MONASH University
Business and Economics

Centre for Health Economics

Research Paper 2005 (2)

RISK FACTOR STUDY

How to Reduce the Burden of Harm from Poor Nutrition, Tobacco Smoking, Physical Inactivity and Alcohol Misuse: Cost-Utility Analysis of 9 Multi-Risk Factor Interventions

Kim Dalziel
Research Fellow, Centre for Health Economics,
Monash University

Associate Professor Leonie Segal
Deputy Director, Centre for Health Economics,
Monash University

Duncan Mortimer
Senior Research Fellow, Centre for Health Economics,
Monash University

April, 2005

ISSN 1833 – 1173
ISBN 1 921187 01 8

ACKNOWLEDGEMENTS

Contributions to the research were also made by Susan Day (CHE), especially in analysing the multiple risk factor interventions, Anita Lal from the Vic Health Centre for Tobacco Control, who worked on the smoking interventions, Andrew Dalton who developed the model for the GutBusters, Swinburn & Halbert studies, Alison Seccull and Jonothan Passmore, public health trainees who worked on the alcohol interventions, and Elizabeth Spence, a health economics trainee.

Funding for the research was provided primarily through the Australian Government Department of Health and Ageing, Population Health Division, with support also through Monash University and the Department of Human Services, Public Health Traineeship Program and VicHealth Centre for Tobacco Control. We also acknowledge the input and advice received from members of the Expert Advisory Group.

The content of the report remans the responsibility of the study team.

Contact Details
For information about the report please contact:
A/Prof Leonie Segal, Deputy Director,
Centre for Health Economics, Faculty of Business & Economics, Monash University
Leonie.segal@buseco.monash.edu.au. Phone: 03 9905 0734

TABLE OF CONTENTS

Associated Documents

CHE Research Paper 2005/1	Executive Report
CHE Research Paper 2005/3	Risk Factor Study: Economic Evaluation of Four **Physical Inactivity** Interventions
CHE Research Paper 2005/4	Risk Factor Study: Economic Evaluation of Eight **Nutrition** Interventions
CHE Research Paper 2005/5	Risk Factor Study: Economic Evaluation of Five Interventions to Discourage **Tobacco Smoking**
CHE Research Paper 2005/6	Risk Factor Study: Economic Evaluation of Six Interventions to Promote Safe Use of **Alcohol**

INTRODUCTION TO RISK FACTOR PROJECT

The risk factor project was commissioned by the Department of Health and Ageing, Population Health Division to determine how best to reduce the burden of harm on the Australian community attributable to physical inactivity, poor nutrition, alcohol misuse and tobacco smoking. The research objective was to establish which interventions are most effective and cost-effective and thus able to make the greatest contribution to harm reduction for resources allocated. This is a technical analysis, focused on health, measured by mortality and quality of life as the primary objective of health policy. While there may be other objectives and other issues relevant to policy decisions, these have not been incorporated into the analysis, due largely to their more subjective nature.

The project has been completed in several stages. It commenced with a literature review of evidence concerning interventions designed to modify these four lifestyle behaviours (Segal, Dalton, Robertson et al 2003). The primary purpose of this task was to identify a set of interventions for economic analysis that met nominated selection criteria related to quality of evidence etc. In practice, in order to achieve comprehensiveness, interventions were also included that did not meet the quality of evidence criteria. The interventions selected through this process for economic analysis are listed in Table 1. We identified 35 interventions for assessment and have been able to report 29 cost-utility (C-U) analyses; 22 based on models developed by the research team, 3 based on published models, 2 'scenario analyses', whilst 2 interventions were dominated. The results of these analyses are reported in 6 volumes; an Executive Report, plus 5 technical volumes covering each of the 4 risk factors, plus one for multiple risk factor interventions.

The relationship between the intervention, behaviour and health outcomes are complex and not necessarily directly observable. We have thus adopted a 2-staged approach to measuring economic performance that distinguishes the impact on behaviour from the consequent impact on health. We have in most cases generated an 'intermediate' measure of performance, a *cost-effectiveness ratio,* in which interventions are analysed in terms of the cost to achieve an observed change in lifestyle, based on trial results. Interventions that target the same lifestyle behaviour can then be directly compared, without having to understand the relationship between behaviour and health. This technique is applicable where behaviour is consistently and simply described. It is less useful where the life style attribute is complex, such as nutrition or physical activity. It also cannot be used to compare interventions which target several behaviours or that address different behaviours. The ultimate approach to performance measurement is the *cost-utility analysis* which we have conducted wherever data allowed.

We have, where data allows, estimated QALYs from observed impact on health outcomes, otherwise using published relationships between lifestyle behaviours and health or clinical parameters and health. In short we draw on a combination of trial evidence and pertinent epidemiological and other data in a standard cost-utility analysis. Most use a markov model structure, with the primary input the probability of moving control and intervention cohorts between pertinent health states. Full details of each model and the assumptions adopted are described in the chapters of this Executive Report and the five Technical Reports, one for each risk factor and are summarised in Table 2. Where possible, consistent assumptions have been used for all interventions. The impact of alternative assumed values for uncertain parameters have been explored via univariate sensitivity analysis or probabilistic sensitivity analysis where data quality allows.

Table 1 Interventions selected for economic evaluation

MULTI-FACTORIAL (Chapter 1-9)

Research Paper 2

Adult Interventions	**School-based Interventions**
▪ Fighting Fit, Fighting Fat Media Campaign	▪ Student TV viewing and obesity
▪ Stanford 5 City media/community Project	▪ Interdisciplinary student intervention and obesity
▪ GutBusters Workplace Program	▪ Cardiovascular disease risk factors in children
▪ Workplace prevention of heart disease *	▪ Cardiovascular disease risk reduction in children
▪ Oxcheck – Primary care nurse health checks	

PHYSICAL ACTIVITY (Chapter 1-4)

Research Paper 3

▪ Australian GP Active Script *	▪ General practice exercise referral for cardiovascular disease risk factors
▪ New Zealand GP Active Script	▪ Physical activity program and individualised advice for over 60 year olds
▪ Community based exercise for over 65 year olds	

NUTRITION (Chapter 1-8)

Research Paper 4

▪ Nutritional counselling in general practice *	▪ Orlistat plus diet for obesity
▪ Mediterranean diet in those with previous myocardial infarction	▪ Lifestyle changes to prevent type 2 diabetes
▪ Reduced fat diet for those with impaired glucose intolerance	▪ Talking computer for nutrition *
	▪ Nurse nutritional counselling in general practice
	▪ Multi-media '2 fruit 5 veg' campaign

SMOKING (Chapter 1-5)

Research Paper 5

▪ US mass media smoking campaign – Massachusetts Tobacco Control Program	▪ Meta-analysis of 34 trials evaluating minimal to intensive advice in general practice
▪ Australian mass media campaign – Phase 1 National Tobacco Campaign	▪ Meta-analysis of 86 trials comparing brief intervs, NRT and behavioural interventions *
▪ Meta-analysis of 16 Bupropion SR trials	▪ Phone counselling as adjuvant therapy for NRT

ALCOHOL (Chapter 1-6)

Research Paper 6

▪ US mass media alcohol campaign *	▪ MOCE and BSCT for moderately dependent drinkers
▪ Meta-analysis of 8 trials evaluating brief interventions in primary care for problem drinking	▪ MET and NDRL for mildly to moderately dependent drinkers
▪ Brief interventions for heavy drinkers	▪ Meta-analysis of 7 trials evaluating Naltrexone and psychosocial therapy

Notes

* Cost-utility analysis not completed due to insufficient evidence, interventions too complex or resource and time constraints.

NRT: Nicotene replacement therapy; MOCE: Moderation-Orientated Cue Exposure. BSCT: Behavioural Self-Control Training.

MET: Motivational Enhancement Therapy. NDRL: Non-directive Reflective Listening.

Table 2 Key assumptions underlying the economic modelling

Description	Details
Base case	
Discount rate	5% for costs and benefits.
Cycle length	1 year for all Markov models except the diabetes Gutbusters model of 5 years and the alcohol model with cycles 3 or 6 months.
Time horizon	Chosen to match the disease process, age of participants and reflecting available evidence; ranging between 5 years & life expectancy.
Evidence of treatment effect	Ideally drawn from meta-analyses or if unavailable from key RCTs.
Length of intervention benefit	Generally in the base case the length of intervention benefit is not extended beyond the duration of the trial evidence.
Direct costs of intervention	Estimated in Australian dollars 2003, based on described resource use or published costs adjusted by health price index and exchange rate.
Indirect costs	Indirect costs such as transportation, waiting times, costs to careers and productivity losses have not been included.
Comparator	Usual care, current practice, placebo or no intervention. If the comparator was inappropriate, an own-control comparison was made of intervention group, comparing final outcomes and baseline values.
Downstream costs	Excluded in base case analysis.
Model structure- Examples	
Smoking interventions	Markov model, containing ex-smoker tunnel sequence. Cohort initially distributed across smoker states according to prevalence in Australian population. Mortality differential commences from age 25 years.
Alcohol interventions	Tunnel sequences used to delay the health effects of moving from one state to another, quality of life gain directly attributable to alcohol moderation varies depending on severity of alcohol problems.
Hypothetical scenario analysis	Was performed for selected multi-factorial school based interventions given gap in key effectiveness data.
Modification of published model	Where a sound published model was available Australian costs were applied, and in some cases model assumptions were modified.
Sensitivity analysis - examples	
Discount rate	0%,3% and 7%
Downstream costs	Included for interventions targeted at specific disease such as diabetes or heart disease. Otherwise a threshold analyses was performed to show the downstream cost offset associated with intervention dominance.
External effects	Health effects for family members are considered for alcohol interventions
Other variables frequently varied	Time horizon, length of intervention benefit, utilities, costs, treatment effect, characteristics of starting population, relapse rates.

1. Community-wide education for obesity prevention – 'Fighting Fat, Fighting Fit' campaign

1.1 Description

Type of intervention

In this chapter we analyse a community-wide media based intervention for reducing obesity. This intervention type is illustrated by the 'Fighting Fat, Fighting Fit' (FFFF) which was a multi-media health education campaign targeted at rising obesity levels in the UK. The aim of this intervention was to educate and encourage people to 'eat more healthily and become more physically active' on the assumption that this would impact on obesity. The FFFF campaign emphasised the achievement of 'small and permanent changes to lifestyle rather than dieting and short-term weight loss'. (Wardle et al., 2001)

References/sources of evidence

The descriptions of the intervention and its effectiveness are based on two published articles:

- *Using the mass-media to target obesity: an analysis of the characteristics and reported behaviour change in participants in the BBC's 'Fighting Fat, Fighting Fit' campaign* (Miles et al., 2001)
- *Mass education for obesity prevention: the penetration of the BBC's 'Fighting Fat, Fighting Fit' campaign* (Wardle et al., 2001).

Intervention description

Recruitment and target population: The campaign explicitly targeted groups with a higher prevalence of obesity. The socioeconomic groups 3M and 4 (skilled, semi-skilled and unskilled manual) tend to have higher levels of obesity but do not directly correspond with the BBC's groups who typically watch BBC television programmes. The overlap between those most likely to be obese, and the BBC's typical audience was considered to be people aged 21 to 45 in the 3NM and 3M groups (skilled non manual and manual groups).

During the campaign, people were encouraged to register with the campaign. They could do this by telephoning or writing for a registration pack. This cost £2, and included a 22 page self-help guide to making lifestyle changes; three registration cards to return over a five-month period to chart progress in weight loss, activity levels and eating habits; information about the health benefits of making lifestyle changes; charts for assessing current activity levels; a selection of potential exercise and eating goals to choose from alongside advice on how to achieve them; and money-off vouchers for the FFFF book and exercise video up to a total value of £3. People who sent back the second registration card received a voucher for a free exercise session at a fitness centre. Registrants who showed the greatest improvement in eating and activity habits over the six-month period had a chance to win prizes such as a year's supply of fruit and vegetables or a home visit by a health and fitness expert. (Wardle et al., 2001) The baseline characteristics of the registrants indicated that it was not only the overweight and obese who registered with the campaign (Table 1.1).

Intervention: The intervention had a number of elements: a media campaign, the registration scheme described above, a website, Ceefax pages ("see-facts" a BBC teletext service), a 192 page FFFF book, and a video.

The media campaign lasted for seven weeks and involved television, radio and the print media. The principal television programmes were:

- An hour long, prime time special that launched the campaign (*Weight of the Nation*);
- A six part documentary series broadcast in the evening (*Fat Free*);
- A trilogy broadcast later in the evenings in a science slot which looked at the science of obesity (*Fat Files*); and
- A broadcast on weekday afternoons (*Body Spies*). (Wardle et al., 2001)

The launch on BBC Radio ran for three days and included interviews with celebrities about their eating and activity habits, followed by advice from a TV chef on healthy eating alternatives. The Health Minister was interviewed. Listeners' queries about how to be healthier were answered in a phone-in. In addition, there were a number of other local radio slots and interviews. BBC Radio 2 Action Line had experts available to advise listeners on their eating and activity habits as well as to provide details of how individuals could join the campaign. (Wardle et al., 2001) The campaign was mentioned briefly in more than 60 magazines, nine national newspapers and 120 regional newspapers. Articles regarding the campaign appeared in the national press about 28 times in total before and during the seven weeks of the campaign.

The basic cost of the 192-page book was £4.99 and the video £10.99 but it was somewhat less if people registered with the campaign and received money-off vouchers.

1.2 Quality of evidence

Evaluation description

Design: The evaluation of the effectiveness of the FFFF campaign was based on a before-and-after observational study and did not involve a control group.

Methodology: The methodology involved a mail survey of a random sample of 6,000 people who registered with the campaign. At baseline 3,661 people responded to the mail survey (61%). At follow-up 2,112 responded (35% of the random sample and 58% of baseline responders). Compared to registrants, evaluation participants were more likely to be female, older and have a BMI of 30 kg/m^2 (Table 1.1).

Table 1.1 Comparison of baseline characteristics of registrants and evaluation participants

		Evaluation participants		Statistical significance tests†
		Baseline Questionnaire	Baseline and Follow-up	
Participants	33,474	3,661	2,112	
Sex (Responses*) Female Male	 (30,099) 88% 12%	 (3,649) 87% 13%	 (2,106) 87% 13%	χ^2 = 6.306 df = 2 p = 0.043
Age (Responses*) < 25 25-34 35-64 65+	 (24,374) 8% 28% 58% 6%	 (3,657) 6% 25% 62% 7%	 (2,110) 4% 20% 67% 10%	χ^2 = 209.070 df = 6 p = 0.000
BMI (Responses*) < 25 kg/m^2 25<30 kg/m^2 30+	 (30,099) 14% 37% 49%	 9% 33% 58%	 (2,070) 9% 36% 54%	χ^2 = 168.881 df = 4 p = 0.000

Source: (Miles et al., 2001) p. 363

Notes:

* Responses ≠ number of participants because of missing data

† Chisq tests undertaken at the Centre for Health Economics and based on non-missing data

Outcome measures: The outcome measures were self-reported behaviour (exercise, diet and alcohol consumption) and clinical parameters (BMI calculated from self-reported weight and height). The mail questionnaire included items taken from a number of existing questionnaires: Dietary Instrument for Nutrition Education (Roe et al., 1994), International Physical Activity Questionnaire (Booth, 2000), SF-36 (Ware et al., 1993), and the Dutch Eating Behaviour Questionnaire (van Strien et al., 1986). Some items were modified to suit the current survey. In only two cases were there any indications of the reliability of the items in their modified form (Cronbach's α for the three items from the SF-36 was 0.66 and 0.86 for the items from the Dutch Eating Behaviour Questionnaire).

Analysis: Intention-to-treat in which baseline measures for the non-completers were carried forward to follow-up, and completers-only analyses were undertaken, point estimates (means) and measures of variability (95% CIs) were presented for the primary outcome measures and significance tests (t-tests for interval data and χ^2 for ordinal data) were undertaken for the comparison of changes between baseline and follow-up.

Assessment

Sources of bias: The major sources of bias in the evaluation were selection (non-response) and measurement. In terms of the selection bias, the response rate for the baseline questionnaire was 61%, there were significant differences between the evaluation participants and the campaign registrants and there is no comparison of campaign registrants with the target group for the intervention. In relation to the measurement bias, there was no external validation of the self-reports (eg self-reports have been shown to understate weight and overstate height, particularly in women, Wardle et al., 2001), baseline measure were taken some weeks after the campaign had begun and relied upon accurate recall of events prior to the trial, and the reliability of the many of the items in the mail questionnaire to measure behaviour was not established.

Establishing causality: The lack of a control group makes causality difficult to establish. As Miles *et al*. state: 'Whilst the reported changes in weight and exercise … are encouraging the findings must be evaluated cautiously'. Levels of lifestyle change were not evaluated within a control community so no firm inferences can be made that the … change came as a result of the FFFF campaign' (Miles et al., 2001).

Other: The assessment of outcomes occurred some 19 weeks after the end of the media campaign (approximately 6 months after the start). There is no evidence as to how long the observed changes in the evaluation participants would be maintained.

1.3 Outcomes – as reported

Analysis of the pre-post measures was done on an intention-to-treat basis for the 3,661 participants, with baseline values carried forward for the 1,549 evaluation participants who did not return the follow-up questionnaire, and on a completers-only basis for the 2,112 who returned the second questionnaire.

Behavioral parameters

Among evaluation participants there was an increase in the average servings fruit and vegetable per person per day, an increase in the proportion of people reporting that they ate more than five serves of fruit and vegetables per day and a decrease in the average alcohol units per person per day. (Table 1.3)

Table 1.2 Outcome measures, data sources and analyses for the evaluation of the FFFF campaign

Outcome	Data source	Analyses
1. BEHAVIOUR CHANGE		
Diet ▪ Number of servings of fruit and vegetables per day ▪ Percentage of people eating 5 or more portions of fruit and vegetables per day	Self-report in a mailed questionnaire at baseline and follow-up five months later	Intention to treat and completers-only comparison of baseline and follow-up self-reports
Physical activity ▪ Percentage of people classified as active	Activity levels were assessed using an adapted version of the International Physical Activity Questionnaire. This assessed the frequency and duration of three types of exercise: brisk walking, moderate exercise and vigorous exercise. People were asked to indicate whether they had done the types of exercise daily, 2-6 times a week, 2-3 time a week, once a week, never or cannot exercise. And where they had done the type of exercise approximately how long they had done each for in hours and minutes.	As above
Alcohol consumption	Participants were asked how many 'units' of alcohol they drank over the whole week. (The definition of a unit was not provided in the reference.)	As above
2. CLINICAL PARAMETERS		
Weight	Self-report in a mailed questionnaire at baseline and follow-up five months later	
Height	As above	
BMI	Calculated from self-reports of height and weight	As above

Source: (Miles et al., 2001)

Table 1.3 Baseline measures and changes in behavioral parameters for evaluation participants

	FFFF Registrants	Evaluation participants†	
		Intention-to-treat	Completers only
Participants	33,474	3,661	2,112
Fruit and vegetables servings: Baseline: mean Change: mean (95% CI)	Unknown	 2.86 per day + 0.75 (0.69 – 0.80)*	 3.1 per day + 1.3 (1.2 – 1.4)*
People eating more than 5 serves of fruit and vegetables per day Baseline percentage Follow-up percentage Change	Unknown	 20.9% 33.9% +13.0%*	 24.2% 46.9% +22.7%*
Alcohol consumption: Baseline: mean Change: mean	Unknown	 6.82 units per week -0.9*	 6.84 units per week –1.6*
Classified as 'active': Baseline Follow-up Change	Unknown	 29.9% 46.8% +16.9%*	 30.2% 62.2% +32.0%*

Source: (Miles et al., 2001) p.367

Notes:

* Difference between baseline and follow-up statistically significant at p<0.001.

† Responders to a random sample of 6,000 registrants.

Clinical parameters

The average loss in BMI for evaluation participants was 0.88 kg/m^2 in the intention-to-treat analysis and 1.6 kg/m2 for the study completers (Table 1.4).

Table 1.4 Baseline measures and changes in clinical parameters for evaluation participants and registrants

	FFFF Registrants	Evaluation participants†			
		Intention-to-treat		Completers only	
Number of Participants	33,474	3,661		2,112	
Average BMI‡					
Baseline	30.5 kg/m^2	32.2 kg/m^2		31.7 kg/m^2	
Change: mean (95%CI)	Unknown	- 0.88 (0.82 – 0.94)		- 1.6 (1.5 – 1.7)	
BMI Categories‡	Baseline	Baseline	Change*	Baseline	Change*
Normal <25 kg/m^2	14.3%	8.7%	+4.2%	9.3%	+7.6%
Overweight 25 < 30 kg/m^2	36.6%	33.5%	+1.8%	36.3%	+3.2%
Obese ≥ 30 kg/m^2	49.1%	57.8%	-6.0%	54.4%	-10.8%

Source: (Miles et al., 2001) pp. 363, 366

Notes:

* Difference between baseline and follow-up statistically significant at p<0.001.

† Responders to a random sample of 6,000 registrants.

‡ Based on non-missing data.

1.4 Program costs

As reported by trial

There were no costs reported in the literature.

Based on resource use

It is difficult to estimate costs based on resource use. However, the Transport Accident Commission spent $6m in Victoria on its 1993/94 media campaign (Cameron, 1993) and the Victorian WorkCover Authority's statewide media campaign 'Back Pain – don't take it lying down' which began in 1997 cost $5.8m in its first year, $2.8m in the second and $2m in the third[1]. The FFFF media campaign ran for seven weeks and follow-up of the evaluation participants finished six months after the start of the campaign. The cost of the campaign has been estimated to be UK£2m (AUS$5,367,574) in the 1998-1999 financial year and AUS$6,034,754 in June 2003 dollars.[2]

These costs translate into a cost per person of $323.67 if conservatively dividing the total cost by the cost per registrant. If the total cost is divided by the total number of potential registrants (total number of people in Victoria aged over 18) then the cost reduces to $0.56 per person.

1.5 Performance

Cost effectiveness

The preliminary estimates of cost-effectiveness are based on the demographic profile of Victorians in the 2001 Health Survey, and the registration rates for each BMI category achieved in the FFFF campaign (Table 1.4). Based on an attribution of one-third of the total cost to each risk factor the cost per changer AUS$ at June 2003 was estimated to be $2,504/diet changer, $1,920/physical activity changer and $5,426/BMI changer. (Table 1.5)

[1] Private communication from Dr Rachelle Buchbinder, 2002

[2] Average exchange rate for 1998-1999 financial year 2.6214 (Source: http://www.x-rates.com/cgi-bin/hlookup.cgi accessed 26/02/2004), % increase in CPI as at Jun: 2000=2.4%, 2001=6.0%, 2002=2.9%, 2003=3.1% (Source: {ABS, 2004 #251})

Table 1.5 Preliminary cost-effectiveness estimates an FFFF type campaign in Victoria (AUS$ at June 2003)

Risk Factor and target	Change in percentage of people achieving risk factor target*	Estimated number and % of registrants who changed (Registrants = 6,179†)	Cost per changer based on:	
			Total cost attributed to each risk factor‡	One-third total cost attributed to each risk factor
Diet: Eating >5 serves of fruit and vegetable per day	+13.0%	803 (13.0%)	$7,513	$2,504
Physical Activity: Classified as active	+16.9%	1,044 (16.9%)	$5,779,	$1,920
BMI: Not obese (< 30 kg/m^2)	+6.0%	371 (6.0%)	$16,277	$5,426

Notes:

* Based on the intention-to-treat analysis (see Tables 1.3 and 1.4).

† Population = 3,556,760 persons aged 18+ in Victoria in 2001; 38.8% in normal BMI range, 31.7% in overweight range, 6.3% in obese range (Sources: Tables 1 and 36 in Companion Data to ABS Cat No 4364.0 National Health Survey 2001, available http://www.abs.gov.au accessed 17th March 2004)). Registration rates were based on rates for the FFF campaign (Table 1.4) – 0.1% for normal BMI, 0.2% for overweight and 0.5% for obese. Registrants in each BMI category calculated as 3,566,760 x % in BMI range x % of registrants.

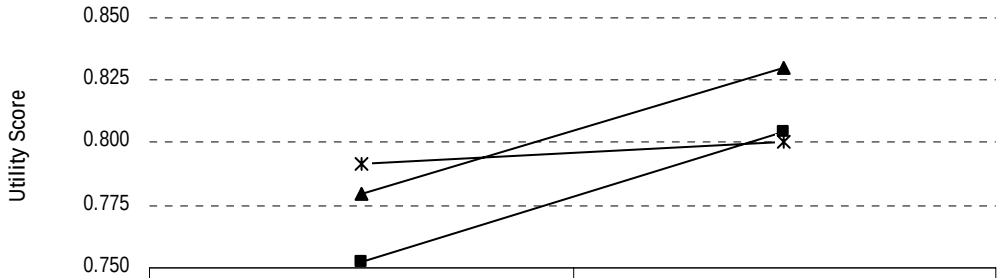
‡ Assumed to be AUS$6,034,754 in Jun-2003 dollars.

Cost-utility analysis

Utility scores for changers prior to, and after, changing were based on the application of the Brazier transformation (Brazier, 2002) to the SF-36 (Ware, 1993) scores for people who participated in the Australian Bureau of Statistics National Nutrition Survey and the National Health Survey in 1995. The changes are shown in Figure 1.1.

The preliminary cost-utility estimates are shown in Table 1.6. These estimates are conservative in that they apply to registrants only and there is no assumption that the campaign resulted in benefit to the rest of the community. The preliminary estimate of cost per QALY gain for diet changers was $575,681, for physical activity changers it was $75,542 and for BMI changers it was $208,284. (Table 1.6)

Figure 1.1 Changes in utility* for people who changed their diet, who became physically active, and people who changed from obese to non-obese



	Baseline	Six Months
Diet	0.7916	0.8003
Physical Activity	0.7792	0.8302
BMI	0.7526	0.8047

Data Source: National Nutrition Survey (1995) CURF (Ref No 691)

Notes: -

* Utility scores based on a Brazier transformation [Brazier, 2002 #108] of the SF-36 scores for people in the National Nutrition Survey (1995) CURF (Ref No 691)

Table 1.6 Costs of increasing the utility of changers by one unit at six months (AUS$ at June 2003)

Risk Factor	Utility Scores per Changer			Cost Per:	
	Baseline (a)	Six Months (b)	Increase (c)	Changer§ (d)	Unit increase in utility = Cost/QALY gain lasts 12 months ¶
Diet	0.7916	0.8003	0.0087*	$2,504	$575,681
Physical Activity	0.7792	0.8302	0.0510†	$1,920	$ 75,542
BMI	0.7526	0.8047	0.0521‡	$5,426	$208,284

Notes:

* Person who changed from eating less than five serves of fruit and vegetables at baseline to at least five serves at six months.

† Person who changed from physically inactive at baseline to physically active at six months.

‡ Person who changed from a self-reported BMI of ≥30 kg/m^2 at baseline to <30 mg/m^2 at six months.

§ Based on one-third of the estimated budget (see Table 1.5)

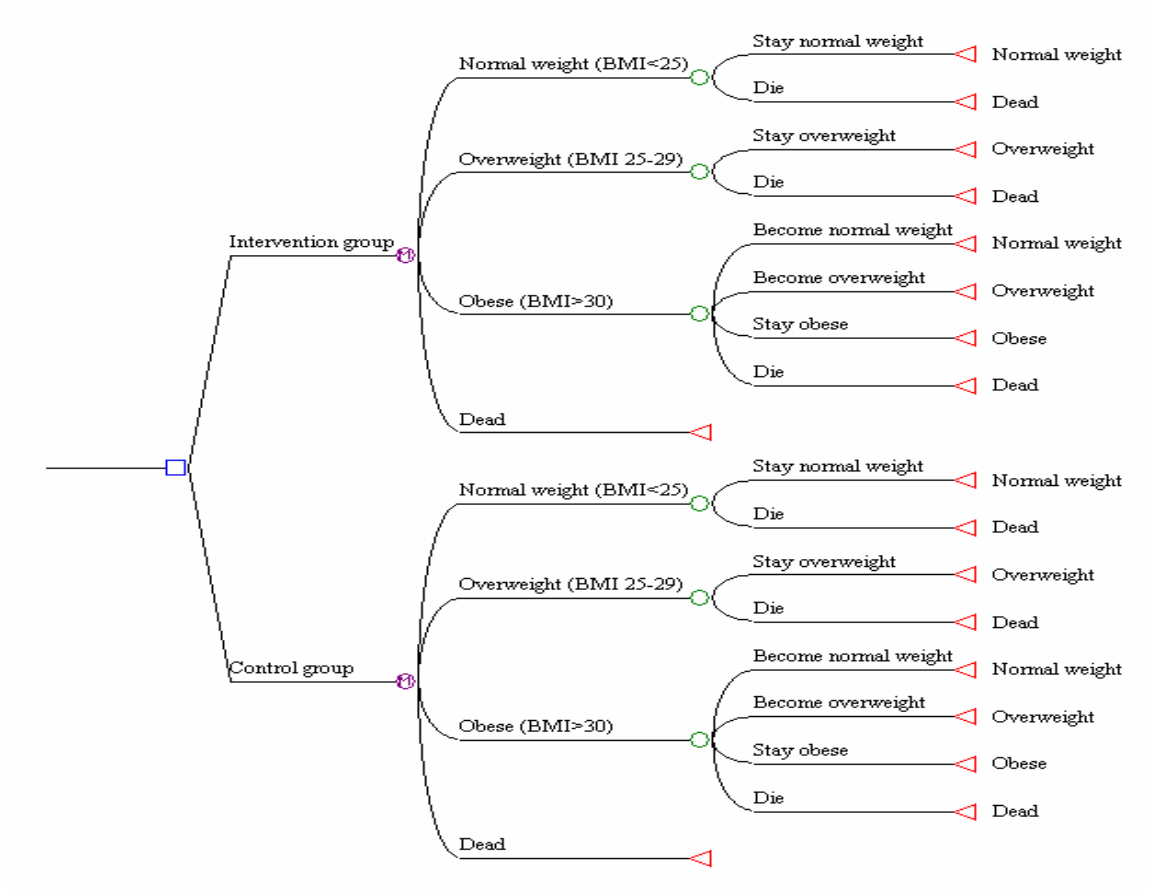
¶ Cost per unit increase in utility = (d) ÷ [(c)*0.5]

1.6 Modelling

Methods

A modelling approach was used to enable the short term outcomes (reported by Miles et al, 2001) to be extrapolated longer term and translated into life-years saved and QALYs gained. A Markov process structure was developed comprising 1 year cycles. The time horizon of the model was 20 years. The model includes the health states normal weight (BMI<25kg.m^2), overweight (BMI 25 to 29.9kg/m^2), obese (BMI>30kg/m^2) and dead. The transitions that are permitted are illustrated in Figure 1.2

Figure 1.2 Representation of states and permitted transitions in Markov model



We determined the progression, costs and utilities of a cohort of 1000 people receiving the FFFF media campaign compared to no intervention.

The cohort progressed annually between health states over a 20 year time horizon according to transition probabilities derived from the published literature. The model commences with 9% of people with normal weight, 33% overweight and 58% obese in both the intervention and control groups (Miles et al, 2001).

The following transition probabilities are incorporated into the model:

- In the first year of the model 4.2% of people of obese weight become a normal weight in the intervention group (Miles et al, 2001)
- In the first year of the model 1.8% of people with obese weight become overweight in the intervention group (Miles et al, 2001)
- Over the 20 years of the model death is time dependent and is different for each category of weight. Probabilities of death for each year are determined by fitting a Weibull curve to survival curves in the paper by Peeters et al (2003). The probabilities of death are weighted for a population that is 50.7% female (ABS 2002) with 27.3% of males assumed to smoke and 21.4% of females (ABS National Health Survey 2001)

In addition the following assumptions have been made:

- The control group have same weight as baseline measures in intervention group and do not change
- The intervention effect is assumed to last for 1 year after which a relapse rate of 50% is applied in the second year

The cost per person for the FFFF media campaign was reported in section 1.4 and was estimated as $323.67 per person for the base case analysis. The cost per potential registrant of $0.56 is included in sensitivity analysis. The downstream costs of being overweight or obese are not included in the base case analysis but are considered in sensitivity analysis (threshold analysis).

Utilities are assumed to be 0.85 for those with normal weight, 0.82 for those overweight and 0.78 for those with obese weight (McNeil & Segal, 1999). Costs and benefits are discounted at 5% per annum (Australian Treasury).

Extensive univariate sensitivity analyses were performed for the assumptions and values described in Table 1.7.

Table 1.7 Sensitivity analysis: attributes, base case and alternative assumed values

Assumptions	Base case	Alternative Values	Source
Time horizon	20 years	5 and 10 years	Researcher judgment
Discount rate	5%	0% and 3%	Researcher judgment
Length of intervention benefit	1 year	3 and 5 years	Researcher judgment
Utility of overweight	0.82	0.79	Utilities from Hakin et al, 2002
Utility of obese	0.78	0.76	Utilities from Hakin et al, 2002
Cost divided by number of potential registrants	$232.67	$0.56	See section 1.4

In addition a threshold sensitivity analysis was conducted to determine the downstream cost associated with the obese state which would lead to the intervention being dominant.

Results

Table 1.8 presents the economic performance of the FFFFs media campaign, and an incremental cost-utility ratio of $5,642 per QALY gained (for base case assumptions, see Table 1.7).

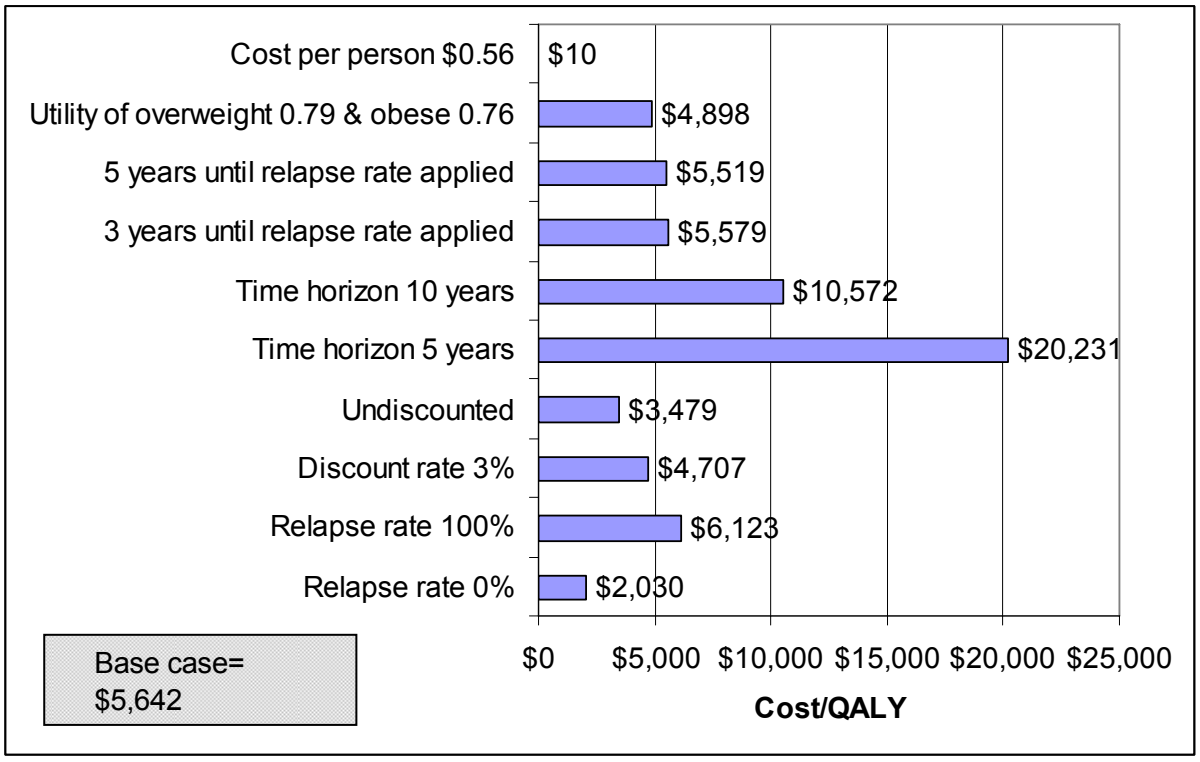
Table 1.8 Modelled cost-utility base case results

	FFFFs media campaign	Control group (no intervention)	Difference
Total costs	$308.20	$0	$308.20
Total life years	12.2134	12.2016	0.0118
Total QALYs	9.8119	9.7572	0.0546
Discounted $/LY gained			$26,071
Discounted $/QALY gained			$5,642

Sensitivity analyses

Sensitivity analyses ranged from $10 per QALY to $20,231 per QALY (Figure 1.3). Results were most sensitive to the time horizon of the model and the costs of the intervention.

Figure 1.3 Results of sensitivity analyses



Inclusion of downstream costs

A threshold analysis was performed for inclusion of downstream costs for the obese health state. It was found that if downstream costs associated with being obese were greater than an average of $265 per person per year (over a 20 year time period) then the FFFFs media campaign intervention would dominate the control group.

2. Community-wide education for CVD risk factors – Stanford Five-City Project

2.1 Description

Intervention type

The section is concerned with establishing the cost-utility of long term, multi-factor, community wide health education programs. The Stanford Five-City Project (SFCP) was used as the exemplar for this type of program. The SFCP aimed to reduce the overall risk of cardiovascular disease (CVD) in the treatment population by 20% and was designed to test two major hypotheses:

1. 'Community-wide education can achieve a lasting reduction in the prevalence of cardiovascular disease risk factors within a general population, leading to a 20 per cent decline in the Framingham multiple logistic measure of risk in representative samples of persons.' (Farquhar et al., 1985)
2. 'The risk decline will lead to a decline in cardiovascular disease morbidity plus mortality in persons aged 30-74 that is significantly greater in the education cities than in reference cities.' (Farquhar et al., 1985)

The reduction in risk was to be achieved in the following risk factors.

1. *Cigarette smoking:* 9% reduction in cigarettes smoked per day,
2. *Weight:* 2% change in relative weight by increasing exercise and physical activity and reducing dietary energy intake,
3. *Blood pressure*: 7% reduction in systolic blood pressure through regular blood pressure checks and full adherence to anti-hypertensive medication regime, weight reduction, and increased exercise, and
4. *Cholesterol:* 4% reduction in total plasma cholesterol through changes in diet. (Fortmann and Varady, 2000, Farquhar et al., 1985)

The project was also designed to 'analyze the secular trends in cardiovascular disease risk factors, morbidity and mortality during a time of declining cardiovascular disease mortality, the cause of which is uncertain'. (Farquhar et al., 1985)

References/sources of evidence

A large amount of literature has been published in relation to the Stanford Five-City Project. The description and analysis contained in this chapter were based on the following:

- *Effects of communitywide education on cardiovascular disease risk factors: The Stanford Five-City Project* (Farquhar et al., 1990)
- *The Stanford Five-City Project: design and methods* (Farquhar et al., 1985)
- *Community intervention trials: reflections on the Stanford Five-City Project Experience* (Fortmann et al., 1995)
- *Effects of community health education on plasma cholesterol levels and diet: the Stanford Five-City Project* (Fortmann and Varady, 2000)
- *Effects of long-term community health education on blood pressure and hypertension control. The Stanford Five-City Project* (Fortmann et al., 1990)
- *Effect of long-term community health education on body mass index. The Stanford Five-City Project* (Taylor et al., 1991)
- *The long-term effects of a cardiovascular disease prevention trial: the Stanford Five-City Project* (Winkleby et al., 1996)

Intervention description

Recruitment and target population: The cities recruited to the project had to be located in northern California, with populations greater than 30,000, be reasonably independent from other cities and have similar ethnic, socioeconomic and demographic characteristics. There was to be shared media markets between the intervention cities but independence of newspaper and electronic media markets between the intervention and controls. (Farquhar et al., 1985) The cities participating in the trial were selected from all northern California cities and the selection criteria included size, distance from Stanford, and independence of the media markets in the control cities. The selection criteria precluded random assignment to the treatment and control groups. (Fortmann and Varady, 2000) There were two intervention and three control sites, but the control site of Santa Maria was included for monitoring morbidity and mortality trends only. (Table 2.1)

Table 2.1 Intervention and control sites in the Stanford Five-City Project

Population	Intervention Cities		Control Cities		
	Monterey	Salinas	Modesto	San Luis Obispo	Santa Maria
1980*	43,400	80,500	132,400	34,300	39,700
Est 1983 †	40,600	82,200	161,600	35,900	Not reported‡

Sources:
* (Fortmann and Varady, 2000)
† (Farquhar et al., 1990)
‡ Monitored for morbidity and mortality only, no surveys undertaken in the site.

Intervention: The intervention consisted of a six-year multi-factor, risk reduction education program that was 'coordinated, comprehensive and community wide'. Each component of the intervention had 'multiple target audiences' and used 'multiple communication channels and settings' including newspapers, television and radio, mass-distributed print media, classes, contests and correspondence courses. (Fortmann and Varady, 2000) Special programs were developed for Spanish-language radio, newspaper and mass-distributed print materials. The school-based programs for grades 4, 5, 7, and 10 included special sessions on nutrition, exercise and smoking, distribution of quit-smoking contest brochures, multi-factor risk reduction classes for teachers and administrators, and materials on exercise and nutrition for parents. (Fortmann and Varady, 2000, Farquhar et al., 1990)

It was estimated that each adult in the intervention sites was exposed to and average of 527 educational episodes fairly evenly distributed over 5 years.

- 7% involved direct educational sessions such as classes, lectures, or training workshops (although this component became more important over time and accounted for 27% of the exposure in the final year of the intervention);
- 34% involved television and radio;
- 41% involved booklets and kits; and
- 18% of newspapers and newsletters. (Farquhar et al., 1985, Farquhar et al., 1990)

The total hourly exposure per adult was estimated to be 26 hours over the life of the intervention and about five hours per year. (Farquhar et al., 1985, Farquhar et al., 1990)

Data from baseline population surveys in the treatment communities were used to develop an overview of knowledge, attitudes and behaviour for the target audience for each component of the intervention. This audience was segmented by age, ethnicity, socioeconomic status, overall cardiovascular risk, media use, organization membership and motivation to change behaviour (including diet). Formative evaluation with these audience subgroups was used to refine educational strategies, programs and materials. (Fortmann and Varady, 2000)

2.2 Quality of evidence

Evaluation description

Design: A quasi-experimental design with matched control cities was used to evaluate the project.

Methodology: Changes in intermediate measures (behaviour and clinical parameters) were estimated from four independent surveys and a panel study in four of the five cities (Table 2.2). All persons aged 12 to 74 years who lived in from households randomly chosen from commercial directories[3] for at least six months of the year, were eligible to participate in the survey. Invitation to participate was by mail, telephone, and in person. Invitees were asked to attend survey centres located in the cities where trained health professionals collected the data. (Farquhar et al., 1990) Estimates of the change in all-cause mortality risk were based on the data collected in the independent surveys.

Table 2.2 Data collection methods and response rates in the Stanford Five City Project

Data collection methods	Timing of data collection		Response Rates	
			Treatment	Control
Biennial Independent Samples Survey	I.1	Apr 1979 – Feb 1980	64%	66%
	I.2	May 1981 – Mar 1982	70%	69%
	I.3	May 1983 – Mar 1984	65%	65%
	I.4	May 1985 – Mar 1986	53%	58%
	I.5	Follow-up 1989/1990	65%	64%
Longitudinal Cohort Survey	C.1*	Apr 1979 - Feb 1980	39% of I.1	38% of I.1
	C.2	Sept 1980 - Jun 1981		
	C.3	May 1982 - Mar 1983		
	C.4	May 1984 - Mar 1985		
Epidemiologic Surveillance	Originally January 1979 to June 1987 extended to 1992			

Source: (Fortmann et al., 1995)
Notes:
* The sample for the cohort study was the first group of respondents for the Independent Samples Survey (I.1)

A 'community surveillance system' was implemented to validate fatal CVD events and count non-fatal events for residents aged 30-74 years at the time of the event in all five cities. Specific criteria were established to identify 'definite events' and less stringent 'sensitive criteria' were established to identify possible events which allowed for the detection of 'diagnostic drift and other threats to the validity of the process'. Only definite events were included in the analysis and, where a person experienced more than one definite event in any category, only the first such event was recorded. Fatal events included deaths due to myocardial infarction and stoke and other fatal coronary heart disease (mainly out-of-hospital sudden death). Non-fatal events included events for which residents were hospitalized (myocardial infarction and acute stoke). (Fortmann and Varady, 2000)

Analysis

Outcome measures and analysis: Details of the outcome measures and the analytic methods are shown in Table 2.4.

Assessment

Sources of bias: The potential sources of bias in the evaluation of the project were selection bias, measurement bias and attrition bias.

Selection bias: According to Farquhar, et al (1990) 'causality inferences are weakened since city selection and allocation to treatment and control conditions were unavoidably nonrandom', and

[3] According to Farquhar, et al (1985) these were 'found to provide a relatively complete listing of households in each community (approximately 97 per cent complete) and revised every two years'. (p. 328)

'adjustment for baseline differences' in terms of age, sex and household size 'only partly compensate for this'. (Farquhar et al., 1990) Differences between the cities in terms of the baseline survey are shown in Table 2.3.

The sampling frame appears to have been relatively complete, and the procedures for eliciting volunteers thorough. However, the participants were people who volunteered to undertake a fairly extensive data gathering procedure through a central agency in each city, therefore the generalisability of the results to non-volunteers was difficult to gauge.

Table 2.3 Baseline characteristics of the independent samples (I.1) in the control and intervention cities for those evaluation participants with blood pressure data

Characteristic	Control Cities (n=1,176)		Treatment Cities (n=1,188)		P Value
	Mean	SE	Mean	SE	
Age (years)	35.6	0.49	37.4	0.49	<0.01
Sex (% male)	48.0	1.46	47.4	1.45	0.75
Ethnicity (% white, non-Hispanic)	87.7	0.96	77.9	1.21	<0.01
Education (% ≤ high school)	49.0	1.46	54.2	1.45	0.01
Family income (% < $20,000 pa)	51.4	1.46	42.8	1.44	<0.01
Parental history CVD (%)	16.1	1.07	20.0	1.16	0.01
BMI (kg/m^2)	23.9	0.12	24.6	0.13	<0.01
Alcohol intake (oz/wk)	2.9	0.15	2.9	0.15	0.79
Physical activity (kcal/kg/day)	38.3	0.24	38.4	0.24	0.69
Knowledge (17 point scale)	6.4	0.09	5.9	0.09	<0.01
Prevalence					
Hypertension (%)*	24.1	1.25	26.7	1.29	0.14
Severe hypertension (%)†	5.6	0.67	7.9	0.79	0.02
Mean blood pressure					
Systolic (mmHg)	122.3	0.50	125.2	0.48	<0.01
Diastolic (mmHg)	72.8	0.33	77.1	0.32	<0.01
Hypertensives (n=599)					
% aware	53.0	2.97	44.3	2.8	0.03
% under treatment	31.8	2.77	26.6	2.49	0.16
% under control‡	27.2	2.65	14.6	1.99	<0.01

Source: (Fortmann et al., 1990), p. 634
Notes:
* Hypertension ≥140 or ≥ 90 mmHg or current use of antihypertensive medications
† Severe hypertension, ≥160 or ≥ 95 mmHg
‡ Under control, <140/90 mmHg

Measurement bias: All survey data were collected in the survey centres located in four of the five cities. The centres were staffed by a nurse/supervisor, an interviewer/ abstractor, a laboratory technician and a medical office assistant, all of whom were full-time employees (Farquhar et al., 1990). It is unlikely that those collecting the data were 'blind' to the intervention status of the participants in the evaluation. Behavioural measures were self-report, and only cigarette smoking was checked by blood assay. The estimated all-cause mortality and CHD risk changes must be interpreted with caution. The risk function used produces an average aggregate of risk change weighted according to the relative contribution of each factor to individual risk. The changes noted do not represent actual changes in disease incidence or mortality. (Farquhar et al., 1990)

Attrition bias: In the cohort sample, the dropout rates by the fourth survey were 50% which further limits the generalisability of the results for this group. (Fortmann et al., 1995)

2.3 Outcomes – as reported

Independent cross-sectional population samples were surveyed over an 11-month period in each city at baseline, 25, 51 and 73 months. All persons 12 through 74 years of age who resided in randomly selected households at least 6 months of the year were invited to participate but the data analysis and outcomes reported in the literature involved only persons aged 25 through 74 years. Those who participated in the baseline surveys were resurveyed 17, 39, and 60 months later to form a panel study. Only those who participated in all four cohort surveys were included in the cohort analyses presented in the published literature. This subgroup represented ~39% of the first independent sample. (Farquhar et al., 1990)

Table 2.4 Outcome measures, data sources and analyses

	Data Sources	Analysis
BEHAVIOURAL MEASURES AND CLINICAL PARAMETERS		
Smoking behaviour (Farquhar et al., 1990, Winkleby et al., 1996)	Questionnaire and measurement of plasma thiocyanate and expired air CO.	***1. Unit of analysis = the individual*** Comparison of the differences in risk factor changes of treatment vs control groups using a one-tailed, two-sample t test of the difference scores for the cohort sample. For the independent samples they used a contrast among the survey means and the appropriate one-tailed t-test. ***2. Unit of analysis = the city*** Regression lines were fitted to each city using mean values for each survey. The rate of change in treatment cities compared to the control cities tested by comparing the average slope of the regression lines for the two treatment cities with the average slope of the regression lines for the two control cities. ***3. Adjustment for baseline differences between cities*** Because gender, age and education level were known to affect most of the variables of interest, the variables were adjust within gender, age and education categories to common means and SDs,
Arterial blood pressure (Winkleby et al., 1996)	Three measurements using a semi-automatic recorder which produced tracings on a graduated paper disk. Blood pressure was obtained by reading the beginning and end of each tracing which corresponds to infrasonic energy transmitted at systole and diastole. Blood pressure was estimated by averaging the second and third readings. All disks were read after the conclusion of the fourth survey by a team of 10 trained coders at Stanford University.	
Obesity (Taylor et al., 1991)	Weight was obtained to the nearest ¼ lb (0.1kg) by a balance scale and height to the nearest 1/4 inch (0.6 cm) by metal rule.	
Plasma cholesterol Plasma high density lipoprotein cholesterol (Fortmann et al., 1993)	Non-fasting venous samples are obtained from participants while seated. The refrigerated plasma samples are shipped to Stanford twice each week and are analysed fresh to lipids and lipo-proteins.	
RISK SCORES		
All-cause mortality (Farquhar et al., 1990, Winkleby et al., 1996)	A risk estimate from the Framingham Study to illustrate the potential impact of risk factor changes on total all-cause mortality.	As above
CARDIOVASCULAR DISEASE INCIDENCE		
Combined fatal plus definite non-fatal coronary disease and stroke events (Fortmann and Varady, 2000)	Epidemiologic surveillance of changes in morbidity (definite non-fatal coronary disease and stroke events) and mortality (from CHD stroke) for the period 1979 through 1982.	***Unit of analysis = the city and the year-specific adjusted rates for both sexes were pooled*** Trends in rates over time were compared for the treatment and control communities. Trends were also contrasted between the first seven years and the second seven years.

Behavioural and clinical parameters

In the independent samples there were statistically significant differences between the control and intervention groups at the third measurement period in blood pressure but these differences were not in evidence at the end of the trial. (Table 10.4) In the cohort sample there were differences between

the groups at the third measurement period in the blood pressure, and at the end of the trial there were differences in blood pressure and all-cause mortality. (Table 2.5)

Table 2.5 Baseline and changes in risk factors for three treatment periods and comparison of net treatment/control differences for respondents aged 25 to 74 years in the independent samples (p values are based on one-tailed significance tests†)

	Baseline‡	Change from baseline to:‡		
	I1	I2	I3	I4
	Mean (± SE)	Mean (± SE)	Mean (± SE)	Mean (± SE)
Mean systolic blood pressure, mm Hg				
Treatment	130.54 (±0.60)***	- 3.15 (±0.60)	- 8.65 (±0.87)***	- 5.04 (±0.86)
Control	127.40 (±0.67)	- 1.50 (±0.90)	- 2.07 (±0.88)	- 3.61 (±0.85)
Mean diastolic blood pressure, mm Hg				
Treatment	81.71 (±0.38)***	- 4.01 (±0.53)**	- 7.49 (±0.55)***	- 3.22 (±0.54)
Control	77.97 (±0.39)	- 2.04 (±0.57)	- 0.15 (±0.55)	- 2.04 (±0.56)
Total cholesterol level, mmol/L				
Treatment	5.36 (±0.04)**	- 0.08 (±0.05)	- 0.04 (±0.06)	- 0.13 (±0.06)
Control	5.22 (±0.04)	- 0.03 (±0.05)	+ 0.01 (±0.05)	- 0.04 (±0.06)
Body mass index, kg/m^2				
Treatment	24.79 (±0.14)*	+0.36 (±0.20)	+ 0.69 (±0.22)	+ 0.49 (±0.21)*
Control	24.38 (±0.14)	+0.55 (±0.22)	+ 0.88 (±0.22)	+ 1.12 (±0.22)
Smokers, %				
Treatment	38.22 (±1.71)	- 1.62 (±2.39)	- 5.17 (±2.33)	- 9.02 (±2.26)
Control	37.62 (±1.79)	- 3.42 (±2.44)	- 3.88 (±2.34)	-10.24 (±2.39)

Source: (Farquhar et al., 1990)

Notes:

† Significance values: *** p ≤ 0.001, ** p ≤ 0.01, * p < 0.05

‡ Adjusted for age, sex, education and household size.

Table 2.6 Baseline and changes in risk factors for three treatment periods and comparison of net treatment/control differences for respondents aged 25 to 74 years in the cohort (p values are based on one-tailed significance tests†)

	Baseline‡	Change from baseline‡ to:		
	C1	C2	C3	C4
	Mean (± SE)	Mean (± SE)	Mean (± SE)	Mean (± SE)
Mean systolic blood pressure, mm Hg				
Treatment	129.65 (±0.91)**	- 2.36 (±0.72)	- 6.45 (±0.88)	- 8.88 (±0.96)***
Control	126.00 (±0.92)	- 3.19 (±0.83)	- 3.80 (±0.87)	- 3.71 (±0.89)
Mean diastolic blood pressure, mm Hg				
Treatment	80.43 (±0.54)***	- 0.94 (±0.53)	- 6.80 (±0.56)***	- 5.13 (±0.66)***
Control	76.99 (±0.51)	- 0.45 (±0.50)	- 2.52 (±0.56)	- 1.41 (±0.54)
Total cholesterol level, mmol/L				
Treatment	5.45 (±0.06)*	- 0.13 (±0.04)	- 0.08 (±0.04)**	+ 0.04 (±0.04)
Control	5.30 (±0.05)	- 0.04 (±0.04)	+ 0.05 (±0.04)	+ 0.07 (±0.05)
Body mass index, kg/m^2				
Treatment	24.63 (±0.20)	+ 0.19 (±0.06)	+ 0.35 (±0.07)	+ 0.51 (±0.09)
Control	24.20 (±0.19)	+ 0.02 (±0.06)	+ 0.44 (±0.08)	+ 0.4 (±0.08)
Smokers, %				
Treatment	28.35 (±2.28)	- 2.87 (±1.26)	- 5.26 (±1.60)	- 7.66 (±1.69)
Control	26.98 (±2.39)	0.00 (±1.07)	- 1.00 (±1.28)	- 3.76 (±1.28)

Source: (Farquhar et al., 1990)

Notes:

† Significance values: *** p ≤ 0.001, ** p ≤ 0.01, * p < 0.05

‡Adjusted for age, sex, education and household size.

All-cause mortality risk scores

Using the individual as the unit of analysis, the decline in the all-cause mortality was greater in the treatment group compared to the control group at the third measurement period but not at the fourth measurement period in the independent samples. For the cohort sample, the decline was greater at both the third and fourth measurement period. (Table 2.7) Based on the results for the cohort sample and using the city as the unit of analysis the estimated 10-year mortality rate decreased risk score decreased an average of 1.78 deaths per 1,000 persons per year in the treatment cities and 0.73 in the control cities ($p<0.02$ one-tailed significance test). (Farquhar et al., 1990)

Table 2.7 Baseline and changes in all-cause mortality risk scores* for three treatment periods and comparison of net treatment/control differences for respondents aged 25 to 74 years in the independent samples cohort sample (p values are based on one-tailed significance tests†)

	Baseline‡	Change from baseline‡ to:		
Independent samples	**I1**	**I2**	**I3**	**I4**
Treatment: Mean(±SE)	55.69 (±2.43)	- 5.52 (±3.29)	-11.82 (±3.15)**	-10.57 (±3.09)
Control: Mean(±SE)	50.80 (±2.24)	- 4.05 (±3.11)	- 0.26 (±3.00)	-10.39 (±2.80)
Cohort sample	**C1**	**C2**	**C3**	**C4**
Treatment: Mean(±SE)	51.97 (±2.98)	- 2.17 (±2.10)	-12.88 (±2.20)***	-11.16 (±2.47)**
Control: Mean(±SE)	48.18 (±2.85)	- 0.72 (±1.32)	- 4.20 (±1.54)	- 3.54 (±1.57)

Source: (Farquhar et al., 1990)
Notes:
* Deaths per 1,000 persons in 10 years.
† Significance values: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p < 0.05$
† Adjusted for age, sex, education and household size.

According to (Winkleby et al., 1996) and based on the independent samples, men and women in both the treatment and control cities showed positive improvements in all-cause mortality. Although the differences between the treatment and control cities were not statistically significant, (using a two-tailed significance test and using analysis of covariance to test the differences) the positive changes were maintained from the end of the trial in 1985-1986 until the follow-up period in 1989-1990.

Cardiovascular disease incidence (mortality and morbidity combined)

Over the full fourteen years of the study into the incidence of cardiovascular disease in both the treatment and control cities, there were no significant trends in any of the cities. During the late period of the study (1986-1992) there were significant downward trends in all except one city. The change in trends between the early and later periods was slightly, but not significantly, greater in the treatment cities. (Fortmann and Varady, 2000) Fortmann and Varady (2000) concluded that 'It is most likely that some influence affecting all cities, not the intervention, accounted for the observed change.'

2.4 Program costs

As reported by trial

The per capita cost of delivering the program was $4 pa (excluding research costs). The total cost in 2002 was US $5.5 million. This amounted to US $45 per capita and AUS $60 per capita. (Table 2.8)

Based on resource use

The descriptions of the trial are not detailed enough to allow a work-up of costs based on resource use.

Table 2.8 Costs of intervention as reported in the literature and converted to 2002 AUS $

Financial Years	**1980-81**	**1981-82**	**1982-83**	**1983-84**	**1984-85**	**1985-86**
Pop[n] in treatment cities, 1980*	123,900	123,900	123,900	123,900	123,900	123,900
Cost per yr per head of pop[n] (US$)†	$4	$4	$4	$4	$4	$4
Cost per year US$	$495,600	$495,600	$495,600	$495,600	$495,600	$495,600
Average Exchange rate (AUS$ per unit of foreign currency) ‡	0.8655	0.9093	1.0713	1.1047	1.2962	1.4175
Cost per year AUS$	$214,462	$450,644	$530,938	$547,468	$642,413	$702,497
		Jun 1982	**Jun 1983**	**Jun 1984**	**Jun 1985**	**Jun 1986**
% increase in Aust annual CPI§		10.53%	11.54%	6.73%	4.31%	8.41%
1980-81 cost inflated to 1986						$319,088
1981-82 cost inflated to 1986						$606,637
1982-83 cost inflated to 1986						$640,788
1983-84 cost inflated to 1986						$619,060
1984-85 cost inflated to 1986						$696,421
1985-86 cost						$702,497
Total cost in 1986 AUS$						$3,584,490
Inflation factor to convert cost to 2003 AUS$§						1.9075
Total cost in 2003 AUS$						**$6,837,355**
Cost per head of population in 2003 AUS$						$55.18
Cost per person aged 25 through 74 years (53.81% of total population)¶						$102.54

Notes:

* 1980 population in treatment cities of Monterey and Salinas, source (Fortmann and Varady, 2000), population assumed to be stable during the intervention.

† Source: (Fortmann and Varady, 2000)

‡ Source Reserve Bank of Australia data 1983-2004, available at http://www.rba.gov.au/Statistics/Historical/index.html accessed 19/04/2004

§ Source: ABS Cat No. 6401.0 Consumer Price Index, Australia, Table 3A.CPI: Groups, Weighted Average of Eight Capital Cities, Index Numbers (Financial Year)

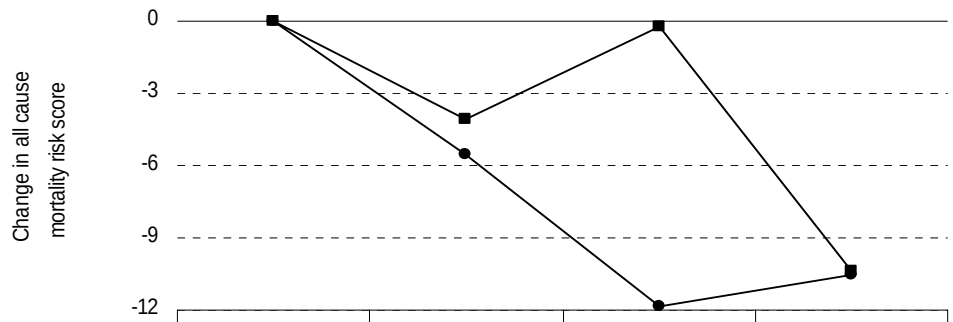
¶ Based on population figures for Victoria at June 1980, 53.81% aged 25 through 74 years. (Source: Australian Bureau Of Statistics population time series file 320102.123 *Table 2. Estimated resident population by single year of age. Victoria*, available at http://www.abs.gov.au accessed 17th March 2004)

2.5 Performance

Cost effectiveness

Compared to baseline the cumulative decrease in average all-cause mortality risk scores (the area under the intervention curve in Figure 2.1) was 22.63 units per 1,000 people. Net of the control group the decrease was 13.12 units per 1,000 people. The cost per 1,000 people aged 25 through 74 years was AUS $102,545 in 2003 and the cost per unit decrease in all-cause mortality risk score was $4,532 to $7,816.

Figure 2.1 Change in all-cause mortality risk scores for Intervention and Control cities for people aged 24 through 74 years



	Independent Sample No.1*	Change IS1-IS2	Change IS1-IS3	Change IS1-IS4
Intervention	0	-5.52	-11.82	-10.57
Control	0	-4.05	-0.26	-10.39

Notes:
* Baseline scores have been set to zero but were not actually zero.
Data Source: See Table 2.7

The cost per death saved was estimated to be $8,504 to $14,664 depending on whether the intervention cities were compared to baseline or the control group cities. (Table 2.9)

Table 2.9 Cost effectiveness of SFCP compared to be control cities and to baseline in the intervention cities (2003 AUS$)

	Compared to control cities	Compared to baseline
Average cumulative decrease in risk score per person	13.12	22.63
Interpretation	Decrease of 13.12 deaths per 1,000 persons over 10 years	Decrease of 22.63 deaths per 1,000 persons over 10 years
Length of intervention	5.3 years	5.3 years
Deaths per 1,000 people averted in response to 5.3 year program	6.99	12.06
Life years averted per 1,000 persons - p		
Cost of intervention per 1,000 people aged 25 through 74 years (2003 AUS$)	$102,545	$102,545
Cost per death averted (2003 AUS$)	$14,664	$8,504

2.6 Modelling

The Stanford 5 city published report (Farquhar et al, 1990) already reports modelled risk equations for 10 year mortality based on Framingham equations. These give results of $14,664 compared to control. Assuming no quality of life gain would lead to upper estimates of $14,664 per QALY.

We have not provided additional modelling for the following reasons:

- We would have needed to rely on a similar technique of translating outcomes into a Framingham risk equation which is already published
- Results from the two separate analyses (cohort and individual) are inconsistent with only one analysis showing a statistically significant reduction in the mortality risk score
- The results from the risk score analysis were not confirmed by actual results which showed no difference in cardiovascular morbidity or mortality over 14 years

3. Reducing children's television viewing to prevent obesity: a randomised controlled trial

3.1 Description

Reference

The evaluation in this intervention is based on a study by (Robinson, 1999). The study involved conducting a randomised school based trial to 'assess the effects of reducing television, videotape, and videogame use on changes in adiposity, physical activity, and dietary intake' (Robinson, 1999 pg. 1561).

Recruitment: target population and participants

The study was conducted in San Jose, California at two public elementary schools, with similar sociodemographic and scholastic characteristics. All third and fourth graders (mean age 8.9 years) at the two schools were invited into the study. Before randomisation, teachers and the principles of the schools agreed to enter the trial. Parents were required to provide written consent to allow their child to participate in the study. Consent was also required from parents for their own participation in telephone interviews. Classroom teachers were told about the 'the nature of the study' but were not told that the primary hypothesis was to test what effect reductions in viewing and videogame use had on childhood obesity. This was to ensure that teachers did not try to reduce the children's body mass index (BMI) in other ways, such as encouraging physical activity. The two schools were randomly assigned as either the intervention group or the control group.

Ninety-two (86.8 %) out of 106 eligible children entered the trial at the intervention school and completed the two assessments. One-hundred (82.6%) out of 121 eligible children entered the trial at the control school and completed the two assessments. As summarised in Table 3., the intervention and control group were comparable with respect to age, sex, the percentage of children with televisions in their room, and the number of televisions and video game players (Robinson, 1999 pg. 1564).

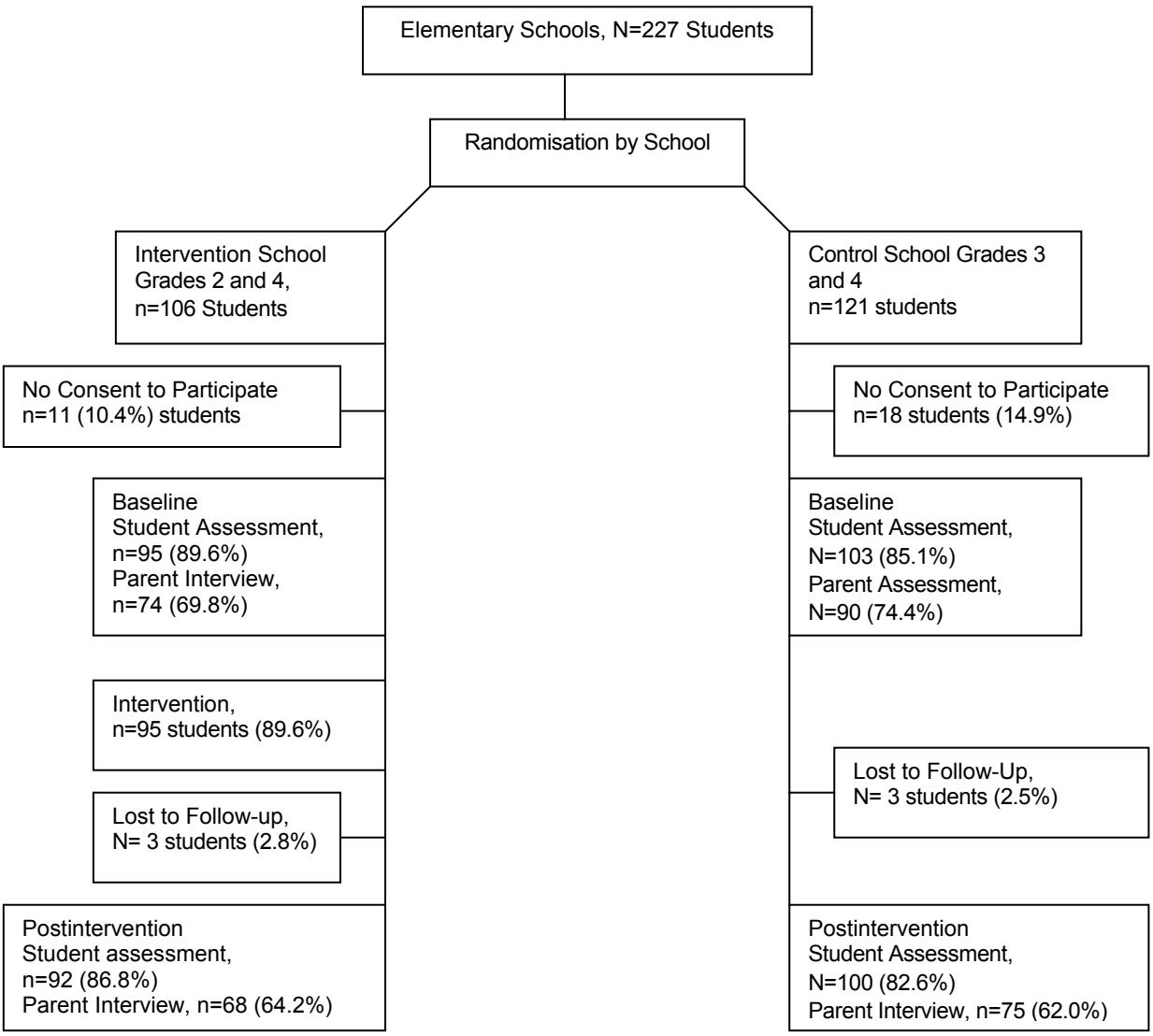
Table 3.1 Demographic comparison of the control and intervention group

	Intervention	Intervention	P Value
Mean Age*	8.95 (0.64)	8.92 (0.7)	0.69
Mean Number of Televisions*	2.7 (1.3)	2.7 (1.1)	0.56
Means Number Video Game Players*	1.5 (2.3)	1.2 (1.7)	0.49
Percentage of Children with Televisions in their Bedroom	43.5%	42.7%	0.92
Percentage of Females	44.6%	48.5%	0.59

*(SD)

Parents were asked to participate in a telephone interview about their child's media viewing, videogame use, physical activity and dietary intake. Sixty-eight parents (71.6%) completed the interviews at baseline and post intervention in the intervention school group. Seventy-five (72.8%) of the parents completed the interviews at baseline and post intervention in the control school group. 'Intervention school parents reported greater maximum household educational levels than participating control school parents (45% vs 21% college graduates, P=0.1) but did not differ significantly in ethnicity (80% vs 70% white, P=.19), sex of respondent (82% vs 88% female, P=.33) or marital status (77% vs 67% married, P=.22) (Robinson, 1999 pg. 1564). The recruitment and participation process is summarised in Figure 3.1.

Figure 3.1 Study design and participant flow
Source: (Robinson, 1999 pg. 1564)



Intervention

It was important to control for other variables that could influence children's body fat, dietary intake and physical activity (other than the amount of television, videotape and video game use). To ensure this happened, reduction in television, videotape and video game use, was specifically encouraged by classroom teaching and television budgeting.

The intervention was based on Bandura's social cognitive theory (Robinson, 1999 pg. 1562) Regular third and fourth grade classroom teachers gave 18 lessons, about 30-50 minutes in length, and were instructed how to give the lessons by the researchers. Teachers gave most of the lessons in the first two months of the year, with the entire curriculum lasting 6 months. The early lessons involved children self-monitoring and self-reporting television, videotape and video game use. A television turn-off was then encouraged, where children were not able to use television, videotape or video games for ten days. After this ten-day period, children were encouraged to watch no more than seven hours of television a week. Ninety-five (90%) children in the intervention group participated in the ten-day television turn-off, while seventy-one (67%) completed the entire television turn-off. Fifty-eight children (55%) handed in a card signed by their parents saying they had kept within their viewing and videogame use budget for the previous week. Later lessons focused on making children

become more selective about viewing and video game time. 'Several final lessons enlisted children as advocates for reducing media. Teachers reported giving all the lessons, however, there was no check to ensure that the lessons were being taught correctly.

Households were allocated a television manager (TV Allowance, Mindmaster, Inc, Miami, Fla). When the television manager was plugged into a power point, it would limit each person (identified by a number) to a certain amount of television, videotape and video game use. Households could order additional television managers for no additional cost. Forty-four parents (42%) said they had installed a television manager (to budget viewing), while 27% (n=29) of families asked for more than one television manager.

Parents were also provided with newsletters that gave advice both on how to encourage children to stay within their 'seven hour' budget, and advice on how to reduce family viewing and video game use.

Table 3.2 Summary of the intervention

Class Room Teaching	▪ 18 lessons about 30-50 minutes in length. Lessons taught children to self-monitor, report and be more selective about viewing and video game use. Final lessons made children advocates for reducing media.
'Television Turn-off Followed by Seven Hour per Week Limit	▪ Children were encouraged not to use television, video tapes or video games for ten days, and then after this ten day period watch no more than 7 hours a week.
Television Manager	▪ Television managers budgeted viewing by controlling the power use of a power socket, and identified each person in the household by an identification number.
Newsletters for Parents	▪ Newsletters provided advice to parents about reducing viewing and video game use of the child involved in the trial, and also the entire household.

3.2 Quality of trial

Recruitment

Over all, the recruitment process was sound. While the sample size was large enough to make formal comparisons between the control and the intervention group, it was too small to formally assess the effects of the intervention within subgroups (Robinson, 1999 pg. 1564).

If the number of schools used in the trial were increased, it would help eliminate the possibility that the results were due to differences in the profile of the students at the two schools. However, it is quite likely that the results were not just due to differences between the schools, as the comparisons made at baseline showed that the two groups were similar with respect to most variables, and both groups came from a similar school district. It is interesting to note, though, that 45% of children in the intervention group, compared to 21% in the control group (p=0,1) had parents who had graduated. It is possible that the children of more educated parents are more likely to be encouraged not to watch television and to be more physically active.

There was a high participation rate in the study, with 86.8% of eligible children in the intervention group and 82.6% of eligible children in the control group, participating in the study. This would suggest that the results are reasonably representative of a school within that distract.

Control group

The randomisation of the trial meant that an unbiased control was established, of a similar size to the intervention group. The control group received that same assessment as the intervention group, but were not told to reduce television, videotape, and vide game use.

Evaluation method

Assessments: There were only two assessments made, one at baseline (September 1996) and the other at post intervention (April 1997). Extra assessment periods could have been used to assess trends over the 6 month curriculum based intervention, to ensure that the results were not just 'one-off fluctuations in the data'. The trial was only over seven months. It would be important to gather data on the long-term effects of the program, particularly since many of the positive health benefits from reduced adiposity only occur in the long term.

Analysis: The evaluation controlled for expected changes in physical measures due to natural growth patterns, by making comparisons between the intervention and control groups. Analysis was conducted to assess the similarity between the intervention group and the control group (Wilcoxon rank sums test for scale variables and chi-square test for categorical variables).

Outcome measures

Measurements were made of television use, videotape use, video game use, adiposity, physical activity, and dietary intake. The trial would have been strengthened with other clinical measurements, such as blood pressure and cholesterol. Also, there was only measurement of moderate and high level physical activity. However, it is possible that the resulting loss of BMI in the intervention group was due to low levels of physical activity. This is particularly since the intervention group did not have statistically significant increases in physical activity, physical fitness or sedentary behaviour.

Having parents and children self-report on the children's viewing, dietary intake and other behaviour could have potentially led to inaccurate results, as it could be difficult to estimate many of the variables accurately, and parents and children might be reluctant to acknowledge or report 'undesirable' results (such as a large amount of time spent watching television). However, asking both parents and children the same questions provided a means of assessing the reliability of the results. Furthermore, a previous study, examining estimates of videotape use, suggested that the estimates could be accurate (Robinson, 1999 pg. 1563).

Bias, confounders, efficacy

When compared to school based lifestyle programs that give individual attention to at risk children, one advantage of this program is that at risk children would be less likely to feel inferior or isolated from their peers. This is because all children in the school undertake the same activities. Such psychological/social benefits of the program would not be directly accounted for by traditional cost-effectiveness analysis.

3.3 Outcomes – as reported

Outcome measures:

Measurements were taken by trained staff at baseline (September 1996) and after the completion of the trial (April 1997). Measurement staff were not made aware of the purpose of the trial. For each of the two measurement periods, children also completed self-assessment questions, at the same time in each of the two schools, on two non-Monday weekdays (as children were asked about their TV viewing yesterday before and after school). Researchers read out the questions, without classroom teachers present. Eleven children, with limited English or learning disability, did not complete this part of the assessment, and only completed the physical assessment. Measurement of physical activity was conducted for each assessment during two physical educational periods, by the same staff at both schools. Trained interviewers interviewed parents by telephone at each of the two assessment periods, according to a standardized protocol. Neither parents, children nor teachers were told that the primary outcome measure was the change in the adiposity.

Clinical measures:

The primary measure of adiposity was the body mass index (BMI), and was calculated by dividing weight in metres by the square of the height. Height was measured using a portable direct-reading stadiometer, while weight was measured according to established guidelines using digital scales. 'Test-retest reliabilities were high (intraclass Spearmen r>0.99 for height, r>0.99 for weight; Robinson, 1999 pg. 1562). Subcutaneous fat was measured according to the skin-fold thickness of the triceps on the right arm (retest reliability r>0.99). There was a high correlation between skin fold thickness and BMI (r=0.82).'Waist and hip circumferences were measured with non elastic tape at the level of umbilicus and the maximum extension of the buttocks (retest reliability r>0.99). Waist and hip circumferences were correlated with BMI (r=0.87, r=0.9 respectively) and triceps skin fold thickness (r=0.72, r=0.78, respectively).

Measurement of viewing and videogame use:

During the first assessment period (September 1996) children were asked how much time they spent yesterday (both before and after school) and last Saturday watching television and videos and playing video games. At the second assessment, children were asked how much time they spent yesterday watching television and videos and playing video games. To ensure that the children's estimates were accurate, exercises were conducted before hand. The instruments used had proved to be successful in a previous study on young adolescences (r=0.94) (Robinson, 1999 pg. 1563).

Parents also assessed children's viewing and video game use on a typical school day and weekend day. There was moderate correlation between children and parents video game use (Spearman r=0.31, P<0.001 for television viewing; r=0.17, P=0.03 for video tape viewing; r=0.49, P<0.001 for video game playing).

Sedentary behaviour:

Parents and children also estimated how much time children spent in other sedentary behaviour (parent-child agreement Spearman r=0.16, P<0.5). Estimates were included on 'the time spent using a computer, doing homework, reading , listening to music, playing a musical instrument, doing artwork or crafts, talking with parents, playing quiet games indoors, and at classes or clubs (Robinson, 1999 pg. 1563).

Physical activity:

Children's out of school physical activity was self reported, based on an activity checklist, at each of the two measurement dates. The average response of the two days was calculated, and weighted, using standard estimates of how much energy each type of activity would use. Parents also estimated the amount of time children spent in both organised and non-organised sport (parent-child agreement Spearman r=0.16, P=0.05).

Diet:

Children's dietary intake was measured at each of the two assessment periods. This involved children recalling 60 types of food in 26 categories. Children, parents and school lunch personnel identified highly fatty foods in focus groups.

According to a four point scale rating (from never to every day), children indicated how often they ate breakfast or dinner with the television on in the past week. On a 3-point scale, children reported how often they had a drink or ate a snack while watching television or videotapes or playing video games. Parents answered similar questions (parent-child agreement for meals Spearman r=0.24, P=0.003; parent-child agreement for snacks Spearman r=0.02, P>0.5).

Fitness test:

To measure cardiorespiratory fitness, 'a maximal multistage, 20-m shuttle run test (20-MST) was used' (reliability on test-retest r=0.73-0.93) (Robinson, 1999). The 20-MST has also been found to

accurately measure oxygen consumption, according to treadmill testing (r=0.73-0.93), and is sensitive enough to monitor change in children.

Statistical analysis:

Comparisons were made at baseline between the control and intervention group using the Wilcoxon test for scaled variables and the chi-squared test for categorical variables.

The intervention was designed to have effects on the entire distribution of adiposity in the sample. This meant comparisons needed to be made between the full distribution of the control and the intervention groups. 'Therefore, to test the primary hypothesis, accounting for the design with school as the unit of randomisation (adjusting for intra call correlation), a mixed-model analysis of covariance approach was used, with post intervention BMI as the dependent variable; the intervention group (intervention vs. control) as the independent variable; and baseline BMI, age, and sex as covariates (SAS MIXED procedure, SAS version 6.12, SAS Institute Inc, Cary, NC) (Robinson, 1999 pg. 1563).

The approach was applied to other secondary dependent variables, such as triceps skinfold thickness, waist and hip circumferences, waist-to-hip-ratio, dietary intake and physical fitness measures. 'Each outcome also was tested for intervention by sex and intervention by age interactions. All analysis were completed on an intention to treat basis, and all tests of statistical significance were 2-tailed with α =0.5 (Robinson, 1999 pg. 1563)'.

Any changes in the intervention group were measured relative to changes in the control group. This was particularly because children of this age are expected to have a natural increase in BMI, triceps skin fold thickness, waist circumference, and hip circumference. Therefore, 'a negative difference is termed a *relative decrease* in comparison with the controls, even if the actual value increased as a result of normal growth and development (Robinson, 1999 pg. 1563)'.

Behaviour change

Table 3.3 and Table 3.4 summarise the child's media use, diet and physical activity, as reported by the child and parent, respectively.

At baseline, both parents and children's estimates were similar. However, children in the intervention group said they ate more meals while watching television, than parents reported. When compared to the estimates by children, parents in the intervention group reported more time in sedentary behaviour, and overall household television use.

There was a significant reduction in the television viewing of the intervention group, according to reports from both children and parents ($p<.001$). There was almost a one third reduction in television viewing when compared to baseline. Children in the intervention group also reported a statistically significant reduction in video game use compared with controls. Parents also reported a decrease in video-game use, and overall household television viewing, but these were not statistically significant. Both parents and children reported a decrease in video viewing, but the results were not statistically significant.

Parents and children both reported a statistically significant decrease in the number of meals watched in front of the television. There were no statistically significant results with respect to reduction in the consumption of high fat-foods, children's physical activity levels, or score on the 20-MST shuttle test.

There were no significant interactions by sex or intervention by age interactions with media use, diet, or physical activity outcomes (Robinson, 1999 pg. 1564-1565).

Table 3.3 Child measures of television viewing, diet, and physical activity and fitness*

	Baseline		Postintervention		Adjusted Change (95%CI)**	P Value
	Intervention	Control	Intervention	Control		
Television (hours per week)	15.35 (13.17)	15.46 (15.02)	8.80 (10.41)	14.46 (13.82)	-5.53 (-8.64 to-2.42)	<0.001
Videotapes (hours per week)	4.74 (6.57)	5.52 (10.44)	3.46 (4.86)	5.21 (8.41)	-1.53 (-3.39 to 0.33)	.11
Video games (hours per week)	2.57 (5.10)	3.85 (9.17)	1.32 (2.72)	4.24 (10.00)	-2.54 (-4.48 to -0.6)	.01
Meals in front of television, 0-3 Scale	2.38 (1.75)	1.84 (1.78)***	1.70 (1.49)	1.99 (1.78)	-0.54 (-0.98 to-0.12)	.01
Frequency of Snacking in Front of the Television, 1-3 scale	2.20 (0.56)	2.15 (0.61)	1.94 (0.51)	2.05 (0.59)	-0.11 (-0.27 to 0.04)	.16
Daily Servings of High-Fat Foods	6.15 (3.63)	6.62 (5.85)	5.14 (3.50)	6.17 (4.88)	-0.82 (-1.87 to 0.23)	.12
Daily Serving of Highly Advertised Foods	1.36 (0.96)	1.55 (1.20)	1.47 (1.10)	1.48 (1.06)	0.06 (-0.24 to 0.36)	.71
Other Sedentary Behaviours, h/d	4.66 (3.81)	4.47 (6.37)	3.81 (2.26)	4.05 (4.53)	-0.34 (-1.21 to 0.52)	.44
Physical activity, metabolic equivalent-weighted, min/wk	396.8 (367.8)	310 (250.7)	362.3 (235.2)	337.8 (277.3)	-16.7 (-78.6 to 45.3)	.60
20-m shuttle test, laps	15.21 (9.60)	14.80 (8.56)	19.72 (11.40)	18.18 (10.72)	0.87 (-1.41 to 3.15)	.45

Source: (Robinson, 1999 pg. 1565)

*Baseline and postintervention values are unadjusted mean (SD).

**Change estimates and 95% confidence intervals (CIs) are the differences between groups after adjustment by mixed model analysis of covariance for the baseline value, age, and sex.

***Groups were significantly different (p<.05) at baseline by a nonparametric Wilcoxon rank sum test.

Table 3.4 Parent reports of children's television viewing, diet, and physical activity*

	Baseline		Postintervention		Adjusted Change (95% CI)**	P Value
	Intervention	Control	Intervention	Control		
Television (hours per week)	12.43 (5.65)	14.90 (7.10)	8.86 (4.91)	14.75 (7.37)	-4.29 (-5.89 to-2.70)	<0.001
Videotapes (hours per week)	4.96 (4.21)	4.41 (3.72)	3.87 (2.87)	3.91 (3.21)	-0.25 (-1.19 to 0.69)	.60
Video games (hours per week)	1.84 (2.73)	2.71 (3.78)	1.44 (1.96)	2.57 (4.41)	-0.76 (-1.75 to 0.22)	.13
Overall household television use, 0-16 scale	7.09 (3.97)	8.60 (3.51)***	6.09 (3.64)	7.76 (3.26)	-0.77 (-1.69 to 0.14)	.10
No. of children's meals eaten in front of the television, 0-14 meals	3.18 (3.69)	3.53 (3.71)	2.19 (2.95)	3.43 (3.64)	-1.07 (-1.96 to-0.18)	0.02
Percentage of Children's Viewing When Snacking	17.28 (20.91)	18.83 (41.24)	19.54 (22.43)	20.25 (22.70)	-1.94 (-9.06 to 5.17)	.59
Children's other sedentary behaviours, h/wk	44.89 (19.76)	39.79 (20.27) ***	41.3 (20.89)	43.37 (26.75)	-4.88 (-11.69 to 1.93)	.16
Children's physical activity, h/wk	11.19 (7.16)	9.19 (5.77)	16.08 (8.45)	17.21 (9.32)	-2.00 (-4.58 to 0.59)	.13

Source: (Robinson, 1999 pg. 1565)

*Baseline and postintervention values are unadjusted mean (SD).

**Change estimates and 95% confidence intervals (CIs) are the differences between groups after adjustment by mixed model analysis of covariance for the baseline value, age, and sex.

***Groups were significantly different (p<.05) at baseline by a nonparametric Wilcoxon rank sum test.

Clinical parameters

Changes in adiposity:

Table 3.5 summarises the anthropometric measures. There were no statistically significant differences at baseline between the control group and the intervention group, with respect to any of the physical measures (P>.10).

Due to the normal growth patterns of children in this age group, the BMI, skinfold thickness, waist circumference and hip circumference increased between the two measurement periods, for both the control and the intervention group. Therefore, the effects of the intervention were measured relative to the changes in the control group. There were statistically significant decreases in the intervention group, when compared to the control, for all the anthropometric measures, except hip circumference. 'There were no significant interventions by sex or interventions by age interactions for any of the body compositions outcomes. The results did not change with ethnicity and parent education were included as additional covariates for children with completed parental interviews' (Robinson, 1999 pg. 1564).

The effects of the intervention were greater, when compared with control, among the middle and higher strata of body fatness.

Table 3.5 Anthropometric measures

	Baseline		Post Intervention			
	Intervention	**Control**	**Intervention**	**Control**	**Adjusted Change (95% CI)****	**P Value**
Body mass index, kg/m^2	18.38 (3.67)	18.10 (3.77)	18.67 (3.77)	18.81 (3.76)	-0.45 (-0.73 to 0.17)	0.002
Triceps skinfold thickness, mm	14.55 (6.06)	13.97 (5.43)	15.47 (5.95)	16.46 (5.27)	-1.47 (-2.41 to -0.54)	.002
Waist circumference, cm	60.48 (9.91)	59.51 (8.91)	63.57 (8.96)	64.73 (8.91)	-2.30 (-3.27 to -1.33)	<.001
Hip circumference, cm	72.78 (8.91)	72.70 (8.78)	76.53 (7.94)	76.79 (8.37)	-0.27 (-1.08 to 0.53)	.50
Waist-to-hip ratio	0.83 (0.05)	0.82 (0.05)	0.83 (0.06)	0.84 (0.05)	-0.02 (-0.03 to -0.01)	<.001

Source: (Robinson, 1999 pg. 1565)

*Baseline and postintervention values are unadjusted mean (SD). At baseline, both groups were compared (P>.10) on measures of body composition.

**Change estimates and 95% confidence intervals (CIs) are the difference between intervention group and control group after adjusting by mixed-model analysis of covariance for the baseline value, age, and sex.

Mortality

Not reported

Morbidity

Not reported

3.4 Program costs

As reported

The study by Robinson (1999) did not report costs associated with the intervention to reduce children's television watching.

Based on resource use

Intervention group:

The intervention consisted of the following general costs: training, assessment, teaching, equipment and consumables (see Table 3.6, Table 3.7, Table 3.8 and Table 3.9).

Table 3.6 Summary of training costs for the 106 children in the intervention group

	Cost	Time	Number required	Cost for study	Average cost per person
Training teachers	$47.23 per hour	4 hours	4.24	$801.02	$7.56

The following assumptions were made when estimating the training and research costs:

- Teachers salary is the NSW middle rate (step 7) of $45 343 per annum plus 25% on costs and assuming 40 weeks per year and 6 hours per day or $47.23 per hour (Australian Education Union, School teachers' salaries Sept 2003).
- The length of training is assumed to be 4 hours
- It is assumed that one teacher is required for every 25 students enrolled

Table 3.7 Teaching costs for the 106 children in the intervention group

	Cost (per hour)	Time	Number of teachers required	Cost for study	Average cost per person
Cost of teaching time (opportunity cost)	$47.23	18 lessons of 40 minutes	4.24	$2,403.06	$22.67

The following assumptions were made when estimating the cost of teaching time:

- That there is an opportunity cost with a value attached to the time taken from the normal teaching curriculum
- That 18 lessons are required with an average time of 40 minutes each (mid point of 30-50; Robinson, 1999)
- A teacher is required for every 25 students
- Teachers salary is the NSW middle rate (step 7) of $45 343 per annum plus 25% on costs and assuming 40 weeks per year and 6 hours per day or $47.23 per hour (Australian Education Union, School teachers' salaries Sept 2003).

Table 3.8 Assessment costs for the 106 children in the intervention group

	Cost (per hour)	Time	Number required	Cost for study	Average cost per person
Interviewing parents	$27.07	1 hour	106 people with 2 interviews each	$5,739.78	$54.15
Assessment of children- baseline	$27.07	40 minutes per lesson for 3 lessons	4.25 researchers required (1 per class)	$229.59	$2.17
Assessment of children- follow up	$27.07	40 minutes per lesson for 3 lessons	4.25 researchers required (1 per class)	$229.59	$2.17

The following assumptions were made when estimating the assessment costs:

- Each parent would require 2 interviews (baseline and follow up) lasting one hour each
- Children would be assessed over three 40 minute lessons by researchers. Assumes one lesson for students to complete self assessment forms with researcher reading out questions, and another 2 lessons for physical education assessment (Robinson, 1999)
- One researcher's time would be required to conduct each class of 25 students
- Researchers are assumed to cost $25.07 per hour (NHMRC level RA4 from www.unimelb.edu.au/ppp/docs/16.html +20% on costs)

Table 3.9 Cost of equipment and consumables for the 106 children in the intervention group

	Cost per unit	Number required	Other costs	Cost for study	Average cost per person
Newsletters	$1.27	106		$134.62	$1.27
TV allowance-television time manager	$141.57 (1st unit) $112.97 (subsequent units)	135	Postage $3.79 per unit	$15,791	$148.97
20 meter shuttle run test	$69	1		$69	$0.30
Digital scale plus stadiometer	$613.17	1		$613.17	$2.70
Skin fold thickness tester	$496.52	1		$496.52	$2.19
Non elastic measuring tape	$5.00	1		$5.00	$0.02

The following assumptions were made when estimating the cost of equipment and consumables:

- The cost of printing is $3.80 per questionnaire (Based on quote for A4 booklet of 8 pages (x3) of 300 for $345 (ex GST) from Melbourne University Design and Print Centre 2003+GST10%= $3.80)
- Eight separate assessments are made (parents at baseline and follow up, children's physical measures at baseline and follow up, children's self-report assessment on television viewing at baseline and follow up and children's food frequency recall at baseline and follow up)
- The cost of each newsletter is $1.27 (Based on quote for A4 booklet of 8 pages: 300 for $345 +10% GST from Melbourne University Design and Print Centre 2003)
- Assumes that TV Allowance units cost $141.57 for the 1st unit and $112.97 for subsequent units and that an average cost of $3.79 is required for the postage of each unit (http://www.tvallowance.com/ Converted from US to AU$ on 27th Oct 03 rate 0.699302)
- Assumes that all children require a TV allowance and that 29 request additional units (Robinson, 1999)
- A digital scale plus stadiometer costs $613.87 (www.medicalresources.com, converted from US to AU$ on 27th Oct 2003 rate 0.701269)

Control group:

The control group are assumed to incur the cost of "researchers (set up, recruitment, designing protocols and forms, conducting training)", the cost of assessment and the cost of the equipment associated with assessment (ie shuffle run test, scales, skinfold tester and measuring tape).

Total costs:

Table 3.10 provides a summary of the total costs incurred for the intervention and control groups and the average cost per person.

Table 3.10 Summary of costs of intervention (n= 106 children)

	INTERVENTION GROUP		CONTROL GROUP	
	Cost for group	Average cost per person	Cost for group	Average cost per person
Recruitment and training	$56,758.47	$500.35	$56,314.80	$496.17
Teaching	$6,198.96	$58.48	$6,198.96	$58.48
Assessment	$19,740.07	$185.86	$3,814.25	$35.61
Equipment and consumables	$1,331	$12.56	$0.00	$0.00
Total	$84,028.52	$757.25	$66,328.01	$590.26

3.5 Performance

Cost effectiveness

The main outcome for the trial was change in adiposity. BMI measurements for the two groups at baseline and follow up are summarised in Table 3.. Changes were adjusted for natural increases in BMI over time in children this age.

Table 3.11 Changes in BMI at baseline and follow up

Group	Baseline (kg/m^2)	Follow up (kg/m^2)	Adjusted Change
Intervention	18.38	18.67	
Control	18.10	18.81	-0.45 (95%CI –0.73 to –0.17)

The costs for the intervention and control groups were summarised in section 3.4.2. The ICER per BMI point reduction for the intervention group compared to the control is as follows:

ICER= costs intervention-cost control/ outcomes intervention- outcomes control
ICER= ($757.25-$590.26)/ 0.45
=$166.99/ 0.45
= $371.09 per point reduction in BMI for the intervention group compared to the control group

The study by Robinson also reported a reduction in hours of television viewed per week as shown in Table 3.12.

Table 3.12 Changes in hours of television viewed per week at baseline and follow up

Group	Baseline (hours/week)	Follow up (hours /week)	Adjusted change
Intervention	15.35	8.80	
Control	15.46	14.46	-5.53 (95%CI –8.64 to –2.42)

This leads to the following ICER per hour reduction in television viewing per week:

ICER= costs intervention-cost control/ outcomes intervention- outcomes control
ICER= ($757.25-$590.26)/ 5.53
=$166.99/ 5.53
=$30.20 per hour reduction in television viewing per week for the intervention group compared to the control group.

Cost-utility

Cost-utility was unable to be calculated due to the lack of reported quality of life or utility measures.

3.6 Modelling – hypothetical scenario analysis

In order to model this intervention using the replicated Wang et al (2003) model the proportion of participants overweight/obese at baseline and follow up would be required for each study group. This information is not provided. We therefore, provide hypothetical scenario analysis based on the replicated Wang et al (2003) model. The following assumptions are made:

- The proportion in population who are overweight or obese at age 9, assuming a boy to girl ratio of 1.05 is 19% (AIHW, 2003)
- Cost of $757.25 for those in intervention group and $590.26 for control group
- 75.4% of those overweight at 14 years are overweight by age 21-29
- 9.8% of normal weight 14 year olds become overweight by age 21-19
- 91.2% of overweight people aged 21-29 are overweight aged 40
- 39.3% of normal weight people aged 21-29 are overweight aged 40
- the QALYs saved per case of overweight prevented is 0.712 with healthy life years only obtained from age 40 to 65

- It is assumed that weight remains unchanged between age 14 and 21 and also unchanged between age 29 and 40
- Discounting of costs and outcomes at 5%
- Exclusion of downstream costs
- 50% of those becoming a normal weight relapse within 7 years to become overweight again
- The final weight of a person between ages 40 and 65 is the only determinant of healthy life years and medical costs (weight at age 14 and ages 21-29 is irrelevant)

Table 3.13 shows estimates for various hypothetical scenarios of the proportion reduction in the number of children average age 9 years who are overweight or obese. This reduction is assumed to occur for a hypothetical intervention group compared to a control group that do not change weight.

Table 3.13 Cost per QALY for hypothetical scenarios of reductions in overweight/obese children

% reduction in overweight/obese	Proportion overweight/obese at end of intervention	Resulting cost/QALY No relapse	Resulting cost/QALY 50% relapse
5%	0.1805	$149,217	$298,630
10%	0.171	$74,609	$149,745
15%	0.1615	$49,739	$103,153
20%	0.152	$37,304	$74,591

Inclusion of downstream costs

The 25 year excess medical costs per overweight person is US$2737. This is converted to Australian exchange rates using 1996 figures and inflated to Australian 2003 values which give a cost per person of $4562. When this cost is included in the hypothetical scenario analysis, the estimates are as presented in Table 3.14.

Table 3.14 Cost per QALY for various hypothetical scenarios including downstream costs of obesity

% reduction in overweight/obese	Proportion overweight/obese at end of intervention	Resulting cost/QALY No relapse	Resulting cost/QALY 50% relapse
5%	0.1805	$136,032	$285,445
10%	0.171	$61,424	$136,561
15%	0.1615	$36,554	$89,968
20%	0.152	$24,119	$61,405

These scenarios are only intended to provide a guide as to what cost/QALYs may be likely if certain results were to be obtained.

3.7 Discussion

The replicated Wang et al (2003) model is structured in such a way that mortality and quality of life benefits are gained for those aged 40-65. This means that the intervention is not cost effective if it is assumed that all of the intervention group revert to baseline values in the few years following the intervention. In reality it is possible that the majority of those losing weight during the intervention will regain it and also that there will be further gains/changes made before age 40.

4. Reducing obesity via a school-based interdisciplinary intervention among youth

4.1 Description

Reference

The evaluation in this section is based on a study by (Gortmaker *et al.*, 1999). The study assessed the impact of a school-based program designed to reduce obesity amongst children in grades 6 to 8. The interdisciplinary intervention was administered over two school years. To determine whether a child was obese, a composite indicator was used, based on both triceps skinfolds (TSF) and body mass index (BMI).

Recruitment: target population and participants

Schools were invited into the trial if they were wiling to adopt the interdisciplinary program, had a multiethnic population, and were willing to be assigned as either a control or an intervention school. There were 10 schools in the study, 5 intervention schools and 5 control schools. A total of 1295 students completed the trial (providing both baseline anthropometry and survey data and follow up anthropometry data). The schools came from four communities in Boston, Massachusetts. Consent was required from the students' parents to permit physical measurements of the students and completion of the surveys. Five of the schools required 'active consent', where written consent was required for the child to be involved in the trial. The other schools required 'passive consent', where a child was involved in the trial unless the parent requested the child not to be. In schools with active consent, 58% of parents consented to the trial - compared to 89% in schools with passive consent. In the intervention group, 56% of potential students required active consent, compared to 36% in the control group.

Before randomisation, schools were divided into pairs, according to either the communities they were in (for eight of the schools) or according to demographic characteristics (for two of the schools). Schools were divided into pairs to reduce the chance of pre-intervention differences between control and intervention schools that might affect the prevalence of obesity. Control schools received no intervention, and received their normal classes. All children in grades 6 and 7 (average age 12 years) in 1995 received the intervention (but only those with parental consent were involved in the assessment).

When baseline data were stratified by sex, there were no 'significance differences among the 1560 intervention and control students in mean values of age, BMI, TSF, or obesity prevalence.' Amongst those students who provided baseline data, there was a higher proportion of African American girls (17% vs 10%) and Hispanic boys (18% vs 12%) in control schools. A lack of parental consent (80%) and school absence (14%) were the main reasons students did not provided baseline anthropometry data.

The intervention group had a lower rate of follow up amongst obese boys compared to the control (87% intervention vs. 94% intervention). 'The main reasons for lack of follow up anthropometric data included school transfer (52%), school absence (27%), and child refusal (10%)'. However overall, the follow up for intervention and control students was similar for both girls (82% control vs 81% intervention) and boys (86% control vs 83% intervention).

Amongst those students who completed the trial, the baseline data were similar - see Table 4.1. There was similar prevalence of obesity at baseline (28% control vs 27% intervention). There was, however, a higher prevalence of African American girls in the control (16% control vs 10% intervention).

Table 4.1 Baseline characteristics of a longitudinal study sample of sixth and seventh grade students in Fall 1995, with anthropometric data in Spring 1997, at the Intervention (I) and Control (C) Schools*

	Students		Girls		Boys	
Characteristics	I (n=641)	C (n=654)	I (n=310)	C (n=317)	I (n=331)	C (n=377)
Background						
Age †	11.7 (0.7)	11.7 (0.7)	11.6 (0.7)	11.6 (0.7)	11.8 (0.7)	11.8 (0.8)
Female %	48	48	-	-	-	-
Ethnicity %						
White	69	63	72	63	67	63
African America	11	15	9	16	12	13
Hispanic	11	16	10	15	12	17
Asian/Pacific Islanders	9	7	8	7	9	7
American	2	2	2	2	2	1
Other	5	9	7	9	4	8
Anthropometric Data						
Obese, % ‡	27	28	24	22	29	35
Height, cm †	152 (8.3)	152 (8.4)	152 (7.8)	152 (7.7)	151 (8.7)	152 (9.0)
Body Mass Index †	20.6 (4.5)	20.7 (4.0)	20.8 (4.6)	20.6 (4.2)	20.5 (4.4)	20.8 (3.9)
Triceps skinfolds †	16.0 (7.2)	15.9 (6.9)	16.9 (6.8)	16.4 (6.7)	15.2 (7.6)	15.5 (7.0)
Body mass index≥85th percentile	34	37	32	32	35	43
Triceps skinfolds ≥ 85th percentile	30	32	26	25	34	40
Females who has completed menarche, %	-	-	28	35	-	-
Baseline Dietary/ Activity Variables						
Television/video, h/dt†	3.4 (2.2)	3.5 (2.2)	3.0 (2.1)	3.1 (2.2)	3.8 (2.3)	3.8 (2.2)
Moderate/vigorous activity, h/dt†	2.2. (1.6)	2.1 (1.6)	1.8 (1.4)	1.7 (1.4)	2.5 (1.7)	2.5 (1.6)
Energy from fat, % †	31.7 (5.2)	31.3 (5.7)	31.3 (5.2)	31.1 (6.0)	32.1(5.2)	31.5 (5.5)
Energy from saturated fat, % †	10.9 (2.4)	10.8 (2.6)	10.7 (2.4)	10.6 (2.6)	11.2 (2.3)	11.1 (2.6)
Fruit and vegetables, servings/d†	3.7 (2.7)	4.2 (2.8)	3.4 (2.4)	4.2 (2.8)	4.0 (2.9)	4.1 (2.8)
Total energy intake, J/d†	8597.4 (4393.2)	8849.4 (4750.2)	7555.8 (3423.0)	8034.6 (4128.6)	9580.2 (49 56.0)	9626.4 (5161.8)
Smoked in last month, %	1	4	1	3	1	5
Physical education, times/wk‡	1.7 (0.6)	2.0 (1.1)	1.6 (0.6)	2.0 (1.1)	1.7 (0.6)	2.0 (1.2)
Knowledge						
Dietary knowledge†	11.4 (3.5)	11.2 (3.6)	12.0 (2.9)	12.0 (3.2)	10.8 (3.9)	10.5 (3.9)
Activity knowledge †	6.9 (1.9)	6.9 (1.8)	7.3 (1.3)	7.3 (1.5)	6.4 (2.2)	6.5 (2.0)
Dieting Behaviour						
Diet to lose weight, %	26	30	29	32	23	29
Exercise to lose weight %	44	43	47	43	41	42
Vomit/take laxatives, %	4	6	2	4	6	7
Take diet pills, %	3	3	2	2	4	4

*Sample size vary slightly due to missing data.

†Values are expressed as mean (SD).

‡Obesity was measured by body mass index and triceps skinfold greater than or equal to the 85th percentile.

Source: (Gortmaker *et al.*, 1999 pg. 414)

Intervention

The intervention was based on the Planet Health program. 'The intervention focused on 4 behaviour changes : reducing television viewing to less than 2 hours per day; increasing moderate vigorous physical activity; decreasing consumption of high-fat food; and increasing consumption of fruit and vegetables to 5 a day or more (Gortmaker *et al.*, 1999 pg.410)'.

The intervention incorporated 'behavioural-choice and social cognitive theories of individual change'. It was hoped that in reducing television viewing children would have 'free space' in their lives to undertake physical activity. The intervention was designed to give children 'cognitive and behavioural skills' to change selected behaviour. The intervention was population based, so did not actively target 'at risk' youth.

The classroom lessons were designed to encourage one of the four behavioural objectives. Each subject (language arts, math, science, and social studies) had one lesson to achieve each of the four objectives per year. This meant there were a total of 16 lessons per year, and 32 lessons over the two school years. Lessons could vary in length from one or two 45 minute periods. The lessons encouraged student interaction, with 'class debates, case studies, group projects, games and student presentations.'

Physical activity lessons encouraged students to replace inactive time with moderate and vigorous physical activity of the students' choice. The students were taught to set goals and self assess their physical activity levels. Fitness funds of $400-600 were given to intervention schools that put forward proposals that were compatible with Planet Health themes. The physical activity program was based around thirty 5 minute long lessons (microunits), which were designed to be repeated with extensions in the second year of the intervention.

'Each intervention school received the Planet Health program of teacher training workshop, classroom lessons, PE materials, and fitness funds.' Training was provided for all teachers and PE staff either in a workshop or by the filed co-ordinator. 'Based on teacher interests, an average of 3 teacher/staff wellness sessions were offered per school, provided at low cost by outside organizations.'

Table 4.2 Summary of the 2 school year intervention

Goals	Reduce television to less than 2 hours a day (this was emphasised) Increase moderate and vigorous physical activity. Decrease consumption of high fat foods. Consumption of fruit and vegetables 5+ a day.
Lessons	There were 16 lessons per year (32 total). Sessions introduced in 4 major subjects (language, maths, arts and social studies), as well as physical education. Lessons involved 'demonstrations, debates, case studies, group projects, games, and student presentations'. Lessons could last for one or two 45-minute periods.
Physical Activity	Physical activity lessons were goal based. Fitness-Funds—monetary incentives of $400-$600 for proposals at intervention schools that fitted with Planet Health themes. Lessons were organized into thirty 5-minute micro units.
Television Reduction	Two week campaign to reduce television viewing (Power Down).

4.2 Quality of trial

Recruitment

One problem in the recruitment process was the low participation rate. Only 65% of students at the schools were involved in the study at baseline. The authors suggest the low participation rate may have been partly because five of the schools required 'active consent', where the parent had to give signed permission for their child to be involved in the trial.

Control group

The randomisation of the trial appeared to be sound. To 'balance factors that could influence study outcomes', before randomisation schools were divided into pairs according to location or demographics. One school was picked as a control and the other as an intervention, for each of the five pairs.

Evaluation method

Assessments:

There were only two assessments made, one at baseline (fall 1995) and the other at post intervention (spring 1997). The follow up period was 2 years which was longer than many school-based trials. The trial would have been improved if the follow-up had been significantly longer. This is because many of the health benefits from lifestyle changes only occur in the long term.

Analysis:

Due to the cluster randomised design a generalised estimating equation method was used for the analysis to adjust for individual level covariates. The software took into account the intraclass correlation of responses within schools and school matching in the study design.

Regressions were performed separately for boys and girls. It is unclear whether the study specified a priori that it would perform this subgroup analysis and whether the study had sufficient power for subgroup analysis. There is a possibility of bias if this decision was made after identifying the non-significance of pooled results for girls and boys. However, there does seem to be a good biological reason for looking at the sexes separately (such as different patterns of growth and differences in maturational tempo).

Regressions were controlled for ethnic category, intervention status, and baseline parameters (with a number of baseline variables tested to see if these added significantly to the regression equations). Analyses were conducted on an intention to treat basis.

Outcome measures

The study classified students as obese or non-obese according to a composite measure of TSF and BMI. This had the advantage that it was possible to monitor whether the intervention was benefiting an at risk group. This is important, as it would be possible for the mean BMI/TSF to fall between baseline and follow-up, largely due to reductions in BMI and TSF amongst people who were not significantly overweight (or even amongst people who were *underweight* at baseline).

Dietary intake and physical activity were based on self-reports. As the authors suggest, it is surprising that when television viewing, fruit and vegetable intake, and total energy intake were controlled, there was only a marginal change in the OR that indicated an intervention effect (Gortmaker *et al.*, 1999 pg. 417). This would seem to suggest that either these variables were not measured accurately or other variables were responsible for changes due to the intervention.

Obesity was defined as when BMI and triceps skin fold were greater than the 85 percentile (gender-age specific). However, this measure does not seem ideal, as it is based on population statistics rather than on the affects of a high body fat content on a person's health (e.g. life expectancy). This is particularly since the anthropometric characteristics of the population are going to change over time, meaning that the BMI/TSF composite value that would classify a person as obese would also change over time. It would seem a superior measure of obesity would define it to reflect a BMI/TSF value that significantly increased a person's morbidity and mortality rate adjusted for age, gender and race specific.

Bias, confounders, efficacy

The sample size was small, with randomisation only occurring amongst ten schools. This increased the 'risk of an unbalanced design and the clustering of outcome observations amongst schools.' However, the 'generalised estimating equation, and SUDDAAN estimating approaches' were used to account for clustering, and the 'cluster randomisation procedure produced comparable intervention and control groups (Gortmaker *et al.*, 1999 pg 417).'

The authors make the point that 'while the analysis of obesity incidence and remission indicates statistically significant intervention effects only for remission, it is important to note that over this 21-month period only 33 incident cases occurred-13 among intervention girls-which limited the statistical power of our study to detect differences. Our findings of significant effect on remission of obesity indicates that intervention effects may be largest among those most at risk for obesity (Gortmaker *et al.*, 1999 pg. 416).

80% percent of class room teachers (n=86), class room teacher lesson evaluations (N=230) showed that teachers completed on average 3.5 lessons a year. The article states that 'teachers were expected to do a minimum of 4 lessons and power down could be one of these (pg. 413). 100% of PE teachers (n=9) completed training, however, the program implementation from PE teachers seemed to be low, with PE teachers taking an average of 8.2 micro units during per year. PE teachers were meant to teach 30 micro units in a year, meaning that PE teachers were conducting less than 30% of the set lessons.

The use of subgroup analyses for girls and boys is of particular concern if this was not specified a priori and if the study was not adequately powered for these analyses.

4.3 Outcomes – as reported

Outcome measures

Anthropometric:

Measurements were made at base-line (fall 1995) and follow-up (spring 1997). Obesity was the primary endpoint and defined as 'a composite indicator based on both BMI and a TSF value greater than or equal to age- and sex- specific 85th percentiles (Gortmaker *et al.*, 1999 pg. 409).' The composite measure was to accommodate for weaknesses in using either BMI or TSF alone. One weakness of BMI as a measure of body fat percentage is that a healthy BMI varies according to sex, race, maturation, and frame size. While a weakness of the TSF is that the results become more unreliable among the most overweight individuals.

BMI was defined as weight in kilograms divided by height in metres squared. Height was measured without shoes to the nearest 0.1 cm using a Shorr stadiometer. Weight was measured in light weight clothing (to the nearest 0.1 Kg) using electronic scales (calibrated using the Seca standard weights set-up test). Holtain callipers were used to measure TSF to the nearest 0.2 mm. Two TSF measurements were made, and if the results differed by more than 2mm then a third measurement was taken, and the average used. Among girls, self-reports on menarche status were used to adjust BMI and TSF measurements.

Food and activity:

Television viewing, physical activity, dietary intake, sociodemographic and behavioural variables were all measured using an optically scannable Food and Activity Survey, completed under teacher supervision (students had an hour long training from teachers).

An 11 item measure was used to gain an estimate of total hours television viewing per day. The measure accounted for television viewing, video use, and video game use over a week period.

A Youth Activity Questionnaire was used to estimate the amount of moderate and vigorous physical activity undertaken (≥ 3.5 metabolic equivalents, walking not included) over the past month.

The Youth Food Frequency Questionnaire measured dietary intake, and excluded French-fried potatoes as vegetables. Also implausible energy intakes were ignored (≤2100 or ≥29 000 Kj; less than 1% of observations).

The students' age, gender and ethnicity (self reported) were also collected. Self reports were conducted to assess whether students were involved in any 'weight loss behaviour; including dieting to lose weight, exercising to lose weight, vomiting or taking laxatives to lose weight, or taking pills to lose weight.

Statistical methods:

'Because schools, not students, were randomised, the generalized estimating equation method for analysis of dichotomous outcomes was used to adjust for individual covariates under cluster randomisation, with schools nested within experimental conditions, using software written for use with SAS data sets (SAS Inc, Cary, NC) (Gortmaker *et al.*, 1999 pg. 412).This generalized estimating equation method 'took into account interclass correlation responses within schools. The generalized estimating equation analysis also took into account the school matching in design, including indicator variables for randomisation pairs' (Gortmaker *et al.*, 1999 pg. 412).

Behaviour change

Table 4.3 summarises the changes in behaviour between baseline and follow up:

- After adjusting for baseline covariates, the intervention group reduced television viewing significantly more than the control between baseline and follow up. This result held for both females (adjusted difference -0.58 hours; P=.001) and males (adjusted difference –0.4 hours; P<.001).
- There was a statistically significant increases in fruit and vegetable intake (0.32 servings/d; 95% CI, 0.14-0.50 servings/d; P = .003) and a smaller increase in estimated intake (-575 J; 95% CI, -1155 to 0 J/d; P = .05), among intervention girls compared to the control.
- 'Reductions in television viewing predicted obesity change and mediated the intervention effect. There was a negative relationship between hours of television viewing and obesity amongst girls (OR, 0.85; 95% CI, 0.75-0.97; P = .02). The intervention effect was then only marginally statistically significant (P =.08).' (Gortmaker *et al.*, 1999 pg. 415)
- Other behavioural variables, such as fruit and vegetable intake and energy intake could not explain the intervention effect among girls.
- 'Among girls who were obese at baseline, when we controlled for baseline television viewing and change in television viewing, each hour of reduction in television viewing was independently associated with increased remission of obesity (OR, 1.92; 95% CI, 1.37-2.70; P = .002) and the estimate of intervention effect was reduced (adjusted OR 2.4-1.6) and was not statistically significant (P =.17).' (Gortmaker 1999, pg. 415)

Table 4.3 Estimating differences in change in behavioural variables from baseline to follow-up for children in Planet Health Intervention vs Control Schools, 1995 to 1997*

Measure	Sample	Baseline†	Follow-up†	Crude Change‡	Adjusted Difference‡	95 % Confidence Interval	P
Girls							
Total T.V./Video/h/d							
Control	304	3.1	2.99	-0.11	-0.58	-0.85 to –0.31	.001
Intervention	289	2.98	2.28	-0.7			
Mod/vig/activ/h/d							
Control	304	1.67	1.74	0.07	0.36	-0.63 to 1.35	.43
Intervention	291	1.76	1.87	0.11			
Total energy from fat, %							
Control	285	31.0	29.8	-1.2	-0.67	-1.43 to 0.09	.07
Intervention	282	31.2	29.4	-1.8			
Fruit and vegetables, servings/d							
Control	284	4.1	3.9	-0.2	+0.32	0.14 to 0.50	.003
Intervention	280	3.4	3.6	+0.2			
Total energy intake, J/d							
Control	285	8122.8	9009	+886.2	-575.4	-1155 to 0	.05
Intervention	282	7526.4	8156.4	+630			
Boys							
Total T.V./Vid/h/d							
Control	319	3.78	3.43	-0.35	-0.40	-0.56 to –0.24	.0003
Intervention	313	3.73	3.03	-0.70			
Mod/vig/activ/h/d							
Control	319	2.47	2.44	-0.03	-0.40	-1.00 to 0.20	.16
Intervention	314	2.54	2.44	-0.10			
Total energy from fat, %							
Control	296	31.5	30.5	-1.0	-0.31	-1.1o to 0.48	.38
Intervention	296	32.0	30.5	-1.5			
Fruit and vegetables, servings/d							
Control	296	4.1	3.6	-0.5	0.18	-0.21 to 0.56	.31
Intervention	297	3.8	3.6	-0.2			
Total energy intake, J/d							
Control	296	9445.8	10147.2	+701.4	-466	-1094 to 164	.13
Intervention	298	9361.8	9815.4	+453.6			

*Restricted to cohort students with paired data. Sample sizes vary due to missing data.
†Baseline and follow-up values are the unadjusted means.
‡Adjusted difference represents the difference in change in scores in the intervention group compared with the control group, after adjustments for baseline value of the dependent variable. Intervention status, randomisation pairs, ethnicity, and baseline measures of obesity, triceps skinfolds, and BMI. Regression estimates were calculated using SUDAAN software to account for cluster randomisation.
Source: (Gortmaker *et al.*, 1999 pg. 416)

Clinical parameters

Table 4.4 and Table 4.5 summarise the changes to clinical parameters between baseline and follow-up:

- After controlling for baseline covariates, there was a significantly higher decrease in the prevalence of obesity amongst girls in the intervention group compared to girls in the control group (odds ratio [OR], 0.47; 95% CI 0.24-0.93; P = .03). The prevalence of obesity between baseline and follow up increased for the control group (21.5% to 23.7) but fell in the intervention group (23.6% to 20.3%).

- After controlling for baseline covariates, there was no significant change in the prevalence of obesity among male students in the intervention group compared to male students in the control group. The prevalence of obesity fell in both groups between baseline and follow up.

There was a significantly greater obesity remission among intervention girls compared to the control (OR, 2.16; 95% CI, 1.07-4.35; P=.04), but no difference for boys. At a school level analysis, four of the five paired schools showed a positive intervention effect in reducing female obesity (though the sample is very small).

The intervention caused a large decrease in obesity amongst African American girls. While the reduction in obesity amongst white girls, was similar to the overall result. The intervention had a small effect on reducing obesity amongst Hispanic girls. There was a similar amount of extreme dieting in control and intervention schools at baseline and follow up.

Table 4.4 Unadjusted changes in obesity prevalence, incidence and remission, for control and intervention groups between baseline and follow-up

Measure	Baseline	Follow-Up	Baseline	Follow-Up
	Female Obesity		**Male Obesity**	
Prevalence Obesity				
Control	68	75	117	107
Intervention	73	63	97	92
Incidence Obesity				
Control	-	20	-	21
Intervention	-	13	-	18
Remission Obesity				
Control	-	13	-	31
Intervention	-	23	-	23

*Note Figures were obtained by using the data in Table 4.5 below, which was replicated from Gortmaker *et al.*, 1999 pg. 415

Table 4.5 Estimated change in obesity from baseline to follow-up for children in Planet Health Intervention vs Control Schools, 1995 to 1997*

Measure	Sample	Baseline % †	Follow-up% †	Crude Change %	Crude Odds **	Adjusted Odds ‡	95% Confidence Interval	P
Female Obesity								
Prevalence %								
Control	317	21.5	23.7	+2.2	1.00	1	0.24-0.93	.03
Intervention	310	23.6	20.3	-3.3	0.59	0.47		
Incidence %								
Control	249	-	8	+8	1.00	1.00	0.23-2.38	.57
Intervention	237	-	5.5	+5.5	0.66	0.77		
Remission %								
Control	68	-	19.1	-19.1	1.00	1	1.07-4.35	.04
Intervention	73	-	31.5	-31.5	2.00	2.16		
Male Obesity								
Prevalence %								
Control	337	34.7	31.8	-2.3	1.00	1.00	0.52-1.39	
Intervention	331	29.3	27.8	-1.5	0.97	0.85		.48
Incidence %								
Control	220	-	9.6	+9.6	1.0	1.00	0.71-1.75	.58
Intervention	234	-	7.7	+7.7	0.79	1.12		
Remission %								
Control	117	-	26.5	-26.5	1.00	1.00	0.44-4.24	.54
Intervention	97	-	23.7	-23.7	0.86	1.37		

*Restricted to cohort students with paired data.
†Baseline and follow-up values are the unadjusted percentages.
‡Adjusted odds control for baseline obesity as well as other baseline covariates, including intervention status, age, ethnicity, indications for randomisation pairs, and baseline measures of triceps skin folds and body mass index. For girls, a variable indicating that the student reported exercising to lose weight at baseline was added to the regression. Regression estimates were calculated using the generalized estimating equation method to account for cluster randomisation.
Source: (Gortmaker *et al.*, 1999 pg. 415)

Mortality

Mortality was not reported by the Gortmaker et al (1999) study.

Morbidity

Morbidity was not reported by the Gortmaker et al (1999) study.

4.4 Program costs

As reported

The costs of the program are reported in an economic analysis paper by Wang et al (2003) and are summarised in Table 4.6.

Table 4.6 Summary of the costs of the 2 year intervention

Item	Quantity	Unit cost	Total cost in 1996 US dollars
Training workshop			
Trainer	1 day/year for each of 5 schools	Annual salary $38,000	$1462
Assistant trainer	1 day/year for each of 5 schools	Annual salary $29,000	$1115
Teacher reimbursement			
Subject teachers	3 hours training for 101 teachers in each of 2 years	$25/hour	$15,150
PE teachers	5 hours training for 9 teachers in 1st year, and 3 hours in 2nd year	$25/hour	$1800
Food	110 teachers each year	$10 per person	$2200
Teacher wellness activities			
Trainer	6x 1 hour session for each of 5 schools	$30/hour	$900
Fitness funds	5 schools	$500/school/year	$5000
Planet Health book	1 copy for each of 110 teachers	$55/book	$6050
Total			US$33,677

(Source Wang et al 2003, page 1317) Values provided by Harvard Prevention Research Centre

If this cost is attributed to the 310 girls in the intervention group then this would cost US$108.64 per person, alternatively if the cost is attributed to the total 641 intervention students then the cost per person would be US$52.54. This evaluation then goes on to model the cost effectiveness of this intervention over a 53 year period. The model is built around the key outcome of BMI and is only constructed for girls as the intervention was dominated for boys based on BMI. The key modelling inputs and assumptions are as follows:

- 20.3% in intervention group are overweight following study verus 25.8% in control group (the intervention group has seen a 21% reduction in overweight assuming a baseline rate of 25.8%)
- The effect of the intervention is assumed to be maintained
- 75.4% of overweight 14 year old girls are overweight by age 21-29
- 9.8% of normal weight 14 year old girls become overweight by age 21-19
- 91.2% of overweight women aged 21-29 are overweight aged 40
- 39.3% of normal weight women aged 21-29 are overweight aged 40
- the QALYs saved per case of overweight prevented is 0.712 with healthy life years only obtained from age 40 to 65
- It is assumed that weight remains unchanged between age 14 and 21 and also unchanged between age 29 and 40
- Discounting of costs and outcomes at 3%
- The 25 year excess medical costs per overweight women is US$2737

- The entire cost of the intervention (US$33,677) is attributed to the 310 girls included in the evaluation
- The final weight of a person between ages 40 and 65 is the only determinant of healthy life years and medical costs (weight at age 14 and ages 21-29 is irrelevant)

The result of this modelling is a cost per QALY of US$4305 for girls (range from sensitivity analysis US$4065 to US$4525) with the model being most sensitive to the medical care costs averted and the discount rate. It should be remembered that the intervention group was dominated by the control group (more expensive and worse outcomes) for boys based on BMI.

The main limitations of this model related to the availability of data linking outcomes in children to adult weight, mortality and quality of life. Little is known about the maintenance of behaviour change in children and the assumption that their weight would be maintained is optimistic.

Based on resource use

Intervention group:

Based on the description of resources used in the study by Gortmaker et al (1999) costs were also estimated by our research team and were assumed to generally fall into the categories of training and research, teaching, and equipment and consumables. Estimates of costs for the intervention group include the opportunity cost of teacher time and are summarised in Table 4.7, Table 4.8 and Table 4.9.

Table 4.7 Costs of training and research for the intervention group

	Cost	Time	Number required	Total cost for study	Average cost per person
Training materials	$3.80	-	95 classroom teachers	$361.00	$0.56
Training time-intervention staff	$53.85 per hour	4 hours	3 sessions at each of 5 schools	$3,231.00	$5.04
Training time-teachers	$47.23 per hour	4 hours	95 classroom teachers	$17,947.40	$28.00

Table 4.8 Costs of teaching (including opportunity cost of teaching time for the intervention group

	Cost	Time	Number required	Number teachers	Total cost for study	Average cost per person
Teaching time-classroom	$47.23 per hour	45 minutes	3.5 lessons	99	$12,273.90	$19.15
Teaching time- PE	$47.23 per hour	5 minutes	8.2 microlessons	9	$290.46	$0.45
Developing lesson units	$47.23 per hour	2 hours	-	10	$944.60	$1.47
Financial incentives to PE	$500	-	5 grants	-	$2,500.00	$3.90
Cost of lesson materials	$8.32	-	3.5 lessons	99	$2,882.88	$4.50

Table 4.9 Costs of equipment and consumable for the intervention group

	Cost per Unit	Number required	Number of students	Total cost for study	Average cost per person
Stadiometer	$2,234.34	1	-	$2,234.34	$1.73
Digital scale	$831.45	1	-	$831.45	$0.64
Skinfold calipers	$496.52	1	-	$496.52	$0.38

Control group:

The control group are assumed to not incur any costs.

Total costs:

The total costs for each group and the average cost per person are summarised in Table 4.10.

Table 4.10 Summary of costs for each group

	INTERVENTION GROUP		CONTROL GROUP	
	Total cost for study	**Average cost per person**	**Total cost for study**	**Average cost per person**
Training and research	$21,539.40	$33.60	$0.00	$0.00
Teaching	$18,891.84	$29.47	$0.00	$0.00
Equipment and consumables	$3,562.31	$5.56	$0.00	$0.00
Total	$43,993.55	$68.63	$0.00	$0.00

4.5 Performance

Cost effectiveness

The main outcome reported by the Gortmaker et al (1999) trial was prevalence of obesity.

Table 4.11 summarises the change in obesity prevalence over the study period. The adjusted change figure takes into account baseline covariates. There was a statistically significant decrease in obesity prevalence for girls only at intervention schools compared to control schools.

Table 4.11 Estimated change in obesity prevalence from baseline to follow up for boys and girls at control and intervention schools

Obesity prevalence	Sample	Baseline (%)	Follow up (%)	Crude change (%)	Crude odds	Adjusted odds (95%CI)	P value
Girls- control	317	68 (21.5)	75 (23.7)	7 (2.2)	1.00	0.47 (0.24 to 0.93)	0.03
Girls- intervention	310	73 (23.6)	63 (20.3)	-6 (-3.3)	0.59		
Boys-control	337	117 (34.7)	107 (31.8)	-10 (-2.3)	1.00	0.85 (.52 to -1.39)	0.48
Boys- intervention	331	97 (29.3)	92 (27.8)	-5 (-1.5)	0.97		

Using the costs estimated in section 4.4 the following costs per case of obesity prevented were obtained

ICER= (cost intervention/ cost control)/ (outcome intervention/ outcome control)
= ($43,993.55-$0.00)/ (7--6)
=$3,384.12 per additional obesity case prevented for girls for the intervention group compared to the control group over the 2 year

For boys the intervention is more costly than control and yields no statistically significant benefit, and is therefore said to be dominated.

The only behaviour change outcome that was statistically significant for both boys and girls was the reduction in hours of television viewed (Table 4.12). The adjusted change figure takes into account baseline covariates. There was a statistically significant reduction in viewing for both boys and girls in the intervention schools compared to the control schools.

Table 4.12 Estimated change in television viewing from baseline to follow up for boys and girls at control and intervention schools

Hours of television viewing	Sample	Baseline (hours)	Follow up (hours)	Crude change	Adjusted difference (95%CI)	P value
Girls- control	304	3.10	2.99	-0.11	-0.58 (-0.85 to -0.31)	0.001
Girls-intervention	289	2.98	2.28	-0.70		
Boys-control	319	3.78	3.43	-0.35	-0.40 (-0.56 to -0.24)	0.0003
Boys- intervention	313	3.73	3.03	-0.70		

Using the costs estimated in section 6.4.2 the following costs per hour of television reduction were obtained:

ICER= (cost intervention/ cost control)/ (outcome intervention/ outcome control)
= ($68.63-$0.00)/ (0.58-0) for girls
= ($68.63-$0.00)/(0.4-0) for boys
=$118.33 per hour reduction per person in television viewing for girls in the intervention group compared to girls in the control group over the 2 year study period
=$171.56 per hour reduction per person in television viewing for boys in the intervention group compared to boys in the control group over the 2 year study period

Cost-utility

Neither quality of life, utility, nor mortality were reported in the study by Gortmaker et al (1999), therefore it is not possible to calculate a QALY measure based on trial results.

4.6 Modelling

The model presented by Wang et al (2003) has been replicated based on the published report. This has enabled the following further sensitivity analyses and alternative assumptions to be modelled with results presented in Table 4.14:

- Exclusion of downstream costs (medical care costs)
- Addition of a 50% relapse rate (of those who changed from overweight to normal weight during intervention) in the 7 years following the trial.
- Discount rate 5%
- Cost per person of $52.54 (attributing cost to intervention girls and boys)
- Inclusion of both girls and boys results and costs in model (see table 4.14)

Table 4.13 Combined results of girls and boys

Total sample	% overweight or obese		
	Baseline	Follow up	Change
Control group	26.5%	26.0%	-0.5
Intervention group	26.5%	24.2%	-2.3%
Total			-1.8%

Table 4.14 Results of further modelling (alternative assumptions and inputs)

Assumptions/inputs	Resulting cost/QALY
Base case scenario presented in paper	**US$4305**
Exclusion of downstream costs	US$8148
Relapse rate of 50% (50% of those who became normal weight, relapse to be overweight by year 7)	US$12,451
Discount rate 5%	US$8858
Cost of intervention US$52.54	US$97
Boys and girls results combined and cost for boys and girls (US$52.54 per person)	US$8197

A multivariate sensitivity analysis was also calculated incorporating all of the following assumptions: results of boys and girls combined, exclusion of downstream costs, cost of Australian $68.63 per person (calculated based on described resource use), 5% discount rate and relapse rate of 50%. The results of this analysis are likely to be most comparable to other evaluations performed for this report. The resulting cost per QALY saved was $50,091 for the intervention group compared to control group.

4.7 Discussion

The main advantage of this trial, over many other school-based studies, was that it measured the number of obese students at baseline and follow up. This meant an assessment could be made as to whether obese, or potentially obese, students benefited from the trial.

However, the trial had a reasonably small sample size (10 schools) and a short follow up period (of less than 2 years). Furthermore, between baseline and follow up, there was not a significant reduction in the prevalence of obesity among males in the intervention group, when compared to the control group.

The model is structured in such a way that mortality and quality of life benefits are gained for those aged 40-65. This means that the intervention is not cost effective if it is assumed that all of the intervention group revert to baseline values in the few years following the intervention. In reality it is possible that the majority of those loosing weight during the intervention will regain it and that the intervention will therefore not be cost effective long term.

5. Effects of a school-based intervention to reduce cardiovascular disease risk factors in elementary school children: Cardiovascular Health in Children (CHIC) Study

5.1 Description

Reference

The evaluation in this section is based on a study by (Harrell *et al.*, 1996). The study aimed to assess the effects of a 'classroom-based intervention to reduce cardiovascular disease risk factors in elementary school children.' Twelve schools participated in the randomised control trial, with intervention schools receiving an '8 week exercise program and 8 weeks of classroom education (on smoking, physical activity and nutrition).' Anthropometric and behavioural variables were measured at baseline and follow-up.

Recruitment: target population and participants

Twelve schools in Northern Carolina were invited into the study. Northern Carolina is in the top quintile of states for deaths from cardio vascular disease or strokes. To join the study the schools had to clearly be in a rural or urban area. Schools were categorised as rural, if located in counties with no cities with populations greater than 50,000 and in towns with populations less than 2,500. Schools were categorised as urban, if located in a city of at least 50,000 people. Schools were also geographically classified, according to whether they were located in the western mountain region, the eastern costal plain or the central piedmont. This geographic stratification occurred because there was a lower heart disease mortality rate in the western mount region than the other two geographic regions. While 108 schools agreed to participate in the study, only 33 met the criteria of being either clearly rural or urban. For each of the three geographic regions, two urban schools and two rural schools were randomly selected (from the 33 schools able to join the study).

At a subject level, 'inclusion criteria were as follows: assignment to the third or fourth grade regardless of age; ability to read and write English; no mental, emotional, or physical handicap identified by parents or teachers; no chronic illness such as diabetes or moderate to severe asthma reported by parents, teachers, or child; ability to participate in an exercise program; and at least one relative available to respond to a questionnaire about family history {Harrell, 1996}. '

Child and parental consent was required, with 58.1% of third and fourth students in the schools participating. There were 1274 children in the sample, aged 7 to 11 years— 616 (48.4%) boys and 658 (51.6%) girls. Data on socio economic status was obtained from 80% of mothers and 96% of fathers. Approximately a third of the sample was from each of the three geographic regions (32.3% costal plains; 35.4% central piedmont; 32.3% western mountains). The percentage of students in rural and urban areas was also similar (49.4% rural; 50.6% urban).

Table 5.1 shows the demographic characteristics of the intervention and control group, with no significant differences at baseline. 'Further, there was no significant differences in race and gender distribution between our subjects and the total population of third and fourth graders at the study schools (Harrell *et al.*, 1996).'

Table 5.1 Demographic characteristics of the study groups at baseline

	Group		
Demographic Variables	Intervention (n=588)	Control (n=686)	Chi-square P
Gender			
Male	48%	49%	0.883
Female	52%	51%	
Race			
Black	20%	21%	0.179
White	76%	73%	
Other	4%	6%	
Grade			
Third	53%	50%	0.145
Fourth	47%	50%	
Parental Education			
Less than high school graduate	16%	19%	0.145
High school graduate, some college	55%	57%	
College graduate+	29%	24%	
Age (yr)*			
8	35%	31%	0.364
9	46%	48%	
10	17%	19%	
11	2%	2%	

*One child in the intervention group was age 7. This child was not included in the percentages.

Source: (Harrell *et al.*, 1996)

Intervention

With each of the six pairs of schools (stratified according to geographic region, and urban/rural classification) one school was randomly assigned to the control group and the other school to the intervention group. Control schools received their normal class room health instruction. As with the intervention group, parents in the control received the results of their child's physical tests, four weeks after each testing period.

Regular classroom teachers at intervention schools gave lessons to third and fourth graders twice a week, for eight weeks, using the American Heart Association Lower and Upper Elementary School Site Program Kits. The kits included information about 'selecting *heart healthy* foods, the importance of getting regular physical exercise, the dangers of smoking, and ways to combat pressure to smoke.' Regular classroom teachers had a training session where they learnt about the program from an 'experienced elementary educator'. The elementary educator was also available to teachers for consultation and guidance and visited a class taught by each teacher to confirm that the program was taught correctly.

Children in the intervention group also received physical activity lessons from physical education teachers, 3 times a week, for eight weeks (total 24 lesson plans). The lessons were especially designed by one of the study team, and 'included a brief warm-up, 20 minutes of various fun, non-competitive aerobic activities designed to work the major muscle groups, and a cool-down period.' Aerobic activities offered included: 'jumping rope to music, "endless relay", parachute and other small-group games, and aerobic dance.'

Table 5.2 Summary of the 8 week intervention

Classroom Lessons	Lessons twice a week. Topics included: 'heart healthy foods', the importance of physical exercise, the dangers of smoking, and how to resist pressure to smoke.
Physical Activity Lessons	Three times a week. Fun aerobic lessons, with warm up and cool down. Activities included: 'jumping rope to music, "endless relay", parachute and other small-group games, and aerobic dance.'

5.2 Quality of trial

Recruitment

- The study recruited schools that could be defined as either clearly rural or urban. The authors wrote that 'initially 108 schools agreed to participate; study schools were randomly selected from 33 of those that met study criteria as clearly rural or urban (Harrell *et al.*, 1996).' Given that less than a third of the schools who agreed to participate were classified as either rural or urban, the study may be non-representative of all schools in Northern Carolina. Different rates of heart disease mortality were a justification as to why the schools were stratified according to geographic region.
- The low participation rate (with only 58.1% of the third and fourth graders at the study schools participating) may have biased the sample. While the authors claim 'there were no significant differences in race and gender distribution between our subjects and the total population of third and fourth graders (Harrell *et al.*, 1996)', it is possible that sample was non-representative with respect to other variables, such as the socio economic status of the student.

Control group

A control and intervention school was randomly chosen for each of the six strata blocks.

Evaluation method

Assessment:

- While the physiological effects of the 8 week intervention were limited (and nonexistent using school level analysis) even two weeks after the intervention ended, it would be important to know whether the intervention effect continued into the long term.
- The results could have been influenced seasonally, as the data was not collected simultaneously at each school. However, the authors argue that 'data were collected within a short time span for all groups within each of the three geographic regions. Thus any potential seasonal effect would be similar within each geographic region, which was a stratification factor, and therefore intervention group comparisons would be internally valid (Harrell *et al.*, 1996)'

Analysis:

Analyses were conducted at both the school level and the individual level. School level analysis is more conservative although may be lacking in statistical power. Regression was performed accounting for the stratified cluster randomised design. For individual level analysis multivariate ANOVA was used and analyses of covariance to adjust for design, demographics and baseline values. The study does not state whether analyses were performed on an intention to treat basis, although some children appear to be missing from some analyses.

Outcome measures

- The study does not nominate a primary outcome measure of interest.
- The authors make the comment that: 'a limitation of the study is the method used to measure cholesterol and the number of measures made…studies have indicated that results obtained with the Reflotron device are reliable…however, other studies have shown a bias, sometimes positive, and sometimes negative…We would have preferred to used venous blood samples obtained in the fasting state…however, study size, the school-based data collection method, and the use of multiple settings across the entire state made it more practical to use a single finger stick measure. In additional, after talking with many school principles, we were convinced that requiring a venipuncture would drastically reduce participation in the study (Harrell *et al.*, 1996).' The potential unreliability of the cholesterol measure brings into question the accuracy of the reported statistically significant difference between the intervention and control for cholesterol (individual level analysis).

- Children were not classified into 'high risk groups', such as obese—only the mean value of each group was used for each outcome measure (such as BMI). This means it is not possible to determine whether any physiological changes occurred amongst those most at risk.
- There did not appear to be any external validation of the self-reported physical activity measure.

Bias, confounders, efficacy

- Because the students were only required to complete the *Healthy Heart Knowledge* test at posttest (and not at baseline), it is not possible to conclude that the intervention caused the intervention group to achieve higher scores than the control group. It is possible that the intervention group already had greater health knowledge at baseline.
- The study was classroom based, therefore reducing the chance that children with a particular high risk of CVD would be stigmatised.
- The randomised sample was reasonably small, with there only being 12 schools involved in the trial. It does not appear that sample size calculations were performed prior to commencement of the study. It is possible that the study did not have sufficient power to detect a difference between groups (particularly in the analysis at school level).

5.3 Outcomes – as reported

Outcome measures

'The variables studied are consistent with the Burhn and Parcel Development of Positive Health Behaviour model, which considers family influences and child development/personal characteristics as factors in predicting health behaviours and health status outcomes (Harrell *et al.*, 1996).' Baseline and posttest data were both collected following a systematic pattern. The parental questionnaire was mailed. Posttest data were measured 2 weeks after the 8 week intervention program ended. Information from children was obtained at school. The following data were collected:

Family influences:

- Highest education level of parents
- Family history (parents or grandparents) of any of the following CVD conditions: 'angina, angioplasty, heart attack, coronary bypass surgery, high blood pressure, or stroke and if these occurred before age 55 in men and 60 in women.'

Behaviour/knowledge:

- Children were asked in a questionnaire whether they were currently smoking and whether they had ever smoked a whole cigarette (children were told their answers were confidential).
- The Know Your Own Body Health Habits Survey, modified for grades 3 and 4, was used to assess the physical activity habits of the students.
- A health knowledge questionnaire (based on the 'Heart Smart' test, and only conducted at posttest) involved 25 multiple choice questions to test the children's knowledge on a variety of topics: nutrition (9 questions), exercise (8 questions), smoking (5 questions), and general heart health (3 questions).

Physiological outcomes:

- 'Blood pressure (BP) was measured on the right arm by a research assistant who used a Baumanometer mercury sphygmomanometer (W.A. Baum Co. Inc) and appropriately sized cuffs according to American Heart Association recommendations for children. The mean systolic and diastolic BP was used as the final outcome measure, and was calculated from two measurements of the first, fourth and firth Korotkoff phases—with the fourth phase used for diastolic BP.
- A Reflotron (Boehringer Mannheim Diagnostics, Indianapolis, Ind.) device was used to measure total serum cholesterol, in a non-fasting state.

- Aerobic power was tested using a children's version of the 'Eurofit submaximal cycle ergometry test using a Body guard Professional Cycle Ergometer (Monark, Varberg, Sweden).' The heart rate during the test was measured using a 'Polar Pacer heart rate monitor' (Polar CIC, Inc., Washington, N.Y.).
- Lange skinfold calipers (Cambridge Scientific Industries, Cambridge, Md.) were used to measure subscapular and triceps skinfolds on the right side of the body. 'The skinfold sites were located and measured at least twice according to National Health and Nutrition Examination Survey procedures, and mean measures at each site were summed for analysis.' The natural logarithm of each skinfolds sum was used, because the distribution of the skinfolds was skewed.
- Height and weight were measured, with children wearing clothes but no shoes, to calculate the BMI (Kg/m2). A stadiometer (Perspective Enterprise, Inc., Portage, Mich) was used to measure height to the nearest cm. A balance beam scale (Detecto Scales, Inc., Jericho, N.Y.) was used to measure weight to the nearest 0.1kg

Outcomes

Multivariate ANOVA showed that there was a significant association (P<0.001) between the four demographic variables (age, gender, race and parental education) and the outcome variables. 'After adjusting for those demographics and for the study design, the two groups were significantly different (F_8, 1186 = 24.81, P<0.001) when all dependent variables were examined simultaneously. All subsequent individual analyses adjusted for these demographics.

Behavioural/knowledge

Table 5.3 shows the unadjusted outcome measures for the control and intervention groups.

Table 5.3 Behaviour and knowledge: baseline and posttest intervention Means and standard errors (unadjusted)

	Intervention Group (n=588)		Control Group (n=686)	
Measure	**Mean**	**SE**	**Mean**	**SE**
Physical Activity Score				
Baseline	64.52	1.27	62.57	1.08
Posttest	66.71	1.23	62.32	1.11
Absolute change	1.89	1.46	-0.76	1.25
Percent change	22.66	3.24	15.09	2.89
Knowledge (% Correct)	67.27	0.65	58.79	0.57
Exercise	51.45	0.85	41.59	0.69
Nutrition	72.32	0.84	66.06	0.78

Source: (Harrell *et al.*, 1996)

School level analysis:

The school level analysis showed that mean self reported physical activity scores increased significantly more in the intervention group between baseline and posttest (3.73; 95% CI 0.37, 7.08). The posttest, Healthy Heart Knowledge test score was also 7.86% higher in the intervention schools compared to the control schools (95% CI 3.89, 11.83). Current smoking prevalence at baseline was 0.1% in intervention schools and 1.1% in control schools and the proportion who had ever smoked was 3.5% and 4.3% in intervention and control schools respectively.

Individual level analysis:

The individual level analysis mean self reported physical activity scores increased significantly more in the intervention group between baseline and posttest (3.9; 95%CI -0.2 to 8.01). The individual level analysis also shows that the Healthy Heart Knowledge test scores were higher in the intervention group than the control (8.37%; 95% CI 6.36, 10.37).

Physiological

Table 5.4 shows the unadjusted outcome measures for the control and intervention groups.

Table 5.4 Physical variables: baseline and posttest intervention Means and standard errors (unadjusted)

	Intervention Group (n=588)		Control Group (n=686)	
Measure	**Mean**	**SE**	**Mean**	**SE**
Cholesterol mg/dl (mmol/L*)				
Baseline	168.28 (4.36)	1.26 (0.03)	164.99 (4.27)	1.07 (0.03)
Posttest	161.32 (4.18)	1.21 (0.03)	163.58 (4.24)	1.18 (0.03)
Absolute Change	-6.79 (-0.18)	0.85 (0.02)	-1.40 (-0.03)	0.78 (0.02)
Percent change	-3.20	0.52	-0.39	0.47
Systolic Blood Pressure (mm Hg)				
Baseline	103.43	0.41	104.06	0.38
Posttest	107.35	0.37	108.24	0.38
Absolute Change	3.96	0.38	4.18	0.35
Percent Change	4.32	0.38	4.46	0.3
Diastolic Blood Pressure (mm Hg)				
Baseline	68.24	0.35	67.88	0.33
Posttest	72.88	0.39	73.66	0.34
Absolute Change	4.63	0.44	5.74	0.39
Percent Change	8.06	0.72	9.72	0.62
Skin folds (mm)				
Baseline	25.58	0.57	26.14	0.55
Posttest	24.63	0.57	26.43	0.59
Absolute change	-0.90	0.21	0.25	0.17
Percent change	-2.91	0.67	1.14	0.57
Body mass index (Kg/m^2)				
Baseline	18.26	0.14	18.47	0.14
Posttest	18.50	0.14	18.70	0.14
Absolute change	0.24	0.03	0.18	0.04
Percent change	1.40	0.17	1.15	0.21
Height (cm)				
Baseline	135.49	0.29	136.28	0.28
Posttest	137.8	0.30	138.18	0.29
Absolute change	2.23	0.05	1.85	0.03
Percent change	4.76	0.17	3.91	0.22
Weight (kg)				
Baseline	33.90	0.36	34.70	0.35
Posttest	35.51	0.37	36.12	0.37
Absolute change	1.60	0.07	1.32	0.07
Percent change	4.76	0.17	3.91	0.22
PVO_2 (ml/kg/min)				
Baseline	42.59	0.40	41.58	0.39
Posttest	45.22	0.39	42.89	0.41
Absolute change	2.66	0.29	1.34	0.25
Percent change	8.27	0.81	4.44	0.65

Source: (Harrell *et al.*, 1996), * converted to mmol/L using following formula: 1 mg/dL= 0.0259 mmol/L

School level analysis:

At a school level analysis, there were no statistically significant differences between the intervention and control group for changes in physiological measures from baseline to posttest. However, there was a noticeably (but non-statistically significant) larger decrease in cholesterol and a smaller increase in diastolic blood pressure (after controlling for baseline and demographic data) in the intervention group from baseline to posttest.

Individual level analysis:

Individual level analysis shows the intervention group reduced cholesterol significantly more than the control group from baseline to posttest (–4.88 mg/dl, 95% CI –7.65, -2.11). There was also a significant difference between the two groups for the baseline to posttest change in the natural log of the sum of triceps and subscapular skin folds (-0.05, 95% CI –0.07, -0.02). The intervention group also increased predicted aerobic power significantly more than the control (1.73, 95 CI 0.8, 2.66).

Table 5.5 Mean differences of intervention group versus control group* with 95% confidence intervals

	School Level Analysis		Individual Level Analysis	
Measure	**Mean Difference**	**95% CI**	**Mean Difference**	**95% CI**
Cholesterol (mg/dl) Cholesterol (mmol/L) §	-5.27 -0.14	-2.11, 1.57 -0.05, 0.04	-4.88 -0.13	-7.65, -2.11‡ -0.20, -0.05
Systolic Blood Pressure (mm Hg)	-0.65	-2.10, 0.79	-0.50	-1.66, 0.65
Diastolic Blood Pressure (mm Hg)	-1.32	-3.65, 1.00	-0.95	-2.25, 0.34
Skin Folds	-0.04	-0.11, 0.03	-0.05	-0.07, -0.02‡
BMI (kg/m^2)	0.05	-0.07, 0.18	0.11	-0.02, 0.24
Predicted Aerobic Power (ml/Kg/min)	1.76	-0.70, 4.22	1.73	0.80, 2.66‡

*Regression models controlled for demographics (gender, race, grade, parental education) and baseline level (except for knowledge).
† 95% C.I. from covariate ANOVAs (changes from baseline to posttest) and ANOVA (knowledge total). ‡95% CI excludes zero.
§ converted to mmol/L using following formula: 1 mg/dL= 0.0259 mmol/L
Changes in natural logarithm transformations of sum of triceps and subscapular skin folds.
Source: (Harrell *et al.*, 1996)

Mortality

Death was not reported in the study by Harrell et al (1996).

Morbidity

Morbidity was not reported in the study by Harrell et al (1996).

5.4 Program costs

As reported by the trial

The study by Harrell et al (1996) does not report costs associated with the intervention to reduce cardiovascular risk factors in school children.

Based on resource use

Intervention group: Costs for the school based intervention to reduce cardiovascular disease risk factors included research and training, teaching (including opportunity cost of teaching time), assessment and consumables, and equipment (Table 5.6, Table 5.7, Table 5.8 and Table 5.9). The average cost per person is based on the 588 students from the intervention group (from the total sample of 1274 students)

Table 5.6 Research and training costs associated with the intervention

	Cost	Time	Number required	Total cost for intervention	Average cost per person
Training sessions-teachers	$47.23 per hour	1 day (8 hours)	6 sessions (one per school) for 7.84 teachers	$17,773.59	$30.23
Training sessions–educators	$61.45 per hour	1 day (8 hours)	6 sessions (one per school)	$2,949.60	$5.02
School site visits	$61.45 per hour	1 day (8 hours)	12 sessions (two per school)	$5,899.20	$10.03
Research educator	$73,740 per year	1 year	1	$73,740.00	$125.41

Table 5.7 Teaching costs associated with the intervention

	Cost	Time	Number required	Total cost for intervention	Average cost per person
School site program kit	$2,000		1 per school (total of 6 schools)	$12,000.00	$20.41
Teaching time (opportunity cost)	$47.23 per hour	45 minutes per class	16 lessons conducted by 23.5 teachers	$13,318.86	$22.65
Physical education sessions (opportunity cost)	$47.23 per hour	45 minutes per class	24 lessons conducted by 23.5 teachers	$19,978.29	$33.98
Development of PE lesson plans	$40.45 per hour	4 weeks		$6,472.00	$11.01
Teaching resources	$8.32 per lesson		16 lessons conducted by 23.5 teachers	$3,128.32	$5.32

Table 5.8 Costs of assessment and consumables for the intervention group

	Cost	Number required	Number of students/ parents	Total cost for intervention	Average cost per person
Parent report at baseline and follow up- consumables	$1.27 per unit	2 per parent	588	$1,493.52	$2.54
Parent report at baseline and follow up- time to prepare	$27.07 per hour	2 per parent, 30 minutes to prepare each	588	$15,919.76	$27.07
Prize for children undergoing testing	$0.87 per unit	1 per student	588	$511.56	$0.87

Table 5.9 Equipment costs associated with the intervention

	Cost	Number required	Total cost for intervention	Average cost per person
Sphygmomanometer	$90.57	1	$90.57	$0.15
Reflotron cholesterol testing	$12,297.92	1	$12,297.92	$20.91
Ergometer	$1,943.55	1	$1,943.55	$3.31
Heart rate monitor	$250.92	1	$250.92	$0.43
Skinfold calipers	$496.52	1	$496.52	$0.84
Stadiometer	$1,180.38	1	$1,180.38	$2.01
Scales	$551.57	1	$551.57	$0.94

Control group:

The control group are assumed to incur the costs of the parent report at baseline and follow up.

Total costs:

The summary of total costs for the intervention and control groups and the average costs per person are presented in Table 5.10.

Table 5.10 Summary of costs for the intervention

	INTERVENTION GROUP		CONTROL GROUP	
	Total cost for group	**Average cost per person**	**Total cost for group**	**Average cost per person**
Research and training	$100,362.39	$170.68	$0.00	$0.00
Teaching	$54,897.47	$93.36	$0.00	$0.00
Assessment and consumables	$17,924.84	$30.48	$17,413.28	$29.61
Equipment	$16,811.43	$28.59	$0.00	$0.00
Total	$189,996.13	$323.12	$17,413.28	$29.61

5.5 Performance

Cost effectiveness

Physical activity score:

As was described in Section 5.3 the mean self reported physical activity score significantly increased by 3.73 more points in the intervention group compared to control between baseline and post test (according to school level analysis). This leads to the following ICER:

ICER= costs intervention-cost control/ outcomes intervention-outcomes control

ICER= ($323.12-$29.61)/ 3.73

=$293.51/ 3.73

= $78.69 per point reduction in self reported physical activity score for the intervention group compared to the control group between baseline and post test

Cholesterol level:

Section 5.3 reported that there was a mean difference in the cholesterol level of 0.14 mmol/L for the intervention and control groups between baseline and follow up (according to school level analysis). The intervention group had lower cholesterol levels. This leads to the following ICER:

ICER= costs intervention-cost control/ outcomes intervention-outcomes control

ICER= ($323.12-$29.61)/ 0.14

=$293.51/ 0.14

= $2 096.50 per point reduction in cholesterol level for the intervention group compared to the control group between baseline and post test

BMI:

Section 5.3 reported that there was a mean difference in BMI of 0.05 kg/m2 for the intervention and control groups between baseline and follow up (according to school level analysis). The control group had a lower BMI than the intervention group and was less expensive so therefore dominates the intervention group.

Cost-utility

The study by Harrell et al (1996) did not report quality of life or utility measures, therefore cost-utility was unable to be calculated based on the trial results.

5.6 Modelling

Modelling is not performed for this intervention due to the intervention group being dominated by the control group for the key outcome BMI. The control group had lower BMI scores and less costs.

5.7 Discussion

At a school level (the level where randomisation occurred), there were no statistically significant differences between the intervention and the control for any of the physiological variables. However, the results of the individual analysis showed there were some significant differences between the two groups. The study was weakened by the short follow up period, the relatively small sample size, and there being no categorisation of students into CVD risk groups.

6. Cardiovascular disease risk reduction for tenth graders: a multiple-factor school-based approach

6.1 Description

Reference

The evaluation in this chapter is based on a study by (Killen *et al.*, 1988). The study aimed to assess the affects of a school-based intervention designed to reduce the incidence of cardiovascular disease. Four schools were involved in the randomised trial—two control schools and two treatment schools. Treatment schools received 'a special 20 session risk reduction intervention.'

Recruitment: target population and participants

All tenth graders enrolled in four Northern Californian high schools (N=1447) were involved in the trial. Within each of the two school districts, one school was randomly assigned as the treatment school and the other as the control school. Before randomisation, 'schools were matched for size and distribution of ethnic groups.' The demographic profile of the four schools (N=1477) is displayed in Table 6.1 below.

Table 6.1 Demographic profile of the four schools (N=1477)*

Age	
15 Years	70%
14 Years	14%
16 Years	14%
Ethnic Group	
White	69.0%
Black	2.0%
Asian	13.1%
Hispanic	6.4%
American Indian	0.3%
Pacific Islander	0.4%
Other	8.9%
Fathers who Completed Four or More Years Of College	50%

*Percentages do not add up because of rounding.

The control and treatment groups were similar at baseline with respect to ethnicity (P =.17), planning to enrol in college (P =.16) and gender (treatment group: 55.5% boys; 44.5% girls; control group: 52.5% boys and 47.5% girls [P =.26]). Parents in the control group had more years of education (χ^2 = 15.8; P<.008).

At baseline, the self-reported behaviour measures of the two groups were similar with respect to cigarette and alcohol consumption, food choice and knowledge. Boys in the two groups also had similar self-reported physical activity scores (treatment group, 30.9%; control group, 32.9% [P=.60]). However, more girls in the control group reported regular aerobic physical activity (treatment group, 35.8%; control group, 53.9% [P = .0001]). 'Regular exercises had significantly lower mean resting heart rates than non-regular exercises (regular exercises, 76.1 beats per minute; non regular exercises, 78.4 beats per minute [t=2.9; P<.003]).'

Dropouts in the treatment group had a lower mean tricep skin-fold thickness (14.7 mm) compared to controls (17.2 mm) [P = 0.02]. There did not appear to be any other significant associations between dropout rates and knowledge or self-reported measures.

Intervention

Special intervention sessions were incorporated into the regular physical education curriculum. All tenth graders at treatment schools received the special intervention sessions three times a week, for seven weeks. There were twenty, 50 minute classroom sessions, 'divided among five program modules: Physical Activity, Nutrition, Cigarette Smoking, Stress, and Personal Problem Solving. Bandura's social-cognitive theory served as a guide in the development of the intervention program. Each module provided students with (a) information on the effects of different health practices designed to increase the attractiveness of healthful lifestyles (b) cognitive and behavioural skills enabling them to change personal behaviour (c) additional specific skills for resisting social influences to adopt or readopt unhealthy habits, and (d) specific practice in using skills to improve performance. As part of the sessions devoted to problem-solving training, each student was asked to carry out a self-change project (Killen *et al.*, 1988).'

Eight 'special full time-teachers', from the Stanford Center for Research in Disease Prevention, taught the lessons. The teachers were aged in their twenties, and 'had previous training in health studies and/or previous experience in healthcare/health research settings.' Each school also had a coordinator/backup teacher who monitored and implemented the program.

6.2 Quality of trial

Recruitment

The recruitment process appeared sound. The researchers did not appear to require parental consent to conduct the physiological measures, and 'all tenth graders in four senior high schools (N = 1477) from two school districts participated in a cardiovascular disease reduction trial.'

Randomisation may not have been adequate with differences reported between groups at baseline for education of parents, body mass index, body fat, heart rates, blood pressure and exercise. These differences indicate that the groups may have differed in ways potentially affecting the trial outcomes.

Control group

To try to achieve a similar control and treatment group, before randomisation schools were divided into pairs according to ethnic makeup, school size and district. It is possible that not all control schools received the same treatment.

Evaluation method

Outcome measures: Many of the outcome measures were objective and measured according to protocols which lessons the potential for bias. Some outcomes such as the skin fold test are known to be associated with an error margin which will influence the clinical importance of any differences observed.

It is unclear if outcome assessors were blinded to group allocation. The self-report measures are subject to bias particularly if participants are not blinded to group allocation. This is a particular issue with this trial as the self-reported outcomes are those with the greatest statistical difference between groups (see results).

One of the outcome measures was food choice pairs where individuals chose which they would be more likely to consume. This instrument is of questionable face validity as a method of evaluating consumption of unhealthy foods.

Follow-up period: Data was only collected over a four-month period. Given that many of the health effects from life style changes only occur in the long term, extending the data collection period would have greatly enhanced the study. Of particular concern is that only subjects with results at both baseline and follow up were included in the analysis. This is a potential source of bias as those who dropped out were known to differ in terms of skin fold measurement results.

Analysis: 'To examine the equivalence of the treatment and control groups at baseline, a one-way analysis of covariance was conducted with continuous variables and X^2 tests were conducted with categorical variables. To examine program effects, a two-way (treatment times sex) analysis of covariance was conducted with continuous variables. Baseline values were used as covariates. X^2 tests were conducted with categorical variables. Analyses were conducted using the individual as the unit of analysis (Killen *et al.*, 1988).'

The analysis was not conducted on an intention to treat basis, with those dropping out omitted from the analysis. The article states that: 'analyses were conducted using the individual as the unit of analysis (Killen *et al.*, 1988).' It is not stated if analyses were appropriately adjusted to account for the cluster randomised study design. It is unclear if the trial was designed to have adequate power to detect statistical differences between the groups based on a cluster randomised design.

Outcome measures

- Students were not categorised into high and low CVD risk groups. This meant it was not possible to determine whether physiological changes occurred amongst those most at risk of CVD. However, students were categorised according to the number of cigarettes they smoked or the amount of exercise they undertook.
- An effort was made to ensure that the self-reported data was reliable. There was a correlation between expired-air carbon monoxide levels and self-reported cigarette consumption; and between self-reported physical activity and resting heart rate.

Bias, confounders, efficacy

The limitations of the trial include uncertain adjustment for cluster randomised design, lack of blinding, different characteristics of groups at baseline and uncertainty as to whether subjects were treated equally in all ways other than the intervention. The length of follow up was possibly too short to observe important treatment affects and drop outs were omitted from the analysis.

6.3 Outcomes – as reported

Outcome measures

Trained staff (no regular school staff were involved) collected data at baseline and two months after the seven-week intervention (follow-up). At each school, assessment occurred over a two-day period, with girls and boys separated into different rooms. Assessment involved both physical measures and self-administered questionnaires. Only students for which there were both baseline and follow-up measurements were included in the results. 'Of the 1447 students responding to the baseline survey, 1130 (78%) were available at the follow up' (treatment n=622; control n=508). The following variables were measured:

Demographic variables:

- Parents Education-'the higher of the mother or fathers education level'.
- College Plans-the student's intention of attending college was measured on a 5-point scale.
- Ethnicity-self-reported by students.

Knowledge of cardiovascular disease risk concepts:

- A multi-choice test assessed the children's knowledge of different areas of the curriculum (maximum possible scores: physical activity, 30; nutrition, 30; and smoking, 8).

Self-reported behaviour:

- Physical Activity- Students were asked to indicate which activities, out of a list of 19, they undertook for 20 minutes non-stop, and the frequency they preformed the activities. Students were classified as an 'aerobic exerciser' (or regular exerciser) if the student undertook one of 5 selected aerobic activities (though not necessarily the same activity each time), for at least 20 minutes non-stop, at least 3 times a week.
- Nutrition/Diet—32 food pairs, one superior with respect to reducing CVD, were presented to students. Students said which food they would usually chose to eat if given the choice.
- Smoking/Drug Use—Students were asked the frequency of their smoking, alcohol, and marijuana use (response rate between 90% and 92% for each of the substances). Smokers were classified as: '(a) those who had never smoked, (b) experimental smokers (those smoking on a monthly basis or less often), and (c) regular smokers (those smoking weekly or more often).'Note: reduction of alcohol and marijuana use was not targeted in the intervention, but was recorded to compare the control and treatment groups. The accuracy of the self-reported smoking was tested against a carbon monoxide measurement, and 'has been found to increase the accuracy of self-reporting of drug use. 'Expired-air carbon monoxide levels correlated (r =.44) with reported daily or almost daily cigarette smoking.'

Anthropometric/physiological variables:

- Body Mass Index (BMI)—Weight divided by height squared. Weight and height were 'measured on a standard balance beam scale. Students wore lightweight gym clothing with overgarments and shoes removed.'
- Subcutaneous skin fold thickness—measured using callipers on the right side of the body at the triceps and subscapular muscles.
- Resting Heart Rate and Blood Pressure—Students sat quietly for three minutes before measurements were made with an automated blood pressure device. Mean arterial diastolic and systolic blood pressures and heart rate were measured three times at one-minute intervals (from the right arm at heart level). The means of the second and third measurements were used.

Behaviour change

Baseline and follow-up measures for boys and girls are shown in Table 6.2.

Change in knowledge:

There was an increase in knowledge in the treatment group compared to the control. 'In the treatment group boys increased their combined score by an average of 11.1 points and girls by an average of 14.2 points. By contrast, in the control group, boys scores decreased an average of 1.4 points, and girls scores increased an average of only 0.8 points (Killen *et al.*, 1988).' For each of the three areas in the knowledge test, the treatment group obtained significantly higher scores than the control: nutrition diet (main effect, F[1, 946] = 369.2 [P<.0001]; sex effect F[1,946] = 27.8 (P<.0001), physical activity (main effect, F[1, 1078] = 371.8 [P<.0001]; sex effect, F[1,1078] = 33.8 [P<.0001]), cigarette smoking (main effect, F[1, 965] = 177.2 [P<.0001]; sex effect, F[1, 965] = 10.9 [P=.001]).

Change in self reported activities:

Thirty percent of non-regular exercisers in the treatment group, at baseline, became regular exercises by follow-up, compared to 20% in the control ($\chi^2(1) = 8.6$; P<.0003).[4] There was also a significant difference between the two groups with respect to the number of healthy foods chosen in the food pairs test—with the number of healthy foods chosen by males and females increasing in the treatment group and decreasing in the control group (main effect, F[1, 850] =56.6 (p<.0001); sex effect, F[1,850] = 10.4 [P=.001]).

[4] However, there was no data showing the number of regular exercises at baseline no longer exercising regularly at follow-up.

Cigarette smoking:

There were no significant differences between the treatment and control group with respect to: (i) the proportion of people who had never smoked at baseline taking up smoking by follow-up (treatment group, 9.7%; control group 14.5% [P =.25]) (ii) the proportion of regular smokers giving up smoking by follow-up (treatment group, 3.5%; control group, 9.3% [P =.39]). However, 'in the treatment group, more of those students who, at baseline, were experimental smokers reported quitting at follow-up (treatment group, 28.5%; control group 17.6%).' In addition, 5.6% of experimental smokers (at baseline) in the treatment group graduated to regular smoking, compared to 10.3% in the control. 'The overall χ^2 for the analysis examining change in status of baseline experimental smokers was significant (X^2 (2) = 9.4; P = .009).'

Table 6.2 Mean values (+/- SD) for measures of CVD* risk at baseline and at two-month follow-up

	Boys				Girls				
	Treatment Group		Control Group		Treatment Group		Control Group		
Measure of CVD Risk	**Baseline**	**Follow-Up**	**Baseline**	**Follow-Up**	**Baseline**	**Follow-Up**	**Baseline**	**Follow-Up**	**Treatment Group vs Control Group**
Exercise Score	12.8 (5.3)	17.4 (6.8)	13.0 (5.0)	11.4 (5.9)	13.6 (4.4)	19.6 (5.1)	13.9 (5.2)	13.9 (5.5)	.0001
Nutrition Score	6.4 (4.6)	11.3 (6.7)	6.5 (4.9)	6.0 (4.6)	7.4 (4.3)	14.5 (6.2)	8.2 (4.8)	8.2 (4.9)	.0001
Smoking Score	3.1 (1.3)	4.6 (2.1)	3.2 (1.5)	3.3 (1.5)	3.1 (1.3)	5.0 (1.7)	3.2 (1.4)	3.6 (1.7)	.0001
Food choice	10.7 (5.4)	12.8 (6.5)	11.4 (5.9)	10.9 (5.1)	13.2 (5.5)	15.6 (6.1)	13.6 (5.4)	12.7 (5.1)	.0001
Body mass index	21.6 (3.5)	21.7 (3.6)	20.9 (2.7)	21.3 (2.7)	22.1 (3.9)	21.9 (3.8)	21.4 (3.0)	21.4 (3.1)	.05
Heart rate, beats/min	75.2 (12.2)	72.9 (11.3)	75.9 (11.4)	76.3 (11.7)	82.7 (12.9)	78.6 (11.4)	78.2 (11.3)	78.6 (10.6)	.0001
Triceps skin fold thickness, mm	11.3 (5.1)	11.2 (5.3)	11.2 (4.9)	10.68 (4.8)	20.4 (6.5)	20.0 (6.3)	18.8 (5.2)	20.3 (5.6)	.004
Subscapular skin fold thickness, mm	9.7 (4.3)	9.6 (4.6)	9.3 (4.1)	9.1 (3.5)	13.9 (6.0)	13.4 (5.6)	12.1 (4.7)	13.0 (4.7)	.01
Systolic blood pressure, mm Hg	119.0 (12.1)	123.0 (12.0)	122.2 (12.6)	124.1 (12.8)	116.0 (11.0)	114.2 (11.3)	113.4 (9.6)	113.7 (9.5)	.84
Diastolic blood pressure, mm Hg	58.5 (9.1)	59.5 (8.9)	59.5 (8.5)	59.7 (8.3)	60.6 (7.5)	60.1 (9.2)	59.2 (7.7)	57.2 (8.3)	.009

*CVD indicates cardiovascular disease.

Source: (Killen *et al.*, 1988)

Clinical parameters

- In the treatment group, the resting heart rate decreased an average of 2.3 and 4.1 beats per minute for boys and girls, respectively. In comparison, the resting heart rate increased by 0.4 beats per minutes for boys and girls in the control group (main effect F[1, 1065]=19.9 [P<.0001]; sex effect F[1, 1065]=5.8 [p<0.02]).
- 'Reductions in body fatness were also achieved, though the impact of the program was only consistent for girls.' The BMI actually increased for boys in the treatment group, but less so than in the control group (main effect F[1,1060] =3.7 [P=0.5]; sex effect F[1,1060] = 21.9 [P<.0001]).

- There were significant beneficial effects from the intervention with respect to triceps skin fold thickness (main effect F[1, 1059] = 8.4 [P=.004]; sex effect, F[1,1059] = 80.9 [P<.0001]; sex times treatment effects, F[1,1059] = 36.3 [P<0.0001])
- There were also significant beneficial effects from the intervention with respect to subscapular skin fold thickness (main effect F[1,1058] = 6.4 [P=0.01]; sex effect F[1,1058] = 37.5 [P<.0001]; and sex times treatment effect F[1, 1058] = 15.9 [P=.0001]).
- There were no significant changes between the two groups for systolic blood pressure, however, there were changes in diastolic blood pressure that favoured the control group (F [1, 1065] = 6.7 [P=.009]). However, there is significant variation in blood pressure readings at this age.

It is important for these clinical parameters to consider the clinical significance of these changes as well as the statistical significance, taking into account generally observed measurement error margins.

Mortality

The study by Killen et al (1988) did not report mortality.

Morbidity

The study by Killen et al (1988) did not report mortality.

6.4 Program costs

As reported by the trial

The study by Killen et al (1988) did not report any costs associated with the school based intervention to reduce risk of cardiovascular disease in tenth graders.

Based on resource use

Intervention group:

Costs for the intervention group consisted of intervention costs and equipment costs (Table 6.3, Table 6.4).

Table 6.3 Intervention costs

	Unit cost	Number required	Length of time	Cost for group	Average cost per person
Development of program	$61.45 per hour	2 staff	8 weeks	$39,328.00	$63.23
Program materials	$8.32	622		$5,175.04	$8.32
Teachers time	$47.23 per hour	8 teachers	20x 50 minute sessions	$6,297.33	$10.12
Coordinators time	$61.45 per hour	2 staff	20x 50 minutes sessions	$2,048.33	$3.29

Table 6.4 Equipment costs for the intervention group

	Unit cost	Cost for study	Average cost per person
Breath carbon monoxide monitor	$573.10	$573.10	$0.40
Disposable mouth pieces	$0.28 for 1447 people	$405.16	$0.28
Balance beam scale	$551.57	$551.57	$0.38
Skin fold calipers	$496.52	$496.52	$0.34
Blood pressure monitor	$90.57	$90.57	$0.06
Heart rate monitor	$250.92	$250.92	$0.17

Control group:

The control group are assumed to not incur any costs.

Total costs:

The total costs for the intervention and control groups along with the average costs per person are shown in Table 6.5.

Table 6.5 Total costs

	INTERVENTION GROUP		CONTROL GROUP	
	Cost for group	Average cost per person	Cost for group	Average cost per person
Intervention	$52,848.71	$84.97	$0.00	$0.00
Equipment	$1,183.92	$1.64	$0.00	$0.00
Total	$54,032.63	$86.60	$0.00	$0.00

6.5 Performance

Cost effectiveness

Costs have been summarised in Table 6.5.

The study reported that 30% of non regular exercisers at baseline in the intervention group became regular exercisers by follow up compared to 20% in the control group.

The incremental cost effectiveness ratio is therefore calculated as follows:
ICER= costs intervention-cost control/ outcomes intervention- outcomes control
= ($86.60-$0.00)/ (0.3-0.2)
= $86.60/ 0.1
= $866 per additional non regular exerciser at baseline who became an exerciser for the intervention group compared to the control group over the study period

The study also reported BMI split for boys and girls at baseline and follow up as summarised in Table 6.6.

Table 6.6 Summary of change in BMI between baseline and follow up for boys and girls

	Boys				Girls			
	Treatment Group		Control Group		Treatment Group		Control Group	
Measure of CVD Risk	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up
Body mass index	21.6 (3.5)	21.7 (3.6)	20.9 (2.7)	21.3 (2.7)	22.1 (3.9)	21.9 (3.8)	21.4 (3.0)	21.4 (3.1)
Change in BMI		+0.1		+0.4		-0.2		0

This leads to the following ICERs for boys and girls:
ICER= costs intervention-cost control/ outcomes intervention- outcomes control
= ($86.60-$0.00)/ (0.4-0.1) for boys
= ($86.60-$0.00)/ (0--0.2) for girls

= $288.67 per BMI point reduction for boys between baseline and follow up for the intervention group compared to the control group
= $433 per BMI point reduction for girls between baseline and follow up for the intervention group compared to the control group

6.6 Modelling – hypothetical scenario analysis

In order to model this intervention using the replicated Wang et al (2003) model the proportion of participants overweight/obese at baseline and follow up would be required for each study group. This information is not provided. We therefore, provide hypothetical scenario analysis based on the replicated Wang et al (2003) model. The following assumptions are made:

- The proportion in population who are overweight or obese at age 15, assuming a boy to girl ratio of 1.05 is 20% (AIHW, 2003)
- Cost of $86.60 for those in intervention group
- 75.4% of those overweight at 14 years are overweight by age 21-29
- 9.8% of normal weight 14 year olds become overweight by age 21-19
- 91.2% of overweight people aged 21-29 are overweight aged 40
- 39.3% of normal weight people aged 21-29 are overweight aged 40
- the QALYs saved per case of overweight prevented is 0.712 with healthy life years only obtained from age 40 to 65
- It is assumed that weight remains unchanged between age 14 and 21 and also unchanged between age 29 and 40
- Discounting of costs and outcomes at 5%
- Exclusion of downstream costs
- 50% of those becoming a normal weight relapse within 7 years to become overweight again
- The final weight of a person between ages 40 and 65 is the only determinant of healthy life years and medical costs (weight at age 14 and ages 21-29 is irrelevant)

Table 6.7 shows estimates for various hypothetical scenarios of the proportion reduction in the number of youth of average age 15 years who are overweight or obese. This reduction is assumed to occur for a hypothetical intervention group compared to a control group that do not change weight.

Table 6.7 Cost per QALY for hypothetical scenarios of reductions in overweight/obese children

% reduction in overweight/obese	Proportion overweight/obese at end of intervention	Resulting cost/QALY No relapse	Resulting cost/QALY 50% relapse
5%	0.19	$73,514	$147,678
10%	0.18	$36,757	$72,356
15%	0.17	$24,505	$48,814
20%	0.16	$18,379	$37,053

Inclusion of downstream costs:

The 25 year excess medical costs per overweight person is US$2737. This is converted to Australian exchange rates using 1996 figures and inflated to Australian 2003 values which give a cost per person of $4562. When this cost is included in the hypothetical scenario analysis, the estimates are as presented in Table 6.8.

Table 6.8 Cost per QALY for various hypothetical scenarios including downstream costs of obesity

% reduction in overweight/obese	Proportion overweight/obese at end of intervention	Resulting cost/QALY No relapse	Resulting cost/QALY 50% relapse
5%	0.19	$60,329	$134,493
10%	0.18	$23,572	$59,171
15%	0.17	$11,320	$35,629
20%	0.16	$5,194	$23,869

These scenarios are only intended to provide a guide as to what cost/QALYs may be likely if certain results were to be obtained.

6.7 Discussion

The replicated Wang et al (2003) model is structured in such a way that mortality and quality of life benefits are gained for those aged 40-65. This means that the intervention is not cost effective if it is assumed that all of the intervention group revert to baseline values in the few years following the intervention. In reality it is possible that the majority of those losing weight during the intervention will regain it and also that there will be further gains/changes made before age 40.

7. Workplace intervention for overweight males: 'GutBusters'

7.1 Description

Intervention type

This analysis is based on a weight loss program known as GutBusters. It was available to the target population as a commercial operation through Weightwatchers until 2002, at which time it was closed in favour of a program developed in the USA, and for which few details have been identified. GutBusters was started by Dr Garry Egger in Australia in 1991 with the NSW Health Department and was the first large-scale program in the world to use waist measurement, and not weight, as an indication of fat loss in men. GutBusters was a 'no gimmicks' program designed to reduce waist size by 1% per week. In doing so, it aimed to fit in with, and not drastically alter, men's lifestyles. Hence, it did not require participants to give up alcohol, start dieting or exercise vigorously.

Recruitment and target population

Results of two studies of the GutBusters Program have been reported in a single publication by Egger et al. (1996). Study 1 comprised 51 retirees who had all completed the standard course. The second study followed men for 1 year after having completed the initial 6 week programme (n = 83), or the initial course plus an additional six fortnightly "advanced" sessions (n = 37). The standard program is evaluated here as this has been the most commonly applied program in Australia, Additionally, the advanced version of the course focuses upon waist / hip ratios (WHR) whereas results from the standard course study still report weight loss and BMI. Reductions in WHR are known to be beneficial to health outcomes, but the correlation between different WHR ratios and the health outcomes of interest (life-years and QALYs) have not yet been quantified, and thus do not lend themselves to modelling. Therefore, only the results of the standard GutBusters program are evaluated here.

The key characteristics of Study Group 1 participants were:

- Mean age (range) : 55 years
- Mean weight (range) : 95 kilograms
- Mean BMI : 31.5

These men were followed-up by telephone at approximately 6-monthly intervals for 2 years at which time 42 were still contactable.

Intervention

The GutBuster program formed part of a collaboration between Deakin University, the Victorian Food and Nutrition Program, and the National Heart Foundation. It was first trialled in 1992 and was subsequently established nationally. It remained a weight loss program targeting overweight males. It specifically targeted abdominal weight as measured by the waist-hip ratio (WHR), with objectives set in terms of centimetres to be lost rather than kilograms. An initial program goal was normally set at a 5% reduction in girth to be lost at the rate of 1% per week. Participants were encouraged to weigh themselves to reinforce their achievements, but this was not the primary goal of the program.

The standard course was run over six weeks and comprised a 1½ hour session each week. Sessions were held in the workplace or in the nearest suitable location (gymnasiums, hospitals, clubs). Conducted under the guidance of a trained leader, small groups were encouraged provide mutual support (social support is an important aspect of the program). The health risks associated with WHRs greater than 0.9 were explained.

At the first session, a computerised assessment questionnaire was completed to assist education in relation to key influences of body fat status, namely: motivation, stage of behaviour, genetic influences, metabolism, dietary fat behaviour, fluid consumption, habitual exercise, preferred exercise and diet, preferred weight loss method and medical conditions. The responses to the questionnaire enabled tailored reports to be returned. Each participant also received a tape for waist measurement and a standard pair of publications:

- *The GutBuster Weight Loss Guide*, 1993 Egger G, Stanton S.
- *Fat & Fibre Counter* Stanton S.

The education and behaviour modification component of the program aimed to fit in with existing lifestyles using the following principles:

1. Modifying habits; eating differently (less fat, more fibre)
2. Moving more (organised and incidental exercise)
3. Trade-offs (movement for food and drinks)

The program promoted walking as a means of exercising with an initial distance goal of 2-4 kilometres every day.

7.2 Quality of evidence

Evaluation description

Design: Egger et al, 1996 report a more 'naturalistic' study, rather than a randomised controlled trial. Participants had been recruited from doctors' referrals and were thus self-selected at the time of enrolment rather than randomly selected members of the general population of overweight male workers (refer also comments in Section below: Bias, confounders, efficacy). However, selection of those for follow-up was related to whether or not participants were still contactable which is unlikely to further bias results. A total of 51 subjects were identified as having completed the 6 week GutBuster course. Of these, 46 could be identified two years later, although 2 had died and 2 had moved out of the area, leaving 42 subjects contributing data to the publication.

The study design was a longitudinal study with no control group. Although this weakens the scientific basis of the evidence compared to randomised controlled trial designs, the evidence is still considered very strong with subjects acting as their own 'controls'. This conclusion is strengthened by the fact that weight gain is typically characterised by slow incremental gain over time whereas, for the study by Egger et al. (1996), reported outcomes were compared to baseline measures taken some 2 years previously.

Methodology: Differences in measurement of waist, hips and weight (from which BMI was calculated) comprise self-reported data. *T*-tests for paired samples were used to test for significance in differences between measurements at baseline and two-years.

Outcome measures: A range of outcomes including body mass index (BMI), responders, average weight loss and WHR were reported. There is sound evidence in the literature that WHR is a more reliable measure of clinical benefit than weight loss alone, however, WHR as a measure is difficult to use in modelling of long-term prognosis due to limited long-term data in the literature. Dietary fat, exercise and alcohol intake were also recorded in Study 2 (only) through the use of questionnaires.

Assessment

Bias, confounders, efficacy: In the study by Egger et al. (1996), the men were followed-up by telephone at approximately 6-monthly intervals for 2 years (n=42). They had not previously been advised that they would be contacted, thus avoiding any bias in results by prompting renewed effort at waist reduction and / or weight loss.

Nevertheless, the results are based upon self-reported data. As with all surveys of weight, overweight and obese subjects tend to understate weight. The magnitude of any such under-estimation cannot be estimated. It is also perhaps relevant to note that good correlations between WHR changes and BMI changes were found by Egger et al., thus providing internal consistency of results at least.

An uncertainty of the analysis is how representative a group of males who choose to participate in the GutBuster Program are of a general population of overweight males of the same age. If GutBuster participants are a special sub-group of motivated males, the generalisability of the results is reduced accordingly.

The survey group comprised compliant men by virtue of having completed at least 5 of the 6 sessions to be eligible for inclusion in the study. The results therefore do not reflect an intention to treat cohort. Of relevance to this evaluation is that Egger et al. (1996) reported in their discussion that 86% of men enrolled in the GutBusters course have completed at least 5 of the 6 sessions. Assuming that the results of the survey are otherwise unbiased, the results reported will therefore apply to at least 86% of all enrolled men, and an "intention-to-treat" evaluation can be performed.

7.3 Outcomes – as reported

After Study 1, the focus on the importance of reducing the WHR was increased resulting in failure to record weight reductions at all in Study 2. For the reasons given, the results of Study 2 are not used in this analysis. It is noted, the WHR reductions in Study 2 were at least comparable to those reported for Study 1.

Behaviour change and clinical parameters

Participation: take-up and attendance: By 2001, there were over 70,000 men who had completed the program in Australia. Egger et al. (1996) reported participants to have been effective in reducing waist size by at least 7% and maintaining or increasing these losses in 70% of men over a 1 year period.

Clinical parameters: Summary results from the standard GutBusters course (Study 1 in Egger et al., 1996) were:

- Mean BMI decreased from 31.5 to 28.9 (p<0.001)
- Average weight loss was 5.27kg.
- 64% of the survey group had greater waist loss after 2 years than that achieved after the 6 week course
- 94% maintained some weight loss
- There were no statistically significant changes in waist-to-hip ratios due to reductions in hip size as well as waist.

Mortality

Although Egger et al., 1996 reported that two of the 51 people who had been chosen for follow-up had died, the cause of death was unknown and no comparison with predicted deaths was made. More importantly, analysis of mortality as an outcome would have been inappropriate as the study was under-powered statistically for evaluation of this outcome.

Morbidity

Morbidity is not reported by Egger et al. (1996). However, insights to the incidence and severity of morbidities associated with DM Type 2 can be seen from the pattern of health care utilisation (Table 7.5). Additionally, the impact of these morbidities has been captured by the DiabCost study which used the EuroQol 5-D to assess the utility (quality of life) of patients with and without DM Type 2 (Table 7.1). The DiabCost study (Colagiuri et al., 2003) is a joint publication of the Australian Diabetes Society and the Australian Diabetes Educators Association.

Table 7.1 Quality of Life (Utility Values); DiabCost Study; (Colagiuri et al 2003)

Age	General Population	Complications			
		None	Microvascular	Macrovascular	Both
All ages		0.79	0.69	0.65	0.65
36-50	0.88	0.83	0.71	0.84	0.72
51-65	0.84	0.80	0.71	0.64	0.65
66+	0.79	0.78	0.67	0.64	0.59

7.4 Program costs

As reported

No program costs were reported in Egger et al. (1996).

Based on resource use

GutBuster participants in Study 1 completed the standard 6 week course (Egger er al, 1996). This course is essentially the same as Professor Trim operated through LifeChoice Pty Ltd and which is advertised at an all inclusive cost of $299 per person (refer http://bne111v.server-secure.com/vs81335_secure/fax_order.html). This fee includes recruitment costs, the cost of educational materials provided, and the cost of session leaders (session leaders are first given 40 hours of training). No government or private insurance company subsidy is received such that no 'shadow pricing' is necessary for this evaluation.

7.5 Cost-effectiveness

Data from the trial results alone are first used to analyse economic performance. A successful outcome is attributed to changed behaviour arising from the GutBusters program. The base case analysis provided first therefore only captures health benefits in the two years following the program, and does not consider savings from any reduction in health care utilisation ('downstream costs').

Program costs

The standard cost of GutBusters is taken as the current price of the program "Professor Trim", namely $299 per participant. This covers the weekly session over 6 weeks and all course materials.

An allowance in the cost-estimate used has been made for participants failing to complete at least 5 of the 6 sessions comprising the course (14% of participants; Egger et al., 1996). The corrected figure used is $356 (= $299 $\times$ 100/86). Implicit in this approach is the conservative assumption that none of the 14% of non-completers gain any benefit. In practice, a small proportion may have gained some degree of benefit from attending up to 4 sessions. However, as the extent of any such benefit would be purely speculative, this assumption is necessary.

Cost effectiveness results

Applying three interpretations of 'changing behaviour' (as defined below), to the outcomes reported in Section 7.3 of this chapter, the cost-effectiveness of the program may be expressed as:

- $426 per person changing behaviour (defined as achieving WHR reduction goal set by Egger et al.,)
- $318 per person changing behaviour (defined as maintaining any weight loss)
- $57 per kilogram of weight loss.

These outcomes measures are considered to be intermediate outcomes for the purposes of economic evaluation. The cost-utility estimate (cost per QALY gained) required modelling, as discussed below.

7.6 Cost-utility analysis

Economic modelling – an overview

A modelling approach was used to enable the intermediate outcome measures of weight loss reported by Egger et al. (1996) to be linked to life-years saved and QALYs gained. A Markov process structure was developed comprising three 5-year cycles. In conformity with other evaluations presented in this report, results were simulated over 15 years. Given the importance of DM Type 2 in overweight and obese subjects to health prognosis, the model provides for transitions between three different metabolic states (DM Type 2, IGT and NGT) and death.

Each 5 year cycle required estimation of 12 probabilities in order to model survival and progression towards DM Type 2, as shown in Figure 7.1. A vector of three probabilities (shown in Figure 7.1 as P1, P2, P3) was used to estimate the differing probability of all-cause mortality by metabolic state. For survivors within each cycle, the remaining nine probabilities (P4 – P12) constitute a probability matrix governing the allowable transitions between the three metabolic states.

Figure 7.1 Matrices used in GutBusters model

	Death	DM Type 2	IGT	NGT
DM Type 2	P1	P4	P7	P10
IGT	P2	P5	P8	P11
NGT	P3	P6	P9	P12

The baseline prevalence of DM Type 2, IGT and NGT is estimated from Dunstan et al. (2002). As this study examined a general Australian population, the prevalence of DM Type 2 and IGT could be expected to be marginally higher in a mildly overweight group, and significantly higher in an obese population. However, the study results of Dunstan et al. (2002) were retained for this evaluation in view of they were derived from an Australian setting. Furthermore, the use of these estimates from a general population make the cost-effectiveness and cost-utility estimates conservative, since the greater the prevalence of disease amongst the participants, the greater the number of beneficiaries of the intervention.

Transition probabilities

The values for the metabolic transition matrix (P4 – P12 in Figure 7.1) were derived from Eriksson et al. (1991). This publication is ideally suited for modelling outcomes from GutBusters as the study cohort comprised middle-aged, overweight males who participated in a lifestyle intervention to reduce weight, and whose baseline BMI and subsequent weight loss were similar to that of the GutBusters study group. The choice of a 5-year cycle reflects the follow-up period in Eriksson et al., 1991.

From a large screening program for DM Type 2, Eriksson et al. (1991), recruited 41 consecutive male subjects with early-stage DM Type 2, and 181 consecutive male subjects with impaired glucose tolerance to undergo a controlled study of the impact of a long-term intervention with an emphasis on lifestyle changes. A 5-year protocol, including an initial 6-months (randomised) pilot study, consisting of dietary treatment and/or increase of physical activity or training with annual check-ups, was completed by 90% of subjects. The average age of the males recruited was 48.1 years (±0.7) with an average BMI of 27.7 (±3.7) for DM Type 2 subjects and 26.6 (±3.1) for IGT subjects, and 24.3 (±2.8) for NGT subjects. Two groups (Groups 1 & 2) received the intervention, and the two remaining groups (Groups 3 & 4) received standard care, or no care for the NGT subjects.

- Group 1 : 41 patients with newly detected DM Type 2.
- Group 2 : 181 IGT patients, of whom 161 attended for follow-up screening.
- Group 3 : 79 IGT patients who were not enrolled in the program but who received a variety of interventions comprising 'standard care' (eg. antihypertensive therapy).
- Group 4 : 114 randomly chosen NGT patients also recruited during the screening process.

The mean weight loss at the end of the first year was 6 kilograms, and 2.0 kilograms and 3.3 kilograms in the DM Type 2 and IGT groups respectively by 5 years. Groups 3 and 4, by contrast, increased weight by 0.2 kilograms and 2 kilograms respectively, thus maintaining a significant difference compared to the intervention group ($p < 0.0001$).

Five years after the intervention, a second screening test and measurements were performed (Table 7.2). The weight losses in the intervention groups were associated with more than half (53.8%) of the DM Type 2 patients being reported as in remission (ie, glucose levels under the diagnostic threshold), and 52.2% of the IGT group had returned to NGT levels, although 10.6% had progressed to DM Type 2. These results contrast with the control subjects, where 28.6% of the non-randomised IGT subjects (Group 3) had progressed to DM Type 2.

Table 7.2 Results of second screening test for DM Type 2; Eriksson et al. (1996)

	Outcome of Oral Glucose Tolerance Tests at Follow-Up.		
	≥ 6.7 and/or 2-h ≥11.1	< 6.7 and/or 2-h 7.0 - 11.0	< 6.7 and/or 2-h <7.0
Group 1 (n = 39) DM Type 2 Treatment Group	18 (46.2%)	12 (30.7%)	9 (23.1%)
Group 2 (n=161) IGT Treatment Group	17 (10.6%)	60 (37.3%)	84 (52.2%)
Group 3 (n=56) IGT Control	12 (21.4%)	24 (42.8%)	20 (35.7%)
Group 4 (n=114) Control (NGT)	0 (0%)	8 (7.1%)	106 (93.0%)
Total (n=370)	47 (12.7%)	104 (28.1%)	219 (59.2%)

Source: Adapted from Eriksson et al., 1996: Table 2;p.895.

The percentages in the above table equate to probabilities and were therefore used directly to estimate the values in the metabolic transition matrix for the GutBusters participants (taken from Groups 1 and 2 of Eriksson et al., 1996) and controls (taken from Groups 3 and 4 of Eriksson et al., 1996). Information lacking from these results to complete the metabolic transition matrix are (i) the transition probabilities for control subjects with DM Type 2, and (ii) the transition probabilities for intervention subjects with NGT. It was therefore necessary to assume these values based upon clinical advice. Additionally, in the absence of evidence as to how these values change over time, the values for the metabolic transition matrix were held constant in the model over each of the three cycles. The values used in the metabolic transition matrix are shown in Figure 7.2.

Figure 7.2 Metabolic transition matrix probabilities applied in model (assumed values given in italics)

Control Metabolic Transition Matrix

	NIDDM	IGT	NGT
NIDDM	*0.950*	*0.040*	*0.010*
IGT	0.214	0.428	0.358
NGT	0.010	0.070	0.920

Intervention Metabolic Transition Matrix

	NIDDM	IGT	NGT
NIDDM	0.462	0.307	0.231
IGT	0.106	0.372	0.522
NGT	*0.010*	*0.040*	*0.950*

Estimation of all-cause mortality

The three values for the mortality vector (P1 to P3) were themselves modelled from Australian Bureau of Statistics data. In Table 7.3, the 5-year cumulative mortality rates of the Australian population are shown by age group in Column 1 (Australian Bureau of Statistics, 2000). These mortality rates for the general population were then adjusted to reflect the increased risk of mortality imposed by changes in metabolic state, and for degree of excess weight. The basis for the adjustment of risk was:

- *Adjustment for metabolic state:* Balkau et al., 1993 reported that, compared with NGT, the relative risk of premature mortality was 2.1 for people with DM Type 2 and 1.6 for people with IGT. More recent evidence (Rockwood et al., 2000) reports a relative risk of 1.9. For this evaluation, the mid-point of 2.0 was used to represent the relative of premature mortality due to DM Type 2. These relative risks for DM Type 2 and IGT were used to adjust the mortality rates for the Australian general population to determine mortality rates by metabolic state Column 3. The general approach to this calculation was by solving for '*X*' in the following formula:

$$ABS = [2.0 \times X \times PDM] + [1.6 \times X \times PIGT] + [X \times PNGT]$$

Where:

X is the (unknown) mortality rate for subjects with NGT

ABS is the 5-year cumulative mortality rates of the Australian population (Source: ABS, 2000)

PDM is the prevalence of DM Type 2 in the Australian population (Dunstan et al., 2002).

PIGT is the prevalence of IGT in the Australian population (Dunstan et al., 2002).

PNGT is the prevalence of NGT in the Australian population (Dunstan et al., 2002).

- *Adjustment for degree of overweight:* In order to reflect the GutBuster population who are all overweight, further adjustment of the mortality probabilities was necessary. A literature review was used to examine the relationship of weight to mortality. Under the assumption that the risk gradient for mortality from increasing levels of overweight is linear, the relative risk for the control group was conservatively estimated to be 1.2 (Manson, 1987; Rissanen et al., 1990) and the resulting mortality rates are shown in Column 5. For the intervention group, who achieved minor weight loss, a minimal reduction of 0.1 to an overall relative risk of mortality of 1.1 was assumed for the model (shown in Column 4).

Table 7.3 Estimation of mortality rates

1		2	3	4	5
Age Group	**ABS Mortality Rates (Cat No. 3302.0)**	**Metabolic Status**	**ABS Rates for General Population Adjusted for Metabolic Status**	**Relative Risk of 1.1 (for Overweight)**	**Relative Risk of 1.2 (for Overweight)**
				Intervention Cohort	**Control Analysis**
45-49	0.0126	NIDDM	0.0213	0.0234	0.0255
		IGT	0.0170	0.0187	0.0204
		NGT	0.0106	0.0117	0.0128
50 – 54	0.0200	NIDDM	0.0338	0.0371	0.0405
		IGT	0.0270	0.0297	0.0324
		NGT	0.0169	0.0186	0.0203
55 – 59	0.0337	NIDDM	0.0569	0.0626	0.0683
		IGT	0.0455	0.0501	0.0546
		NGT	0.0285	0.0313	0.0341

Adjustment assumes relative risk of 1.6 for IGT, and 2.0 for NIDDM.

Downstream costs

The DiabCost study (Colagiuri et al., 2003) results are relevant to this evaluation as weight gain is known to be a key risk factor for DM Type 2. Importantly, weight loss is known to prevent IGT and the onset of DM Type 2. Thus DM Type 2 is used here as a proxy for the higher health care utilisation associated with overweight.

Through an NHMRC grant, the DiabCost survey questionnaire was mailed to 25,000 people selected randomly from the National Diabetes Institute. Details of the respondents are shown in Table 7.4.

Table 7.4 Characteristics of DiabCost respondents

Patient Details	Survey Respondents
Number of respondents	10652 (42% of survey)
Male	49.8%
Mean age	65.2 years
Mean duration of diabetes	5.4 years
Smokers	8.7%
Carers	10.0%
Currently managed by diet and exercise	32.7%
Currently treated with tablets	59.6%
Currently treated with insulin ±tablets	6.7%

From Table 7.4, it is evident that the Diabcost survey respondents are a relatively recent group of DM Type 2 patients, the average number of years since diagnosis being approximately 5 years (equivalent to the duration of one cycle of the disease model described below) and with almost one-third of patients being controlled by diet and exercise alone. These observations are important as it suggests the resource use attributions made by the authors are conservative in the sense that a cohort with a longer average duration of DM Type 2, and greater use of tablets and/or insulin, would be expected to have a greater incidence of insulin use, greater incidence of comorbidities, and thus greater use of health care.

DiabCost results were published in December, 2003. The publication focuses specifically upon DM Type 2 in providing estimates of:

- Direct health costs to the health system
- Out-of-pocket expenses borne by people with DM Type 2.
- Community resources used by people with DM Type 2.
- The impact of DM Type 2 upon quality of life.

Health care utilisation was reported as being as shown in Table 7.5.

Table 7.5 Health care utilisation per annum; DiabCost study; (Colagiuri et al., 2003)

Service	Overall	Complications			
		None	Microvascular	Macrovascular	Both
GP Visits (surgery)	10.5	9.5	12.7	12.3	14.4
GP Visits (home)	0.4	0.2	0.6	1.0	1.1
Outpatient visits	1.1	0.8	1.8	1.7	2.1
Emergency service visits	0.1	<0.1	0.1	0.4	0.5
Inpatient nights	2.0	1.2	2.9	3.4	4.2

The DiabCost study methods were designed to report costs to within $250 if a 10% response was received (a 42% response was in fact received). The reported costs are shown in Table 7.6.

Table 7.6 Health costs; DiabCost study; (Colagiuri et al., 2003)

Cost Category	Overall Respondents	Complications			
		None	Microvascular	Macrovascular	Both
Health costs:					
Direct	$5325	$3990	$6990	$8985	$9610
Indirect	$35[1]	$35	$35	$70	$35
Total	$5360	$4025	$7025	$9055	$9645
Govt Subsidies	$5540[2]	$5075	$6200	$6120	$6240

[1] Comprises productivity losses due to ill-health or premature death. These costs are not incurred by the health sector directly.

[2] Includes welfare payments, referred to in economics as 'transfer payments'.

The DiabCost study estimates the average cost of treating a person with DM Type 2 is $5,325 per annum, being the average health care cost per patient with DM Type 2 from all people responding to the survey. Importantly, people in 'normal' health still use health care for a range of conditions and illnesses unrelated to overweight. The Australian Institute of Health and Welfare have estimated the average cost per person in Australia to be $2,817 per annum (AIWH, 2002).

Thus, the incremental cost attributable to DM Type 2 may be considered the difference between these two figures, or **$2,508** per patient per annum.

Cost-utility results

The primary results were based upon modelling that included the economic convention of discounting costs and benefits at the rate of 5% per year. To facilitate transparency of results, undiscounted results are given in the Sensitivity Analysis.

Health outcomes: When analysed for a 'cohort' of one patient, the model results can be interpreted as representing the estimated probabilities of being in a particular state (Type 2, IGT, NGT or dead) in any given cycle. The following tables display the distribution of patients at baseline (Dunstan et al., 2002) and for each of the subsequent 3 cycles.

Table 7.7 Modelling results: distribution of patients in GutBuster cohort by cycle

	Baseline	5 yrs	10 yrs	15 yrs
Type II	0.080	0.061	0.047	0.037
IGT	0.174	0.117	0.092	0.078
NGT	0.746	0.808	0.828	0.819
Dead	0.000	0.013	0.034	0.067

Table 7.8 Modelling results: distribution of patients in control cohort by cycle

	Baseline	5 yrs	10 yrs	15 yrs
Type II	0.080	0.118	0.140	0.152
IGT	0.174	0.127	0.107	0.096
NGT	0.746	0.740	0.713	0.672
Dead	0.000	0.015	0.040	0.080

If the patient dispositions in Tables 7.7 and 7.8, are multiplied by 5 years, the total time (years) in each state can be estimated. Cumulative deaths by the end of the 15th year were 6.7% of the GutBuster group and 8.0% of the control group. Over the 15 years of the model, it is estimated that the average survival time will be:

- 12.50 years in the GutBuster cohort
- 12.40 years in the control cohort

Thus it is estimated the GutBuster program is associated with an average incremental gain of 0.09 life-years[5] over the time-frame of the model.

Multiplying the time in each state (calculated as 5 years times each of the DM Type 2 values in Tables 3.7 and 3.8) by the utility values in Table 7.1 (Colagiuri et al., 2003) provides an estimate of the QALYs over time. The results are shown in Tables 7.9 and 7.10.

Table 7.9 Modelling results for QALYs in GusBuster cohort

QALYs Intervention			
	5 yrs	10 yrs	15 yrs
Type II	0.213	0.163	0.126
IGT	0.415	0.326	0.264
NGT	2.940	3.010	2.800
Sub-Totals	3.568	3.499	3.189

Table 7.10 Modelling results for QALYs in control cohort

QALYs Control			
	5 yrs	10 yrs	15 yrs
Type II	0.407	0.486	0.514
IGT	0.452	0.381	0.326
NGT	2.691	2.592	2.297
Sub-Totals	3.550	3.459	3.138

[5] Rounding errors may be observed in reporting of results.

From a maximum of 15 QALYs (equivalent to 15 years lived in normal health), the QALYs estimated from the model were:

- 10.26 QALYs in the GutBuster cohort
- 10.15 QALYs in the control cohort

Thus it is estimated the GutBuster program is associated with an average incremental gain of 0.109 QALYs over the time-frame of the model.

Estimation of cost-offsets: The years spent in the model in DM Type 2 was multiplied by the annual cost of incremental health care use ($2,508; refer 'Downstream Costs' above) to produce Tables 7.11.

Table 7.11 Model estimated health care use costs

	5 yrs	10 yrs	15 yrs	Total Over 15 years
GutBuster Cohort	$667	$511	$404	$1,582
Control Cohort	$1,277	$1,525	$1,654	$4,456
Savings (=difference)	$610	$1,014	$1,250	$2,874

Cost per life year saved: The result is one of dominance. That is, the GutBusters program results in both lower costs, $1,938 (=health care use cost of $1,582 + per person program cost of $356; Section 7.5) for GutBusters participants compared to $4,456 (Table 7.11) for controls, and better health outcomes, 12.50 life years for GutBuster participants and 12.40 for controls.

Cost per QALY gained: Consistent with expectations given the above cost-effectiveness estimate, the result is again one of dominance. That is, the GutBusters program results in both lower costs and better health outcomes, 10.26 QALYs for GutBusters participants and 10.15 QALYs for controls.

7.7 Sensitivity analysis

Undiscounted results

Over the 15 years of the model, it is estimated that the average survival time in undiscounted years is:

- 14.43 years in the GutBuster cohort
- 14.33 years in the control cohort

Thus it is estimated the GutBuster program is associated with an average incremental gain of 0.107 life-years over the time-frame of the model.

From a maximum of 15 QALYs (equivalent to 15 years lived in normal health), the undiscounted estimate of QALYs from the model were:

- 11.84 QALYs in the GutBuster cohort
- 11.72 QALYs in the control cohort

Thus it is estimated the GutBuster program is associated with an average incremental gain of 0.126 QALYs over the time-frame of the model.

Cost per life year saved (undiscounted): Although the values changes as a consequence of no longer discounting costs and benefits, the result remains one of dominance by GutBuster with both lower costs, $2,183 for GutBusters participants compared to $5,186 for controls, and better health outcomes, 14.43 life years for GutBuster participants and 14.33 for controls.

Cost per QALY gained (undiscounted): Dominance by GutBusters with both lower costs and better health outcomes, 11.84 QALYs for GutBuster participants and 11.72 QALYs for controls.

Cost-offsets are ignored: If the 'downstream' costs (cost-offsets) estimates here from the DiabCost Study are ignored, and the costs of GutBusters are considered against the health benefits of the program, the results become:

Cost per life year saved (undiscounted): $3,317 per life-year saved (= $356 / 0.107)

Cost per QALY gained (undiscounted): $2,836 per QALY (= $356 / 0.126)

Health benefits only realised over years 0-5, and no cost-offsets

If the benefits of the program only extend for 5 years, such as may happen if weight is re-gained, and cost-offsets are not considered, the 5-year results become:

Cost per life year saved (undiscounted): $40,551 per life-year saved (= $356 / 0.01)

Cost per QALY gained (undiscounted): $19,796 per QALY (= $356 / 0.02)

7.8 Discussion

Although the main aim of GutBusters is to capture a broad range of health benefits known to be associated with weight loss in overweight subjects, DM Type 2 alone has a major role in the onset of the key morbidities and premature mortality. Structuring the model around the progression of this disease, based upon the evidence from Eriksson et al. (1991), serves to highlight the impacts of weight (and its interaction with DM Type 2) upon all-cause mortality (refer Methods section this chapter). Consistent with the evidence from other weight loss programs, these publications show programs such as GutBusters can be effective in achieving at least moderate weight loss in overweight to obese subjects. However, unlike many other programs, evidence from long-term follow-up (2 years) is available from Egger et al. (1996).

Contributing to the robustness of the model that links the clinical measures reported by Egger et al. (1996) (intermediate endpoints) to life-years and QALYs is the compatibility of the secondary data sources with that of the Egger et al., 1996. The studies by Egger et al. (1996), Eriksson et al. (1991), and the DiabCost study have all been conducted in overweight subjects in the early years of diabetes[6].

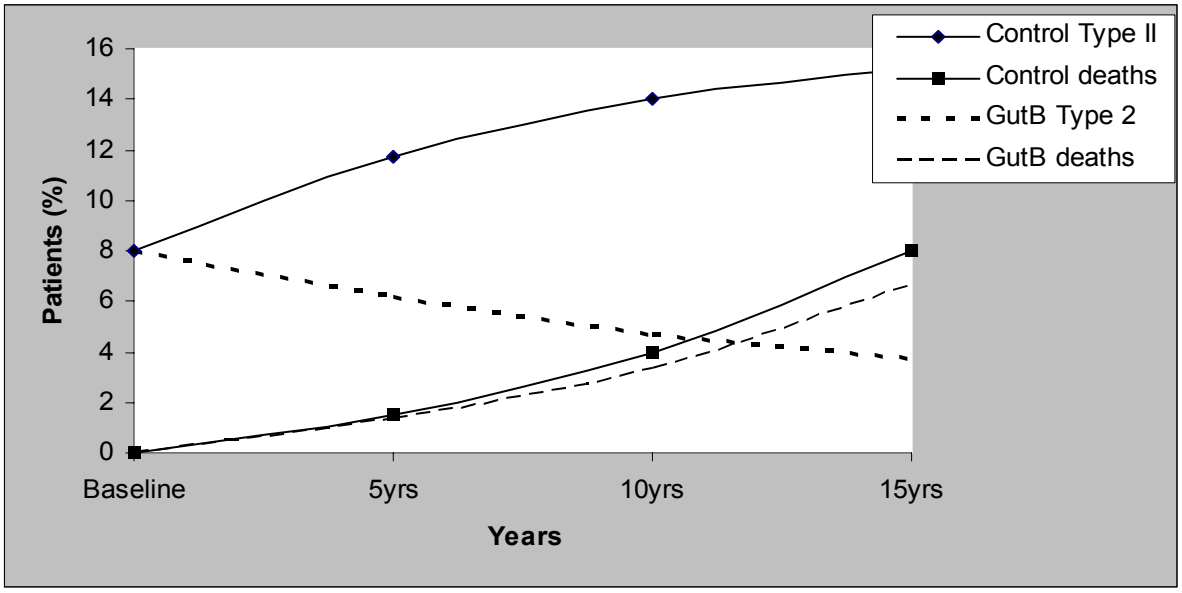
Furthermore, the evidence for the efficacy of weight loss programs in preventing DM Type 2 shown by Eriksson et al., has since been strongly supported such that the role of weight loss in the prevention of DM Type 2 is no longer contentious (Astrup and Finer, 2000; Moore et al., 2002). Of particular relevance are the results of the Framingham Study published by Moore, 2000. Moore, 2000 studied 618 overweight patients (BMI $\geq$ 27) aged 30-50 years and reported that sustained weight loss of between approximately 4-6 kilograms was associated with a relative risk of diabetes of 0.63 (and 0.34 for obese patients), compared to a reference group. Greater weight losses were associated with a 51% reduction in risk of DM Type 2.

[6] The average age of respondents in the DiabCost study was 65.2 years, and thus older than the age range in the model (45-60 years), the high proportion of DiabCost study participants managed by lifestyle change alone suggest the respondents were, on average, a group with mild DM Type 2 and therefore lower health care use than more severe cohorts.

Inevitably in modelling, the use of assumptions is required. The key assumptions in this evaluation are all considered to be conservative. For instance, the limitation of the model to 15 years is likely to favour control patients given the difference in the prevalence of DM Type 2 between the two cohort's increases with age (as do the associated health care costs and mortality). Additionally, the initial prevalence of DM Type 2 and IGT were taken from a survey of a general population and are likely to be higher in the 'at risk' population targeted by GutBusters. The assumed values in the metabolic transition matrix allow for 1% of DM Type 2 patients to spontaneously return to NGT without intervention; an outcome that is perhaps optimistic.

An overview of the performance of the model in capturing the impacts of the GutBusters program upon each of DM Type 2 and mortality is shown in Figure 7.3.

Figure 7.3 Modelled mortality and DM Type 2; GutBuster and control cohorts



Consistent with the literature, the prevalence of DM Type 2 group in the control group rises in this at risk over the 15 year period of the model. The cumulative effect of GutBusters upon the prevalence of DM Type 2 in the intervention cohort, which is estimated from Eriksson et al. to have approximately halved the rate of progression to DM Type 2 each cycle, is similarly shown. The effect upon mortality is modest in the first 5 years, but increases over the life of the model. The predicted deaths in a general population after 15 years, based upon ABS mortality data (ABS, 2000), is almost 5%. The model estimates of deaths being 8.0% and 6.6% of participants in control and GutBuster cohorts respectively are consistent with expectations given some residual relative risk is assumed to exist even in the intervention cohort.

Although not well documented in Australia, the health care cost of treating DM Type 2 as estimated by the DiabCost study, are also consistent with previous findings. Dalton and Segal, 1996, used 1994 data to estimate an attributable annual cost of approximately $2,000. This estimate suggests the DiabCost result may even be conservative given cost escalations that have occurred since.

One uncertainty amongst the assumptions is the constant values in the metabolic transition matrix. How these transition probabilities would change over time is an unknown. In relation to this uncertainty, the sensitivity results provide reassurance as to the value-for-money of the program. What is clear from the sensitivity analysis is that the dominance of GutBusters results over those of the controls is driven largely by the relative size of cost-offsets to the cost of the modest cost of the program. However, if cost offsets are ignored in the analysis, the cost per QALY gained is still excellent value at $3,275 per QALY. Furthermore, a complete convergence of the transition probabilities in the second and third cycles may be considered as equivalent to the effects of weight

being regained in the intervention group. These effects were tested in the sensitivity analysis by assuming the benefits of weight loss (whether due to weight regain or convergence of transition probabilities) are only maintained for 5 years. The cost per QALY under this revised assumption became $19,769, which is still attractive compared to many health interventions currently receiving funding support. Additionally the 5-year results for cost per life-year saved being estimated to as a significantly higher figure of $40,551, whilst still acceptable, suggest that much of the early gains are derived from quality of life gains rather than duration of life.

Another perspective on the program can be seen from calculation of the number needed to treat (NNT) under this program. Based upon the results of Eriksson et al., 1991, one in every 7 patients completing the program would either prevent or delay the onset of DM Type 2. When considering this ratio, given an intention-to-treat cost of $356, and that prevention the disease avoids an annual incremental health care cost of approximately $2,500 for many years, the results here are not surprising.

In conclusion, the model is considered to be robust, and generates results that are strongly supportive of the program.

8. Workplace based trial of a multi-factorial prevention of coronary heart disease

8.1 Description

Intervention type

This chapter was concerned with multifactorial interventions in the workplace that were designed to reduce lifestyle risk factors. The exemplar for this type of intervention was a workplace based, cross-country, randomised control trial (RCT) aimed at reducing the risk factors for coronary heart disease (CHD) in males aged 40-59 years of age. It also examined the impact of risk factors reduction upon the 'incidence of mortality from CHD and total mortality'. (Kornitzer and Rose, 1985)

References/sources of evidence

The descriptions of the intervention and its effectiveness were based on the following articles:

- *Multifactorial trial in the prevention of coronary heart disease: 1. Recruitment and initial findings* (World Health Organization European Collaborative Group, 1980)
- *Multifactorial trial in the prevention of coronary heart disease:2. Risk factor changes at two and four years* (World Health Organisation European Collaborative Group, 1982)
- *Multifactorial trial in the prevention of coronary heart disease:3. Incidence and mortality results* (World Health Organisation European Collaborative Group, 1983)
- *WHO European Collaborative Trial of multifactorial prevention of coronary heart disease* (Kornitzer and Rose, 1985)

Intervention description

Recruitment and target population: The target population for the intervention were 'middle-aged men' employed in the recruited workplace units. (World Health Organization European Collaborative Group, 1980) The trial started in 1971 with the recruitment of 12 pairs of workplace units (18,210 men) in the UK. Under WHO auspices, it was extended to Belgium, Italy, Spain and Poland. In total, 88 workplace units (factories and other occupational groupings), ranging in size from 69 to 2,508 men aged between 40 and 59 years, were recruited. Within each country, the units were arranged in matched pairs and one unit in each pair was randomly allocated to the intervention group and the other to the control group. (Kornitzer and Rose, 1985, WHO European Collaborative Group, 1986) According to Kornitzer (1985), each trial centre was 'basically autonomous and self-sufficient, but standardization of the protocol and data quality control systems permitted pooling of data on risk factor changes, disease incidence, and mortality'. (Kornitzer and Rose, 1985) Spain participated in measuring impact of the intervention on risk factors but not in the monitoring of mortality and morbidity. Warsaw did not complete recruitment until 1977, some two years later than the other centres.

Intervention: The trial began in the intervention units with all eligible men (ie those aged between 40 and 59 years in each workplace unit) being offered a screening examination. This examination was used to identify the men with the highest multivariate risk for CHD who received individual and sustained advice including personal consultation with doctors as well as the general health education campaign. (Kornitzer and Rose, 1985, World Health Organisation European Collaborative Group, 1983) The general campaign of risk factor modification was supported by posters, brochures, personal letters, progress charts and group discussions, and the content consisted of:

- Dietary advice to lower serum cholesterol for all participants;
- Advice about ceasing smoking for participants who smoked 5+ cigarettes per day;
- Weight reduction for those >15% overweight;
- Daily exercise for the sedentary; and

- Treatment of hypertension for those whose systolic pressure averaged 160mm Hg or more. (Kornitzer and Rose, 1985, World Health Organisation European Collaborative Group, 1983)

Within the general framework of the intervention, there were differences between the centres in terms of the percentage of high-risk participants (Table 8.1), the nature of the intervention and the resources used in the intervention (Table 8.2), and the baseline characteristics of the participants and the resources used in the intervention (Table 8.3). Italy had six full-time staff for 3,131 men and the UK had two equivalent full-time staff for 9,734 men. (World Health Organisation European Collaborative Group, 1983) To accommodate the differences between the centres and to ensure that there is comparable data across the centres the analysis in this report focused on the results for the centres in the UK, Belgium and Italy.

Table 8.1 Participants, factory pairs and high risk participants in the intervention and control units in the UK, Belgium and Italy

		UK	Belgium	Italy
Recruitment years		1971-73	1972-74	1973-74
Factory pairs		12	15	2
Intervention	Enrolled	9,734	8,509	3,131
	Screened n (%)	8,398 (86%)	7,399 (87%)	2,642 (84%)
	High Risk n (%)	1,278 (15%)	1,597 (22%)	603 (23%)
Control	Enrolled	8,476	10,900	2,896
	Screened n (%)	786 (9%)	901 (8%)	592 (20%)

Source: (World Health Organization European Collaborative Group, 1980), p. 77 and (World Health Organisation European Collaborative Group, 1982), p. 186

Table 8.2 Brief descriptions of the interventions in the UK, Belgium and Italy

UK	Belgium	Italy
'The intervention was delivered through factory doctors, reinforced after the two year results by the project nurses who combined screening visits with health education and returned for follow-up visits. Initial attention was given to high risk men who were seen four times in the first year for about 15 minutes each. Personal contact was made with non-high risk men after the first two years by the project nurses and some factory doctors.'	'Two half-time physicians working for the project did the bulk of the health education in the high risk men. (High risk men) were … seen three times in the first year for individual counselling, twice in the following year and annually thereafter. Other men were seen as part of the anniversary 5% samples and additional 5% samples were drawn at 30, 42 and 54 months and given individual advice. Hypertensives were referred to their general practitioners but also advised to lose weight to restrict their salt intake'	'For each factory two doctors and one nurse or dietician worked four hours a day on the project … reinforced by the other project personnel. Non-high risk men received individual attention from the start but not as intensively as the high risk group. Blood pressure was treated by project personnel in the factories in most cases.'

Source: (World Health Organisation European Collaborative Group, 1982), pp 189-190

8.2 Quality of evidence

Evaluation description

Design: A quasi-experimental design was used to evaluate the trial. Although the allocation of participants to the intervention and control groups was controlled by the investigators, the method fell short of 'genuine randomisation' of participants with 'allocation concealment'. It was the workplace unit rather than individual who was randomised, and staff delivering the intervention and undertaking the risk factor assessments were aware of the status of the workplace unit. Only in the case of morbidity and mortality were the researchers blind to the group assignment.

Methodology: The methodology for evaluating the intervention involved a comparison of the changes in the risk factors and differences in mortality and morbidity in the intervention and control unit participants. Each outcome measure involved different data collection methods and analyses. (Table 8.3)

Outcome measures: The primary endpoints for comparison between the control and intervention groups were fatal CHD (myocardial infarction and other sudden or non-sudden death presumed to be due to CHD), fatal CHD and non-fatal myocardial infarction and total mortality. (World Health Organisation European Collaborative Group, 1983). The intermediate outcomes involved tracking the risk factors for CHD over time and calculation of a CHD risk score from the risk factor levels.

Table 8.3 Methodology used in evaluation

Risk Factors	Survival Status	Morbidity
Intervention Group *Independent surveys*: Baseline evaluation of 86% of enrollees who volunteered for the screening examination and follow-up at 2 and four years of 5% random samples of participants still employed at the factories. All participants still in employment were offered a final screening examination *Panel study*: Periodic follow-up of all men designated as high risk at baseline and still in employment at each follow-up anniversary. **Control Group** *Panel study*: Baseline evaluation of a random sample of enrollees, with follow-up at two and four years for those still employed. All participants still in employment were offered a final screening examination.	*Enumeration, intention-to-treat analysis*: Survival status of all baseline enrollees determined from institutional records. Coding of the records was undertaken centrally by researchers who were blind to the group to which the person belonged. Participants in control group panel study were excluded from determination of survival status.	*Enumeration, completers only analysis*: Occurrence of non-fatal MI events for all employees 'still in employment' and who were absent from work for 21 days were determined from GP and hospital records Coding of the records was undertaken centrally by researchers who were blind to the group to which the person belonged. Participants in control group panel study were excluded from determination of survival status.

Sources: (World Health Organisation European Collaborative Group, 1982, World Health Organisation European Collaborative Group, 1983)

After the baseline screening, risk factor changes in the intervention group were monitored by annual examination of a fresh 5% random sample at 2 and 4 years. In this intervention group high-risk men were followed for the entire length of the study and therefore constituted a panel study. In the control group risk factor changes were assessed by examination of the same 10% random sample at 2 and 4 years. At the end of the trial all participants still employed in both the intervention and control units were offered a final screening examination. (Kornitzer and Rose, 1985)

The incidence of non-fatal CHD events for continuing employees was based on a search of hospital and general practitioner reports on men who had been absent from work due to illness for 21 days or longer. The definition of non-fatal myocardial infarction was based on WHO Registry criteria, events being classified without knowledge of group assignment. (Kornitzer and Rose, 1985)

At the end of the trial survival status was established in over 99% of eligible participants at baseline, and causes of death were obtained from death certificates. The definition of fatal CHD was based on WHO Registry criteria, events being classified without knowledge of group assignment. (Kornitzer and Rose, 1985)

Analysis: The 'centre' was the unit of analysis. Risk factor and risk score results were presented as a baseline score and the 'net change' at two and four years. (The method used to calculate the net change has been included in Table 8.5.) The results for the final screening exam were not presented. No statistical significance tests were performed in the analysis of the risk factors and risk scores because 'the samples of men were not completely random, being clustered in heterogeneous groups both within the factories and centres'. (World Health Organisation European Collaborative Group, 1982)

The only analysis of the endpoint results for each centre was the cumulative six-year incidence rate of fatal CHD, total CHD and total deaths for the control group and the percentage difference in the intervention group rate compared to the control group. The method of calculating the incidence rates (Table 8.5) enabled significance tests to be performed in the analysis and the calculation of 'approximate confidence intervals'. (World Health Organisation European Collaborative Group, 1983) The confidence intervals were not reported.

Assessment

Sources of bias: The potential sources of bias in the evaluation of the trial were selection bias, performance bias, measurement bias and attrition bias.

Selection bias: This bias involves 'systematic differences between comparison groups'. This multifactorial trial involved randomization of matched workplace pairs. The success of this form of randomization depends upon the degree of match between the units in each pair. An analysis of the baseline characteristics of all screened intervention and control group participants has been used to as evidence that the method of allocation was 'successful in producing two well-balanced groups'. (World Health Organization European Collaborative Group, 1980) However, on the evidence presented there is no way of telling whether this balance was achieved for each centre. Also the baseline characteristics of the control group were based on approximately 10% random samples of control group participants and 'the lack of baseline data for the control group also meant that there was no adjustment in the analysis for differences in baseline characteristics of the participants'. (World Health Organisation European Collaborative Group, 1983)

In terms of the intervention group there were differences between the centres in the baseline characteristics of the participants. (Table 8.4) No baseline characteristics of the control group samples for each of the individual centres were reported in the literature.

Table 8.4 Age and age adjusted baseline characteristics of intervention group participants at the centres in the UK, Belgium and Italy

Centre (work-place pairs)	Age Mean (SD)* years	Cholesterol Mean (SD)* mg/dl	Systolic BP Mean (SD)* mm Hg	BMI Mean (SD)* kg/m^2	% Non and Ex Cigarette Smokers	Mean Cigarettes /Day/ All Men	CHD risk Mean (range)†
UK (12 pairs)	49.9 (±5.5)	216 (±40)	140 (±20)	25.5 (±3.0)	49%	8.6	3.3 (2.9-3.7)
Belgium (15 pairs)	48.3 (±5.6)	142 (±18)	142 (±19)	25.1 (±4.4)	35%	12.3	4.2 (3.5-5.0)
Italy (2 pairs)	48.8 (±5.6)	134 (±19)	134 (±19)	26.4 (±3.4)	37%	11.9	3.6 (3.2-3.8)

Source: (World Health Organization European Collaborative Group, 1980), pp 75-76

Notes: * Between subject standard deviations.

† Range of factory means. CHD risk based on individual men's values of age, cigarettes, systolic blood pressure, cholesterol and BMI.

Performance bias: There were differences between the centres in the way in which the intervention was implemented and the resources devoted to the intervention. (Table 8.2) To overcome this source of bias, the results for each centre were reported and analysed separately and there was no pooling of the data.

The intervention was available to all employees in the intervention sites and this implied 'a mixture of primary and secondary prevention since some men already had a history or signs of myocardial ischaemia at the time of entering the trial'. (World Health Organization European Collaborative Group, 1980) To the extent that there were differences between the sites in terms of the proportion of men with a history of signs of myocardial ischaemia at baseline (Table 8.5) then there would be differences between the sites in the proportion of primary and secondary interventions be implemented.

Table 8.5 Age adjusted prevalence rates of electrocardiographic findings related to CHD, angina, possible infarction and intermittent claudication for intervention group participants

		UK	Belgium	Italy
% Major Q/QS waves*		0.90% (± 0.10)	1.17% (± 0.12)	0.74% (± 0.15)
% Other suspected ischaemia		6.93% (± 0.26)	6.55 (± 0.27)	5.84 (± 0.42)
Prevalence	angina	3.6%	5.0%	3.0%
	possible infarction‡	6.6%	5.2%	7.3%
	intermittent claudication	0.6%	1.5%	2.2%
% with positive ECG		7.1%	6.8%	5.4%

Source: (World Health Organization European Collaborative Group, 1980), pp 76-77

Notes:

* Codes I·I-2, suggesting myocardial infarction

† ST depressed (codes 4·I-3) or T wave inverted/flat (codes 5·I-3) or left bundle branch block (code 7·I).

‡ A history of severe central chest pain lasting for half an hour or more.

§ Ischemic-type electrocardiographic findings, Minnesota codes I·I-2,4; I-3,5; I-3 or 7·I

Measurement bias: There were systematic differences between the control and intervention groups in the measurement of risk factor changes which had the potential to introduce bias into the measurement of the outcomes. The panel study in the control group was used to reflect 'trends that occurred in the study populations independent of the trial'. (Kornitzer and Rose, 1985) This group was excluded from the endpoint estimates 'for fear that screening might alter their behaviour'. (World Health Organisation European Collaborative Group, 1983) To the extent that this did occur, then the risk factor results for the control group as measured by the panel study would not accurately reflect independent trends in the study population.

Attrition bias: In regard to the measurement of risk factors and predicted CHD risk, attrition from the trial occurred because of death, change in employment and non-response to the invitation on the screening day. There is no analysis in the literature used for this review of the nature of the attrition from each of the centres, but panel study attrition rates varied between the intervention and control group within and between each centre. (Table 8.6)

Table 8.6 Attrition rates in the intervention and control group panel studies

	UK	Belgium	Italy
Control Group Panel Study			
Baseline	786 (100%)	901 (100%)	592 (100%)
2 years	614 (78%)	800 (89%)	446 (75%)
4 years	324 (68%)	302 (81%)	256 (65%)
High Risk Intervention Group Panel Study			
Baseline	1,278 (100%)	1,597 (100%)	603 (100%)
2 years	43* (3%)	1,268 (79%)	429 (71%)
4 years	736 (58%)	1,078 (68%)	375 (62%)

Source: (World Health Organisation European Collaborative Group, 1982), pp. 186-187

Notes:

* This is the figure reported in the literature. There is no explanation as to why it is so low.

8.3 Outcomes – as reported

The outcome measures, data sources and analytic methods are shown in Table 8.7. Because the differences between the interventions offered to the men depended on their risk level, the preferred method of reporting the results would have been to have three groupings (high risk, non-high risk and total) for the entire six-year period. However, these data were not available in the literature. Therefore, using the results that were available, behaviour change and clinical parameters have been presented for the whole intervention group and the high risk group at two and four years (Table 8.8) and mortality endpoints presented for the intervention group as a whole at the end of six years (Table 8.9).

Table 8.7 Outcome measures, data sources and analyses for the evaluation

Outcome	Data source	Analyses
1. BEHAVIOUR CHANGE AND CLINICAL PARAMETERS (KORNITZER AND ROSE, 1985)		
Smoking: % Smokers Cigarettes/day	Self report questionnaire at baseline, two and four (Also see Clinical Parameters)	The baseline values at entry to the study were subtracted from the two and four year anniversary data from each man screened at two and four years. Mean values and mean changes were calculated for each centre's intervention and control groups. The mean change observed in control men screened at the same anniversary was subtracted from that obtained in intervention men to obtain a 'net change'. This net change was then divided by the original mean value in the whole intervention group at initial screening (not just those seen at the anniversary) to give a percentage change from baseline. This was done for all intervention men, and for high-risk men separately.
Serum cholesterol Weight Systolic blood pressure	Cardiovascular screening examination during working hours at baseline, two and four years.	
CHD Risk	The risk factor changes were summarized by a multiple logistic function to give an estimate of CHD risk.	
2. MORTALITY (WHO European Collaborative Group, 1986)		
Fatal CHD All deaths	Survival status was established for 99.8% of all men in employment at the start of the trial and causes of death were obtained from death certificates.	Incidence data were analyzed by life-table methods. If a man suffered multiple events, only the first event in the appropriate class was counted. Each paired unit yielded one estimate of the effect of the intervention based on the difference in rates between the intervention and control men; its importance was related to the number of person years on which it was based. A normally distributed statistic was derived by combining each individual estimate of difference, weighted by the inverse of its variance.

Behaviour change and clinical parameters

Whole intervention group: There were differences between the centres in the way the risk factor scores changed over time. In the UK, the increase in cholesterol at the two year mark led to an increase in predicted CHD risk despite reductions in scores on the other behavioural and clinical parameters. At the four year mark there was a net reduction on all parameters and this led to a 12.8% reduction in predicted risk in the UK. In Belgium, there were net reductions on all parameters and, consequently, a decrease in predicted risk at the two year mark. However, an increase in cholesterol at the four year mark meant that the reduction in predicted risk from baseline was not as high at four years as it had been at the two year mark. In Italy, there were reductions in the clinical parameters at two and four years but increases in the proportion of cigarette smokers. Despite this latter result, the predicted CHD risk decreased over the four year period. (Table 8.8)

High risk men: As would be expected, the baseline scores on the behavioural and clinical parameters were higher for high risk men compared to the whole intervention group. There was a net percentage reduction on all the risk factors scores and predicted CHD risk at two and four years. However, the reduction in predicted CHD risk accelerated in the UK and Italy and slowed in Belgium after the two year mark. (Table 8.8)

Table 8.8 Baseline values and percentage net change* from baseline in risk factor levels at two and four years for the whole intervention group in each site

	Whole Intervention Group‡									High Risk Intervention Men								
	UK			Belgium			Italy			UK			Belgium			Italy		
	Base line	2 yrs	4 yrs	Base line	2 yrs	4 yrs	Base line	2 yrs	4 yrs	Base line	2 yrs	4 yrs	Base line	2 yrs	4 yrs	Base line	2 yrs	4 yrs
Men screened	8,398	357	324	7,399	327	307	2,642	268	256	1,278	43§	736	1,597	1,268	1,078	603	429	375
Risk factor results: baseline score and % net change in the score at two and four years																		
Cholesterol (mg dl^{-1})	216	+ 2.0	- 4.1	232	- 3.3	+ 2.0	220	- 4.4	- 4.2	257	- 0.3	- 6.8	265	- 4.3	- 0.9	260	- 0.9	- 8.7
Systolic BP (mm Hg)	140	- 0.4	- 1.8	141	- 4.4	- 3.8	133	- 3.9	- 5.5	154	- 3.4	- 2.1	153	- 4.1	- 4.7	147	- 4.0	- 6.0
Cigs/day (all men)	8.3	-12.3	-15.6	12.2	- 5.3	- 3.9	12.0	- 5.8	- 8.1	14.3	- 8.1	-18.9	18.5	-13.8	-11.2	17.5	-15.5	- 8.9
Cigarette smokers (%)	51	- 3.5	- 1.4	65	- 0.3	- 1.9	63	+ 1.0	+ 7.5	74	-10.4	- 8.5	85	- 7.1	- 4.7	82	- 5.2	- 7.6
Predicted CHD risk†	**3.4**	**+ 4.1**	**-12.8**	**3.8**	**-26.3**	**-15.8**	**3.4**	**-26.0**	**-31.4**	**7.0**	**- 5.2**	**-19.1**	**7.0**	**-28.6**	**-20.0**	**7.0**	**-14.1**	**-37.9**

Source: (World Health Organisation European Collaborative Group, 1982), pp 186-187

Notes:

* Net of control group (Table 8.5)

† Events per 1,000 persons per year.
Est. five year CHD incidence = 1 / (1 + exp(-13,177 + 0.0888 x age (years) + 0.0084 x cholesterol (mg/dl) + 0.0165 x systolic BP (mm Hg) + 0.0720 x cigarette smoking category)
Smoking categories: non/ex smoker = 3; 1 to 4 cigs/day = 4; 5 to 9 cigs/day = 5; 10 to 19 cigs/day = 6, 20 to 29 cigs/day = 7; 30+ cigs/day = 8. (World Health Organization European Collaborative Group, 1980), p. 75)

‡ This includes high risk men. Because of the method of calculation it was not possible to separate the non-high risk men and the high risk men.

§ As indicated in the literature

Mortality

In the UK there was an increase in the cumulative six year incidence rates for mortality in the intervention group compared to the control group. In the other two centres there were declines in the intervention rates compared to the control rates. Only the difference in the all cause mortality rates in Belgium reached statistical significance. (Table 8.7)

Table 8.9 Cumulative six year mortality incidence rates in the control and intervention units

Endpoint	UK		Belgium		Italy	
	Control Rate	Intervention Est Rate (% Δ*)	Control Rate	Intervention Est Rate (% Δ*)	Control Rate	Interv Est Rate (% Δ*)
Fatal CHD	2.0%	2.2% (+ 8%)	1.4%	1.1% (-21%)	2.1%	1.5% (-30%)
All deaths	4.2%	4.8% (+14%)	4.0%	3.3% (-17%†)	4.5%	4.2% (- 6%)

Source: (World Health Organisation European Collaborative Group, 1983), p. 277

Notes:

* Estimated from published control rate and intervention difference: control rate x (1 + int difference/100). Estimates will be somewhat different from the actual rates due to rounding of the 'intervention difference' in the literature

† p value <0.05

8.4 Program costs

As reported by trial

There were no costs reported in the literature.

Based on resource use

According to Kornitzer and Rose (1985), it was 'estimated that intervention staff averaged two doctors and one nurse-nutritionist per 8,000 intervention men. The estimates of the costs of the intervention in the UK, and Italy are based on the number of staff reported in the literature. (World Health Organisation European Collaborative Group, 1982) The UK was reported as having the lowest staff-to-participant ratio and Italy the highest. The reported staffing arrangements were incomplete in the literature reviewed so it has been included at the 'average' rate. Using current year expenditure in the six years from 1998 to 2003 inclusive, the estimated cost per enrollee in the UK was $102, in Belgium $200 and in Italy $511. (Table 8.10)

Table 8.10 Estimated staff costs for delivering the intervention in the UK, Belgium and Italy using current Australian dollars for the six years 1998-2003 inclusive

Country and cost item	EFT§	Costs in 2003 AUS$ at:¶		
		2 Years	4 Years	6 Years
UNITED KINGDOM				
Salaries* Doctors†	1	$217,905	$435,810	$653,715
Nurses‡	1	$143,985	$287,970	$503,947
Salary sub-total		$361,890	$723,780	$1,157,662
Est total costs (salary sub-total +5% loading)		**$379,984**	**$759,969**	**$1,214,545**
Cost per participant (n = 8398)		$45	$90	$145
Cost per high risk participant (n=1,278)		$297	$594	$951
BELGIUM				
Salaries* Doctors†	2.725	$593,791	$1,187,582	$1,781,373
Nurses‡	1.3625	$196,179	$392,359	$686,628
Salary sub-total		$789,970	$1,579,941	$2,468,001
Est total costs (salary sub-total +5% loading)		**$829,469**	**$1,658,938**	**$2,591,401**
Cost per participant (n = 8398)		$112	$224	$350
Cost per high risk participant (n=1,278)		$519	$1,039	$1,623
ITALY				
Salaries* Doctors†	2	$435,810	$871,620	$1,307,430
Nurses‡	1	$143,985	$287,970	$503,947
Salary sub-total		$579,795	$1,159,590	$1,811,377
Est total costs (salary sub-total +5% loading)		**$608,785**	**$1,217,569**	**$1,901,946**
Cost per participant (n = 8398)		$230	$461	$720
Cost per high risk participant (n=1,278)		$1,010	$2,019	$3,154

Notes:

* Nurses employed casually receive 25% loading to compensate for public holidays, recreation leave and recreation leave loading and sick leave, etc. This percentage has been added to the base salaries of the doctors and nurses to cover on-costs.

† Senior Registrar rate in 2003, Australian Medical Association Ltd., July 2003, *Rates Guide,* 3% deflation rate used to estimate current dollars for previous years.

‡ Secondary School Nurse rate in 2001, DHR, Agreement 2001, Nurses, salaries inflated using health inflation rates at June: 2002 = 3.41%, 2003= 6.83% (Source: {ABS, 2004 #251})

§ EFT = equivalent full time. Because of the lack of data, the staffing for the Belgium centre has been included at the 'average rate'.

¶ Assumes costs evenly distributed over the six years of the intervention.

8.5 Performance

Cost effectiveness

Preliminary cost-effectiveness ratios for reductions in predicted CHD risk is shown in Tables 8.9. It was estimated that reducing the predicted CHD risk by 1% from the baseline levels would be $615/participant in the UK, $1,242/participant in Belgium and $1,285/participant in Italy. A 1% reduction in the cumulative incidence of all-cause mortality would cost $28,547/participant in Belgium and $170,196 in Italy.

Table 8.11 Cost per reduction in one predicted CHD event at four years (2003 AUS$)

	UK		BELGIUM		ITALY	
Cost per 1,000 participants at four years	$90,494		$224,211		$460,851	
	All	**High Risk**	**All**	**High Risk**	**All**	**High Risk**
Baseline predicted CHD events	3.4	7.0	3.8	7.0	3.4	7.0
% change at 4 years	-12.8%	-19.1%	-15.8%	-20.0%	-31.4%	-37.9%
Follow-up predicted CHD events	3.0	5.7	3.2	5.6	2.3	4.3
Cost / reduction in one predicted CHD event at four years	$30,523	$15,980	$70,075	$40,038	$197,587	$106,016

8.6 Modelling

Modelling has not been performed for this intervention due to time constraints and the fact that the disease structure would require a completely separate economic model to other interventions.

9. Health checks by nurses in a primary care setting to reduce the risk factors for cardiovascular disease and cancer

9.1 Description

Type of intervention

The chapter is concerned with multifactorial interventions based in general practice. The exemplar was an intervention undertaken in the UK between 1989 and 1993. The aim of the intervention was to determine the effectiveness of health checks by nurses in a primary care setting in reducing risk factors for cardiovascular disease and cancer.

References/sources of evidence

The descriptions of the intervention and its effectiveness were based on the following published articles:

- *Prevalence of risk factors for heart disease in OXCHECK trial: implications for screening in primary care* (Imperial Cancer Research Fund, 1991)
- *Effectiveness of health checks conducted by nurses in primary care: results of the OXCHECK study after one year* (Imperial Cancer Research Fund, 1994)

The costs and modelling were based on the following published articles:

- *Costs and cost effectiveness of health checks conducted by nurses in primary care: the Oxcheck study* (Langham et al., 1996)
- *What can be concluded from the Oxcheck and British family heart studies: commentary on cost effectiveness analyses* (Wonderling et al., 1996)

Intervention description

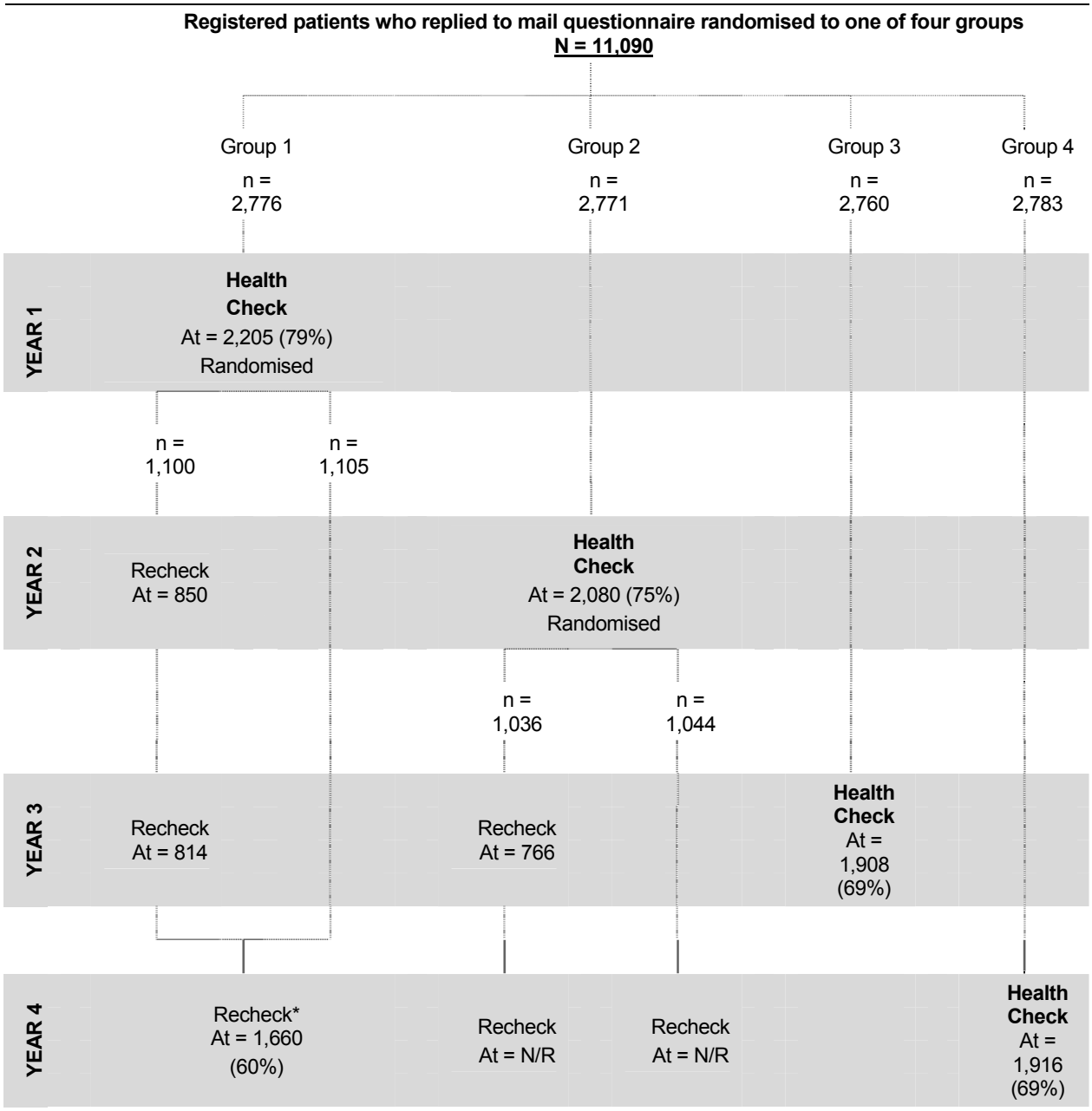
Recruitment and target population: The intervention targeted patients aged between 35 and 64 registered with five general practices in the UK. Recruitment was done in two stages – general practices were recruited first then participants were recruited from the lists of patients enrolled with the participating practices. Potential participants were sent a questionnaire in the mail to obtain information relating to behaviour, social status, attitudes, health and symptoms suggestive of ischaemic heart disease and intermittent claudication[7] and to check the accuracy of the mailing addresses. Patients responding to the mailed questionnaire were then invited to participate in the trial.

Intervention description: Patients participating in the trial were randomised by household to one of four groups, each with a different schedule to receive their first health check. Participants randomised to Group 1 received their first health check in 1989-90, some three years before those randomised to Group 4 were scheduled to receive theirs. (Figure 9.1) Practice Nurses undertook the health checks according to a standard protocol which included assessment of the risk factors pertaining to cardiovascular disease and cancer. (Figure 9.2) No formal risk score was used, but overall risk was assessed from a specially constructed visual chart which showed the interactive effect on the relative risk of blood pressure, smoking status and cholesterol concentrations. Nurses were instructed in the importance of identifying and following-up patients with multiple risk factors.

[7] A complex of symptoms characterized by the absence of pain or discomfort in a limb when at rest, the commencement of pain, tension and weakness after walking is begun, intensification of the condition until walking becomes impossible, and the disappearance of the symptoms after a period of rest. The condition is seen in occlusive arterials diseases of the limbs, such as thromboangiitis obliterans, and in compression of the cauda equina. Called also *Charcot's syndrome* and *angina cruris*. Taylor, E. J. (Ed.) 1988. *Dorland's Illustrated Medical Dictionary,* W.B. Saunders Company, Philadelphia.

There were set protocols for repeat measurement of high blood pressure and hyperlipidemia, but other follow-up was by mutual agreement between the nurse and the patient. Initial health checks were estimated to take 45-60 minutes, follow-up examinations 10-20 minutes, and the annual rechecks 30 minutes. (Figure 9.2) An external audit indicated that an average health check lasted 44 minutes (range 28-68 minutes).

Figure 9.1 Design of the Oxcheck trial



Sources: (Imperial Cancer Research Fund, 1994, Imperial Cancer Research Fund, 1995)

Notes:
At = number attending
N/R = data not reported in the literature
* The results for Group 1 participants who were randomised to annual rechecks (n=1,100) and those who did not have annual rechecks (n=1,105) were combined in the reporting of the three-year follow-up results. This was apparently done because there were no significant differences between the two groups on the main outcome measures in 1992-93, indicating that increasing the 'dosage' of the intervention did not have an impact on the outcomes.

Figure 9.2 Protocols for health checks and rechecks

HEALTH CHECK PROTOCOL	RECHECK PROTOCOL
Introduction (3 mins average, range 1-5 mins)	Briefer than the initial health check but included:
Information gathering (14 mins average, range 7-25 mins): Personal family history relating to ischaemic heart disease, stoke, hypertension, diabetes, cancer, smoking history, reported exercise rates, alcohol consumption and habitual diet	**Information gathering**: Smoking habit and dietary fat intake and alcohol.
Clinical measurement (12 mins average, range 7-24 mins): Measurement of height and weight (Seca scales with three monthly calibration); Blood pressure (Hawksley random zero sphygomomanometer); Blood samples for cholesterol concentration; BMI calculated	**Clinical measurement**: Height, weight, blood pressure, blood samples for serum cholesterol concentration, BMI calculation. Reported smoking cessation was confirmed by measuring serum continine concentration.
T**arget negotiation and health education** (15 mins average, range 4-24 mins)	

Source: (Imperial Cancer Research Fund, 1994)

9.2 Quality of evidence

Evaluation description

Design: Two articles were published relating to the effectiveness of the intervention. One purported to present findings at 12 months and the other at three years. Each study used different sets of participants for comparison across groups. The 12 month evaluation compared data relating to the 3,988 people in Groups 2 and 3 who received a health check in years 2 and 3 and data for the 2,136 and in Groups 1 and 2 who received a recheck in the corresponding period. (Figure 9.1) Because the design of this evaluation appeared to be comprised by the inclusion of the 1,036 people in Group 2 randomised to receive a recheck in Year 3 in both the 'control' and 'intervention' groups it has not been considered. The three-year evaluation which used data for Groups 1 and 4 was used in this report as the basis for examining the effectiveness of the intervention. The design of this evaluation was quasi-experimental with Group 4 being designated the 'control' group and Group 1 the 'intervention' group.

Methodology: Because there was no baseline data for the control group in the evaluation, the methodology involved the following comparisons:

- Pre - post comparisons for the intervention group;
- Comparison of trial end data for the intervention and control groups.

The results presented in this report related to the comparison of the intervention and control groups at trial end. The underlying assumptions of this methodology were that the two randomised groups were comparable at baseline and that the differences between the two groups at trial end reflected the impact of the intervention.

Outcome measures: The outcome measures involved a combination of self-report, anthropometric data collected during interview and pathology reports from blood samples taken during interview. However complete data were not available for all measures for both the intervention and control groups. Therefore discussion of the results has been limited to those measures for which there were adequate data. The impact on service use was examined for a sample of the intervention group but there was no similar analysis for the control group.

Analysis: The analysis for the intervention group was on an attenders-only and an intention-to-treat basis. The intention-to-treat analysis for the intervention group used the last recorded value for those for who did not attend. The Imperial Cancer Research Fund (1994) authors note that:

> 'This assumption is generally conservative, except in the case of cholesterol concentration and blood pressure, which rise with age. The effect of this assumption was modelled by adjusting the means for those with missing values to recheck for age related changes. The effects were negligible (a 0.01 mmol/l difference in cholesterol concentration on women only), and unadjusted values were therefore used in the subsequent analysis.'

Lack of baseline data meant that the only analysis possible for the control group was on an attenders-only basis.

Assessment

Sources of bias: The potential sources of bias in the study included selection bias, measurement bias and attrition bias.

Selection bias: The only information relating to the comparability of the groups at baseline was the statement that: 'There was no significant difference between the four groups … in terms of age, sex, social class or marital status. The mean (SD) age of groups 1 and 4 were 49.3 (8.5) and 49.4 (8.6) years respectively.' (Imperial Cancer Research Fund, 1994) To the extent that randomization by household did not result in comparable groups, then the lack of baseline data for the control group may have resulted in bias being introduced into the study due to a non-comparability of the groups that could not be controlled for in the analysis.

Measurement bias: Measurement error in the objective data was minimised by having the pathology tests undertaken centrally, using the same instruments across all sites, and regular quality checks on the instruments used to gather the data (eg three monthly calibration of the scales). In terms of the self-report data, participants who were classed as smokers at their first health check were only reclassified as non-smokers if their report of having quit was confirmed by measurement of serum cotinine concentration. However, there were factors in the analysis and presentation of results which may have introduced measurement bias.

Firstly, the unit of analysis was the individual. However, participants were randomised by household and were registered patients at five general practices. To the extent that there are similarities between the members of a household and the registered patients at a general practice, the results may be biased by an individual's membership of these different units of analysis. This was addressed in part by a test for heterogeneity between the five practices in the differences between the intervention and control groups' mean serum cholesterol concentration, BMI and diastolic blood pressure. The results for the five practices were pooled because 'no heterogeneity was found'. (Imperial Cancer Research Fund, 1994) Secondly, in the presentation of the results there was no indication of how much 'missing data' there was and how this was handled. In the results reproduced in Table 9.3 in this report, there was no indication of the number of valid responses for each outcome measure.

Attrition bias: The disposition of the participants in Groups 1 and 4 indicated differences in attendance rates at the health checks and rechecks. Seventy-nine percent of randomised Group 1 participants attended their health check in year one, 69% of Group 4 participants attended theirs in year 4 and 60% of Group 1 participants attended their recheck in that same year. (Table 9.1) To the extent that these differences in presentation represent attrition of a particular group of participants, these differences had the potential to systematically bias the results presented in the three-year evaluation.

Table 9.1 Differences in attendance rates between Groups 1 and 4

Disposition	Group 1 (Intervention)				Group 4 (Control)	
	Health Check (Yr 1)		Recheck (Yr 4)		Health Check (Yr 4)	
	Number	Percent	Number	Percent	Number	Percent
Attended	2,205	79%	1,660	60%	1,916	69%
Known to have moved	94	3%	267	10%	426	15%
Non-attendance	477	17%	819	31%	441	16%
Total Randomised	2,776	100%	2,776	100%	2,783	100%

Statistical Significance: χ^2 = 437.374, df = 4, p = 0.000 (not included in the literature, calculated at Centre for Health Economics)

Sources: (Imperial Cancer Research Fund, 1991, Imperial Cancer Research Fund, 1994, Imperial Cancer Research Fund, 1995)

9.3 Outcomes – as reported

Table 9.2 shows the behavioural and clinical parameters that were measured, the data sources and the methods of analyses used in the evaluation. Because of a lack of data for some of the outcomes, notably diet and exercise have not been included in this report. This omission is unlikely to affect the quality of the report as the modelling is based on a calculation of the Dundee Risk Score which does not include diet and exercise in the formula used to calculate the risk score. (Tunstall-Pedoe, 1991)

Table 9.2 Outcome measures, data sources and analyses reported in the study

Outcome measure	Data sources	Analyses
BEHAVIOUR CHANGE		
Smoking	Self-report in interview with Practice Nurse and for those who reported quitting, analysis of serum cotinine at a central location from a blood sample taken by the Practice Nurse.	Number of smokers: Comparisons between baseline intervention group, end of the trial intervention group, end of the trial non-intervention group (Chi sq test).
Alcohol use	Self-report in interview with Practice Nurse	Number using unsafe levels of alcohol: Comparison of end of the trial intervention group, end of the trial non-intervention group (Chi sq test).
Exercise	Self-report in interview with Practice Nurse	Number inactive ie exercising vigorously <1/month. Comparison end of the trial intervention group, end of the trial non-intervention group (Chi sq test).
Diet	Self-report in interview with practice nurse	Number using mainly full cream milk: Comparison end of the trial intervention group, end of the trial non-intervention group (Chi sq test). Number using mainly butter/hard margarine: Comparison end of the trial intervention group, end of the trial non-intervention group (Chi sq test).
CLINICAL PARAMETERS		
Cholesterol levels	Blood sample taken during health check with Practice Nurses and sent to a central location for analysis.	Mean: Comparison of the end of the trial differences between intervention and non intervention groups (t-test) Number with cholesterol level ≥8 mmol/l: Comparison between baseline intervention group, end of the trial intervention group, end of the trial non-intervention group (Chi sq test).
Blood pressure	Taken by the Practice Nurses during the health checks and rechecks.	Mean: Comparison of the end of the trial differences of between the intervention and non intervention groups (t-test) Number with diastolic blood pressure ≥100 mm Hg: Comparison between baseline intervention group, end of the trial intervention group, end of the trial non-intervention group (Chi sq test).

Outcome measure	Data sources	Analyses
BMI	Calculated from height and weight measurements taken by the Practice Nurses during the health checks and rechecks	Mean: Comparison of the end of the trial differences of between the intervention and non intervention groups (t-test) Number with BMI $\geq$30 kg/m ^{2}: Comparison between baseline intervention group, end of the trial intervention group, end of the trial non-intervention group (Chi sq test).

Behaviour change

Smoking: At the end of the trial, the proportion of people smoking in the intervention group was lower than in the control group (5% lower for attenders-only and 1.4% lower for intention-to-treat) but only the attenders-only difference was statistically significant (95% CI 2.2,7.8). (Table 9.3)

Alcohol use: At the end of the trial there were only small differences between the control and intervention groups in the proportion of people with unsafe levels of alcohol use (1.6% difference attenders-only and 0.6% intention-to-treat) and the differences were not statistically significant. The patterns of results were the same for both men and women. (Table 9.3)

Clinical parameters

Cholesterol: At the end of the trial, there were a lower proportion of people in the intervention group than in the control group with elevated cholesterol and the differences between the groups were statistically significant. The results were the same for both women and men. (Table 9.3)

Blood pressure: Compared to the control group, there was a smaller proportion of people in the intervention group with elevated blood pressure but the differences between the groups were small (1.2% attenders-only and 1.1% intention-to-treat) and not statistically significant. (Table 9.3)

BMI: Compared to the control group, the proportion of obese people in the intervention group was lower, but only the difference between the attenders-only group was statistically significant (2.4% 95% CI 0.0, 4.7). When the disaggregated results for women and men were analysed the small differences between the control and intervention groups were not statistically significant. (Table 9.3)

Table 9.3 Patients in high risk categories

	Group 4 Control Group		Group 1 Intervention Group				Control – Intervention Difference			
	Attenders Only		Attenders-only		Intention-to-treat*		Attenders-only		Intention-to-treat	
	Number	Percentage (a)	Number	Percentage (b)	Number	Percentage (c)	Percentage (a) – (b)	95% CI	Percentage (a) – (c)	95% CI
Women and Men	**1,916**		**1,660**		**2,205**					
Smoking†	506	26.4%	356	21.4%	552	25.0%	5.0%	2.2 - 7.8	1.4%	-1.3 to 4.1
Alcohol use‡	210	11.0%	156	9.4%	229	10.4%	1.6%	-0.4 - 0.0	0.6%	-1.3 – 2.5
Total cholesterol ≥8 mmol/l	148	7.8%	49	3.1%	82	3.9%	4.7%	3.2 – 6.2	3.9%	2.4 – 5.3
Diastolic BP ≥ 100 mm Hg	86	4.5%	53	3.3%	73	3.4%	1.2%	-0.1 – 2.5	1.1%	-0.1 – 2.3
BMI ≥ 30 kg/m^2	304	15.9%	220	13.5%	310	14.3%	2.4%	0.0 – 4.7	1.6%	-0.6 – 3.8
Women Only	**1,031**		**922**		**1,218**					
Smoking†	236	22.9%	166	18.0%	256	21.0%	4.9%	1.3 – 8.5	1.9%	-1.6 – 5.3
Alcohol use‡	55	5.3%	44	4.8%	65	5.3%	0.6%	-1.4 – 2.5	0.0%	-1.9 – 1.9
Total cholesterol ≥8 mmol/l	92	9.0%	30	3.4%	48	4.1%	5.6%	3.4 – 7.7	4.8%	2.7 – 6.9
Diastolic BP ≥ 100 mm Hg	37	3.6%	25	2.8%	34	2.9%	0.8%	-0.8 – 2.4	0.7%	-0.7 – 2.2
BMI ≥ 30 kg/m^2	190	18.4%	147	16.3%	207	17.3%	2.2%	-1.2 – 5.6	1.2%	-2.0 – 4.4
Men Only	**885**		**738**		**987**					
Smoking†	270	30.5%	190	25.7%	296	30.0%	4.8%	0.4 – 9.1	0.5%	-3.7 – 4.7
Alcohol use‡	155	17.5%	112	15.2%	164	16.6%	2.3%	-1.3 – 5.9	0.9%	-2.5 – 4.3
Total cholesterol ≥8 mmol/l	56	6.4%	19	2.7%	34	3.6%	3.7%	1.7 – 5.7	2.8%	0.8 – 4.8
Diastolic BP ≥ 100 mm Hg	49	5.5%	28	3.9%	39	4.0%	1.6%	-0.4 – 3.7	1.5%	-0.4 – 3.5
BMI ≥30 kg/m^2	114	12.9%	73	10.1%	103	10.6%	2.8%	-0.3 – 5.9	2.3%	-0.6 – 5.2

Source: (Imperial Cancer Research Fund, 1995)

Notes:
* Last value from health check or recheck used for non-attenders. † Smoking any form of tobacco at least daily. ‡Reported weekly intake of >21 units for men and >14 units for women.

9.4 Program costs

As reported

The cost data reported in the literature were 'extracted from detailed records kept during the trial'. (Langham et al., 1996) Costs related to research staff, conferences, meetings, data handling, equipment that would not be used in normal clinical practice (eg hospital based laboratory estimation of cholesterol and the particular sphygmomanometer used) were excluded because they were considered to be 'research' costs. It was assumed that time spent on administration and recruitment was evenly spread across all visits. The total cost of the trial reported in the literature was £329,686. Rechecks were considered research costs and excluded from the calculation of the cost of implementing the intervention. Therefore the programme costs without the research component of the trial was £237,374 (AUS$531,891). The total cost per person randomised to the trial was AUS$94 and the cost of the intervention per person randomised was AUS$70. (Table 9.4)

Table 9.4 Program costs as reported in the literature

	Total	Health Check*	Follow-up	Intervention Costs¶	Recheck (Research)
Proportion of Total Costs (based on Practice Nurse time)	1.00	0.48	0.24	0.72	0.28
Equipment including quality assurance	£ 10,190	£ 4,891	£ 2,446	£ 7,337	£ 2,853
Consumables	£ 21,198	£ 10,175	£ 5,087	£ 15,262	£ 5,935
Overheads†	£ 40,000	£ 19,200	£ 9,600	£ 28,800	£ 11,200
Nurse support and supervision‡	£ 69,142	£ 33,188	£ 16,594	£ 49,782	£ 19,436
Recruitment and training of nurses*	£ 6,517	£ 3,128	£ 1,564	£ 4,692	£ 1,825
Practice nurses	£ 142,749	£ 68,520	£ 34,260	£102,779	£ 39,970
Administration	£ 39,618	£ 19,017	£ 9,508	£ 28,525	£ 11,093
Total	£ 329,686	£158,249	£ 79,125	£237,374	£ 92,312
A$ Jun 2003§	$1,071,626			$771,570	
Cost per person randomised (n = 11,090)	$ 93.63			$ 69.57	
Cost per person having an initial health check (n = 8,109)	$ 132.15			$ 95.15	

Source: Langham *et al.*, 1996
Notes:
* A discount rate of 6% over five years was used to calculate an annual equivalent cost, which was then multiplied by four to derive a four-year cost
† Based on the cost of a dedicated serviced room (10 m^2) at £2,000 per year
‡ Included salary of one nurse coordinator plus other expenses such as travel and administration
§ Average annual exchange rate for 1992-1993 financial year = 2.304858 (Source: http://www.x-rates.com/cgi-bin/hlookup.cgi accessed 27th February 2004). Annual CPI health rates for financial years ending June: 1994 = 4.66%, 1995 = 5.94%, 1996 = 5.26%, 1997 = 6.32%, 1998 = 3.57%, 1999 = -1.21%, 2000 = -2.88%, 2001 = 3.53%, 2002 = 3.41%, 2003 = 6.83% (Source: (ABS, 2004)
¶ Health checks plus follow-up only. In the literature, the costs of the re-examinations determined by the protocol were considered to be research costs as these were conducted exclusively for the trial.

Based on resource use

In calculating the costs of the intervention based on resource use, it was assumed that Group 1 participants (the intervention group) were recruited, underwent an initial health check, had a series of follow-up visits in year 1 and then underwent a recheck in year 4. The control group (Group 4) participants were recruited then it was assumed that they only underwent a health check comparable to the Group 1 recheck in year 4 for the purposes of comparison with Group 1 participants. The total time that practice nurses spent on each type of visit and the number of participants and visits are shown in Table 9.5. This table was used as the basis for allocating Practice Nurse time between intervention and control group clients.

Table 9.5 Time spent by practice nurses on initial health check, follow-ups and rechecks

	Health Checks		Follow-up		Rechecks		Total
	Hours	Participants	Hours	Visits	Hours	Participants	Hours
Year 1	1,654	2,205*	838	2,514†	0	0	2,492
Year 2	1,560	2,080	790	2,371	425	850¶	2,775
Year 3	1,431	1,908	725	2,175	791	1,582	2,947
Year 4	1,437	1,916‡	728	2,184§	2,433	3,244**	4,598
Total¶	**6,082**	**8,109**	**3,081**	**9,244**	**3,649**	**5,677**	**12,812**

Source: Langham *et al.*, 1996

Notes:

* Group 1 participants

† Group 1 participants attending follow-up examinations based an audit which showed that each health check generated an average of 1.14 follow-up visits.

‡ Group 4 participants attending their first health check.

§ Group 4 participants attending follow-up examinations based on an audit which showed that each health check generated an average of 1.14 follow-up visits.

¶ Group 1 participants – see Figure 9.1.

** Includes 1,660 Group 1 participant attending for final recheck – see Figure 9.1.

Based on an estimate of resources used in the trial, the total cost of implementing the trial without the research component was estimated to be $166,727 for the intervention group and $69,538 in June 2003 dollars. The cost per intervention group completer was $100 and $36 for control group completers, a difference of $64. (Table 9.6)

Table 9.6 Estimated costs of Oxcheck intervention based on resource use (A $ financial year ended June 2003)

	June 2000	June 2003		
		Jun 2000 inflated**	Rechecks	Total
Group 1 – Intervention Group				
Recruitment*	**$ 26,452**	**$ 30,254**		**$ 30,254**
Practice nurse†	**$ 56,847**	**$ 65,018**	**$ 21,570**	**$ 86,587**
Nurse coordinator‡	**$ 13,721**	**$ 15,693**	**$ 5,206**	**$ 20,899**
Overheads§	**$ 1,159**	**$ 1,325**	**$ 421**	**$ 1,746**
Other¶	**$ 17,932**	**$ 20,509**	**$ 6,799**	**$ 27,308**
Intervention Total	**$116,111**	**$132,799**	**$ 33,996**	**$166,795**
Cost per intervention person randomised			(n = 2,776)	$ 60.08
Cost per intervention person completing recheck in year 4			(n = 1,660)	$ 100.48
Group 4 - Control Group				
Recruitment*	$26,519	$30,330		$ 30,330
Practice nurse†			$ 24,896	$ 24,896
Nurse coordinator‡			$ 6,009	$ 6,009
Overheads§			$ 486	$ 486
Other¶			$ 7,848	$ 7,848
Control Total			$ 39,239	$ 69,569
Cost per intervention person randomised			(n = 2,783)	$ 24.99
Cost per intervention person completing health check in year 4			(n = 1,916)	$ 36.29

Source: Langham *et al.*, 1996

Notes:

* Based on total mailout of 37,741 (17,965+ 11,498 + 8278) response rates of 36%, 28% and 17% and total response of 11,090. Costs allocated on basis of number of respondents randomised to each group. Did not include labour costs as assume the administrative aspects of the mailout would be undertaken by the practice nurses over time.

† Source: (Department of Human Services, 2001), p. 18, mid point of HCS 3 as at 1 July 2001 and 1 July 2002, 25% loading for casual rate, 38 hour week, 52 weeks per year. Time allocation to Group 1 consisted of health checks of 1,654 hours in year 1 for initial health checks and 2433 hours x1660/3244 in year 4. Time allocation for Group 4 consisted of rechecks of 2433 hoursx1916/3244 in year 4.

‡ Source: (Department of Human Services, 2001), p. 18, mid point of HCS 5 as at 1 July 2001 and 1 July 2002, 25% loading for on-costs, 38 hour week, 52 weeks per year, costs allocated according to each group according to Practice Nurse time. Fraction for Group 1 in year 1 = 0.19, Group 1 in year 4 0.06, Group 4 in year 4 = 0.07. Does not include any travelling and accommodation costs.
§ Commercial quotation of $6,500 pa for a room with outgoings paid by the owner received in 2003 from O'Donoghues First National Real Estate. Costs apportioned according to Practice Nurse time.
¶ Other included equipment including quality assurance, consumables, nurse recruitment and training, and administration. In the published costs these other items represented 23.5% of the overheads + Practice Nurse + nurse support and supervision costs at the initial checks, 27.5% at follow-up and 23.5% at the recheck. Therefore, assumed to be 25% of the overhead + Practice Nurse cost + Nurse Coordinator costs for each group.
** CPI for health in financial years ending June : 2001 3.53%, 2002 3.41%, 2003 6.83%. (Source: (ABS, 2004))

9.5 Performance

Cost effectiveness

The aim of the intervention was to determine the effectiveness of health checks by nurses in a primary care setting in reducing risk factors for cardiovascular disease and cancer. The effectiveness of the intervention has been reported in terms of a net reduction in the prevalence of each of the risk factors based on published costs and resource use costs in June 2003 Australian dollars. In those terms, the most cost effective impacts of the intervention were on smoking ($49,759 to $47,273) and cholesterol ($52,935 to $50,290). (Table 9.7)

Table 9.7 Cost effectiveness of the Oxcheck trial in reducing the prevalence of the risk factors based on published and resource use costs (2003 AUS$)

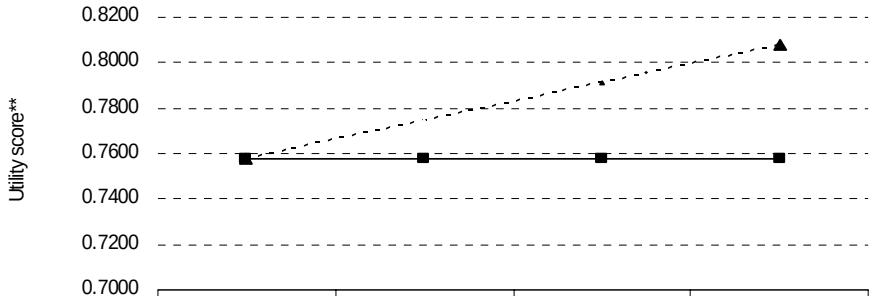
Risk Factor	Baseline * (a)	Follow-up		Number of changers		Cost per changer‡	
		No Inter-vention† (b)	Inter-vention (c)	(b)-(c) (d)	(a)-(c) (e)	Based on (d)	Based on (e)
Smoking	629	582	552	-30	-77	$ 5,560	$2,166
Alcohol Use	N/A	242	229	-13	N/A	$12,830	
Cholesterol ≥ 8 mmol/l	165	172	82	-90	-83	$ 1,853	$2,010
Diastolic BP ≥ 100mm Hg	50	99	73	-26	+23	$ 6,415	
BMI >30 kg/m^2	290	351	310	-41	+20	$ 4,068	

Notes:
* Source: Imperial Cancer Research Fund (1991), p 1058
† Percentages in the control group at year 4 applied to 2,205 intervention group participants at baseline (see Table 9.3)
‡ Total cost for intervention group = 2003 AUS$166,795 (see Table 9.6)
N/A not available in the literature

Cost-utility

The preliminary cost-utility analyses were based on the difference in utility levels for people changing from a BMI ≥30 kg/m^2 to a BMI less than this amount (ie for those becoming non-obese). The utility scores for obese and non-obese persons were based on a Brazier (Brazier, 2002) transformation of the Sf-36 scores in the National Nutrition Survey (1995) for people aged 18+ years. (Figure 9.3) Over the four years of the intervention, net increase in utility is 0.1000 and based on a cost per BMI changer of $4,068 (Table 9.7), the cost per unit increase in utility is $40,680.

Figure 9.1 Cost of increasing utility levels by one unit based on



	Year 1	Year 2*	Year 3*	Year 4
Changer	0.7575	0.7742	0.7908	0.8075
Non Changer	0.7575	0.7575	0.7575	0.7575

BMI

Notes:

* Estimated as a continuous increase from Year 1 to Year 4

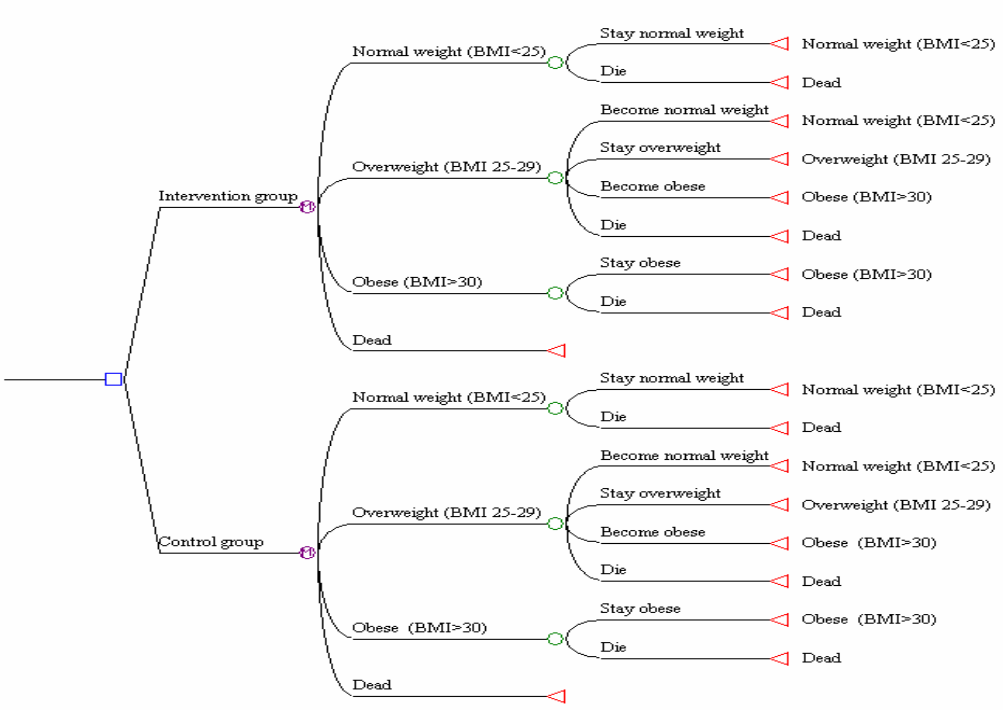
** Data Source: National Nutrition Survey (1995) CURF, Ref No 691, Brazier {Brazier, 2002 #108} transformation of the SF-36 scores for people with BMI 18.5 -30 kg/m^2 and people with BMI ≥30 kg/m^2

9.6 Modelling

Methods

A modelling approach was used to enable the short term outcomes (reported by Imperial Research Fund, 1995) to be extrapolated longer term and translated into life-years saved and QALYs gained. A Markov process structure was developed comprising 1 year cycles. The time horizon of the model was 20 years. The model includes the health states normal weight (BMI<25kg.m^2), overweight (BMI 25 to 29.9kg/m^2), obese (BMI>30kg/m^2) and dead. The transitions that are permitted are illustrated in Figure 9.4.

Figure 9.4 Representation of states and permitted transitions in Markov model



We determined the progression, costs and utilities of a cohort of 1000 people receiving the Oxcheck intervention compared to a control group.

The cohort progressed annually between health states over 20 year time horizon according to transition probabilities derived from the published literature. The model commences with 49% of people with normal weight, 38% overweight and 13% obese in both the intervention and control groups (Imperial Cancer Research Fund, 1991).

The following transition probabilities are incorporated into the model:

- Each year for the first 4 years of the model 3.7% of overweight people become obese in the intervention group. This is calculated from 290/2205 having a BMI>30kg/m^2 at baseline minus 310/2205 at 4 year follow up (Table 7.9). The four year rate is translated into a 1 year probability using the formula 1-EXP(LN(1-number of cycles)/four-year probability)
- Each year for the first 4 years of the model 4.2% of overweight people become obese in the control group. This is calculated from 304/1916 having a BMI>30kg/m^2 at 4 year follow up, which is assumed not to have changed during the study period (Table 7.9). The four year rate is translated into a 1 year probability using the formula 1-EXP(LN(1-number of cycles)/four-year probability)
- Over the 20 years of the model death is time dependent and is different for each category of weight. Probabilities of death for each year are determined by fitting a Weibull curve to survival curves in the paper by Peeters et al (2003). The probabilities of death are weighted for a population that is 50.7% female (ABS 2002) with 27.3% of males assumed to smoke and 21.4% of females (ABS National Health Survey 2001)

In addition the following assumptions have been made:

- The control group do not change their weight
- The intervention effect is assumed to last for 4 years after which no additional weight gain is permitted

The cost per person for the Oxcheck intervention are reported in Table 9.6 and are estimated as $100.48 per person in the intervention group and $36.29 for the control group for the 4 years of the intervention. These figures are based on the cost per person completing their four year health check (most conservative for the base case analysis). This equates to $25.12 per year for the intervention group and $9.07 per person per year in the control group for the base case analysis. The cost per person randomised is included in sensitivity analysis. The downstream costs of being overweight or obese are not included in the base case analysis but are considered in sensitivity analysis (threshold analysis).

Utilities are assumed to be 0.85 for those with normal weight, 0.82 for those overweight and 0.78 for those with obese weight (McNeil & Segal, 1999). Costs and benefits are discounted at 5% per annum (Australian Treasury).

Extensive univariate sensitivity analyses were performed for the assumptions and values described in Table 9.8.

Table 9.8 Sensitivity analysis: attributes, base case and alternative assumed values

Assumptions	Base case	Alternative Values	Source
Time horizon	20 years	5 and 10 years	Researcher judgment
Discount rate	5%	0% and 3%	Researcher judgment
Length of intervention benefit	4 year	5 and 10 years	Researcher judgment
Utility of overweight	0.82	0.79	Utilities from Hakin et al, 2002
Utility of obese	0.78	0.76	Utilities from Hakin et al, 2002
Cost per person randomised,	I=$25.12 C=$9.07	I=$15.02 C=$6.25	Table 9.6

I- intervention group, C- control group

In addition a threshold sensitivity analysis was conducted to determine the downstream cost associated with the obese state which would lead to the intervention being dominant.

Results

Table 9.9 presents the economic performance of the Oxcheck intervention compared to a control group, and an incremental cost-utility ratio of $12,613 per QALY gained (for base case assumptions, see Table 9.8).

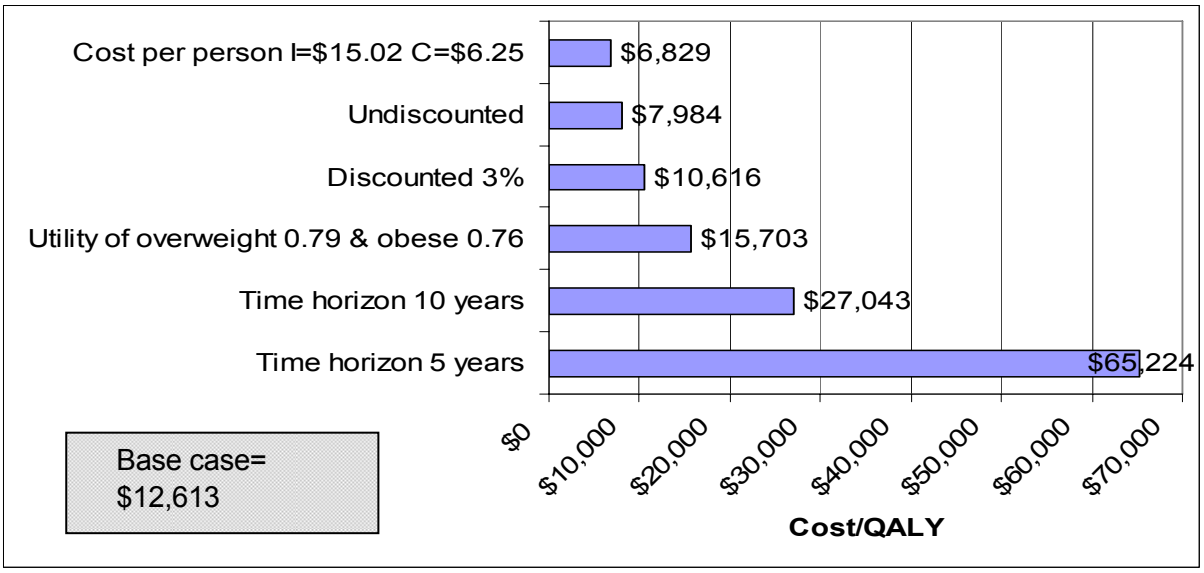
Table 9.9 Modelled cost-utility base case results

	Oxcheck intervention	Control group	Difference
Total costs	$89.10	$32.20	$56.90
Total life years	12.2792	12.2778	0.0014
Total QALYs	10.1599	10.1554	0.0045
Discounted $/LY gained			$41,459
Discounted $/QALY gained			$12,613

Sensitivity analyses

Sensitivity analyses ranged from $6,829 per QALY to $65,224 per QALY (Figure 9.5). Results were most sensitive to the cost of the intervention and the time horizon of the model.

Figure 9.5 Results of sensitivity analyses



Inclusion of downstream costs

A threshold analysis was performed for inclusion of downstream costs for the obese health state. It was found that if downstream costs associated with being obese were greater than an average of approximately $405 per person per year (over a 20 year time period) then the Oxcheck intervention would dominate the control group.

References

Astrup A, Finer N. Redefining type 2 diabetes: 'diabesity' or obesity dependent diabetes mellitus. *Obes Rev* 2000; **1**(2):57-59.

Australian Bureau of Statistics (ABS). Deaths, 3302.0. Canberra, 2000.

Australian Bureau of Statistics (ABS). Census of population and housing: selected social and housing characteristics, 2015.0. Canberra, 2002.

Australian Bureau of Statistics (ABS). Consumer Price Index, Australia. Canberra, 2004.

Australian Bureau of Statistics (ABS). National Health Survey. Canberra, 2001.

Australian Bureau of Statistics (ABS). National Nutrition Survey CURF. Canberra, 1995.

Australian Institute of Health and Welfare (AIHW). Australia's children: their health and wellbeing 2002, Cat No. PHE 36. Canberra.

Australian Institute of Health and Welfare (AIHW). Australia's Health 2000. Canberra.

Australian Institute of Health and Welfare (AIHW). Australia's young people: their health and wellbeing 2003, Cat No. PHE 50. Canberra.

Balkau E, Eschwege E, Papoz L, Richard JL, Claude JR, Warnet JM, Ducimetiere P. Risk factor to early death in non-insulin dependent diabetes and men with known glucose tolerance. *BMJ* 1993; **307**:295-299.

Booth M. Assessment of physical activity: an international perspective. *Res Q for Exerc Sport* 2000; **71**:114-120.

Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002; **21**:271-292.

Cameron M, Haxworth N, Oxley J, Newstead S, Le T. Evaluation of transport accident commission road safety television advertising, Report no 52. Melbourne: Monash University, Accident Research Centre, 1993.

Colagiuri S, Colagiuri R, Conway B, Grainger D, Davey P. *DiabCost Australia: assessing the burden of type 2 diabetes in Australia*. Diabetes Australia: Canberra, 2003.

Dalton A, Segal, L. The health care cost of diabetes in Australia, Chapter 5: *The Rise and Rise of Diabetes in Australia, 1996.* A Report to the National Action Plan Committee: McCarty D, Zimmet P, Dalton A, Segal L, Welborn T. International Diabetes Institute: Victoria, Melbourne.

Department of Human Services. Nurses (Department of Human Services) Agreement 2001. Melbourne.

Dunstan DW, Zimmet PZ, Welborn TA, de Courten MP, Cameron AJ, Sicree RA, Dwyer T, Colgiuri S, Jolley D, Knuiman M, Atkins R, Shaw JE. The rising prevalence of diabetes and impaired glucose tolerance: the Australian Diabetes, Obesity and Lifestyle study. *Diabetes Care* 2002; **25**:829-834.

Egger G, Bolton A, O'Neill M, Freeman D. Effectiveness of an abdominal obesity reduction program in men: the GutBuster 'waist loss' program. *Int J Obes* 1996; **20**:227-231.

Eriksson KF, Lindegärde K. Prevention of Type II (non-insulin dependent) diabetes mellitus by diet and physical exercise: the six year Malmo feasibility study. *Diabetologia* 1991; **34**:891-898.

Eriksson KF, Lindgarde F. Poor physical fitness, and impaired early insulin response but late hyperinsulinaemia, as predictors of NIDDM in middle-aged Swedish men. *Diabetologia* 1996; **39**(5):573-579.

Farquhar JW, Fortmann SP, Flora JA, Taylor CB, Haskell WL, Williams PT, Maccoby N, Wood PD. Effects of communitywide education on cardiovascular disease risk factors: the Stanford Five-City Project. *JAMA* 1990; **264**:359-365.

Farquhar JW, Fortmann SP, Maccoby N, Haskell WL, Williams PT, Flora JA, Taylor CB, Brown BW, Jr, Solomon DS, Hulley SB. The Stanford Five-City Project: design and methods. *Am J Epidemiol* 1985; **122**:323-334.

Fortmann SP, Varady AN. Effects of a community-wide health education program on cardiovascular disease morbidity and mortality: the Stanford Five-City Project. *Am J Epidemiol* 2000; **152**:316-323.

Fortmann SP, Flora JA, Winkleby MA, Schooler C, Taylor CB, Farquhar JW. Community intervention trials: reflections on the Stanford Five-City Project Experience. *Am J Epidemiol* 1995; **142**:576-586.

Fortmann SP, Taylor CB, Flora JA, Winkleby MA. Effect of community health education on plasma cholesterol levels and diet: the Stanford Five-City Project. *Am J Epidemiol* 1993; **137**:1039-1055.

Fortmann SP, Winkleby MA, Flora JA, Haskell WL, Taylor CB. Effect of long-term community health education on blood pressure and hypertension control: the Stanford Five-City Project. *Am J Epidemiol* 1990; **132**:629-646.

Gortmaker S, Peterson K, Wiecha J, Sobol A, Dixit S, Fox M, Laird N. Reducing obesity via a school-based interdisciplinary intervention among youth. *Arch Pediatr Adolesc Med* 1999; **153**:409-418.

Hakin Z, Wolf A, Garrison LP. Estimating the effect of changes in body mass index on health state preferences. *Pharmacoeconomics* 2002; **20**:393-404.

Harrell J, McMurray R, Bangdiwala S, Frauman A, Gansky S, Bradley C. Effects of a school-based intervention to reduce cardiovascular disease risk factors in elementary-school children: the cardiovascular health in children (CHIC) study. *J Pediatr* 1996; **128**:797-805.

Imperial Cancer Research Fund, O. S. G. Prevalence of risk factors for heart disease in OXCHECK trial: implications for screening in primary care. *BMJ* 1991; **302**:1057-1060.

Imperial Cancer Research Fund, O. S. G. Effectiveness of health checks conducted by nurses in primary care: results of the OXCHECK study after one year. *BMJ* 1994; **308**:308-312.

Imperial Cancer Research Fund, O. S. G. Effectiveness of health checks conducted by nurses in primary care: final results of the OXCHECK study. *BMJ* 1995; **310**:1099-1104.

Killen J, Telch M, Robinson T, Maccoby N, Taylor B, Farquhar J. Cardiovascular disease risk reduction for tenth graders: a multiple-factor school-based approach. *JAMA* 1988; **260**:1728-1733.

Kornitzer M, Rose G. WHO European Collaborative Trial of multifactorial prevention of coronary heart disease. *Prev Med* 1985; **14**:272-278.

Langham S, Thorogood M, Normand C, Muir J, Jones L, Fowler G. Costs and cost effectiveness of health checks conducted by nurses in primary care: the Oxcheck study. *BMJ* 1996; **312**:1265-1268.

Manson A, Stampfer MJ, Hemekens CH, Willett WC. Body weight and longevity. *JAMA* 1987; **257**:353-358.

McNeil H, Segal L. Quality of life and obesity, Research Report 17. Melbourne: Monash University, Centre for Health Program Evaluation, 1999.

Miles A, Rapoport L, Wardle J, Afuape T, Duman M. Using the mass-media to target obesity: an analysis of the characteristics and reported behaviour change of participants in the BBC's 'Fighting Fat, Fighting Fit' campaign. *Health Educ Res* 2001; **16**:357-372.

Moore L. Sustained weight loss reduces diabetes mellitus risk in overweight individuals *Epidemiology* 2002; **11**:269-273.

Peeters A, Barendregt JJ, Willekens F, Mackenback, JP, Mammun AA, Bonneux L. Obesity in adulthood and its consequences for life expectancy: a life table analysis. *Ann Intern Med* 2003; **138**:24-32

Rissanen A, Heliovaara M, Knekt P, Reuanen A, Aroma A, Maatela J. Risk of disability and mortality due to overweight in a Finnish population. *BMJ* 1990; **301**:835-837.

Robinson TN. Reducing children's television viewing to prevent obesity: a randomised controlled trial. *JAMA* 1999; **282**:1561-1567.

Rockwood K, Awalt E, MacKnight C, McDowell I. Incidence and outcomes of diabetes mellitus in elderly people: report from the Canadian Study of Health and Ageing. *Can Med Assoc J* 2000; **162**(6):769-771.

Roe L, Hunt P, Bradshaw H, Rayner M. Dietary intervention in primary care: validity of the DINE method for diet assessment. *Fam Pract* 1994; **11**:375-381.

Taylor CB, Fortmann SP, Flora J, Kayman S, Barrett DC, Jatulis D, Farquhar JW. Effect of long-term community health education on body mass index: the Stanford Five-City Project. *Am J Epidemiol* 1991; **134**:235-249.

Taylor EJ. (Ed) Dorland's Illustrated Medical Dictionary. W.B. Saunders Company: Philadelphia, 1988.

Tunstall-Pedoe H. The Dundee coronary risk-disk for management of change in risk factors. *BMJ* 1991; **303**:744-747.

van Strien T, Frijters JE, Bergers GP, Defares PB. The Dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional and external eating behaviour. *Int J Eat Disord* 1986; **5**:295-315.

Wang LY, Yang Q, Lowry R and Wechsler H. Economics analysis of a school-based obesity prevention program. *Obes Res* 2003; **11**:1313-1324.

Wardle J, Rapoport L, Afuape T, Duman M. Mass education for obesity prevention: the penetration of the BBC's 'Fighting Fat, Fighting Fit' campaign. *Health Educ Res* 2001; **16**:343-355.

Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Boston, MA: The Health Institute New England Medical Center, 1993.

Winkleby MA, Taylor CB, Jatulis D, Fortmann SP. The Long-Term Effects of a Cardiovascular Disease Prevention Trial: The Stanford Five-City Project. *Am J Public Health* 1996; **86**: 1773-1779.

Wonderling D, Langham S, Buxton M, Normand C, McDermott C. What can be concluded from the Oxcheck and British family heart studies: commentary on cost effectiveness analyses. *BMJ* 1996; **312**:1274-1278.

World Health Organisation European Collaborative Group. European collaborative trial of multifactorial prevention of coronary heart disease: final report on the 6-year results. *Lancet* 1986; **1**:869-872.

World Health Organization European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 1. Recruitment and initial findings. *Eur Heart J* 1980; **1**:73-80.

World Health Organisation European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 2. Risk factor changes at two and four years. *Eur Heart J* 1982; **3**:184-190.

World Health Organisation European Collaborative Group. Multifactorial trial in the prevention of coronary heart disease: 3. Incidence and mortality results. *Eur Heart J* 1983; **3**:184-190.