other 41 patients declined to have their Green Prescriptions faxed to the regional Sports Foundations, or they did not receive the intervention. The regional Sports Foundation was also receiving Green Prescription referrals for patients who were not part of the study during the twelve months of rolling recruitment for the study. Therefore, exercise specialists at the regional Sports Foundation received 912 faxed Green Prescriptions, of which 410 were for study participants (45% of total).

The proportion of the Sports Foundation’s wages, overheads, toll-calls and mail-out costs attributable to the 410 study patients, representing 45% of their Green Prescription programme costs, are presented in Table 4-28. The average cost of intervention support from the Sports Trust was **$68.81** per patient ($31,032.65/451 study participants).
Table 4-28 Costs of Sports Trust attributable to the Green Prescription Study Participants for 12 months (2001/2002)

<table>
<thead>
<tr>
<th>Cost Category</th>
<th>Details</th>
<th>Cost per item in NZ$</th>
<th>Total Cost in NZ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Wage Rates:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff 1</td>
<td>$21.27 per hr x 7.2 hrs/week</td>
<td>7,965</td>
<td></td>
</tr>
<tr>
<td>Staff 2</td>
<td>$20.19 per hr x 7.2 hrs/week</td>
<td>7,560</td>
<td></td>
</tr>
<tr>
<td>Staff 3</td>
<td>$14.00 per hr x 11.25 hrs/week</td>
<td>8,190</td>
<td></td>
</tr>
<tr>
<td>Staff training</td>
<td></td>
<td>900</td>
<td></td>
</tr>
<tr>
<td>Course fees</td>
<td></td>
<td>900</td>
<td></td>
</tr>
<tr>
<td><strong>Total staff costs</strong></td>
<td></td>
<td></td>
<td><strong>25,515</strong></td>
</tr>
<tr>
<td>Office Space</td>
<td>Rental/ Cleaning</td>
<td>1,512</td>
<td>1,512</td>
</tr>
<tr>
<td>Admin. Support</td>
<td>Reception/Admin</td>
<td>1,350</td>
<td>1,350</td>
</tr>
<tr>
<td>Tolls</td>
<td>Staff 1</td>
<td>108</td>
<td>1,350</td>
</tr>
<tr>
<td></td>
<td>Staff 2</td>
<td>216</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staff 3</td>
<td>1,026</td>
<td></td>
</tr>
<tr>
<td>Mail-outs</td>
<td>Newsletter</td>
<td>200</td>
<td>1,305.65</td>
</tr>
<tr>
<td></td>
<td>Postage</td>
<td>491.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Envelopes</td>
<td>122.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Photocopying</td>
<td>491.40</td>
<td></td>
</tr>
<tr>
<td><strong>Total overheads</strong></td>
<td></td>
<td></td>
<td><strong>5,517.65</strong></td>
</tr>
<tr>
<td><strong>Total cost of intervention support for study patients</strong></td>
<td></td>
<td></td>
<td><strong>31,032.65</strong></td>
</tr>
</tbody>
</table>

*Figures supplied by the regional Sports Foundation accounting department

4.4.2.3 General Practice Staff Costs

At participating practices, the average charge for patients with a community services card (A1) or high-users card (AZ), during the 2000-2001 period, was $20. The average for those without a card (A3) was $35. The usual general practitioner consultation time within the study was 15 minutes. The average fee, received from patients’ charges and government subsidies, was $35 ($140/hour revenue from general practitioner consultations). The recommended standard practice nurse wage was $19.12 per hour assuming at least five years of practice experience (personal communication with the national Nurses Organisation). The average recorded time to give a Green Prescription was seven minutes for general practitioners and thirteen minutes for nurses.

Amongst the 451 intervention study patients, 385 received their Green Prescription from a general practitioner. The other 66 patients received the intervention from the practice nurse—
either by way of referral from the general practitioner, who did not have time to give the intervention, or within independent practice-nurse consultations. The opportunity cost for the Green Prescription delivered by a general practitioner was just under one consultation for every two Green Prescriptions delivered. The costs involved in delivering the intervention are estimated to be $14.59 per patient (Table 4-29).

Table 4-29 General Practice Staff Costs of Delivering the Green Prescription Intervention

| 385 Green Prescriptions given by general practitioners @ 7min/script: 0.117hr x $140 = $16.38 and $16.38 x 385 patients = $6,306.30. |
| 66 Green Prescriptions given by practice nurses @ 13min/script: 0.217hr x 19.12/h = $4.15 and $4.15 x 66 patients = $273.90 |
| $6,306.30 + $273.90 = $6,580.20 / 451 patients = $14.59 per patient |

Very few patients recalled receiving follow-up advice from the general practitioner or nurse with respect to the Green Prescription programme over the following 12 months. Of the 389 intervention patients who attended follow-up, 87 said that they recalled receiving further advice from their health professional throughout the year after the initial intervention (Table 4-30). The added cost spread across all patients receiving the intervention was an extra $4.60 per patient over the 12 months after intervention delivery.

Table 4-30 Average Cost of Follow-up Advice from General Practitioner or Practice Nurse during the 12 months following Intervention Delivery

<table>
<thead>
<tr>
<th>Doctor or nurse who gave follow-up</th>
<th>Number of subjects receiving advice throughout the year</th>
<th>Average duration of advice/session*</th>
<th>Average number of episodes of follow-up/yr</th>
<th>Total Cost</th>
<th>Cost per patient **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctor</td>
<td>64</td>
<td>2.3 minutes</td>
<td>2.9</td>
<td>$995.84</td>
<td>$15.56</td>
</tr>
<tr>
<td>Nurse</td>
<td>14</td>
<td>7.8 minutes</td>
<td>6.1</td>
<td>$212.27</td>
<td>$15.16</td>
</tr>
<tr>
<td>Not stated</td>
<td>9</td>
<td>5.8 minutes</td>
<td>4.8</td>
<td>$584.64</td>
<td>$64.96</td>
</tr>
<tr>
<td><strong>Total average cost/patient #</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>$4.61</strong></td>
<td></td>
</tr>
</tbody>
</table>

*As estimated by patient. **Assuming cost of nurse time $19.40/hr and opportunity cost of GP time $140/hr. For those not stated, it was assumed that the advice had come from the GP (more conservative assumption than assuming the nurse saw them). # Total cost/389 patients who had received intervention and had returned for follow-up. It was assumed that non-attendees had the same rate of follow-up advice from health professionals.
4.4.3 Offset Costs

4.4.3.1 Patient Exercise Costs

Personal costs incurred as a result of exercise are presented in Table 4-31. Three hundred and thirty intervention participants responded to the question on whether they found their increased physical activity pleasurable, not pleasurable or neither. Of these respondents, 271 (82%) said it was pleasurable, 18 (5.5%) said it was not pleasurable and 41 (12%) said it was neither. As most participants found these activities pleasurable and because the increased exercise was generally not done during work hours, zero opportunity cost for the time spent on increased activity was assumed. Furthermore, 48% of the participants were retired or not in current employment.

Only 85% of the study participants attended the 12-month follow-up. The percentage not attending follow-up was the same for intervention and control groups. It was assumed that the costs were as high for non-attendees in each group as attendees of the same group, to obtain the most conservative costing intention-to-treat analysis (i.e. intervention and control groups). The incremental cost difference of exercise-related costs incurred by the intervention compared with the control group was $26.96 per patient per year.

Table 4-31 Offset Patient Costs Associated with Exercising for the Year between Baseline and Follow-up

<table>
<thead>
<tr>
<th>Description</th>
<th>Intervention total cost (cost per patient) in $NZ *</th>
<th>Control total cost (cost per patient) in $NZ *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 389</td>
<td>n = 361</td>
</tr>
<tr>
<td>Exercise/ sports shoes</td>
<td>$16,823 ($43.25)</td>
<td>$12,380 ($34.29)</td>
</tr>
<tr>
<td>Exercise group, sports club or gym</td>
<td>$23,004 ($59.14)</td>
<td>$17,986 ($49.82)</td>
</tr>
<tr>
<td>membership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise or physical activity</td>
<td>$15,871 ($40.80)</td>
<td>$9,673 ($26.80)</td>
</tr>
<tr>
<td>equipment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other costs associated with exercise</td>
<td>$3,991 ($10.26)</td>
<td>$3,688 ($10.22)</td>
</tr>
<tr>
<td>Travel cost #</td>
<td>$32,228 ($82.85)</td>
<td>$31,843 ($88.21)</td>
</tr>
<tr>
<td>Total</td>
<td>$91,917 ($236.30)</td>
<td>$75,570 ($209.34)</td>
</tr>
</tbody>
</table>

* Number of participants that incurred this cost in the 12 months between baseline and follow-up, multiplied by the average cost incurred, gives the total cost incurred by the group.

# Travel cost equals total km per week x 16.6 cents/km x 52 week = cost per year
4.4.3.2 Changes in Primary and Secondary Health Care Costs (Year2 - Year1)

Table 4-32 shows the number of general practice visits of intervention and control patients for the year preceding baseline and the year between baseline and follow-up. This data were collected on 677 study participants (74% intervention participants [n = 332] and 81% control participants [n = 346]). Eight practices (5 intervention and 3 control) were not able to, or chose not to provide the information on general practitioner visits. In addition, data were missing on a few individuals who left the area soon after enrolling in the study.

Community services or high-users subsidy cards were held by 54% of control patients (n=231) and 49% of intervention patients (n=219). Community services card status could not be ascertained for four of the 878 participants. It was assumed that these patients did not have a community services subsidy card for the intention-to-treat analysis.

An average charge to patients for general practitioner consultations in the region was approximately $35.00 for an A3 consultation and $20.00 for an A1 or AZ consultation. An average cost to the patient of $10.00 was used for accident-related visits to a physiotherapist, chiropractor or osteopath. (Personal communication with local health providers)

The number of physiotherapy, osteopathy and chiropractic accident-related patient visits for the year prior to baseline and the year following baseline, are also included. Information was available on 750 of the 878 participants because these data were collected at follow-up. An intention-to-treat analysis was also conducted that assumed those who did not attend follow-up had no change in their rate of visits over the two years.
### Table 4-32 Change in Primary Health Care Use by Study Participants for the Year-Before and the Year-After Baseline (adjusted for inflation and discounted at 5%)

<table>
<thead>
<tr>
<th>Description</th>
<th>Intervention: Yr 1 (n*) [average no. visits/patient]</th>
<th>Intervention: Yr 2 (n) [average no. visits/patient]</th>
<th>Change Yr2-Yr1 av. no. visits/patient</th>
<th>Control: Yr 1 (n) [average no. visits/patient]</th>
<th>Control: Yr 2 (n) [average no. visits/patient]</th>
<th>Change Yr2-Yr1 av. no. visits/patient</th>
<th>Incremental Change average no. visits/patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total GP visits: A1</td>
<td>1227 (n=145) [8.46]</td>
<td>1139 (n=146) [7.80]</td>
<td>-0.66</td>
<td>1294 (n=157) [8.24]</td>
<td>1296 (n=157) [8.25]</td>
<td>0.01</td>
<td>-0.67</td>
</tr>
<tr>
<td>GP visits: A3 [cost per patient]</td>
<td>846 (n=170) [4.98]</td>
<td>808 (n=171) [4.73]</td>
<td>-1.46</td>
<td>858 (n=172) [4.99]</td>
<td>851 (n=170) [5.01]</td>
<td>0.02</td>
<td>-1.48</td>
</tr>
<tr>
<td>GP visits: AZ [cost per patient]</td>
<td>191 (n=15) [12.73]</td>
<td>189 (n=15) [12.60]</td>
<td>-0.13</td>
<td>233 (n=15) [15.53]</td>
<td>257 (n=15) [17.13]</td>
<td>1.6</td>
<td>-1.73</td>
</tr>
<tr>
<td>GP Visits Total [cost per patient]</td>
<td>2264 (n=330) [6.86]</td>
<td>2136 (n=332) [6.43]</td>
<td>-0.43</td>
<td>2385 (n=346) [6.93]</td>
<td>2404 (n=344) [7.03]</td>
<td>0.10</td>
<td>-0.53</td>
</tr>
<tr>
<td>ACC GP visits: A1 [cost per patient]</td>
<td>168 (n=145) [1.16]</td>
<td>171 (n=145) [1.18]</td>
<td>0.02</td>
<td>143 (n=157) [0.91]</td>
<td>164 (n=157) [1.04]</td>
<td>0.13</td>
<td>-0.11</td>
</tr>
<tr>
<td>ACC GP visits: A3 [cost per patient]</td>
<td>101 (n=170) [0.59]</td>
<td>107 (n=171) [0.62]</td>
<td>0.03</td>
<td>169 (n=172) [0.98]</td>
<td>137 (n=170) [0.81]</td>
<td>-0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>ACC GP visits: AZ [cost per patient]</td>
<td>17 (n=15) [1.13]</td>
<td>21 (n=15) [1.40]</td>
<td>0.27</td>
<td>29 (n=15) [1.93]</td>
<td>59 (n=15) [3.93]</td>
<td>2.00</td>
<td>-1.73</td>
</tr>
<tr>
<td>ACC GP Visits Total [cost per patient]</td>
<td>286 (n=330) [0.87]</td>
<td>299 (n=331) [0.90]</td>
<td>0.03</td>
<td>341 (n=346) [0.99]</td>
<td>360 (n=344) [1.05]</td>
<td>0.06</td>
<td>-0.03</td>
</tr>
<tr>
<td>#Physio, chiroprac, osteopath visits</td>
<td>518 (n=366) [1.42]</td>
<td>602 (n=367) [1.64]</td>
<td>0.18</td>
<td>521 (n=337) [1.54]</td>
<td>577 (n=333) [1.73]</td>
<td>0.19</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

* 'n' refers to the number of patients in each category

** Adjusted cost multiplies total cost by 1.0398 to allow for inflation (which adjusts according to the ratio of December 2001 consumer price index (1,065) to the December 2000 equivalent (1,046)) The discount rate used is 5%.

# This category includes visits to a physiotherapist, chiropractor, or osteopath.
Approximately 34% (299/878) of study patients reported being admitted to a public hospital or attending a hospital outpatients clinic in the year prior to baseline or the year between baseline and follow-up. Table 4-33 shows the public hospital costing-data, obtained for 282/299 (94%) of these patients (144 intervention and 138 control patients). Unique identifying numbers (NHIs), and therefore costing data, could not be obtained for 17 participants (10 intervention patients and 7 control patients) who had used secondary health care services within this period. A total of $993,387 ($1,630,694 adjusted to 2001 values) was spent on public secondary care for the study participants over the two-year period.

With respect to the rate of private hospital use, thirty-seven patients said they had attended a private hospital in the year between enrolment and 12-month follow-up (19 intervention and 18 control patients). Actual costs for private hospital admission could not be obtained. Therefore, the analysis does not include these costs.

Table 4-33 Change in Costs of Public Hospital Admissions and Outpatient Use for the Year-before and the Year-after Study Enrolment (costs adjusted for inflation)

<table>
<thead>
<tr>
<th>Cost description</th>
<th>Intervention Yr 1</th>
<th>Intervention Yr 2</th>
<th>Intervention Yr2-Yr1</th>
<th>Control Yr 1</th>
<th>Control Yr 2</th>
<th>Control Yr2-Yr1</th>
<th>Incremental change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cost ($NZ)</td>
<td>$370,189</td>
<td>$514,889</td>
<td>$144,700</td>
<td>$267,119</td>
<td>$478,497</td>
<td>$211,378</td>
<td>-$66,678</td>
</tr>
<tr>
<td>Mean cost per patient who used sec care</td>
<td>$2.571 (n = 144)</td>
<td>$3.576 (n = 144)</td>
<td>$1.005 (n = 144)</td>
<td>$1.936 (n=138)</td>
<td>$3.467 (n=138)</td>
<td>$1.531 (n=138)</td>
<td>-$526</td>
</tr>
</tbody>
</table>

4.4.3.3 Changes in Loss of Productivity

These data were collected at follow-up, which included information from 85% (n = 750) of those enrolled in the study. Fifty-two percent of those who attended follow-up (n = 393) were in paid employment during the year between baseline and follow-up. Complete sets of data on the estimated number of days off work due to sickness or accident were available for 382/393 (97%) of patients who reported being employed, at follow-up. The change in the number of days of non-accident-related and accident-related sick leave taken from the year before and the year after enrolment in the study, for the intervention and the control groups is included in Table 4-34.
### Table 4-34 Change in the Cost of Days of Work due to Sickness or Accident during the Year-before compared with the Year-after Study Enrolment for Intervention and Control Study Participants

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sick days leave</td>
<td>842.5 [4.15]</td>
<td>887.5 [4.25]</td>
<td>0.10</td>
<td>721.5 [4.05]</td>
<td>749.5 [4.16]</td>
<td>0.11</td>
<td>-0.01</td>
</tr>
<tr>
<td>Total ACC days leave</td>
<td>69.5 [0.34]</td>
<td>153 [0.74]</td>
<td>0.40</td>
<td>187 [1.05]</td>
<td>267 [1.50]</td>
<td>0.45</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

#### 4.4.3.4 Incremental Offset Costs

Incremental offset costs were calculated for intervention patients compared with control patients, adjusting for clustering using a generalised least square random effects regression model in STATA, and using an intention-to-treat approach (Table 4-35) (Stata Corporation 1985-2001). There were no statistically significant differences between the groups in any of the hypothesised offset cost categories. However, all incremental offset costs were included in the cost-effectiveness analysis to avoid the risk of a type-2 error, which could result if there was an insufficient sample size to show a real difference. There were trends evident, particularly in visits to the general practitioner and hospital costs, which decreased in the intervention group relative to the control group in the year following the intervention, compared with the year before the intervention. Large 95% confidence intervals were evident for the offset costs, reflecting the uncertainty of estimates with the short time span and small sample size. Alternatively, they may reflect the fact that the offset costs were not significantly different between the groups. This was particularly likely for the differences in loss of productivity costs, which were different by only $1.21 (95% CI: -$522.06, $524.49) per patient, between the groups, over the two years.
Table 4-35 Offset Costs per Patient for the Intervention Group compared with the Control Group (Intention-to-treat Analysis)

<table>
<thead>
<tr>
<th>Cost Variable (NZ$)</th>
<th>Intervention # change (Yr2-Yr1)</th>
<th>Control group # change (Yr2-Yr1)</th>
<th>Between-group difference (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health-funder costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC referrals</td>
<td>1.21 (-8.08, 10.50)</td>
<td>1.56 (-9.13, 12.20)</td>
<td>-0.36 (-14.43, 13.72)</td>
<td>0.96</td>
</tr>
<tr>
<td>GP visits</td>
<td>-4.01 (-7.98, -0.04)</td>
<td>-0.05 (-7.15, 7.05)</td>
<td>-4.39 (-15.41, 6.62)</td>
<td>0.44</td>
</tr>
<tr>
<td>ACC GP visits</td>
<td>0.34 (-5.05, 5.73)</td>
<td>0.78 (-6.07, 7.63)</td>
<td>-0.45 (-9.09, 8.20)</td>
<td>0.92</td>
</tr>
<tr>
<td>Hospital costs</td>
<td>320.85 (-69, 711)</td>
<td>495.03 (108, 882)</td>
<td>-174.19 (-722.75, 374.38)</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>Patient costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC referrals</td>
<td>0.84 (-4.01, 5.69)</td>
<td>1.04 (-4.48, 6.56)</td>
<td>-0.20 (-7.50, 7.10)</td>
<td>0.96</td>
</tr>
<tr>
<td>GP visits</td>
<td>-7.24 (-16.80, 2.37)</td>
<td>0.89 (-11.00, 12.80)</td>
<td>-8.21 (-27.75, 11.32)</td>
<td>0.41</td>
</tr>
<tr>
<td>ACC GP visits</td>
<td>0.24 (-1.82, 2.30)</td>
<td>0.44 (-2.18, 3.06)</td>
<td>-0.20 (-3.51, 3.11)</td>
<td>0.91</td>
</tr>
<tr>
<td>Costs of exercise</td>
<td>236.29 (192, 281)</td>
<td>209.37 (152, 267)</td>
<td>26.95 (-45.08, 98.98)</td>
<td>0.46</td>
</tr>
<tr>
<td><strong>Productivity costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sick-days off work</td>
<td>42.19 (-166, 251)</td>
<td>37.47 (-78.20, 153)</td>
<td>$1.21 (-522.06, 524.49)</td>
<td>0.99</td>
</tr>
</tbody>
</table>

*Adjusted for clustering; # Not adjusted for clustering

Table 4-36 shows the incremental cost differences, for the intervention compared with the control participant. The cost from the programme-funders’ perspective was $170.45/patient/year if all those who received the Green Prescription were included in the analysis. A sensitivity analysis was performed that included only those referred for on-going support from exercise specialists, in which the programme cost was $304.97/patient/year. The incremental cost from a societal perspective was $37.16/patient/year if all those who received the Green Prescription were included in the analysis, and $171.68/patient/year in the sensitivity analysis. Again, there are large 95% confidence intervals when offset costs are included in the analysis.
<table>
<thead>
<tr>
<th>Description of costs</th>
<th>Incremental costs per patient (NZ$)</th>
<th>Sensitivity analysis (including only those Green Prescription recipients that were referred to exercise specialists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green Prescription set-up and co-ordinating costs</td>
<td>$82.45</td>
<td>$216.97</td>
</tr>
<tr>
<td>Regional sports trust support costs</td>
<td>$68.81</td>
<td>$68.81</td>
</tr>
<tr>
<td>General practice delivery of intervention costs</td>
<td>$14.59</td>
<td>$14.59</td>
</tr>
<tr>
<td>General practice follow-up support costs</td>
<td>$4.60</td>
<td>$4.60</td>
</tr>
<tr>
<td><strong>Total programme costs</strong></td>
<td><strong>$170.45</strong></td>
<td><strong>$304.97</strong></td>
</tr>
<tr>
<td>Total patient offset costs</td>
<td>$18.62 (-55.63, 92.88)</td>
<td>$18.62 (-55.63, 92.88)</td>
</tr>
<tr>
<td>Total health funder offset costs</td>
<td>-$178.94 (-728.58, 370.70)</td>
<td>-$178.94 (-728.58, 370.70)</td>
</tr>
<tr>
<td>Productivity offset costs (accident- and non-accident-related)</td>
<td>$1.21 (-522.06, 524.49)</td>
<td>$1.21 (-522.06, 524.49)</td>
</tr>
<tr>
<td><strong>Total incremental cost difference</strong></td>
<td><strong>$37.16 (-945.21, 1019.53)</strong></td>
<td><strong>$171.68 (-810.69, 1154.05)</strong></td>
</tr>
</tbody>
</table>

### 4.4.3.5 Sensitivity Analysis Excluding Outliers

Although offset costing data were generally distributed normally, mean values were influenced by outlying values, particularly within hospitalisation data and loss of productivity data. This distribution of data also threatened the stability of the regression model and tended to produce large 95% confidence intervals around point estimates. Figure 4-23 shows the distribution of data for total offset costs for the intervention group and the control group.
In a sensitivity analysis of total costs of the intervention from a societal perspective (including offset costs), nine outliers were excluded. These included all those that had more than a change over the two years in offset costs, of more than +$30,000 or -$20,000, because there was a natural step in costs at these levels in the distribution of the data. The data without these outliers had a more even spread, although most patients still had minimal change in offset costs between the year before and the year after baseline (Figure 4-24).

This sensitivity analysis of total costs, adjusting for clustering using a generalised least square random effects regression model significantly changed the total programme and offset costs per patient. Instead of $37.16 (95% CI: -$945.21, $1019.53) per patient, without outliers, the total cost was $272.25 (95% CI: -$199.39, $743.89).
Figure 4-24 Distribution of Change in Programme and Offset Costs (Offset) for Control Patients (0) and Intervention Patients (1) in New Zealand Dollars for the Year before Baseline compared with the Year following Baseline with Outliers* Removed

* Nine outliers (9/878) were removed from the dataset if total offset costs were greater than +$30,000 or -$20,000 change between the year before and year after study baseline

4.4.4 Incremental Cost-effectiveness Ratios and Cost-effectiveness

The incremental changes of physical activity in the Green Prescription intervention group compared with the ‘usual care’ control group, were included from the previous cluster randomised controlled trial. These changes were 34 minutes/week (95% CI: 2.4, 64.2) or 0.38 kcal/kg/day (95% CI: 0.07, 0.69) for moderate/vigorous-intensity leisure-time physical activity and 1.34 kcal/kg/day (95% CI 0.57, 2.12) for total energy expenditure. Cost effectiveness ratios from a programme funder perspective and a societal perspective are presented in Table 4-37. Sensitivity analyses are also presented in Table 4-37, which exclude Green Prescription recipients who were not referred to exercise specialists, from the analyses.
Table 4-37 Incremental Monthly Cost-Effectiveness Ratios (CER) for the Green Prescription vs. Usual Care from Programme-Funder and Societal Perspectives with Sensitivity Analysis

<table>
<thead>
<tr>
<th>Monthly Incremental Cost Categories</th>
<th>#CERs from programme-funder perspective</th>
<th>#CERs from a societal perspective (95% confidence interval)</th>
<th>*Sensitivity analysis of CERs from programme-funder perspective</th>
<th>*Sensitivity analysis of CERs from a societal perspective (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of programme per participant per month</td>
<td>$14.20</td>
<td>$3.10 (-78.76, 84.96)</td>
<td>$25.41</td>
<td>$14.31 (-67.56, 96.17)</td>
</tr>
<tr>
<td>Cost of increase in energy expenditure per kcal/kg/day</td>
<td>$10.59</td>
<td>$2.31 (-58.78, 63.40)</td>
<td>$18.96</td>
<td>$10.68 (-50.42, 71.77)</td>
</tr>
<tr>
<td>Cost of increase in moderate or vigorous leisure activity per kcal/kg/day</td>
<td>$37.37</td>
<td>$8.16 (-207.26, 223.58)</td>
<td>$66.87</td>
<td>$37.65 (-177.79, 253.08)</td>
</tr>
<tr>
<td>Cost of increasing moderate or vigorous leisure activity by one hour per week</td>
<td>$25.36</td>
<td>$5.47 (-138.90, 149.84)</td>
<td>$45.38</td>
<td>$25.24 (-119.15, 169.61)</td>
</tr>
</tbody>
</table>

# Allocating programme set-up costs to all Green Prescription recipients * Allocating programme set-up costs only to Green Prescription recipients that were referred to exercise specialists

Table 4-38 presents a cost-effectiveness analysis of the Green Prescription intervention compared with the ‘Lifestyle’ Project Active exercise intervention and the ‘Structured’ Project Active exercise intervention (Sevick, Dunn et al. 2000). The Project Active interventions were community-based and evaluated amongst less-active adults, although the age group differed slightly (35-60 years) and the follow-up was 24 months. Even so, the Green Prescription intervention was more cost-effective than the Project Active interventions.

The results of the Green Prescription randomised controlled trial also showed that for every 10.3 Green Prescriptions written and referred, one extra person reached the target of 2½ hours of at least moderate leisure-time physical activity per week. Therefore the cost of converting one ‘sedentary’ adult to an ‘active’ adult, over a twelve-month period, was NZ$1,756 in programme costs. This meant that the Green Prescription intervention was also more cost-effective than the ‘Prescription for Exercise’ programme in primary care, evaluated in the United Kingdom (UK) (Stevens, Hillsdon et al. 1998). The cost of
converting one ‘sedentary’ adult to an ‘active’ adult, in the UK programme, was NZ$8,663 \(^3\) (UK£2,500) in programme costs. The UK evaluation included a similar age group (45-74 years) although the follow-up was slightly shorter (eight months).

Table 4-38 Cost Effectiveness Ratios for the Green Prescription Compared with Project Active ‘Lifestyle’ and ‘Structured’ Physical Activity Promotion Programmes \(^a\)

<table>
<thead>
<tr>
<th>Monthly Incremental Cost Categories</th>
<th>CERs for Green prescription programme at 12 months</th>
<th>CERs for Project Active ‘Lifestyle’ Programme at 24 months*</th>
<th>CERs for Project Active ‘Structured’ Programme at 24 months*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of programme per participant per month</td>
<td>$14.20</td>
<td>$41.26</td>
<td>$118.62</td>
</tr>
<tr>
<td>Cost of change in energy expenditure per kcal/kg/day (95% CI)</td>
<td>$10.59</td>
<td>$48.11</td>
<td>$170.80</td>
</tr>
<tr>
<td>Cost of change in at least moderate intensity activity per kcal/kg/day (95% CI) (^a)</td>
<td>$37.37</td>
<td>$43.31</td>
<td>$358.43</td>
</tr>
</tbody>
</table>

\(^a\) (Sevick, Dunn et al. 2000)

All costs were converted to New Zealand dollars using the December 2001 exchange rate, $NZ1 = $US0.4157 or $US1 = $NZ2.4056.

* Comparisons with the Project Active 6-month results were not used, as these values were even less cost-effective than at 24 months (Sevick, Dunn et al. 2000).

\(^a\) The Green Prescription study used moderate or vigorous leisure-time physical activity; whereas the Project Active study included all moderate or vigorous physical activity. This may produce an underestimate of the comparative cost-effectiveness of the Green Prescription intervention.

4.4.5 Summary

The monthly cost per participant of the Green Prescription programme was $14.20 from a programme funder’s perspective and $3.10 (95% CI: -78.76 to 84.96) from a society’s perspective, if all Green Prescription recipients were included in the analysis. In a sensitivity analysis, it was found that if only those who were also referred to the exercise specialists at Sports Foundations, were included, then the monthly costs per participant were $25.41 and $14.31 (95% CI: -$67.56 to $96.17), respectively.

When offset costs of intervention and control groups were compared, there were no significant differences. However, there was a trend towards more exercise-related costs to the patient (such as exercise shoes, equipment or membership fees) and lower hospitalisation costs for intervention patients compared with control patients. All offset costs were included

\(^3\)The exchange rate used was UK£ 0.2886 = NZ 1.00 from December 2001
in the calculation of cost effectiveness ratios from a societal perspective. However, these estimates have large confidence intervals showing the uncertainty surrounding any possible offset of costs.

The cost of increasing moderate or vigorous leisure-time physical activity in one person by one hour per week, using the Green Prescription programme, was $25.36 per month from a programme funder’s perspective and $5.47 (95% CI: -$138.90 to $149.84) per month from a societal perspective. The total cost of moving one ‘less-active’ person to the recommended level of activity over a one-year period, was $1,756 in programme costs. The Green Prescription intervention was more cost-effective than the Project Active community exercise interventions in the United States and the ‘Prescription for Exercise’ primary care scheme in the United Kingdom.
5 Discussion

5.1 Introduction

This chapter provides a discussion of the three inter-related studies that make up this thesis. The first part refers to the findings of the validity and reliability study of the two physical activity questionnaires. The second part discusses the findings of the pilot study and cluster randomised controlled trial of the Green Prescription intervention. Lastly, the findings of the cost-effectiveness analysis of the Green Prescription are discussed. Chapter 6 will present the conclusions of the thesis.

5.2 Discussion of the Physical Activity Questionnaire Reliability and Validity Study

5.2.1 Main Findings

The AHS questionnaire has demonstrated acceptable test-retest reliability and adequate validity when compared with the reference standards of 7-day diary and pedometer recordings for moderate-intensity physical activity as well as total energy expenditure. Results from the GSS questionnaire tended to underestimate all activity variables compared with the diary, although correlation coefficients were acceptable when compared with the 7-day diary and pedometer recordings. Both questionnaires, and the AHS physical activity questionnaire in particular, would be appropriate for use within epidemiological study in primary care. The results from this study have been published (Elley, Kerse et al. 2003a).

5.2.2 How the Findings relate to Previous Research

For the purposes of this study, accepted levels of reliability and validity reported in the literature from studies using comparable physical activity measurement techniques have been used as a benchmark to measure the adequacy of results achieved here (Appendix 3). The correlation coefficients of validity for the two physical activity questionnaires appear to be as good as that achieved for other physical activity questionnaires used in epidemiological study.
5.2.2.1 Test-retest Reliability of the Three-month Recall AHS Questionnaire

In this study, Spearman's correlation coefficients for test-retest reliability ranged from 0.71 for total energy expenditure estimates, to 0.48 for leisure-time moderate-intensity activity (Chapter 4, section 4.2.4). Like many studies of physical activity questionnaire reliability, total energy expenditure tended to have higher test-retest reliability than other activity sub-categories as discussed in Chapter 2, section 2.3 (Bharathi, Sandhya et al. 2000; Wareham, Jakes et al. 2002).

Although the correlation coefficients found in this study are not high, they are in the same range as those achieved with other physical activity questionnaires used in large population-based studies (Table 7-1, Appendix 3). Examples of reliability correlation coefficients from other questionnaires for different activity categories include the EPIC questionnaire (0.47-0.89) (Pols, Peeters et al. 1997), a modified Baecke questionnaire (0.73-0.82) and the pre-EPIC questionnaire (0.42-0.60) (Pols, Peeters et al. 1996). Correlation coefficients of reliability in the same range were also found with the College Alumnus questionnaire (0.34-0.72) (Ainsworth, Jacobs DR et al. 1993), and a new physical activity questionnaire used in epidemiological study (0.37-0.62) (Suzuki, Kawakami et al. 1998). The correlation coefficients found in this study are also similar to those of recognised physical activity questionnaires used amongst the elderly. Examples include the Yale physical activity survey for the elderly (0.42-0.65) (Dipietro, Caspersen et al. 1993) and the PASE questionnaire (0.68-0.84) (Washburn, Smith et al. 1993). In addition, correlation coefficients were in the same range as in reliability studies using well-recognised questionnaires amongst other subgroups. Examples of this include the 7-day recall questionnaire (0.33-0.86), Godin-Shephard survey (0.37-0.75), Harvard Alumni Survey (0.23-0.68), National Health Interview Survey (0.31) and Baeke questionnaire (0.25-0.87) amongst Latinos (Rauh, Hovell et al. 1992).

Many of these studies have used Pearson correlation coefficients, which sometimes produce higher values than non-parametric statistics, such as Spearman's correlation coefficients, when data are skewed, which is frequently the case with physical activity data. In addition, the interval between questionnaire administrations has varied markedly, as reported in Chapter 2, section 2.3.6. Despite the variability in methodologies used, correlation coefficients of test-retest reliability for different activity categories achieved in the present study, are acceptable when compared with those found in the literature.
Intraclass correlation coefficients for reproducibility of 0.52 to 0.81 achieved in this study were also variable but in the same range as those achieved by other physical activity questionnaires used in epidemiological research. Examples of these include the physical activity questionnaire used in the WHO-MONICA project (0.45-0.92) (Roeykens, Rogers et al. 1998), and the Canadian fitness survey (0.28-0.86) (Weller and Corey 1998). The Tecumseh Community Health Study questionnaire, the Five City Project questionnaire and the Baeke questionnaire (0.47-0.95) also achieved similar intraclass correlation coefficients for reliability (Philippaerts and Lefevre 1998).

5.2.2.2 Validity of the Three-month Recall AHS Questionnaire and the Two-week Recall GSS Questionnaire

There was reasonable correlation between the two questionnaires and the 7-day diary for total energy expenditure, total moderate activity and leisure moderate activity. Spearman's correlation coefficients ranged from 0.50 to 0.74 for the AHS questionnaire and 0.55 to 0.99 for the GSS questionnaire (Chapter 4, section 4.2.5). Again, the highest correlation coefficients tended to be found for total energy expenditure (r = 0.59-0.74), (except for the GSS questionnaire where leisure-time vigorous activity was very highly correlated with estimates from the 7-day diary (r = 0.99)).

Other validity studies of physical activity questionnaires have used activity diaries as a reference and achieved correlation coefficients in the same range as those found in the current study (Arroll, Jackson et al. 1991; Ainsworth, Leon et al. 1993; Pols, Peeters et al. 1997). For example, correlation coefficients of 0.25-0.86 were achieved with the College Alumnus physical activity questionnaire (Ainsworth, Jacobs et al. 1993), and 0.63-0.73 for heavy activity and 0.23-0.39 for light activity with the ARIC/Baeke questionnaire (Richardson, Ainsworth et al. 1995). Correlation coefficients of 0.39-0.55 for total and heavy activity were achieved with the Minnesota leisure time physical activity questionnaire (Richardson, Leon et al. 1994), and 0.28-0.81 with the EPIC core questionnaire (Pols, Peeters et al. 1997) when compared with 48-hour activity records and 3-day activity diaries, respectively. Although these results suggest the questionnaires are probably adequate for population-based estimation of physical activity, the authors of the latter study recommended that this level of validity was not adequate for estimating individual energy expenditure (Pols, Peeters et al. 1997). Again, many of these studies used Pearson's rather than Spearman's correlation coefficients in their analyses, as discussed in Chapter 2.
Spearman’s correlation coefficients found for the AHS and GSS questionnaires compared with pedometer recordings were 0.37-0.61 for total moderate activity. These coefficients are comparable to those found for other physical activity questionnaires that used motion sensors as a reference such as the questionnaires reviewed by LaPorte (LaPorte, Montoye et al. 1985). Other studies found similar results. Correlation coefficients of 0.49-0.64 were found with the PASE questionnaire (Washburn and Ficker 1999) and 0.52 with a Scottish physical activity questionnaire (Lowther, Mutrie et al. 1999). Moderate-intensity physical-activity as measured by the Minnesota Leisure Time physical activity questionnaire had a correlation coefficient of 0.23 when compared with motion sensors (Richardson, Leon et al. 1994).

The use of correlation coefficients alone to assess agreement between two measurement techniques may be misleading (Bland and Altman 1986). The scale or absolute values from the two measurement techniques may be quite different but still produce high correlation coefficients. For example, mean estimates of activities were very similar for the AHS and the 7-day diary, but significantly lower for the GSS, yet correlation coefficients were as high for the GSS questionnaire as the AHS questionnaire. To assess agreement, Bland and Altman developed a graphical depiction whereby the difference between the values obtained by each measurement technique are plotted for each individual against the mean of the two values (Bland and Altman 1986). The Bland-Altman graphs for total moderate activity and total energy expenditure for the GSS questionnaire show mean discrepancies below zero, indicating underestimation of activity compared with the diary.

The Bland-Altman graphs of total moderate activity and total energy expenditure for the AHS questionnaire compared with the 7-day diary show mean discrepancies of close to zero, but large standard deviations. Consequently, the AHS questionnaire is adequate for epidemiological study but may not be suitable for individual estimation of activity levels. Bland-Altman plots are rarely presented in physical activity questionnaire reliability and validity studies. When they are presented, results are often disappointing (Pols, Peeters et al. 1997). This may reflect the fact that physical activity questionnaires are not precise due to the real variability in peoples’ activity patterns on a daily, weekly or seasonal basis. The reporting of activities is also open to recall bias and variability in individuals’ judgement as to the intensity of activities. In addition, most physical-activity assessment methods estimate
different aspects of activity, resulting in less than optimal agreement of estimates using different methods for measuring physical activity.

5.2.3 Strengths and Limitations of the Study

The 75% response rate and 67% completion rate achieved in the current study, means that study participants are likely to be representative of the target population of ‘less-active’ adults in primary health care, particularly those targeted for physical activity interventions. Validity of the questionnaires has also been assessed for individuals of a low average educational background, not highly educated volunteers, as in other studies (Roeykens, Rogers et al. 1998; Suzuki, Kawakami et al. 1998). Ethnicity was representative of the region. Furthermore, the reliability and validity study was conducted in the every-day general practice setting. These features will aid the generalisability of findings.

The study used a combination of recall and objective activity measures to assess validity of the questionnaires. Test-retest reliability was carried out over one to two weeks. This was an appropriate interval as the interval was not long enough to be asking about different periods, and not short enough to risk remembering answers from the previous administration of the questionnaires. If the time period of recall is different for the two administrations of the questionnaire, correlation coefficients are likely to be lower, reflecting real changes in activity, not weaknesses of the instrument (Booth, Owen et al. 1996).

Limitations of this study include the small sample size, which limits both the power of the analysis and the generalisability of findings. Furthermore, the validation study was undertaken alongside the intervention pilot study. There is a risk that the intervention caused change in activity levels over the validity study period, potentially reducing correlation of activity estimates taken at different times. However, to avoid potential change in physical activity behaviour during the validity study, intervention patients did not undergo the reliability and validity component of the research until after the intervention study was complete. Even so, the patients may have been more aware of their activity levels than usual following the pilot study, which may have biased results.

During the validity study, the physical activity diary recorded the activities of a week, while the AHS questionnaire asked about usual activity over the last three months. Therefore, you would expect some variability in responses from these two tools, which may reflect real
differences in activity, as periods of recall differed in length. The GSS questionnaire had a more similar period of recall to the diary, asking about activity over the previous two weeks.

Activity diaries and pedometer recordings were chosen for validation comparisons in this study because of their practicality and accepted use as validation tools in the literature (LaPorte, Montoye et al. 1985; Wilson, Paffenbarger et al. 1986; Pols, Peeters et al. 1998). However, there is also a potential for correlated errors between diary estimates and questionnaire estimates, because both methods rely on subjective recall of activities. Furthermore, pedometers measure the number of steps while questionnaires ask about all activities and their intensity. Pedometers do not take into account intensity of activity and do not measure upper body activity. Consequently, this and other studies have found only modest agreement between self-reported physical activity (diaries, records and questionnaires) and motion sensors (LaPorte, Montoye et al. 1985; Washburn and Ficker 1999).

Physical activity levels assessed by heart rate monitoring and doubly labelled water techniques are recognised as more accurate reference measurements (Montoye, Kemper et al. 1996; Rennie and Wareham 1998). Even so, when these methods have been used in physical activity questionnaire validation studies, they often produce similar results to studies using self-report validation tools (Philippaerts, Westerterp et al. 1999). The use of doubly labelled water as a validation reference, for total energy expenditure, would have been an optimal choice, although the expense and availability of doubly labelled water limited this as an option. Other objective measures such as heart rate monitoring or VO2 max could have been used as a reference comparison. However, the older age of the participants, and the complications such as heart rate changes due to emotion, or medication, limit their use in older populations. VO2 max is an objective measure, but correlates more closely with physical fitness or vigorous-intensity activity, than light- or moderate-intensity activity (Hopkins, Wilson et al. 1991). The participants of the current study were selected for their low activity status, so such objective measures may have been of limited use in assessing levels of mostly moderate and light activity. Consequently, there are some limitations to the use of objective measures as reference standards to assess the validity of self-reported physical activity. Therefore, a combination of subjective and objective reference standards was used in this study.
5.2.4 Implications and Recommendations for Future Research

The current study has addressed the shortage of physical activity questionnaires that are valid and practical for use in primary care research (Little and Margetts 1996). In addition, total energy expenditure, as well as the component leisure-time, domestic and occupationally related activity, can be estimated from these two short questionnaires. Levels of reliability and validity for the questionnaires were as high as those reported for most other physical activity questionnaires used in epidemiological research. This was found, despite the fact that the questionnaires were filled out by an educationally diverse and 'less-active' population who participate in mostly light or moderate activity.

Because the reliability and validity were judged to be adequate, the AHS physical activity questionnaire was chosen for use in the subsequent randomised controlled trial of the Green Prescription. The AHS questionnaire had already been validated amongst a random sample of 113 adults in New Zealand (Arroll, Jackson et al. 1991). It has now been revalidated amongst 34 systematically selected less-active adults in a primary care setting (Elley, Kerse et al. 2003a).

In general, physical activity estimation from questionnaires will produce only an approximation of reality. However, there is a real need for activity measures that are suitable for use in primary care research. It is also important that valid tools are used to measure outcomes for trials, which has been a priority for the Green Prescription trial in this thesis. Even so, the physical activity questionnaires, to date, are not useful for estimating individual activity levels in clinical practice. The investigation of measures useful for clinical practice is a whole new area of research.

5.3 Discussion of the Green Prescription Cluster Randomised Controlled Trial

5.3.1 Pilot Study Findings

The pilot study demonstrated good acceptability of the study protocol and intervention. High rates of screening, study participation, and two-month follow-up, were achieved and the study participants were drawn from representative ethnic and socio-economic backgrounds.
The pilot study fulfilled the aims of assessing the acceptability to the practice staff and patients, of the screening process, intervention delivery, and study protocol. The pilot study also predicted the likely recruitment and participation rates. Some trends were evident in outcome measures, which reinforced the need for a randomised controlled trial. The pilot study also provided estimates of outcome variable means and standard deviations that could be used to check the sample size calculations for the cluster randomised controlled trial.

5.3.2 Baseline Analysis of the Randomised Controlled Trial

There were high participation rates in the randomised controlled trial, with 74% (117/159) of general practitioners in the region, and 66% of ‘less-active’ patients (878/1322) participating in the study. An analysis of the baseline results from this study has been published (Elley, Kerse et al. 2003b).

This study population of ‘less-active’ adults, visiting their general practitioner, had similar demographic characteristics to the age-matched general population. However, they had a high prevalence of cardiovascular risk factors compared with the general population. Over half of the present study population was hypertensive. Nearly all participants (93%) had at least one risk factor for cardiovascular disease, 79% were overweight and 43% were obese. The prevalence of known diabetes among study participants was 10.5%, which is considerably higher than in the general population. Similar high rates of cardiovascular risk factors have been found in previous studies of less active adults in primary health care (Simons-Morton, Hogan et al. 2000), although the study by Simons-Morton and colleagues involved mainly adults from a higher socio-economic background.

In the present study, even within a ‘less-active’ population, low levels of total energy expenditure were associated with higher cardiovascular risk after controlling for gender, age, body mass index, ethnicity and economic status. Again, these results are similar to those found in the Simons-Morton study (Simons-Morton, Hogan et al. 2000). Therefore, targeting physically inactive patients has the potential to detect individuals at high risk of cardiovascular disease. Changes in other cardiovascular risk factors, as well as inactivity, could follow such systematic screening and subsequent intervention.

Baseline levels of physical activity were lower amongst those who were female, older, of lower socio-economic status, recently hospitalised, taking more medications, or at high cardiovascular risk. These findings support previous research, (Simons-Morton, Hogan et al.
200a) and may help primary care health professionals to identify the ‘at-risk’ individual for physical activity counselling.

In the present study, the baseline SF-36 quality of life scores of this ‘less-active’ population were significantly lower for almost every quality of life parameter, when compared with age-similar and gender-matched norms for New Zealand (Scott, Tobias et al. 1999). This was demonstrated in Chapter 4, Figures 4-14 to 4-17. The lower quality of life may have been associated with the ‘less-active’ and high pre-existing morbidity status of this population. Alternatively, it may have been representative of patients visiting their general practitioner or feeling unwell on the day of recruitment.

The single screening question used to identify sedentary patients as they entered the practice was efficient and effective. Indeed, this question could be easily replicated both in research and in practice. The study population identified was more sedentary than the general population, as only 19% were achieving 2.5 hours of moderate or vigorous leisure-time activity per week compared with at least 60% of New Zealand adults (Ministry of Health 1998). The total energy expenditure of 33.8 kcal/kg/day and 11.6 minutes per day of moderate or vigorous leisure-time activity, are typical of sedentary populations found in other studies (Simons-Morton, Hogan et al. 2000). These findings reflect the sedentary nature of this study population.

The positive predictive value of the physical-activity screening question was encouraging (81%). Although this question was developed empirically, it has not been validated against more objective physical-activity measures. Sensitivity and specificity could not be estimated because the prevalence of ‘inactivity’ amongst adults visiting their general practitioner is not known, and was not measured in this study. It could be assumed that the adult population visiting the general practices had the same prevalence of ‘inactivity’ as the New Zealand population (e.g. 40%). If this were the case, the activity-screening question used in this study would have a sensitivity of 92.5% and specificity of 85.5%. However, these figures are only theoretical. There is a need to assess criterion-related validity and test-retest reliability of this single screening question in the future.

5.3.2.1 Strengths and Limitations of Baseline Analysis

Many prior studies of sedentary adults in primary health care have been carried out within only one or two family practices (Bull and Jamrozik 1998; Stevens, Hillsdon et al. 1998;
Halbert, Silagy et al. 2000). Alternatively, study populations have comprised higher socio-economic well-educated patients from urban or University-linked family practices (The Writing Group for the Activity Counseling Trial Research Group 2001). These characteristics of previous studies limit generalisability of their findings. By contrast, the present study had the advantage of sampling from a wide range of urban and rural general practices (n=42) with a high rate of participation from a large geographical region. These features enhance the external validity of the findings.

Demographic characteristics of study subjects appeared to be representative of the general population. The proportion of Māori in the study (17%) was similar to regional (18%) and national (14%) figures (Department of Statistics New Zealand 2002). Levels of formal education were low but were similar or slightly lower than national levels for similar age groups (Department of Statistics New Zealand 2002). Socio-economic status was similar or slightly lower than national figures using the proportion of community services cardholders as an indicator (Department of Statistics New Zealand 2002; Work and Income New Zealand 2002). Therefore, study participants were from diverse backgrounds. Populations with low socio-economic or low educational status were well represented.

There are, however, some limitations to this study. The characteristics of the 33% of sedentary patients who declined to participate are not known. Some commented they did not have time for baseline measurements on the day of recruitment. This constraint may introduce a bias by excluding busy people. It is not possible from this study to determine how many declined because they were not willing to receive a lifestyle intervention. Nor could it be determined how many did not have time on the day, how many did not feel well enough, and how many did not want to take part in the research. However, it is interesting to note that two thirds of all 'less-active' adults identified in the waiting rooms of medical practices were willing to receive a lifestyle intervention from their general practitioner.

A small percentage of enrolled patients did not complete baseline measures owing to lack of time. It could be argued that recruitment of patients from the waiting room, measurements taken immediately, and interventions delivered at the general practice, are more representative of the real-life circumstances of screening for risk factors and delivering an intervention. However, the conditions and time available in such circumstances are less easily controlled than in a laboratory or pre-arranged setting.
There is also the possibility that the high rates of cardiovascular risk and morbidity were characteristic of adults visiting their general practitioner rather than of 'less-active' adults. However, there is evidence that the levels of morbidity in a random sample of the population are similar to those of a random sample of patients who visit their general practitioner (O'Toole, Driver et al. 1991).

**5.3.2.2 What the Baseline Findings Mean**

As a large at-risk 'less-active' population can be easily identified in general practice, the public health issue of physical inactivity may justifiably be seen as a clinical problem within primary health care. This is particularly so because nearly 80% of all adults visit their general practitioner at least annually, including more than 90% of over 65 year olds (Ministry of Health 1998). Furthermore, two thirds of all those adults identified as 'less-active' were willing to participate in a study that involved a clinician-based lifestyle intervention. Screening for inactivity can be justified because physical inactivity is an important risk factor for cardiovascular disease, and is a marker for other risk factors. Furthermore, inactivity is amenable to certain clinician-based physical activity interventions (The Writing Group for the Activity Counseling Trial Research Group 2001).

**5.3.3 Main Findings of the Cluster Randomised Controlled Trial**

This study has shown that the "Green Prescription" intervention in general practice is effective in increasing physical activity and improving health-related quality of life over 12 months without evidence of adverse effects (Elley, Kerse et al. 2003c). In particular, statistically significant increases were found in leisure-time moderate- or vigorous-intensity physical activity, as well as total energy expenditure amongst the intervention group compared with the control group. Improvements were also found in the SF-36 scores of 'general health', 'role physical', 'vitality', and 'bodily pain' amongst the intervention group compared with the control group. There was no increase in falls, injuries or hospitalisations amongst the intervention group compared with the control group. In fact, there was a trend towards fewer hospitalisations. A trend towards improved blood pressure was also demonstrated amongst the intervention group compared with the control group, with or without controlling for antihypertensive medication change. However, no significant change in 4-year risk of coronary heart disease was detected.
For every ten Green Prescriptions written, one person achieved and sustained 2½ hours of moderate- or vigorous-intensity leisure activity per week, at 12 months. Achieving this amount of activity, approximately 1000 kcal/week, is associated with a 20-30% risk reduction in all-cause mortality compared with persistently sedentary individuals (Lee and Skerrett 2001).

The systematic review, of physical activity counselling in primary health care, that was summarised in Chapter 2, section 2.3 of this thesis, was inconclusive about the longer-term effectiveness of physical activity interventions in primary care. This study has added substantially to the literature, by demonstrating and quantifying an effect of such an intervention over 12 months.

5.3.4 Methodological Strengths and Limitations of the Study

This subsection discusses the techniques used in this trial to address the methodological challenges of conducting effectiveness trials in the community. Both the strengths and the limitations of these techniques and the results achieved in this study are discussed below.

5.3.4.1 Participation Rates

As stated above, the findings of this study have widespread generalisability, due to high participation rates, recruitment of consecutive patients, and the socio-economic diversity of the sample population. In addition, there were high rates of participation at follow-up (750/878, 85%).

Participant recruitment and follow-up rates were higher in this study than in previous primary-care-based physical-activity intervention studies with long-term follow-up, reported in the literature (Hillsdon, Thorogood et al. 1995; The Writing Group for the Activity Counseling Trial Research Group 2001). These high rates may have been due to a variety of factors. Baseline and follow-up measurements were carried out at usual practices to improve participation and follow-up rates, by minimising the inconvenience to the participants. All study participants were supplied with summaries of their results at baseline and follow-up, which may have increased their interest. A record of ‘change in outcome variables’ over the year was not supplied to participants at follow-up, as assessors did not refer back to baseline values after they were collected, until follow-up data collection and data entry was complete. However, patients could look up their own copies of results from the previous year,
independently, or if requested, the research team offered to send out comparative results after data entry. In addition, a $10 petrol voucher was given at follow-up to cover any travel expenses of returning for follow-up, although participants were not usually informed of this prior to the visit. Personalising the approach may also have helped. Wherever possible, the same researcher conducted the baseline measures, re-contacted the patient after 12 months, and conducted the follow-up measures at their usual practice.

The inclusion criteria of 40-79 years of age was chosen as management of cardiovascular risk in New Zealand uses the Framingham equation to assess risk. Change in risk was identified as an outcome measure and the Framingham equation is valid for ages 35-75 years (Anderson 1990; D'Agostino 2000). In addition, recent studies have measured cardiovascular risk for those 35-75 years (Tyre 1994; Bullen 1998) and those 45-75 years (Ministry of Health 1998). Other international trials have used similar age groups, such as the ACT trial (King 1998). These studies provided comparisons for this study. Evenso, there is evidence that older age groups can increase physical activity and reduce risk as a result of primary care interventions (Kerse 1998). Therefore inclusion of participants up to the age of 80 years also allowed subgroup analysis of those under 65 years with those over 65 years. This analysis is being undertaken, presently, but is beyond the scope of this thesis.

5.3.4.2 Adequacy of Randomisation and Intention-to-Treat Analysis

Randomisation was carried out at the level of general practice. Stratification by practice size, ensured intervention and control group sizes were comparable. The systematic screening and invitation of patients to participate in the study meant that the number of patients recruited from each practice was roughly proportional to the number of general practitioners in each practice.

The attrition of four practices (seven general practitioners) in the control group did mean that there were slightly fewer rural-solo, and urban-larger, practices in the control group than the intervention group. This loss also meant that there were fewer patients recruited into the control group than the intervention group. However, both groups had well over 400 participants (427 control and 451 intervention), which was adequate to detect outcome differences as statistically significant according to prior calculations of sample size. There was also no reason to believe that the slightly higher proportion of solo or large practices in the intervention group compared with the control group, would have influenced the response
to the intervention. However, there is a possibility that the characteristics of the practices that withdrew from the study, or their patients, were in some way systematically different from those that did not withdraw, which may have biased results.

The long delay between recruitment of practice and participation of the last practices during the twelve months of rolling recruitment probably contributed to the practice attrition. Continued contact with practices or incentives to participate may have reduced the practice attrition rate. However, in most cases the reasons given for withdrawal were unanticipated and unavoidable.

Even so, patient characteristics were well balanced at baseline. Statistical testing has been considered inappropriate when judging balance of baseline characteristics and adequacy of randomisation (Roberts and Torgerson 1999). However, additional tests have been recommended, such as adjusting for baseline variables in sensitivity analyses of change for each outcome variable, to ensure that no significant baseline imbalance has occurred (Raab and Butcher 2001). Although these sensitivity analyses did not change the outcome for most variables in this study, improvement in 'vitality' was no longer significantly different in the intervention group from the control group. Furthermore, sensitivity analyses of physical activity outcomes, controlling for other possible confounding factors, such as age, gender, ethnicity, and socio-economic status found no effect on results. These findings provide reassurance that randomisation was likely to have been adequate, that the results were robust, and that the attrition of seven general practitioners prior to patient recruitment was unlikely to have caused significant selection bias of study patients.

The randomisation of practices prior to patient recruitment was necessary to allow enrolment and delivery of the intervention to the patient at the same visit to the practice. This strategy not only maximised the delivery of the intervention and the participation of patients enrolling and minimised the time between enrolment and intervention delivery, but it also kept study costs to a minimum. It has been recommended that the time between screening, enrolment and baseline readings be kept to a minimum to reduce the potential for behaviour change and distortion of baseline results (Friedman, Furberg et al. 1985).

However, there is an increased risk of selection bias if allocation of intervention is not concealed prior to enrolment (Schulz, Chalmers et al. 1995). Therefore, patients were kept blind to allocation prior to enrolment, and systematic screening and recruitment methods
were used independent of the researcher, in order to try and minimise the potential for this bias. There were also very few exclusion criteria, and no subjects were excluded post enrolment. (Less than 2% (15/878) of subjects were found to have exclusion criteria after enrolment.) Data from subjects were analysed according to the original randomisation, regardless of actual intervention delivery or compliance with physical activity.

In addition, an intention-to-treat approach was employed, where no change was assumed for participants who did not attend follow-up, for all variables except coronary heart disease risk, where a conservative estimate of deterioration was assigned. This approach has been encouraged in community based effectiveness trials, to mimic real-life situations (Hollis and Campbell 1999). However, others suggest that such an approach risks a type-2 error and is less likely to show a positive outcome (Fergusson, Aaron et al. 2002). Therefore, this study conducted a ‘per protocol’ sensitivity analysis (excluding those that did not attend follow-up) to assess the effect of using an intention-to-treat approach. The results of the sensitivity analyses are presented in Chapter 4, section 4.3.12. The ‘per protocol’ effect estimates tended to be greater, but were sometimes less precise than with an intention-to-treat approach, presumably due to the reduced sample size. The similarity of results with both techniques shows that the results were reasonably stable.

### 5.3.4.3 Lack of Blinding

The main reason the study could not be blinded was that it involved a lifestyle intervention. The doctors and nurses from the practices were aware of the allocation of randomisation, as they were the ones who initiated the intervention. Patients were usually aware that they had or had-not received a lifestyle programme. The investigators were also aware of randomisation, as all measurements were taken at the practice. It would have been difficult to keep an independent assessor blinded to allocation of randomisation, when the subject and practice staff were all well aware. One comment from any of those individuals would ‘unblind’ the allocation of all patients from that practice, because of the clustered design.

However, a review of randomised controlled trials showed that a lack of double blinding was associated with a 17% exaggeration of estimates of treatment effects and odds ratios (Schulz, Chalmers et al. 1995). To minimise the risk of recruitment and assessor bias due to the lack of blinding, self-administered questionnaires, objective and electronic health measures, and signed witness statements of results were used (Friedman, Furberg et al.
The fact that baseline characteristics were well balanced suggests that recruitment and assessor bias were minimised.

Another potential for bias was the fact that, sometimes, baseline measurements could not be completed prior to the patient seeing their general practitioner or nurse. Physical activity status was taken first to ensure that this status was established prior to initiation of intervention. Any outstanding measurements were taken directly after the consultation. On a very few occasions, the SF-36 questionnaire had to be sent home with the patient to complete and send back, due to lack of time at the practice. Within any trial where there is a lifestyle intervention, as soon as the subject knows about the nature of the intervention, their behaviour and therefore baseline measurements may be altered. Therefore, there was a small risk of bias due to some baseline measures being recorded directly after a few patients knew to which group they had been allocated.

5.3.4.4 Bias Due to Self-report

As with other physical activity trials, there was a risk of recall or ‘social desirability’ bias when using self-reported physical activity as an outcome measure, as was discussed in Chapter 2. However, the physical activity questionnaire used in this trial was validated against pedometers and 7-day diaries prior to the trial amongst ‘less-active’ adults in general practice. In addition, the use of physical activity questionnaires has been shown to be both practical and valid for epidemiological study when measured against more objective measures such as motion sensors, heart rate monitoring and doubly labelled water (Philippaerts, Westerterp et al. 1999; Sobngwi, Mbanya et al. 2001).

In addition, the use of clustering and randomising at a practice level meant that study participants were less aware of differences between intervention and control procedures. However, both groups were aware that they were part of a study that asked about exercise. As such, any ‘social desirability’ bias may well have applied to the control and intervention groups in similar ways. Furthermore, ‘change’ rather than absolute values were used as efficacy variables, which may also lessen the chance of any ‘exaggeration’ of outcomes due to recall-bias.
5.3.4.5 Cluster Design

The principal reason for using a clustered design in this study was to reduce the risk of contamination of intervention. If individual randomisation had been used, general practitioners and practice nurses would have been asked to 'turn on' and 'turn off' advice, and to treat their patients differently according to randomisation. This strategy would have been difficult to achieve. The clinicians may have inadvertently given physical activity advice to all patients, because of their heightened awareness of the issue of physical inactivity or their difficulty turning on and off the advice. Alternatively, the clinicians may have changed their usual practice and advice in an attempt not to 'contaminate' the intervention to control patients. Either way, 'usual practice' could have been influenced, which may have biased results of the effectiveness of the intervention compared with 'usual care'.

To assess the degree of contamination of intervention by clinicians at control practices within this trial, all patients were asked at follow-up if they recalled receiving a 'Green Prescription' in the previous year. While 95% of intervention patients recalled receiving the intervention, only 2.8% of control patients recalled receiving a Green Prescription. Therefore, contamination of the intervention as a whole, to the control group was low. However, this does not preclude contamination of part of the intervention. For example, it became apparent at some control practices that the presence of the research personnel heightened practitioner awareness about advising exercise. There was anecdotal evidence that some clinicians increased their physical activity advice to control patients during the week of study enrolment. The fact that the control group also increased their leisure time moderate and vigorous activity during the study, but to a lesser degree than the intervention group, may reflect this. Alternatively, the increase may have been due to patients knowing that they were participating in a study about physical activity. This apparent Hawthorne effect has occurred in previous studies (Arroll 1992). Another possibility was that other influences in the community increased activity levels across the population, affecting both groups in a similar way. In which case, the change as measured in this trial is an appropriate estimate of the incremental effect of the Green Prescription intervention, over and above background influences.

Clustering also reduced the expectation of patients to receive the intervention, as all patients from each practice were treated similarly and contact between control and intervention
groups was kept to a minimum. Therefore, control patients were less likely to seek a similar programme or to object to not receiving the intervention at enrolment. (There were very few complaints about this.) In addition, study patients were informed at enrolment that control patients would be offered the intervention, if appropriate, after the 12 months. This meant that they were not missing out on an opportunity to receive the intervention.

It also became obvious that many of the patients forgot, or were not aware, that they were in the control group of an intervention trial. Some thought the study was designed to follow their physical activity and health characteristics over a year. Others thought that the baseline measures were the intervention. These reactions were found, despite full information about the trial being provided to them at commencement of the study. These beliefs may have worked in favour of the trial producing an unintended pseudo-blinding. The ethics committee had not allowed blinding as to the nature of the intervention, but had agreed for the research team to inform all patients at enrolment that the research involved a lifestyle intervention, without specifying the details. This is an example where ethical issues must be weighed up against scientific rigour. A reasonable compromise was reached in this case.

The risk of co-intervention using medications was also checked. Antihypertensive and lipid lowering medication types and doses were monitored at baseline and follow-up. No significant difference was found between the groups that might have exaggerated or diluted the intervention effect. It is interesting to note that the adjustment for antihypertensive and lipid-lowering medication change had very little effect on the analysis of change in blood pressure and coronary heart disease risk (see sections 4.3.11 and 4.3.12). This is because medication changes were fairly balanced in both groups, which suggests that this was not a source of co-intervention or contamination bias. Most results were unaltered in all the sensitivity analyses.

There are some drawbacks with using a cluster randomised controlled design. Substantially more power could have been achieved by using individual randomisation. In Chapter 2, design effects, based on the intraclass correlation coefficients for each variable and average cluster size, produced large inflation of sample sizes required to detect significant change (e.g. physical activity and blood pressure). More power may have been attained by individually randomising and inflating the sample size to allow for the predicted level of contamination. If the level of contamination had been less than 30%, then it is likely that the inflation factor would have been less than if clustering was used (Slymen and Hovell 1997;
Torgerson 2001). Furthermore, if individual randomisation had been used, the achieved sample size would have been able to detect as statistically significant, changes in variables that were small and not considered statistically significant using a clustered design. For example, differences in blood pressure change found in this study, were statistically significant if clustering was not accounted for. When large trials are expensive to conduct, finding the most efficient way to demonstrate effect is important.

The analyses in this study were adjusted for clustering by practice to account for the difference of the between- and within-practice variances of participant characteristics. In addition, such an approach also allowed for the influence of the particular environment of the practice, and the medical management styles of the general practitioners and practice staff. It also allowed for the variation of intervention delivery at different practices.

### 5.3.4.6 Sub-group Analyses

Post-hoc sub-group analyses were carried out by gender and rurality in Chapter 4, section 4.3.11.5. These analyses found that greater leisure-time physical-activity gains were achieved by men compared with women (68 minutes compared with 20 minutes). The reasons for this are not clear. Urban participants tended to increase their overall total energy expenditure by 7.55 kcal/kg/week more than their rural counterparts. The geographical context may have influenced the response. Anecdotally, there were comments from rural subjects about the limited facilities and sidewalks available in the country for increasing activity. Future studies may investigate these differences more closely.

However, the results of these subgroup analyses should be viewed with caution. This study was not powered to undertake such analyses. In addition, the hypotheses tested by these sub-group analyses were not determined prospectively, and therefore risk a charge of data dredging to find significant differences. It was for these reasons that the only sub-group analyses conducted were those based on gender and rurality, as these are commonly used categories and because the sub-sample sizes were relatively large. Sub-group analyses based on ethnicity, for example, were not done because the study was not powered to do so and sub-group sample sizes were small.
5.3.4.7 Multiple Outcome Measures

The use of multiple comparisons to assess the three primary outcomes (physical activity, cardiovascular risk and quality of life) may risk detecting an effect where one does not exist (type I error) (Freemantle 2001). However, if pre-defined multiple outcomes are needed to demonstrate several effects, these can be justified, without the need for adjustment (Rothman 1986; Perneger 1998). While a significant increase in leisure energy expenditure could have been shown in isolation, it was important to ensure that total energy expenditure had not decreased overall. In addition, an attempt was made to detect different potential health and quality of life benefits from a physical activity intervention.

5.3.4.8 Summary of Methodological Issues

Most of the methodological issues discussed here were due to the nature of conducting a randomised controlled trial of a lifestyle intervention within the context of every-day general practice. While there are obviously multiple methodological difficulties with this approach, the findings may have more external validity than controlled trials conducted in artificial environments, using interventions that may not be sustainable or have the same effect in every-day practice.

5.3.5 How the Findings relate to Previous Research

5.3.5.1 Primary Outcomes

This study builds on the findings of the initial randomised controlled trial of the Green Prescription, which demonstrated increases in self-reported physical activity after 6 weeks (Swinburn, Walter et al. 1998). There was some indication that many intervention patients sustained these increases when surveyed at 12 months (ibid). These assertions have been confirmed by the present randomised controlled trial, which also demonstrated improvements in health parameters such as quality of life and an indication that blood pressure may also be improved.

The present study used a true control group and the patients prompted their usual general practitioner or nurse to deliver the intervention. Previous studies have found it difficult to demonstrate changes in physical activity, some possibly because there was no 'true control' group included in the study (The Writing Group for the Activity Counseling Trial Research Group 2001; Dubbert, Cooper et al. 2002). Both these studies involved brief physical
activity advice given to all three randomised groups, although the nature of follow-up varied. Others have lacked the statistical power to demonstrate effect due to poor rates of intervention delivery (Harland, White et al. 1999). This problem was avoided in the present study because of routine screening and patient prompting.

Furthermore, this study differed from studies that used a visiting activity specialist rather than the usual general practitioner or nurse to initiate the intervention (Stevens, Hillsdon et al. 1998; Harland, White et al. 1999; Halbert, Silagy et al. 2000). This difference may suggest the importance of the role of the usual practitioner, as an intervention agent within a continuing therapeutic relationship, to achieve behavioural change.

Ten percent more intervention patients than control patients moved from being 'less-active' to 'active' (achieving at least 2.5 hours of moderate- or vigorous-intensity leisure activity per week) over the year. This figure is similar to previous studies. For example, a 'Prescription for Exercise' scheme undertaken in the United Kingdom, found that 12% more intervention patients than control patients achieved this status eight months after the intervention (Stevens, Hillsdon et al. 1998). In the present Green Prescription trial, the proportion of study participants that had achieved the 'active' status was 34% of intervention patients and 26% of control patients, as depicted in figure 4-22 of Chapter 4. These figures are also very similar to the INSURE multibehavioural intervention, where 34% of intervention patients and 24% of control patients were active at 12 months (Logsdon, Lazaro et al. 1989). When the control group also increase their activity, it makes it difficult to know if they would have achieved this amount without the study, or whether this was real background improvement. The latter has to be assumed, because participation in the study may have enhanced the response of intervention patients, also. The consistency in the proportion of behaviour change that is achievable by brief physical-activity interventions, across studies, suggests that this may be a realistic expected behaviour change with such interventions.

Most of the leisure-time physical-activity gains in this study were found in moderate-intensity activity. Recent evidence has suggested that increases in moderate-intensity physical activity are not adequate for health benefit and that vigorous-intensity activity is necessary (Yu, Yarnell et al. 2003). However, numerous intervention studies and meta-analyses have shown significant improvements in cardiovascular risk factors, such as lower blood pressure, with increases in moderate activity, as well as vigorous activity (Halbert,
Silagy et al. 1997; Whelton, Chin et al. 2002). The inability of some longitudinal studies to show associations between moderate-intensity activity and health outcomes may be due to the imprecision of measuring moderate and lower intensity activities, or due to other methodological issues. For example, one study used baseline physical-activity participation as the independent variable for a follow-up period of ten years of health outcomes, ignoring potential change in activity levels over that time (Yu, Yarnell et al. 2003).

The present study did not show a statistically significant impact from the Green Prescription intervention on 4-year coronary heart disease risk. This finding may have been partly due to the limited sensitivity of the Framingham equation as a surrogate measure of future cardiovascular events. While exercise is likely to have an effect on blood pressure and lipid profile, it will have less effect on smoking status and diabetic status in the medium term, and no effect on age. The benefits of exercise on other physiological aspects such as coagulation profiles and cardiac contractility (Chapter 2, section 2.2.1.9) will not be measured by the Framingham equation. These omissions may limit the ability of the equation to reflect the full cardiovascular benefit of physical activity. The non-significant reduction in relative risk of coronary heart disease (1.7%) in the intervention group, compared with the control group, was lower than that predicted in sample size calculations (Chapter 3, section 3.3.6). The estimates that were used to predict an achievable reduction in risk were based on reductions found from other lifestyle interventions reported in the literature.

The methods of one such study were quite different from those of the present study, which may have accounted for the difference in reductions in coronary risk achieved in each study. For example, the British Family Heart study found that the overall reduction in coronary risk score was 12% in the intervention practices at one year (Anonymous 1994). The greatest component of this reduction was attributed to a reduction in blood pressure. However, the comparison control group did not have a baseline blood pressure reading taken. Therefore, there was a risk of accommodation bias (Friedman, Furberg et al. 1985), although this was adjusted for in a subsequent analysis (Wonderling, Langham et al. 1996). If the follow-up blood pressures of the intervention group in the Green Prescription trial had been compared with the first set of readings of blood pressure from the control group, the difference would have been statistically significant (3/2 mm Hg) (Chapter 4, sections 4.3.8 and 4.3.11).

In addition, the British intervention included screening for cardiovascular risk and addressed several lifestyle behaviours (physical activity, diet, alcohol consumption and smoking)
The British intervention also resulted in an increase in antihypertensive and lipid lowering medication use amongst the intervention group. Therefore, aspects of the intervention, other than physical activity advice, may have led to the drop in blood pressure, cholesterol and coronary risk. Lastly, the Dundee coronary risk score, rather than the Framingham equation, was used in the British study. Consequently, the coronary risk effect estimate from the Green Prescription study was substantially lower than that predicted from the British Family Heart study.

5.3.5.2 Stage of Change

There was a trend towards greater physical activity gains amongst ‘contemplators’ compared with ‘pre-contemplators’, which has been found for other types of behaviour change (Dunn, Deroo et al. 2001). Even so, ‘pre-contemplators’ tended to be older than ‘contemplators’, which may have limited their ability to increase their activity. This finding may suggest that it was the age and the possible physical limitations of the ‘pre-contemplator’ group, rather than their ‘attitude’ that acted as a barrier to behaviour change.

A relationship between stage of change and actual level of activity has been demonstrated elsewhere, and there have been indications that changes in activity are in the direction predicted by the transtheoretical theory (Marshall and Biddle 2001). However, it is not clear whether physical-activity behaviour change occurs in a series of qualitatively different stages (Marshall and Biddle 2001). One systematic review of the use of motivational interviewing and stage of change techniques for lifestyle counselling found that there was evidence that these techniques were better than other techniques for substance abuse counselling (Dunn, Deroo et al. 2001). However, this was not the case for physical activity counselling (ibid). Even so, many physical activity interventions are based on these techniques (Petrella and Lattanzio 2002). In addition, several physical activity intervention trials have used stage of change as an outcome (Calfas, Long et al. 1996; Pinto, Lynn et al. 2001; Steptoe, Kerry et al. 2001; Woods, Mutrie et al. 2002). Surveys have also used stage of change as an indicator for participation in exercise (Booth, Macaskill et al. 1993). However, these studies have often assumed that stage of change either reflects or predicts activity adherence. More research into the reliability and validity of these categories is needed.
Intraclass correlation coefficients (ICCs) were low for baseline health characteristics, such as cholesterol \((r = 0.004)\) (Chapter 4, section 4.3.13). However, ICCs were high for demographic characteristics such as socio-economic status \((r = 0.12)\), educational level \((r = 0.12)\), and ethnicity, as measured by classification into Māori and non-Māori \((r = 0.38)\). The higher ICCs for demographic characteristics are understandable, as general practices are located in regions that differ in terms of socio-economic characteristics. In addition, seven of the practices that participated were Māori health providers that served mostly Māori, while other practices were mostly attended by non-Māori. Even so, it seems that there was less variability between practices in clinical and quality of life characteristics than expected but more variability amongst demographic variables.

For example, baseline ICCs for self-rated general health, \((0.000)\), role emotional \((0.007)\), mental health and bodily pain \((0.000)\), and physical function \((0.009)\), were all lower than the ICCs used for the quality of life variable in sample size calculations \((0.01)\) (Chapter 3, section 3.3.6). In addition, physical activity variables \((0.01-0.02)\) were lower in this study than those reported previously in the literature \((0.05)\) and used in the sample size calculations. However, baseline ICCs for vitality \((0.01)\), systolic blood pressure \((0.018)\) and cholesterol \((0.004)\) were very similar to those used in the sample size calculations \((0.01, 0.016\) and \(0.0036\), respectively). Therefore, sample size inflation was adequate to detect the changes that were predicted prior to the study.

The publication and availability of ICCs will facilitate appropriate cluster randomised trial sample calculations for future studies in primary health care. Reviews of randomised trials using cluster designs have highlighted the problem of inappropriate sample-size calculations and analysis in cluster randomised trials, and the lack of information about levels of intraclass correlation coefficients (Donner, Brown et al. 1990; Simpson, Klar et al. 1995). Publication of intra-class correlation coefficients from cluster randomised trials has been recommended by others (Donner, Brown et al. 1990; Kerry and Bland 1998).

The intraclass correlation coefficients were calculated, primarily, to assess how much variability there was within practices compared with that between practices. The observed values may influence the decision to use cluster-randomised controlled trials or individually-randomised controlled trials in the future. The degree of inflation \((d.eff)\) needed for a
clustered design using known ICCs, can be calculated if cluster size \((m)\) is also estimated \((d'_{\text{eff}} = 1 + (m-1) \times ICC)\). This inflation factor can be compared with the inflation necessary to allow for 'dilution' of effect due to contamination of intervention within an individually-randomised controlled trial design, as discussed in Chapter 2, section 2.6.3 (Slymen and Hovell 1997; Torgerson 2001). This strategy also assumes that it is possible to estimate the likely degree of contamination of intervention, which is problematic. Even so, the efficiency of using a cluster- or individually-randomised design can be compared, when deciding on the design of a trial.

Of interest was the fact that the intraclass correlation coefficients of 'change', in some variables, such as cholesterol and self-rated 'general health', were higher by a factor of 10 compared with ICCs of baseline values \((r=0.04 \text{ vs. } r=0.004)\). This finding reflects the fact that people at some practices changed these variables more than at other practices. This tendency may have been due to the intervention, as there was a trend towards greater drops in cholesterol and improved 'general health' amongst the intervention group, compared with the control group, although the drop in cholesterol was small. For variables where there was not a systematic change in intervention compared with control practices (eg: coronary heart disease risk, and body mass index), the ICCs of 'change' were actually higher than baseline ICCs. Alternatively, these observations may reflect the greater ability of some practices to help patients change their cholesterol and 'general health' compared with their ability to change body mass index and overall coronary heart disease risk over a 12-month period.

5.3.6 Implications of the Findings

This study has demonstrated that prompting the usual general practitioner for brief advice, coupled with on-going telephone support, can change physical activity behaviour and improve self-rated general health, vitality, role-physical and bodily pain parameters for at least a year. Increases in physical activity and better quality of life have been linked to improved health outcomes, particularly in the elderly (Spirduso and Cronin 2001). If implemented widely, such a strategy could result in major health benefits for sedentary people.

This study did not have sufficient statistical power to detect a change in blood pressure of 1.4 mm Hg as statistically significant, because sample size calculations used larger efficacy estimates from previous reviews of exercise and blood pressure (Halbert, Silagy et al. 1997).
The clinical significance of such a small blood pressure change across a population is also questionable. However, a reduction of 2 mm Hg diastolic blood pressure amongst an adult population could lower the prevalence of hypertension by 17%, the risk of coronary heart disease by 6%, and the risk of strokes and transient ischaemic attacks by 15% (Cook, Cohen et al. 1995). Blood pressure changes in this study resemble long-term changes achieved by other lifestyle interventions such as weight-loss or salt-reduction programmes (Anonymous 1997).

Ten percent more of the intervention patients than control patients achieved and maintained at least 2 ½ hours of moderate- or vigorous-intensity leisure-time physical at 12 months, in the Green Prescription study. A review of 44 longitudinal observational studies found that achieving physical activity levels of at least 1000 kcal/week (approximately equal to 2.5 hours/week of moderate to vigorous intensity physical activity) was associated with a 20-30% reduction in all-cause mortality (Lee and Skerrett 2001). In a conservative analysis, therefore, there is likely to be an average reduction in risk of all-cause mortality of at least 2-3% from the Green Prescription intervention. A similar reduction in 4-year coronary heart disease risk of 1.7% was found in this study, although this result did not reach statistical significance (Chapter 4, sections 4.3.8 and 4.3.11). This analysis does not take into account the intervention patients that increased their physical activity but did not reach the recommended 2 ½ hours per week. There is some evidence to suggest that levels as low as 500 kcal per week (approximately half the recommended amount) may also have positive health outcomes (Paffenbarger, Hyde et al. 1986; Lee and Paffenbarger 1996).

Furthermore, the average increase in total energy expenditure amongst the intervention patients, compared with the control patients, was 975 kcal/week (Chapter 4, section 4.3.11). In addition, it is likely that physical activity increases achieved at 12 months will be maintained, as has been shown in previous research (unpublished data, Ngaire Kerse). Therefore, in the best possible scenario, an average of a 20-30% reduction in all-cause mortality with the Green Prescription intervention could be achieved.

A sub-sample of general practitioners estimated that the time spent giving the Green Prescription intervention to one patient was 7 minutes. A sub-sample of practice nurses estimated that the time they spent was an average of 13 minutes. This amount of time is probably too long to expect a general practitioner to spend on a brief intervention because it makes up nearly half an average consultation time. It is possible that the general
practitioners over estimated the time spent advising exercise, based on anecdotal comments from study patients. A study that filmed or audiotaped real consultations may be able to clarify this. Alternatively, if this was the time required, reimbursement for such preventive measures may be required as an incentive for general practitioners.

Both the control and the intervention groups showed improvements in every outcome measure except coronary heart disease risk, which tends to deteriorate with age. It is possible that a portion of the improvement in self-rated quality of life at follow-up was found because some of the study participants were unwell at the time of their baseline measures, but not at follow-up. (All study participants were visiting their general practitioner on the day of baseline measures. Pre-arranged appointments were made with participants for follow-up measures.) This possibility may also explain why there was an overall improvement in physical activity. Alternatively, participation in the study may have prompted health behaviour change, increasing physical activity and improving quality of life. Changes in attitudes or initiatives in the community or in general practices may also have contributed to a background increase in physical activity levels. There was also a general improvement in blood pressure over the 12 months (statistically significant amongst the intervention group but not amongst the control group). This finding may represent, in part, regression to the mean. Alternatively, the patients may have been feeling better than at baseline, and be more accustomed to the researcher (accommodation), producing lower blood pressures in general.

It was also interesting to note that several variables, including blood pressure, weight, and cholesterol serum concentration showed consistent trends towards improvement amongst the intervention group, compared with the control group. The one inconsistent finding was the fact that HDL serum concentration tended to decrease when compared with the control group, where the reverse may have been expected. It is possible that some of the benefit derived from the intervention was due to dietary choices rather than change in physical activity, although diet was not explicitly mentioned in the intervention.

There was no significant difference found in the change in likelihood of having an adverse event such as a fall, injury or hospitalisation, from the beginning to the end of the study, between the intervention and the control groups. However, when analysed separately, the intervention patients were less likely to have an injury, or be hospitalised, in the year following intervention delivery, compared with the year before intervention delivery. The change in likelihood of injury or hospitalisation did not reach statistical significance for
control patients. This before-after analysis is not as conservative as analysing incremental change between the groups, but may indicate another trend towards improved health outcomes as a result of a physical activity intervention. Other studies have also demonstrated trends towards fewer injuries and fewer hospitalisations with physical activity interventions amongst the elderly (Robertson, Devlin et al. 2001; Robertson, Gardner et al. 2001). These studies were also able to demonstrate significant reductions in risk of falls.

5.3.7 Recommendations for Future Research and Potential Implementation

To evaluate cardiovascular benefits, larger samples capable of detecting smaller blood pressure changes, and longer follow-up periods are suggested. Alternatively, more intensive on-going support may improve compliance and physical fitness, which is linked to health benefit, as has been shown elsewhere (The Writing Group for the Activity Counseling Trial Research Group 2001).

Although this study was not powered to perform gender sub-analyses, men tended to increase their activity more than women (Chapter 4, section 4.3.11.5). More research is needed to clarify the reasons behind gender differences in the success or failure of particular interventions. In addition, there was some indication that the responses may have differed between rural and urban areas. Barriers to increasing activity, such as geographical isolation, need to be investigated in more depth.

Other potential barriers to implementation are the characteristics and skills of the practitioner giving the advice, and the circumstances and motivation of the patients receiving the advice. For example, general practitioners are three times more likely to give advice about physical activity if they are physically active themselves, although time also played a part in this decision (McKenna, Naylor et al. 1998). Nurses were found to be four times as likely to prescribe physical activity if they were physically active, themselves (ibid).

A potential barrier to patient participation in physical activity schemes, has been 'cost', particularly for schemes that refer to leisure centres, as in the United Kingdom. However, one study found that even when the cost was very low, participation rates did not improve, even amongst those who had previously stated that 'cost' was a significant barrier (Tai, Gould et al. 1999). A more significant barrier was 'not knowing about local exercise facilities' (ibid). In schemes that use referral to exercise centres, some people may be intimidated by the centre environment or competitiveness, or they may be bored with the
exercise regimes (Illifie, Tai et al. 1994). Further barriers include the lack of knowledge about initiating exercise, the lack of self-discipline or the lack of transport to exercise venues (ibid). Insufficient time, lack of motivation and responsibilities of child care, have been cited as common barriers to increasing physical activity amongst younger adults, while poor health is more of a barrier to older adults (Booth, Bauman et al. 1997). Other research has found that older adults participate in physical activity if there is a social component to the activity, and also if there is a friend or partner who is exercising with them (Chalip, Thomas et al. 1996). The motivational state of the patient, the ability of the advisor to take this factor into account, and the ability to tailor the advice to the person’s lifestyle and circumstances, may also be important (Chalip, Thomas et al. 1996). Therefore, more research is needed into strategies to overcome these barriers, and into the implementation and maintenance of interventions such as the Green Prescription programme in primary health care.

The Green Prescription intervention has been shown to be sustainable and has been used by 65% of general practitioners in New Zealand, according to a survey conducted amongst 427 general practitioners in 2002 (Figures supplied by Sport and Recreation, New Zealand). However, rates of use per practitioner continue to be low. Of those prescribing Green Prescriptions, an average of 4.1 Green Scripts were written per month. The use of patient prompting and systematic screening in this study was able to increase the rate of recruitment of ‘less-active’ patients to this intervention. At least eight Green Prescription referrals, per week per general practitioner, were achieved. However, the practicality of routine screening on an on-going basis needs to be investigated. In addition, the rate of eight Green Prescriptions per week would not be sustainable for general practitioners who have many other demands on their time. A compromise could be attained, particularly if screening systems are put in place and practitioners’ time initiating the intervention is kept to a minimum. In addition, assessment of the time required and compensation for such preventative action may help sustain effective lifestyle interventions, such as the Green Prescription in primary health care.
5.4 Discussion of the Cost-Effectiveness Study

5.4.1 Main Findings

This study represents one of the most comprehensive cost-effectiveness analyses of a physical-activity promotion programme in primary health care, to date. Incremental costs and cost-effectiveness ratios for physical activity outcomes, from a funder’s perspective and a societal perspective, have been presented.

The cost of delivering the Green Prescription was $170.45 per participant if only the costs of the programme were considered (funder’s perspective). If offset costs of changes in health care utilisation, productivity and expenses associated with exercising, were taken into account the cost of the Green Prescription intervention was $37.16 (95% CI: -$945.21, $1019.53) per participant (societal perspective).

A sensitivity analysis was carried out that allocated programme ‘set-up’ costs to those Green Prescription recipients who were referred to exercise specialist telephone follow-up, only, excluding Green Prescription recipients who were not referred. This analysis estimated that the programme costs were $304.97 and $171.68 (95% CI: -$810.69, $1154.05) per participant, from a funder’s perspective and a societal perspective, respectively.

The monthly cost-effectiveness ratio for increasing leisure-time moderate- or vigorous-intensity physical activity by one hour per week was $25.36 from a programme funder’s perspective and $5.47 (95% CI: -$138.90, $149.84) from a societal perspective. Although 95% confidence intervals for estimated offset costs were large, the trend indicates that there is a potential for substantial offset savings in changes to health care utilisation, particularly secondary care, and to a lesser extent, primary care. Such large offset savings, particularly from reduced hospitalisation costs, have also been found in studies of other physical activity interventions amongst older people in primary care (Robertson, Devlin et al. 2001). The cost-effectiveness ratios are also likely to improve over time if participants maintained their increased level of physical activity, as the major expense of the intervention was in the delivery.

The large 95% confidence intervals around the offset costs indicate that there is a great deal of uncertainty surrounding these estimates. It is quite likely that there is no real difference in these offset costs between the intervention and control groups and that the intervention had
no effect on these cost variables in the short-term. If that were the case, then an inclusion of only the programme costs would be appropriate. Excluding offset costs that do not show a statistically significant difference between the intervention and control groups, in the final analysis, has been carried out elsewhere (Robertson, Devlin et al. 2001; Robertson, Gardner et al. 2001). However, to do so, may risk a type-2 error, which would occur if the sample size was inadequate to show a difference as statistically significant. If this were the case, then larger sample sizes or longer follow-up periods would be required.

Alternatively, the large 95% confidence intervals may reflect the fact that the regression model involving offset costs was not stable. This was because most individuals had very little change in overall offset costs before and after the intervention, while a few individuals either had very large increases or very large decreases in offset costs, which unduly influenced results (Chapter 4, section 4.4.3). Use of the median or categorical logistic regression analysis may have been more appropriate for such a distribution. However, to attain a point estimate of the difference of change in cost per patient and to allow for clustering, a generalised least squares random effects regression model was used with continuous data.

A sensitivity analysis of the costs of the programme from a societal perspective (including offset costs), which excluded nine outliers, changed the point estimate of the incremental cost of the intervention from $37.16 per patient to $272.25 per patient (Chapter 4, section 4.4.3.5). The large change in point estimate may again, reflect the instability of the regression model due to the distribution of offset costs. Of course, it may be inappropriate to exclude these outliers. The outlying costs that were excluded were actual cost changes, mainly from hospitalisation and productivity costs, of individuals in the study. Therefore, their exclusion is difficult to justify, and the results from such a sensitivity analysis may be of limited value. The results of the analysis do suggest, however, that most of the offset savings in the intervention group, compared with the control group, came from a few individuals. These individuals incurred costs and savings many times greater than the majority of the study participants.

Estimates of costs from a programme funder's perspective (without offset costs) were more precise. This precision was largely because the total cost of programme delivery divided among the total number of intervention patients did not involve the same degree of uncertainty as calculating incremental changes over time in offset costs for each participant.
5.4.2 Strengths and Limitations of the Study

This cost-effectiveness study was conducted prospectively, alongside a randomised controlled trial, which is a strength of the study. The costing data collected are very comprehensive, follow-up rates were high, and in almost all cases, actual costs, rather than estimated costs, were used. Accordingly, few assumptions had to be made. This is in contrast to many of the previous cost-effectiveness studies conducted of lifestyle interventions, which estimated costs retrospectively (Langham, Thorogood et al. 1996; Sevick, Dunn et al. 2000).

This study has fulfilled the criteria set out by Drummond and colleagues as important for cost-effectiveness studies as described in Chapter 2, section 2.5.2.4 (Drummond and Jefferson 1996; Drummond, O'Brien et al. 1997). For example, a clear research question was stated and the perspective of the analysis was explicit. Only primary outcomes that had been found to be statistically significant in the effectiveness trial were included in the calculation of cost effectiveness ratios. This is in contrast to previous studies that have erroneously used non-significant changes to calculate cost-effectiveness ratios (Sevick, Dunn et al. 2000). The choice of the two comparative programmes was appropriate and explained. Furthermore, the sources of economic information were stated, and quantities, as well as unit costs, were reported. Appropriate discount rates and adjustment for inflation were carried out and currencies were converted to equivalent values. Productivity changes were included in the analysis, incremental costs were used, and a sensitivity analysis was undertaken. The conclusions made about the cost-effectiveness of the Green Prescription intervention were justified and conservative.

This study has a number of limitations, however. Primary care utilisation costing data were not available from five of the 42 general practices. Secondary care costs were not available for the 17/299 patients who used public hospital services and for the 38 patients who used private hospital services. These omissions may have biased the results. However, a conservative intention-to-treat analysis was conducted, where no change in rate of health care utilisation was assumed where data were missing. In addition, data were missing from intervention and control groups in similar proportions (Chapter 4, section 4.4.3.2). Therefore, the missing data are unlikely to have led to exaggerated differences between intervention and control results.
In the absence of a health price index, the consumer price index (CPI) was used to adjust for inflation for costs from previous years. However, the general inflation rate may differ from the inflation rate experienced in the health sector. The set-up costs of the programme from 1996 until 2001 were discounted using 5%, which is a commonly accepted discount rate (Drummond, O'Brien et al. 1997).

The cost of loss of productivity was calculated based on the average daily income for New Zealand. However, as incomes tend to increase with age, the incomes of study participants who were employed, may have been higher than the average New Zealand income. Even so, a daily income rate adjusted for age is unlikely to have altered the final result. This was because there was virtually no difference between the intervention and control groups in change in loss of productivity over the two years ($1.21 (95% CI: -$522.06 to $524.49), p = 0.99).

The costs of medications were not included in the analysis. However, there is evidence from the trial that alterations in medications, such as antihypertensive medication, did not differ significantly between the groups. If anything, there was a trend towards more medication use in the control group. Therefore, this cost-effectiveness analysis is likely to be conservative, if any bias exists from not including costs of medications.

Physical activity rather than health outcomes are used in this cost-effectiveness analysis. It is also difficult to assign quality-adjusted life-years or a monetary value to increases in physical activity for a cost-utility or a cost-benefit analysis. However, the cost-effectiveness analysis carried out in this study, does allow a comparison with other physical activity interventions that have calculated cost effectiveness ratios for physical activity outcomes.

A theoretical cost-utility analysis could be carried out, as was the case in a study carried out in 1988 (Hatziandreu, Koplan et al. 1988). This 1988 study was a theoretical analysis of 1000 ‘active’ versus 1000 ‘inactive’ 35 year-old men followed over 30 years. The authors concluded that there would be 78.1 fewer coronary heart disease events and 1,138 Quality Adjusted Life Years (QALYs) gained in the ‘active’ group compared with the ‘inactive’ group (Hatziandreu, Koplan et al. 1988). They also predicted the cost per QALY to be $11,313. However, this analysis defined ‘active’ as achieving 2,000 kcal/week of relatively vigorous leisure activity, whereas the definition of ‘active’ used in the current study was approximately 1,000 kcal/week of mostly moderate leisure activity. The relative risk
assumed by the Hatzandreu study for 'inactive' versus 'active' men was 2.0 for coronary heart disease events. As discussed in Chapter 2, the relative risk is likely to be lower than this.

Alternatively, the change in quality of life scores attained in this study could have been converted into the preference-based measure of health from the SF-36 quality of life scores. This could be calculated, using the SF-6D scoring system, which takes some of the components from seven of the eight dimensions of the SF-36 results, to produce a preference-based measure (Brazier, Roberts et al. 2002). This measure would allow a cost-utility analysis and calculation of cost per quality adjusted life year (QALYs) gained, although it would ignore the potential cardiovascular and other health benefits that may accrue from the intervention.

In addition, there are some limitations to the use of the SF-6D measure. This measure does not include the 'general health' score in the calculation, which was the dimension with the most significant gain in the Green Prescription study of this thesis. Furthermore, the measure was derived from preference scores provided by 611 members of the United Kingdom public, using the standard gamble technique. Therefore, the measure may not be valid for a New Zealand population. Lastly, there are some inconsistent estimates and over prediction of the value of the poorest health states, according to the authors (Brazier, Roberts et al. 2002).

5.4.3 How the Findings relate to Previous Research

The Green Prescription was more cost-effective than the primary care-based 'Prescription for Exercise' scheme in the United Kingdom that used a visiting exercise specialist to deliver the advice, in moving 'less-active' adults into the 'active' category. The cost of moving one 'less-active' adult to an 'active' state (achieving at least 2 ½ hours of moderate- or vigorous-leisure-time physical activity per week), over a one-year period using the Green Prescription was $1,756 in programme costs. The programme cost of moving one person to the recommended level of activity in the UK programme was NZ$8,663 (UK£2,500) (Stevens, Hillsdon et al. 1998). The large cost difference suggests that initiation of the exercise prescription by the usual practitioner, with screening and patient prompting, is more cost effective than initiation of the exercise prescription by a visiting exercise specialist.

The exchange rate used was UK£ 0.2886 = NZ 1.00 from December 2001 (Ibid).
This conclusion should be qualified, because the programmes were conducted in different countries where the costs of programme components may have differed, either diminishing or accentuating the real difference in cost-effectiveness.

The Green Prescription intervention was also more cost-effective than the Project Active interventions. In Project Active, the monthly programme-cost of increasing moderate- or vigorous-intensity physical activity was NZ$358.43/kcal/kg/day with the ‘Structured’ (facility-based) programme, and NZ$43.31/kcal/kg/day with the ‘Lifestyle’ (home-based) programme (Sevick, Dunn et al. 2000). The monthly programme-cost of increasing moderate- or vigorous-intensity leisure-time physical activity was NZ$37.37/kcal/kg/day with the Green Prescription programme.

There were some differences in the methods used in the Project Active and Green Prescription studies. The follow-up times differed (24 months compared with 12 months) (Sevick, Dunn et al. 2000). The monthly cost effectiveness ratios of the Project Active programmes at six-months were not compared with those of the Green Prescription, because the six-month Project Active costs were even greater. In addition, the Project Active programmes were conducted amongst 35 to 60 year-old people in the United States, while the Green Prescription study involved 40 to 79 year old people. Apart from the slight difference in age, there were also some differences in the categories of moderate and vigorous activity used in the Project Active study and the Green Prescription study. The cost of increasing leisure-time moderate- and vigorous-intensity energy expenditure using the Green Prescription was compared with the cost of increasing all moderate-intensity energy expenditure using Project Active interventions. However, this comparison is likely to underestimate rather then overestimate the cost-effectiveness of the Green Prescription intervention compared with the Project Active programmes. Again, different countries may have different costs for the same programme components. As such, comparisons must be made with caution.

Even so, this comparison suggests that brief physical activity counselling by usual clinicians and an exercise prescription, with telephone follow-up from exercise specialists, appears more cost-effective than face-to-face group counselling about exercise or facility-based supervised exercise sessions. The main costs in the Project Active interventions were for intervention personnel, which were lower, due to the shorter face-to-face contact time with participants and cheaper telephone contact, in the Green Prescription intervention.
Continuing this telephone support for longer than three months may help to increase compliance and improve the effect of the intervention at minimal cost. However, longer follow-up telephone support does not always increase the long-term physical activity gains compared with a brief intervention (Dubbert, Cooper et al. 2002).

The risk of falling, sustaining serious injury and hospitalisations have been reduced by other physical activity interventions delivered in the community to older people (>80 years of age) (Robertson, Devlin et al. 2001). Robertson’s cost effectiveness of this community intervention delivered to older people included the cost of falls or hospitalisations averted. The Green Prescription physical activity intervention did not significantly reduce the rate of falling, injuries or hospitalisations, although this was not a primary outcome measure of the trial. The positive outcome of reduced falls, found in the Robertson study, may have been due to the nature of the intervention, which included muscle strengthening and balance exercises. The positive outcome may also have been because of the older age group of the subjects in the study. Falls are less frequent amongst middle age compared with elderly people, so the incidence of falls in the 12 months before and after the intervention may not have been high enough to show a difference, in the Green Prescription trial. Even so, there was a trend in the Green Prescription study towards fewer injuries and hospitalisations in the intervention group.

The British Family Heart Study involved practice nurse screening for cardiovascular risk and a multi-component lifestyle intervention (plus an increase in medication in the intervention group) (Anonymous 1994). Costs of the programme were estimated in a similar manner to those of the Green Prescription cost-effectiveness analysis (Wonderling, McDermott et al. 1996). In the British study, primary health care utilisation and medication offset costs were also estimated for the practice as a whole, then mean offset cost per subject estimated, rather than actual costs per individual obtained. It was only when overall difference in mean costs was calculated by pooling together the differences across the 13 participating practices, using a random effects meta-analysis, did the estimates produce 95% confidence intervals. Although none of the offset costs were statistically significant, the 95% confidence intervals were a lot smaller than in the Green Prescription study. These smaller confidence intervals may have been because costs from the British study were based on mean values from practices, rather than actual values for individual patients, as in the present study. It is quite likely that the actual offset costs of the individuals in the British
study were as variable as those found in the Green Prescription study. The use of actual individual offset costs in the Green Prescription study, and the inclusion of change in secondary health care utilisation and changes in productivity, limited the precision of the estimates in the current study. However, this level of uncertainty is probably closer to reality, than that presented in the British study.

The British Family Heart study included incremental offset costs that were not statistically significant, when calculating the cost of the programme and cost effectiveness ratios, as was done in the Green Prescription study. The cost of the British programme per patient (including offset costs) was substantially higher than that of the Green Prescription programme. These costs were £51.63 (95% CI: £12.37, £90.90) or NZ$179.16 (95% CI: $42.92, $315.42) for the British programme compared with $37.16 (95% CI: -$945.21, $1019.53) for the Green Prescription programme (Wonderling, McDermott et al. 1996). However, the effect estimate of coronary risk reduction was statistically significant in the British study. Therefore, the cost effectiveness ratios were calculated per 1% reduction in coronary risk, and could not be compared with the cost-effectiveness ratios from the Green Prescription study, which were calculated per increase in physical activity. As with the present study, the authors of the British cost-effectiveness study state that a much larger sample would be needed to estimate the cost-effectiveness, reliably.

5.4.4 Implications and Recommendations for Future Research

Chapter 4, section 4.3.11 describes how ten percent more intervention patients than control patients, increased and maintained their leisure-time moderate- or vigorous-intensity activity to the recommended levels after 12 months. The fact that the intervention has the potential to reduce a sedentary population by ten percent has potential economic implications.

A study in 1987 of the cost of inactivity in New Zealand, predicted that if the number of ‘active’ individuals in New Zealand could be increased by 10%, then $24.75 million could be saved in direct and indirect costs (health care and offset costs) associated with ischaemic heart disease and hypertension (Russell, Worsley et al. 1987). This figure has since been updated to $55 million (Hillary Commission 1998). Potential savings to the New Zealand economy from reduced morbidity and mortality associated with cardiovascular disease, diabetes, colon cancer, obesity and osteoporosis, were estimated to be $48 million, if the 31% of the population that were sedentary, was reduced to 21% (Jensen, Sullivan et al.
Another report predicted that $24.28 million could be saved in the direct and indirect (including productivity) costs of cardiovascular disease, colon cancer and fractured neck of femur, if there was a 5% increase in levels of physical activity (Russell, Berkeley et al. 1993).

The cost of cardiovascular disease due to physical inactivity is now likely to be substantially higher than $55 million. The high levels of physical inactivity in New Zealand continue to account for a significant proportion of cardiovascular morbidity and mortality, as discussed in Chapter 2, section 2.2. The most recent survey of physical activity levels in New Zealand, predicts that 878,000 adults over 18 years of age in New Zealand are not achieving 2 ½ hours of leisure-time activity per week (Sport and Recreation New Zealand 2003). The Green Prescription intervention has been estimated by the study in this thesis to cost $37.16 per person in programme and offset costs. Therefore, if all less-active adults were to receive a Green Prescription, the total cost would be $32.6 million to save at least $55 million per year in costs associated with cardiovascular disease. It is likely that the annual cost of the intervention would be a lot lower than $32.6 million because the individuals would not need the Green Prescription intervention delivered to them every year. In addition, set-up costs would decrease proportionately each year. It has been demonstrated in other studies, that the costs of physical activity interventions decrease in subsequent years, so they become more cost-effective over time (Sevick, Dunn et al. 2000).

Much of the potential long-term savings from increased physical activity would be due to increased productivity and reduced health-care utilisation from avoided cardiovascular morbidity and mortality (Russell, Worsley et al. 1987; Russell, Berkeley et al. 1993). Therefore, perhaps changes in productivity and health care utilisation should not be included as offset costs in calculating the cost of the Green Prescription intervention. It may be more appropriate to use the direct costs of delivering the intervention without offset costs, when comparing cost of intervention with potential savings from improved health outcomes. In the most conservative analysis, it could be assumed that the effect of the intervention lasted only one year, and that the cost of the programme was equivalent to the cost of programme delivery, alone, without offset costs. In this scenario, the annual cost of delivering the intervention to all sedentary people would be $150 million, which would not be cost-saving, in terms of reduced cardiovascular morbidity and mortality. It would be interesting to value the ‘cost per life-year gained’ using different scenarios.
However, the offset costs included in the present study are 'short-term' savings, which are unlikely, at this stage, to be due to reduced cardiovascular events. These savings are more likely to be due to other shorter-term benefits of increased physical activity, and may be appropriately included in the analysis.

The cost and effectiveness data from the Oxcheck and British Family Heart studies were modelled to produce a range of estimates for 'cost per life-year gained' (Wonderling, Langham et al. 1996). The estimates were calculated and discounted, assuming the effect of the intervention lasted from one to 20 years. If the effect of the intervention lasted for 20 years, the cost per life-year gained was 3% to 4% of the estimated cost if the effect of the intervention lasted for only one year. It could be conjectured, therefore, that if the effect of the Green Prescription intervention lasted for 20 years instead of one year, then the annual cost of delivering the intervention to all adults would be lower than $32.6 million. In this circumstance, the annual cost of the intervention would be closer to $1.3 million if offset costs were included in the analysis, or $6 million if offset costs were not included. In these scenarios, the Green Prescription intervention would be even more 'cost-saving', if $55 million could be saved annually from reduced cardiovascular disease.

Modelling of these figures in two cost-utility analyses is underway, although such an analysis goes beyond the scope of this thesis. Even so, there are indications from this study that the Green Prescription intervention may be 'cost-saving'. The gains from the intervention seem small, but the cost of the intervention is low and the potential health gains from increases in physical activity across a population, are great.

Furthermore, the discussion above does not take into account the potential savings from the demonstrated improvements in health-related quality of life with the Green Prescription intervention. Nor does the discussion consider the potential savings from health benefits, other than cardiovascular benefits, that may have occurred from the increased physical activity (Russell, Berkeley et al. 1993). Therefore, the savings may be even greater if these quality of life benefits, and other health benefits, were considered.

There is ample evidence that physical activity improves health outcomes. However, physical activity interventions for primary prevention of disease usually have poor adherence rates of around 8-12% at 6-12 months (Sevick, Dunn 2000; Petrella 2002; Elley, Kerse 2003c; Bull 1998) These rates are similar to other lifestyle interventions such as smoking cessation
dietary or alcohol interventions (Ashenden, Silagy 1997). A systematic review by Ebrahim and Davey-Smith in the BMJ in 1997 and a Cochrane review by the same authors in 1999, found that although there were reductions in cardiovascular parameters, such as blood pressure, there were no reductions in mortality (although follow-up may have been too short). Ebrahim et al. also noted that most trials had multiple interventions including dietary, exercise and pharmaceutical and that much of the benefit was from pharmaceutical intervention (e.g. Finnish Businessman's study, Oxcheck, MRFTT, Family Heart Study) and many had design faults and may have overestimated benefit. The authors concluded that primary care lifestyle interventions should not be encouraged, but rather the resources should be put into legislative changes to salt/fat content of foods and areas to smoke, taxes on smoking, and facilities to allow exercise.

However, although the gains from physical activity interventions are small and adherence is low, the costs are also small. The cost-effectiveness of the Green Prescription intervention has been demonstrated to be good, with recent cost-utility analyses demonstrating $10,959/QALY taking into account the immediate small quality of life gains and predicted reduced mortality from cardiovascular disease, diabetes and cancer (bowel and breast) and assuming that the 10% incremental adherence only lasts the year and declines each year until all become sedentary by four years (Dalziel unpublished data). Another cost utility modelling of the cost effectiveness data in this thesis estimates the cost per DALY to be $2,727 for males, and $4,229 for females (O'Dea, unpublished data). These compare favourably with the approximate threshold of $20,000/QALY that New Zealand is sometimes quoted as using to fund pharmaceutical interventions. It also compares reasonably well with the estimated cost-utility of screening and treating with statins of $5,043/QALY for people with over 10% 5-year cardiovascular risk (Milne 2003). Even if 75% of those that had increased their activity with the Green prescription intervention reverted to a sedentary status by the end of a further 12 months, O'Dea found that this would still produce a ratio of $20,000-$25,000/DALY for women (lower ratios for men). While Pharmac spend $61 million subsidising cardiovascular medications annually, mostly for primary prevention in primary care, only $825,000 is spent by Pharmac and SPARC on the Green Prescription physical activity promotion programme nation-wide (SPARC and Pharmac, personal communication, 2003). Although multiple criteria are used to inform funding of interventions, it may make economic sense to reconsider this imbalance (Metcalf, 2003).
Longer-term follow-up within a trial would clarify whether increases in physical activity are maintained permanently. At least two-year follow-up may also demonstrate more definitive health outcome change. Previous reviews of intervention trials have found that it takes 2-4 years for cardiovascular outcomes (e.g. hospitalisation/mortality) to be affected by blood pressure and lipid-lowering pharmaceutical interventions (Stephenson 1999; Ebrahim, Davey-Smith 1999).

In conclusion, the Green Prescription is an effective and inexpensive way of increasing physical activity and improving health-related quality of life, amongst 40-79 year-old adults in primary health care. In addition, the Green Prescription is likely to improve health outcomes and may be ‘cost-saving’. This study represented a cost-effectiveness analysis, using cost per physical activity unit gained as its primary outcome, to allow comparison with previous cost-effectiveness studies of community-based physical activity interventions. However, modelling of the potential savings from health outcomes related to the increased proportion of active adults in New Zealand, and a cost-utility analysis taking into account the quality of life gains, are needed. Research is being undertaken at present. Such research would allow a comparison of cost-effectiveness of the Green Prescription intervention with other lifestyle and pharmacological interventions (Wonderling, Langham et al. 1996).

Even so, this study represents the first step in assessing the efficiency of an intervention designed to address the risk factor of physical-inactivity, which has been identified as a significant cause of morbidity and mortality in New Zealand (Russell, Worsley et al. 1987; Russell, Berkeley et al. 1993; Ministry of Health 1999).
6 Conclusion

Physical inactivity has been associated with an increase in risk of cardiovascular disease, diabetes, osteoporosis, several cancers, and depression, amongst other conditions. Increasing physical activity can improve health outcomes and has been associated with better health-related quality of life. Achieving 2 ½ hours of moderate- or vigorous-intensity leisure-time physical activity per week has been found to be adequate to produce significant health gains (Centers for Disease Control and Prevention 1996). Between 32% and 42% of all adults in New Zealand are not achieving this amount (Ministry of Health 1998; Sport and Recreation New Zealand 2003). Furthermore, 24% of cardiovascular mortality in New Zealand has been attributed to physical inactivity. Therefore, physical inactivity has been identified as a significant risk factor by the Ministry of Health (Ministry of Health 1999).

Consequently, the Hillary Commission (now called Sport and Recreation New Zealand) in conjunction with New Zealand primary health care, developed the Green Prescription intervention to assist clinicians encourage less-active adults to become more active for the benefit of their health. This intervention involved brief physical activity counselling using motivational interviewing techniques, and written advice, given by the usual general practitioner or practice nurse. Follow-up telephone support was offered from exercise specialists for a further three months.

The Green Prescription intervention has been implemented into primary health care throughout New Zealand. In addition, most health professionals have been offered training in motivational interviewing and these techniques are now taught in all New Zealand medical schools. Furthermore, the new ‘Primary Healthcare Organisation’ (PHO) structures in New Zealand, have set aside funding for health promotion initiatives and are in a good position to attain systematic coverage, lifestyle risk factor screening, and patient prompting of interventions such as the Green Prescription. Therefore, there is evidence that the intervention can be incorporated into routine practice and could achieve a systematic coverage of those at risk.

Although physical activity interventions for use in primary health care have been developed in many countries, few rigorous evaluations have been carried out. This thesis systematically
reviewed the evidence for the effectiveness and cost-effectiveness of such interventions (Chapter 2, sections 2.4 and 2.5). The reviews showed that there was insufficient evidence of effectiveness or cost-effectiveness of such interventions to recommend physical activity interventions in primary care, which was also the finding of other organisations (U. S. Preventive Services Task Force 2002). However, further rigorously conducted trials have been recommended to clarify whether these interventions may be effective and cost-effective (Simons-Morton, Calfas et al. 1998; Eden, Orleans et al. 2002). In response to this recommendation, the main study of the thesis was designed to assess the effectiveness and cost-effectiveness of the Green Prescription intervention, using a cluster randomised controlled trial design.

Conducting randomised controlled trials of physical activity interventions in a primary health care setting, is a challenge. It is difficult to blind participants or assessors in such trials, and there is a risk of contamination of intervention. These factors may well bias the results of an evaluation. Techniques to reduce this bias were used in the present study, such as the use of objective measures, self-administered questionnaires, and a cluster-randomised design. Sample size issues were properly investigated, and taken account of, and appropriate intention-to-treat analyses were employed. In addition, the implications of using a cluster-randomised, rather than an individually-randomised trial design, were considered. These techniques were used to ensure the internal and external validity of the results and the most efficient and accurate assessment of effectiveness.

The validity of findings from this research could be further threatened by the imprecision of measurement of physical activity. While physical activity questionnaires are the most practical instruments for use in large epidemiological studies, results are open to recall bias. Therefore, it is important to validate the results, or the questionnaires, against other measures such as activity diaries or motion sensors. It is also important to ensure that the instruments used are valid for the relevant population in the appropriate setting.

A systematic review of reliability and validity studies of physical activity questionnaires did not reveal an existing physical activity questionnaire that had been validated in a primary care setting amongst less-active adults (Chapter 2, section 2.3). Therefore, two physical activity questionnaires were adapted by the present author, for use amongst less-active adults in primary care. The reliability and validity of the questionnaires were assessed within a general practice setting against 7-day diaries and pedometer readings. The reliability and
validity of the Auckland Heart Study (AHS) physical-activity questionnaire were found to be adequate, when compared with levels accepted in the literature. Therefore, the AHS questionnaire was used to assess levels of physical activity in the evaluation of the Green Prescription intervention.

The cluster randomised controlled trial of the Green Prescription intervention was conducted amongst 878 ‘less-active’ 40-79 year-old patients from 42 rural and urban general practices in the Waikato. Patients in the intervention practices received the Green Prescription within usual consultations, while patients in control practices received ‘usual care’. This study found that the intervention was successful in increasing physical activity amongst less active adults in primary health care over a 12-month period. The intervention also produced improved health-related quality of life measures, including SF-36 measures of ‘general health’, ‘role physical’, ‘bodily pain’ and ‘vitality’. There were trends towards improved blood pressure, by 1.3mm Hg systolic and 1.4mm Hg diastolic, in the intervention group compared with the control group. The difference in 4-year coronary heart disease risk was not significant, although a non-significant relative risk reduction of 1.7% was observed. There was no increase in reported falls, injuries or hospitalisations from the intervention. In fact, there was a trend towards reduced hospitalisations in the intervention group.

These findings add to the literature, which has been unable to demonstrate a convincing positive effect of clinician-based physical activity interventions in primary health care. This study also provides evidence that should prompt reconsideration of current recommendations asserting that there is inadequate evidence to support counselling for physical activity in primary care (U. S. Preventive Services Task Force 2002). To the author’s knowledge, the present study is the first primary-care based trial of a physical activity intervention, to demonstrate long-term health-related quality of life benefit in addition to physical activity gains. It is also one of the few to examine adverse events (Robertson, Gardner et al. 2001; The Writing Group for the Activity Counseling Trial Research Group 2001). This study also included a cost-effectiveness analysis, which has been recommended by several previous systematic reviews (Simons-Morton, Calfas et al. 1998; Eden, Orleans et al. 2002; Petrella and Lattanzio 2002).

The average increase in leisure physical activity was 250 kcal/week, while the average increase in total energy expenditure amongst the intervention patients compared with the control patients, was 975 kcal/week. Achieving 1000 kcal/week in moderate- or vigorous-
intensity physical activity per week compared with a sedentary lifestyle, has been associated with a 20-30% reduction in all-cause mortality (Lee and Skerrett 2001). Consequently, it is reasonable to infer that this intervention has the potential to produce significant health gains.

For every ten Green Prescriptions written, one person can be expected to attain and maintain 2½ hours of leisure-time moderate- or vigorous-intensity physical activity per week as a result of the intervention. In other words, 10% more intervention patients than control patients achieved and maintained this level of activity after one year. An increase of 10% in the physically active population would save at least $55 million per year in costs associated with cardiovascular disease. This figure was estimated in the 1990’s, derived from estimates from a 1987 report (Russell, Worsley et al. 1987; Hillary Commission 1998). These savings are greater than the estimated total cost of delivering the Green Prescription to all 878,000 less-active adults in New Zealand ($32.6 million). Therefore, the Green Prescription intervention could be cost-saving, even without considering the other health and quality of life benefits from the intervention. If the effect of the intervention lasted more than one year then the cost savings would be even greater. Modelling of the potential cost-benefit and cost-utility of the Green Prescription is needed.

The programme under investigation was acceptable to clinicians and patients. Two thirds of less active adults routinely visiting their general practitioner were willing to receive a lifestyle intervention from their general practitioner or practice nurse. Other studies have shown that the most preferred source of physical activity advice, for older adults particularly, is the general practitioner (Booth, Bauman et al. 1997). The Activity Counselling Trial examined the acceptance of brief physical activity counselling in routine practice and the rates that the 54 physicians in the trial gave the advice to study patients (The Writing Group for the Activity Counseling Trial Research Group, 2001). Ninety-one percent gave the advice, over 60% said it did not add to the length of the consultation and over 80% said it enhanced the consultation. Most clinicians delivered the advice in 3-4 minutes. Likewise, there is wide acceptance of the Green Prescription intervention amongst general practitioners in New Zealand. A postal survey of 423 general practitioners throughout New Zealand in 2002 indicated that 65% of New Zealand general practitioners had given Green Prescriptions to their patients. The average number of Green Prescriptions given per month was 4.1 (Statistics provided by Sport and Recreation, New Zealand). However, the number of adults eligible and willing to receive such an intervention was found to be 8-10 people per
week per general practitioner in the Green Prescription trial presented in this thesis. Therefore, while the intervention is feasible and acceptable, widespread dissemination of the intervention would have to be carefully planned and well resourced to ensure that all those eligible are offered the intervention.

The Green Prescription intervention has the potential to be implemented in a systematic way. This policy could be achieved by the new Population Health Organisation (PHO) structures, which have funding formulae that could prioritise the management of population health and allow for the delivery of effective prevention programmes. In addition, the present study has shown that such an intervention can be delivered in a systematic way, within a general practice setting. This outcome was achieved using screening of patients at the reception area and patient prompting of doctors and practice nurses to deliver the intervention, rather than relying on clinician-initiation of the intervention. It is likely that the intervention will need to be delivered in the same way as in the present trial to achieve the potential population health gains. It is not reasonable to expect the current Green Prescription use (in the drawer, waiting to be thought of by the general practitioner) to have similar success to this trial. Systematic screening and patient prompting is likely to increase the 'reach' and effectiveness of the intervention.

Furthermore, if the adherence to the advice could be increased, the cost-effectiveness of the Green Prescription intervention would be further improved. However, changing behaviour is difficult, with both external and internal patient barriers inhibiting change (Booth, Bauman et al. 1997; Ziebland, Thorogood et al. 1998). Future research into overcoming these barriers and into the matrix of the patient-doctor interaction is needed, to improve the effectiveness and cost-effectiveness of such interventions.
References


Freemantle, N. (2001). "Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic?" BMJ 322 989-91.


Appendices