Health Technology Assessment of Medical Interventions in the Prevention and Treatment of Disease
-
Directions of Further Research and Policy Implications

Health technology assessment voor medische interventies bij de preventie en behandeling van ziekte
Richtlijnen voor nader onderzoek en beleidsconsequenties

Proefschrift

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam op gezag van de rector magnificus Prof.dr. S.W.J. Lamberts en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
donderdag 12 februari 2009 om 11.00 uur

door

Carmen Mihaela Galani
geboren te Bucharest, Romania

Promotiecommissie

Promotor

Prof.dr. F.F.H. Rutten

Overige leden

Prof.dr. CA Uyl de Groot

Prof.dr. J.J. van Busschbach

Prof.dr. J. Brug

Copromotor

Dr. M.J. Al

Contents

Chapter 1	Introduction	**1**
Chapter 2	Systematic review and meta-analysis in health technology assessment	**11**
Chapter 3	Regulatory requirements: validation of electronic data capture of quality of life instruments	**55**
Chapter 4	Decision analytic model of management strategies for Paget's disease of bone	**69**
Chapter 5	Probabilistic cost-effectiveness model of lifestyle intervention in the prevention and treatment of obesity	**83**
Chapter 6	Direction of further research: value of additional information in cost-effectiveness analysis	**107**
Chapter 7	Health policy implications: decision-makers' attitude towards economic evaluations of medical technologies	**127**
Chapter 8	Discussion	**143**
References		**153**
Summary		**177**
Samenvatting		**183**
List of abbreviations		**189**
Portfolio Summary		**191**
Acknowledgements		**193**

Publications[1]

Chapters 2 to 7 are based on the following articles:

Chapter 2

Galani C, Schneider H. Prevention and treatment of obesity with lifestyle interventions: review and meta-analysis. Int J Public Health. 2007; 52(6):348-59

Chapter 3

Bushnell DM, Reilly MC, Galani C, Martin ML, Ricci JF, Patrick DL, McBurney CR. Validation of electronic data capture of the Irritable Bowel Syndrome – Quality of Life measure, the Work Productivity and Activity Impairment Questionnaire for Irritable Bowel Syndrome and the EuroQol. Value Health. 2006; 9(2):98-105

Chapter 4

Al M, Galani C, Selby P, Engbersen A, Mesenbrink P. Costs and effects of long term management strategies for Paget's disease of bone: a UK study. Submitted

Chapter 5

Galani C, Schneider H, Rutten FF. Modelling the lifetime costs and health effects of lifestyle intervention in the prevention and treatment of obesity in Switzerland. Int J Public Health. 2007; 52(6):372-82.

Chapter 6

Galani C, Al M, Schneider H, Rutten FF. Uncertainty in decision-making: value of additional information in the cost-effectiveness of lifestyle intervention in overweight and obese people. Value Health. 2008;11(3):424-34.

Chapter 7

Galani C, Rutten FF. Self-reported health care decision-makers' attitudes towards economic evaluations of medical technologies. Curr Med Res Opin. 2008; 24(11):3049-5

[1]*Reprinted with kind permission of SpringerLink (chapter 2 and 5), Blackwell Publishing (chapter 3 and 6) and Informa UK Ltd (chapter 7)*

Chapter 1

Introduction

Health technology assessment (HTA) originated from the spread of costly medical equipment and growing concerns over the ability and willingness of taxpayers and health insurers to pay for them (Banta 2003). There was greater public scrutiny of health care rationing decisions and a growing consumerist approach, both of which called for decision-making processes that were more accountable, transparent, and legitimate. A more comprehensive approach was needed to help decision-makers set priorities and obtain maximum benefit from limited resources, without compromising the ethical and social values underpinning health systems (Hutton 2006). The growth and development of HTA reflected this demand for well-founded information to support decisions regarding the development, uptake, and spread of medical technologies.

After the 1970s, HTA broadened to include a range of medical interventions, including drugs, medical devices, medical and surgical procedures, and the organizational and support systems used in care provision (Jonsson 2002). However, most HTAs so far have been directed at pharmaceuticals, rather than at other technologies such as medical devices and surgical procedures (Hutton 2006).

HTA can be defined as 'the systematic evaluation of the properties, effects, and/or other impacts of health care technology' (Jonsson 2002). It

involves evaluating an intervention through the production, synthesis, and/or systematic review of a range of scientific and non-scientific evidence. The evidence typically considered includes safety, efficacy, cost, and cost-effectiveness, as well as the social, organizational, legal, and ethical implications (Jonsson 2002). For example, HTA often considers the macro-economic impacts of health technologies on national healthcare budgets, resource allocation among different health programs, regulation, and other policy changes on technological innovation, investment, technology transfer, and employment (Goodman 1998). As well as identifying valuable technologies, HTA can reduce or eliminate interventions that are unsafe and ineffective, or whose cost is too high compared to the benefits. HTA can also help to identify technologies that are underused (e.g. preventive screening or smoking cessation interventions) and identifying the reasons for lack of use.

The main aim of HTA is to provide a range of stakeholders, typically those involved in funding, planning, purchasing, and investing in healthcare, with accessible, useable, and evidence-based information that will guide decisions about technology and the efficient allocation of resources. It has been called 'the bridge between evidence and policy making', because it provides information for healthcare decision-makers at macro-, meso-, and micro-levels (Battista 1999). In particular, the increased use of medical technologies has encouraged decision-makers to rely on HTA to help determine the reimbursement status and pricing of interventions. HTA also contributes in many ways to the knowledge base for improving the quality of care, especially in supporting the development of clinical practice guidelines and health service standards.

In Europe, the first organizations dedicated to evaluating healthcare technologies were set up in the 1980s. Over the following 25 years most countries set up HTA programs, either by providing new agencies or institutes, or by setting up academic units or governmental and non-governmental entities. The diversity of HTA activities in the EU reflects the different healthcare and political systems, with different mandates, funding mechanisms, and roles in policy formulation (Banta 2003). The use of HTA in decisions that influence the spread and uptake of technologies can be influenced by several factors, such as income levels, reimbursement mechanisms, regulatory environments and behavioral determinants i.e. demand for new technologies. As HTA strives to bring together policy and evidence, it reflects the specific needs of decision-makers within a specific system, which explains the variation between countries.

Economic evidence is needed to restrict the use of products, especially innovative and expensive technologies where there may be uncertainty. Reimbursing such technologies can be confined to certain indications, patient populations, treatment settings, and therapeutic positioning i.e. first- or second-line therapy (Zentner 2005). In the Netherlands, for example, if expensive inpatient drugs meet certain criteria after an initial assessment (e.g. projected sales higher than 0.5% of total drug sales in the hospital) then they will be given conditional reimbursement for three years. In this time additional information on the drug's 'real world' cost-effectiveness is collected. If the evidence does not show value for money then reimbursement will be withdrawn. Conditional approvals have an important role because they allow use of the technology under

limited conditions in an attempt to minimize uncertainty. However, the usefulness of conditional approvals depends on further data collection and the subsequent re-evaluation of the product. These arrangements are known as a “risk-sharing agreements” (Drummond 2007). Technologies are generally reimbursed without conditions when cost-effectiveness and marginal therapeutic and patient benefits have been established (Anell 2005). But some drugs with poor cost-effectiveness are covered if the disease is severe (with a small patient population) or there is a lack of treatment alternatives such as in the case of orphan drugs. Payers favor such an approach because it reduces the risk that they will pay too much for a drug, at least in the long term. Manufacturers welcome the prospect of faster market access for their products as a result of accelerated pricing and reimbursement negotiation prior to launch. However, these deals can present certain practical difficulties. For example, who will decide on the additional data requirements? Who will pay for the studies? How will any “performance criteria” be determined? (Drummond 2007).

The results of health technology assessments are also used to develop clinical or practice guidelines. Guidelines typically include recommendations on priority-setting, and provide national support to help decision-makers. However, health economic evidence is not used as well as it could be when developing guidelines, with only a few recommendations grounded in HTA. Berg et al. (2004) suggest that this could be caused by a gap between the data generated and the requirements of clinical practice; aversion among the doctors to combine economics and health; and the fact that guidelines rely more on data on effectiveness, rather than on cost-effectiveness. Therefore guidelines are of limited value for influencing the use or uptake of new health technology (Berg 2004). This situation is probably made worse by the lack of coordination between the bodies that produce guidelines and those that set priorities and fund HTA studies. However, guideline development and HTA are beginning to come together.

There remains a lack of evidence on the ’real world‘ effectiveness of economic evaluation in terms of healthcare planning, clinical practice, spread of technologies, or overall healthcare costs. Decision-makers continue to ignore the principles of economic evaluation, despite the advances in techniques and methodology (Goddard 2006). In addition, the available evidence on the impact of HTA and research development is relatively weak, with an explicit link only in the Netherlands and the UK (Jonsson 2002). Many factors may prevent decision-makers from using strict cost-effectiveness criteria when setting priorities and other stakeholders from using HTA products (e.g. reports, practice guidelines) in healthcare decisions. Goddard et al. (2006) argue that the lack of impact is not so much caused by methodological shortcomings than by the wider context of public-sector decision-making. While decision-makers may value health economics information, other aspects of the public policy process result in sporadic and unsystematic application of HTA.

The place of HTA in the decision-making process can affect the extent to which evidence is used to inform policy and priority-setting. Countries often disagree on the use of HTA recommendations (Draborg 2005). Some support recommendations on the grounds that experts are the best people to provide

them, while others prefer decision-makers to make recommendations in the light of political context and other country-specific circumstances. However, decision-makers may not have the technical expertise to understand the methodological strengths and weaknesses of an assessment. Improvements are still needed, but much has been done by assessment bodies to enhance the accessibility and usability of HTA among different audiences such as policy-makers, health professionals, and general public. Although different decision structures provide policy-makers with a wide range of discretion, not employing HTA evidence may lead to inefficient, ineffective, and inequitable healthcare.

The present thesis will address some of the methodological challenges of the health technology assessment and will assess the impact of economic evaluations in healthcare decision-making process.

Methodological and practical challenges of health technology assessments

The use of meta-analysis in systematic literature review

Systematic reviews have a central role in evidence based-medicine. The quantitative systematic review, also known as meta-analysis provides a logical structure for quantifying the existing evidence (Berman 2002). Meta-analysis now offers the opportunity to critically evaluate and statistically combine results of comparable studies or trials. The aim is to get a consistent estimation of the global effect of a procedure on a specified outcome by increasing the number of observations and statistical power. There are several reasons to perform a meta-analysis in a systematic literature review (Cochrane Manual 2008):

a) To increase power. Power is the chance of detecting a real effect as statistically significant if it exists. Many individual studies are too small to detect small effects, but when several are combined there is a higher chance of detecting an effect.
b) To improve precision. The estimation of a treatment effect can be improved when it is based on more information.
c) To answer questions not posed by the individual studies. Primary studies often involve a specific type of patient and explicitly defined interventions. A selection of studies in which these characteristics differ can allow investigation of the consistency of effect and, if relevant, allow a reason for differences in effect estimates to be investigated.
d) To settle controversies arising from apparently conflicting studies or to generate new hypotheses. Statistical analysis of findings allows a degree of conflict to be formally assessed, and reasons for different results to be explored and quantified

There are some methodological issues that need to be addressed before performing a meta-analysis such as quality of original studies, combinability of the selected studies, publication bias and bias in the abstraction of data. However, if used appropriately, meta-analysis is a powerful tool for deriving meaningful conclusions from data and can help physicians and health policy makers to answer specific questions. Although there has been some controversy about its

validity (Bailar 1997), meta-analysis has become increasingly popular, as the number of studies with similar protocols has grown.

Measurement of patient reported outcomes

Patient reported outcomes (PRO) provide the patient perspective on the effectiveness of treatment, and for many diseases the patient is the only source of health outcome data (Revicki 2000). Patients, clinicians, pharmaceutical industry, decision-makers, payers and regulatory authorities acknowledge the need to understand the impact of symptoms and diseases on patients' lives and to evaluate how treatment affects patient functioning and well-being as a criterion for licensing new medications and for policy decisions. During the past decade a number of products have been submitted to regulatory bodies, such as the Food and Drug Administration (FDA) in the United States and the European Agency for the Evaluation of Medicinal Products (EMEA) for approval to communicate patient benefits. Along with these submissions have come questions about the underlying methodology of patient based data and health-related quality of life research and its application to the drug approval process. For PRO endpoint data to be accepted as evidence of treatment effectiveness there must be evidence documenting the instrument's conceptual framework, content validity, and psychometric qualities, including reliability, validity and responsiveness. Health outcomes researchers throughout the world have actively debated diverse conceptual and methodological issues related to all types of PROs, resulting in the publication of various "best practices" documents (Wildd 2005, Aquadro 2003). Regulatory agencies have recently developed statements to guide the development and use of these measures, especially by the pharmaceutical industry in the drug approval process (FDA 2006, EMEA 2005).

Traditionally, health-related quality of life research has predominately been conducted with the use of paper-and-pencil questionnaires. However, as computers become smaller and hand-held units more prevalent, a greater number of studies are utilizing/developing technology in what is termed as electronic data capture to eliminate the need for lengthy or cumbersome paper surveys. Electronic data capture offers many benefits, including personalization of questionnaires both to study protocols and specific populations, automatic data stamping, programmable skip patterns, and immediate data entry which eliminates the possibility of entry errors that may be made manually. Several studies have been conducted to test for differences between the traditional paper-and-pencil and electronic data capture, all concluding that data collected electronically is more complete (Bushnell 2003, Drummond 1995), equivalent (Pouwer 1998), reliable (Velikova 1999), and cost-effective (Johannes 2000). For health technology manufacturers it is important to provide evidence of the validity of the quality of life instruments in the electronic format as compared to paper version.

The use of decision analytic models

Mathematical modeling is used widely in economic evaluations of medical interventions. Health economic models represent an important analytic framework

to generate estimates of cost-effectiveness based on a synthesis of available data and explicit representation of uncertainty (Claxton 2002). The purpose of modeling is to structure evidence on clinical and economic outcomes in a form that can help to inform decisions about clinical practices and healthcare resource allocations. Models synthesize evidence on health consequences and costs from many different sources, including data from clinical trials, observational studies, insurance claim databases, case registries, public health statistics, and preference surveys. The use of decision-analytic modeling for health technology assessment has increased exponentially in recent years, and therefore practice guidelines have been defined (Philips 2006).

The use of models in the economic evaluation is also controversial. One aspect is related to the use of models to synthesize comparisons of two treatments in the absence of head-to-head trials. Because of the difficulties in conducting comparative trials for all available treatment options, models are often used as a proxy for head-to-head comparisons in economic evaluations of health technologies. However, indirect comparisons are potentially subject to bias. It can be very difficult to ensure that all the studies used in the model were equivalent in terms of variables such as patients' baseline risk, treatment settings, and the measurement of clinical outcomes. The apparent superiority of one therapy over another may be attributable to differences in trial design rather than inherent differences between therapies. Nevertheless, indirect comparisons are sometimes necessary. The standard approach is to find trials of the therapies of interest again a common comparator.

Another controversial aspect in decision-analytic models is related to the extrapolation of treatment beyond the duration observed in clinical studies. Extrapolation is usually required in reimbursement applications because clinical trials generally do not last long enough to demonstrate the full benefits and disadvantages of a therapy. Nevertheless, cost-effectiveness models offer a structured, rational, quantitative approach to help decision-makers to achieve better value for health resources expended. Consensus groups and experts have recommended the technique as an aid to guide resource allocation decisions in healthcare (Drummond 2008, Neumann 2008). As a consequence, many countries have incorporated decision analytic models into their health technology assessment and reimbursement procedures.

Characterizing uncertainty in decision analytic models

Most review bodies conduct or require sensitivity analyses on all variables that could potentially influence the overall results, or on a subset of inputs. This is because of the uncertainty inherent in conducting economic evaluations, specifically over the value of particular estimates and their relative effect on costs and benefits. The stipulation for sensitivity analyses comes from the need to test or verify the robustness of the findings. Different countries have different requirements for sensitivity analysis (e.g. univariate or multivariate) so the choice of parameters and methods must be substantiated and well documented. Most countries recommend or require this, and it is particularly important in the case of assessments for new technologies, where the necessary data for evaluations is seldom clear.

The use of sensitivity analysis and modeling as well as subgroup analysis may also be used to predict the effect of certain patient characteristics (e.g. age, sex, and ethnicity) on cost-effectiveness and equity (Michaels 2006). Some review bodies (e.g. National Institute for Health and Clinical Excellence, NICE) suggest that modeling for subgroups of patients might be appropriate, but there are no recommendations as to which variables would be considered ethical. Outlining clear criteria for subgroup analysis, based on specific variables, could help to incorporate social values into decision-making in an explicit, transparent, and consistent way.

Traditionally, the uncertainty has been examined using sensitivity analysis. In the recent years there has been considerable emphasis on the development of appropriate statistical methods for handling uncertainty in economic evaluations of medical interventions, with a tendency to move from univariate sensitivity analysis towards probabilistic descriptions of uncertainty e.g. cost-effectiveness planes, cost-effectiveness acceptability curves and distributions of incremental net benefit (Briggs 1999).

Value of information analysis is a natural methodological extension of Bayesian decision theory that quantifies the existing level of uncertainty and estimates the impact on the expected net benefit of alternative decision options through obtaining perfect information on model parameters. Information is valuable because it reduces the expected costs of uncertainty surrounding a clinical decision. The expected costs of uncertainty are determined by the probability that a treatment decision based on existing (prior) information will be wrong and by the consequences if the wrong decision is made (loss function). The expected costs of uncertainty can also be interpreted as the expected value of perfect information, since perfect information (as an infinite sample) can eliminate the possibility of making the wrong decision (Claxton 2001). It is also the maximum a decision maker should be willing to pay for additional evidence to inform this decision in the future. In the last decade, there has been a huge interest in developing and applying value of information methods within health economic decision analysis and clinical trial design (Felli 1999, Claxton 2004, Schulpher 2005).

The approaches to quantify the uncertainty are useful for several reasons. First, the uncertainty analysis helps to test the hypotheses about the sign and magnitude of costs, effectiveness and cost-effectiveness ratio. Second, decision-makers are informed about the confidence they should place on the result of the economic evaluation of medical interventions. Third, uncertainty enables continuous update of the model by identifying the sensitive parameters that require more information. In addition, analysis of uncertainty can assess the benefits and costs of obtaining or requiring more information and, therefore, help guide decision about future research.

The use of economic evaluations in decision-making

The importance of health economics research utilization in policy-making and of understanding the mechanisms involved is increasingly recognized. The existence of relevant research, though necessary, is not sufficient. Evidence-based policy is difficult to achieve and it is widely agreed that health policies do

not reflect research evidence to the extent that in theory they could (Davis 1996, van Velden 2005). Understanding the reasons behind the resistance of policy-makers to use health economics research has been the source of numerous scholarly papers, books and conferences (Nuemann 2004, Tunis 2004). Examination of the policy-making process confirms be complex, with many genuine obstacles to evidence-based policy-making at the same time as there are factors that could increase research utilization. Examinations of the use of economic evaluations and HTA in policy-making have considered the importance of the quality, reliability, timeliness and comprehensiveness of research in influencing the level of utilization (Drummond 2000, Jonsson 2002)

The purpose of this thesis is to address some of the challenges discussed above by giving practical examples of how these methodological and practical challenges can be addressed in the economic evaluations of medical interventions.

Outline of this thesis

Chapter 2 presents the results of a systematic literature review with meta-analysis on the effectiveness of lifestyle intervention in the prevention and treatment of obesity. The global rise in obesity prevalence continues to be a threat to people's health. This review is part of a HTA in the prevention and treatment of obesity performed for Swiss Ministry of Health to support the decision to allocate funds in the prevention of obesity programs in Switzerland in 2007. The systematic review provides new information on the effectiveness of lifestyle interventions by assessing the mid- to long-term effects on weight and cardiovascular risk profile in overweight and obese people.

Chapter 3, using data collected from a stand-alone study, establishes the validity of the electronic versions of three quality of life measures in comparison to the existing paper versions. As a part of Novartis Pharma development program, a selective 5-HT4 receptor partial agonist was developed for the treatment of functional gastrointestinal dysmotility disorders. A multinational clinical trial has been completed using electronic data collection. In order to evaluate changes in patient reported outcomes, standard instruments were included to assess change in patients' quality of life (IBS-QOL), productivity (WPAI:IBS-C), and utility (EQ-5D). For the manufacturer, it is important to assess these measures to provide evidence of validity in the electronic format that would meet satisfactory claims of equal performance in an electronic mode compared with paper version.

Chapter 4 presents the development and results of a cost-effectiveness model of zoledronic acid versus risedronate in Paget's disease of bone. Zoledronic acid is a third generation of nitrogen containing bisphosphonate seeking approval for Paget's disease of bone indication. For registration purposes two six-month randomized clinical trials were performed in order to evaluate the safety and efficacy of intravenous zoledronic acid for the treatment of Paget's disease of bone using oral risedronate as a comparator. Our economic model, based on the two clinical trials, goes beyond six-month and evaluates the cost-effectiveness of zoledronic acid versus standard therapy, risedronate, for the first two years of market access of zoledronic acid in the United Kingdom.

Chapter 5 presents the development and results of an economic model that evaluates the lifetime effects of three-year lifestyle intervention in the prevention and treatment of obesity. The model estimates the cost-effectiveness of lifestyle intervention versus standard treatment in overweight and obese people in Switzerland. The model is a key component of the HTA that was carried out for lifestyle intervention in the obesity area. Probabilistic sensitivity analysis was undertaken to establish the uncertainty associated with the decision to adopt the lifestyle intervention program in the prevention and treatment of obesity.

Chapter 6 addresses the uncertainty in decision-making by evaluating the value of additional information. We used the value of information analysis on a probabilistic cost-effectiveness model of lifestyle intervention in overweight and obese people to evaluate the uncertainty. Our analysis quantified the uncertainty surrounding the decision to adopt lifestyle intervention.

Chapter 7 assesses the use of research evidence relating to economic analyses in healthcare decision-making. We conducted a literature review to summarize and synthesize published literature on self-reported attitudes of healthcare decision-makers towards economic evaluations of medical technologies. The aims of this literature review was to determine the extent to which economic evaluations are used in health policy decision-making, and to consider factors associated with the utilization of such research findings.

Chapter 2

Systematic Review and Meta-Analysis in Health Technology Assessment

Summary

Systematic review has a central role in health technology assessment. The aim of the present chapter is to assess the mid- to long-term effectiveness of lifestyle interventions in the prevention and treatment of obesity. A systematic literature review with meta-analysis was performed. Electronic databases, reference lists, books and reports covering topic of obesity were searched. The included studies were randomized clinical trials of lifestyle interventions in overweight and obese subjects that had a minimum observation period of one year. Outcomes evaluated were measurements of body weight, body mass index, waist circumference, systolic and diastolic blood pressure, blood lipids: total cholesterol, low density lipoprotein, high density lipoprotein, triglyceride, blood glucose control: two-hour plasma glucose, fasting plasma glucose, and glycosylated haemoglobin. Thirteen studies have been selected in the prevention and seventeen in the treatment of obesity. Compared with standard care, lifestyle intervention reduced significantly body weight, body mass index, waist circumference, blood pressure, blood lipids and blood glucose in overweight and obese people. The favorable effects were maintained up to three years. Lifestyle interventions were efficacious in the mid- to long-term prevention and treatment of obesity leading to a significant reduction in body weight and cardiovascular risk factors.

Introduction

Obesity is a chronic disease whose prevalence is reaching epidemic proportions around the world (World Health Organization 2000). Obesity is associated with a high risk of morbidity, mortality as well as reduced life expectancy. The major health consequences of overweight and obesity are type 2 diabetes mellitus, hypertension, coronary heart disease, gallbladder disease, psychosocial problems and certain types of cancers (Fontaine 2003).

The increasing prevalence of overweight and obesity highlight the need for improved prevention strategies to overcome this significant public health problem. Many government initiatives and awareness campaigns have been initiated worldwide to combat obesity i.e. International Obesity Task Force (IOTF). The best strategies to prevent and treat obesity have not been settled yet. A wide variety of obesity treatments are available, including diet, exercise, behavioral modification, pharmacological treatment and surgery. Among several strategies, lifestyle intervention has been documented to lead safely to improvements in metabolic abnormalities such as increased body weight, dyslipidemia, elevated blood pressure, glucose control, pro-coagulant and pro-inflammatory activity that are linked to the development of obesity, diabetes, metabolic syndrome and cardiovascular disease (Pritchett 2005). Lifestyle programs are multi-factorial interventions that are designed for each patient or group of patients according to their risk factor status and the needs of the subjects. These include promoting healthy lifestyle habits, dietary counseling, physical exercise training, and behavioral change targets. Individuals at risk for obesity, diabetes and cardiovascular disease may be influenced trough learning process to allow the lifestyle changes to control risk factors such as body weight, blood pressure, blood cholesterol and blood glucose levels. Several individual lifestyle interventions proved to be efficacious in the prevention of diabetes (Pan 1997, Lindström 2003). The aim of the present study was to systematically assess the mid- to long-term effectiveness (1 – 6 years) of lifestyle interventions in the prevention and treatment of obesity.

Methods

Titles and abstracts were obtained from systematic searches of electronic databases: Medline, CINHAL, MBASE, and PubMed. The searches were carried out using search terms ‘obesity prevention’, ‘overweight treatment’, ‘obesity treatment’, ‘lifestyle intervention’, ‘weight change’, ‘weight control’, ‘body mass index’, ‘cardiovascular disease’, and ‘diabetes mellitus’. The search was restricted to the period from 1995 to 2005 due to advancements in research towards conducting high quality studies and/or better reporting. The rapid changes in the obesity environment made the last ten years more suitable for study selection. Books and reports covering the topic of obesity were searched. The reference lists of all included studies were checked and all potentially appropriate studies were obtained and assessed for additional evidence to be used in this study. In addition, several individual-based studies that focused on prevention of type 2 diabetes and cardiovascular disease were reviewed.

For the purpose of this study, we defined 'prevention of obesity' as interventions that target overweight individuals with body mass index (BMI) between 25 and 29.9 kg/m^2 aiming to prevent the transition from overweight to obesity and 'treatment of obesity' as interventions that target obese individuals with a BMI $\geq$ 30 kg/m^2 aiming to reduce the progression of obesity and associated co-morbidities.

The inclusion criteria consisted of randomized controlled trials of lifestyle interventions performed in overweight or obese subjects over 18 years of age that had a minimum observation period, including treatment and follow-up, of at least one year. Lifestyle interventions had to include dietary counseling and physical exercise associated or not with behavioral modification techniques. The goals of lifestyle interventions were to achieve and maintain a weight reduction through consumption of a healthy low-calorie, low-fat diet and to engage in regular physical activities.

The study population was classified according to the type of intervention: lifestyle intervention and control group regarded as standard care. The common characteristic of all selected studies is that interventions were carried out in overweight or obese people with or without associated co-morbidities. Additional subgroup analyses were performed in overweight subjects with cardiovascular risk factors and overweight or obese subjects with impaired glucose tolerance at risk of developing type 2 diabetes.

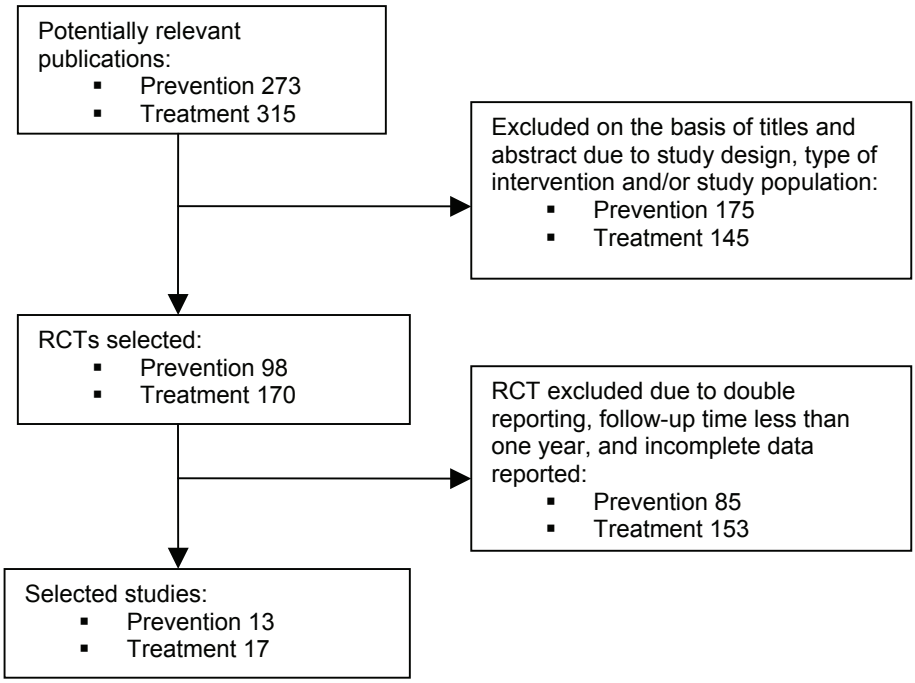
For studies that fulfilled the inclusion criteria, the following outcomes were evaluated: body weight, BMI, waist circumference, systolic blood pressure (SBP) and diastolic blood pressure (DBP), blood lipids: total cholesterol (TC), low density lipoprotein cholesterol (LDL), and high density lipoprotein cholesterol (HDL), triglyceride (TG), and blood glucose control: two-hour plasma glucose (2h-PG), fasting plasma glucose (FPG), and glycosylated haemoglobin (HbA1c). Meta-analysis technique was used to combine the results from distal follow-up (the last follow-up reported) from independent studies. The summary outcome measure calculated was the difference in means between lifestyle intervention and standard care. Effects were combined using a random effects model. The pooled estimates of the effect were obtained using Comprehensive Meta-Analysis software (CMA 2005).

Methodological quality of the selected studies was assessed by an adjusted Jadad scale (Jadad 1996). Since lifestyle interventions are usually not blinded, an adjustment in the five point Jadad scale was made excluding the double blinding score. A quality checklist assessed the overall quality of the study including the sample size, conduct of the study, follow-up, analysis and the interpretation of the results (Avenell 2004). Sensitivity analysis was performed in high quality studies i.e Jadad score $\geq$ 2 points and quality score > 80 points, to test the robustness of the study results. One reviewer abstracted the relevant study population and intervention characteristics using a standardized template. Two reviewers assessed the methodological quality of evaluated studies independently. Discrepancies were resolved by consensus discussion.

Results

Thirteen studies were selected for evaluation of the prevention of obesity approach and 17 studies investigated the treatment of obesity approach. The flowchart provides an overview of all included and excluded studies (Figure. 1).

Figure 1. Flowchart



Prevention of obesity

Table 1 presents the descriptive data of the 13 studies selected investigating the prevention of obesity. Lifestyle intervention components of each study are presented in Appendix 1. The studies included a total of 3566 participants with an average BMI of 28 kg/m^2 and an average body weight of 81 kg. The study participants had a mean age of 49 years and were predominantly of female gender.

Table 1. Characteristics of included studies in overweight people

Author year	Country	Population	Duration (years)	N participants	N[a] Follow-up	Age	BMI	Adj. Jadad Score[b]	Quality Score[c]
Anderssen 1996	NOR	Ov CV risk	1	219	95%	45	28.8	1	68
Burke 2005	AU	Ov HT	1	241	85%	56	29.9	2	73
Carr 2005	USA	Ov IGT	2	64	97%	56	26.2	1	73
Dyson 1997	UK, FRA	Ov D risk	1	227	50%	50	28.5	2	70
He 2000	USA	Ov HT risk	7	208	87%	43	28.9	2	75
Kastarinen 2002	FIN	Ov HT risk	2	715	82%	43	28.7	1	78
Ketola 2001	FIN	Ov CV risk	2	150	95%	-	27.8	3	65
Liao 2002	USA	Ov IGT	2	74	72%	54	26.1	3	70
Mensink 2003ab	NL	Ov IGT	2	114	77%	57	29.7	3	85
Pan 1997	CHI	Ov IGT , D	6	530	92%	45	25.8	2	60
Simkin-Silverman 2003	USA	Ov	4.5	535	95%	47	25	2	83
Stefanick 1998	USA	Ov CHD	1	377	97%	52	28	2	83
Trento 2001	ITA	Ov D	4	112	71%	62	28.9	3	88

[a] N, percentage of subjects finishing the study; [b] Adjusted Jadad score: 1 Low, 2 Moderate, 3 High, [c] Quality score: 1-50 Low, 51-80 Moderate, 81-100 High; Ov, overweight; CV, cardiovascular; D, diabetes; HT, hypertension; IGT, impaired glucose tolerance; CHD, coronary heart disease; N, number

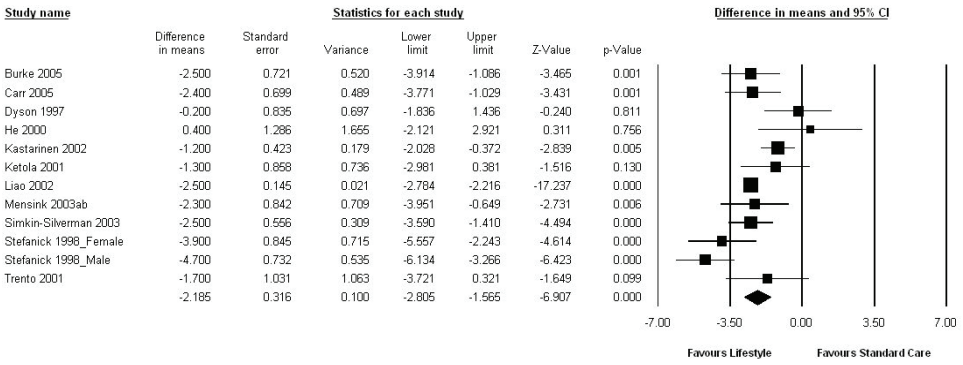
The results of the meta-analysis are summarized in Table 2. At an average follow-up of three years, the pooled effect size showed significance in favor of the

lifestyle intervention compared with standard care in reducing body weight (-2.2 kg, Figure 2) and cardiovascular risk factors with the exception of HDL and HbA1c.

Table 2. Meta-analysis results in overweight people

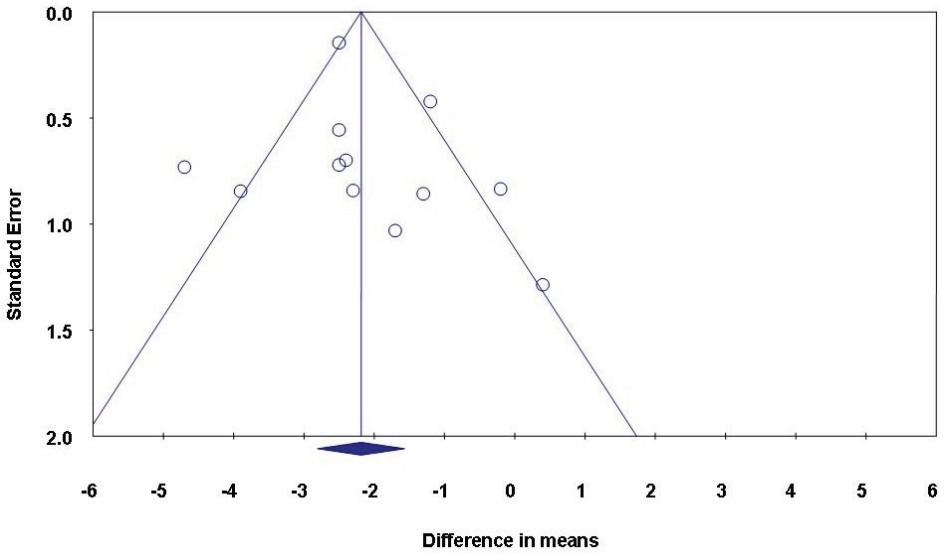
Outcome	Number studies	Number participants	Difference in means	Standard error	95% CI Lower limit	95% CI Upper limit	p-Value
Weight (kg)	11	2373	-2.19	0.32	-2.81	-1.57	<0.0001
BMI (kg/m^2)	5	926	-1.11	0.23	-1.56	-0.66	<0.0001
Waist (cm)	3	208	-2.12	0.23	-2.56	-1.68	<0.0001
SBP (mmHg)	9	2239	-2.08	0.61	-3.28	-0.89	0.001
DBP (mmHg)	9	2239	-1.59	0.55	-2.67	-0.51	0.004
TC (mmol/l)	7	1516	-0.26	0.07	-0.41	-0.12	<0.0001
HDL (mmol/l)	7	1875	0.01	0.01	-0.22	0.04	0.640
LDL (mmol/l)	5	1690	-0.16	0.06	-0.28	-0.03	0.013
TG (mmol/l)	7	1875	-0.23	0.08	-0.38	-0.08	0.003
HbA1c (%)	3	397	-0.50	0.52	-1.52	0.52	0.339
FPG (mmol/l)	6	804	-0.28	0.09	-0.45	-0.11	0.001
2h-PG (mmol/l)	2	284	-0.63	0.24	-1.10	-0.16	0.009

Figure 2. Meta-analysis weight change (kg) in overweight people



A funnel plot of the mean difference in body weight reduction plotted against the study size, represented by standard error, is shown in Figure 3. The vertical line indicates the pooled mean difference of all trials (-2.2 kg). Usually, studies with larger sample size appear toward the top of the graph and a distributed symmetrically around the combined effect size. Smaller studies usually appear toward the bottom of the graph, and since there could be more sampling variation in the effect size estimates in the smaller studies, they are dispersed across a range of values. Visual inspection implies that the evaluated lifestyle intervention studies participated equally to the pooled mean difference i.e. the studies are dispersed symmetrically around the combined effect size.

Figure 3. Funnel plot of the mean difference in body weight in overweight people plotted against standard error



A sensitivity analysis was performed on high quality studies (Mensink 2003ab, Simkin-Silverman 2003, Stefanick 1998, Trento 2001). The studies included a total of 1168 participants with an average age of 51 years and an average BMI of 27 kg/m^2. The results of the sensitivity analysis confirmed the results of the main analysis: compared with standard care, lifestyle intervention reduced significantly body weight and cardiovascular risk factors in overweight people with the exception of SBP, HDL and HbA1c. The difference in means was -3.1 kg in body weight (p=0.0001), -1.6 mmHg in SBP (p=0.068), -2 mmHg in DBP (p=0.03), -0.32 mmol/l in TC (p=0.0001), 0.001 mmol/l in HDL (p=0.96), -0.22 mmol/l in LDL (p=0.006), -0.21 mmol/l in TG (p=0.002), -0.75 % in HbA1c (p=0.37), and -0.35 mmol/l in FPG (p=0.002).

A subgroup analysis was performed in overweight people with cardiovascular risk factors (Anderssen 1996, Burke 2005, He 2000, Kastarinen 2002, Ketola 2001, Stefanick 1998). The studies included a total of 1910

participants with an average BMI of 27 kg/m^2. The results of the meta-analysis are presented in Table 3. At an average follow-up of three years, compared with standard care, lifestyle intervention reduced significantly body weight and cardiovascular risk factors in overweight people with identified cardiovascular risk with the exception of TG and HDL.

Table 3. Subgroup analysis in overweight people with cardiovascular risk factors

Outcome	Number studies	Number participants	Difference in means	Standard error	95% CI Lower limit	95% CI Upper limit	p-Value
Weight (kg)	5	1343	-2.30	0.70	-3.67	-0.92	0.001
SBP (mmHg)	6	1419	-2.43	0.78	-3.96	-0.91	0.002
DBP (mmHg)	6	1419	-2.16	0.77	-3.67	-0.67	0.005
TC (mmol/l)	4	1120	-0.35	0.09	-0.53	-0.16	<0.0001
TG (mmol/l)	3	970	-0.24	0.14	-0.52	0.04	0.087
HDL (mmol/l)	3	970	0.01	0.02	-0.04	0.05	0.798
LDL (mmol/l)	2	897	-0.27	0.08	-0.43	-0.43	0.001

Treatment of obesity

Table 4 presents the descriptive data of the 17 studies selected in the treatment of obesity. The lifestyle intervention components of each study are presented in Appendix 1.

Table 4. Characteristics of included studies in obese people

Author, year	Country	Population	Duration (years)	Number participants	N[a] Follow-up	Age	BMI	Adj. Jadad score[b]	Quality score[c]
DPP 2005ab	USA	Ob D risk	2.8	2161	93%	51	34	2	93
Esposito 2003	ITA	Ob	3	120	93%	35	35	2	85
Harvey-Berino 2004	USA	Ob	1	255	76%	46	32	2	58
Jeffery 1995	USA	Ob	2.5	202	88%	37	31	1	53
Lindhal 1999	SWE	Ob D risk	1	186	96%	56	30	2	78
Lindstrom 2003	FIN	Ob D risk	3	522	83%	55	31	2	95
Messier 2004	USA	Ob OA	1.5	316	80%	69	34	2	73
Moore 2003	UK	Ob	1.5	991	62%	48	37	2	65
Narayan 1998	USA	Ob	1	98	95%	34	35	1	55
Sbrocco 1999	USA	Ob	1	24	88%	42	33	2	60
Stevens 2001	USA	Ob D risk	3	1191	92%	43	31	1	68
Tate 2003	USA	Ob D risk	1	92	84%	49	33	3	63
Wylie-Rosset 2001	USA	Ob CV risk	1	588	81%	52	36	1	75
Whelton 1998	USA	Ob HT	2.5	886	86%	46	36	2	85
Wing 1998	USA	Ob D risk	2	154	84%	46	36	1	83
Wolf 2004	USA	Ob D	1	147	80%	53	37	3	58
Yeh 2003	USA	Ob	2	80	66%	50	37	2	75%

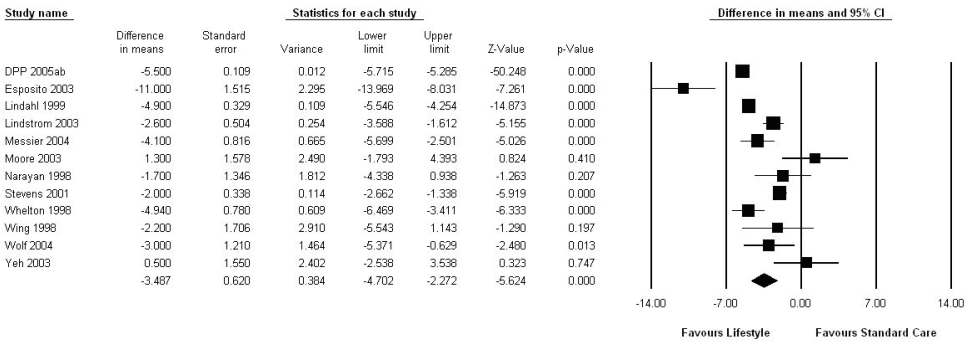
[a]N, percentage of subjects finishing the study; [b]Adjusted Jadad score: 1 Low, 2 Moderate, 3 High;[c] Quality score: 1-50 Low, 51-80 Moderate, 81-100 High, Ob, obese; D, diabetes; HT, hypertension; OA, osteoarthritis; CV, cardiovascular

The studies included 8013 participants, predominantly females, with an average age of 49 years and an average BMI of 34 kg/m^2. The meta-analysis results are presented in Table 5. Compared with standard care, lifestyle intervention reduces significantly body weight (-3.49 kg, see Figure 4) and cardiovascular risk factors in obese people with the exception of FPG and HbA1c. The average follow-up time of interventions was three years.

Table 5. Meta-analysis results in obese people

Outcome	Number studies	Number participants	Difference in means	Standard error	95% CI Lower limit	95% CI Upper limit	p-Value
Weight (kg)	12	5124	-3.49	0.62	-4.70	-2.27	<0.0001
BMI (kg/m^2)	7	3522	-1.33	0.31	-1.93	-0.72	<0.0001
SBP (mmHg)	6	4182	-2.78	0.82	-4.38	-1.18	0.001
DBP (mmHg)	6	4063	-1.42	0.43	-2.23	-0.57	0.001
TC (mmol/l)	5	893	-0.14	0.05	-0.24	-0.03	0.011
HDL (mmol/l)	4	2778	0.04	0.02	0.004	0.08	0.028
TG (mmol/l)	4	2964	-0.15	0.06	-0.27	-0.04	0.011
FPG (mmol/l)	5	2934	-0.15	0.08	-0.31	0.02	0.079
2h-PG (mmol/l)	2	692	-0.54	0.16	-0.84	-0.24	0.001
HbA1c (%)	2	497	-0.09	0.03	-0.40	0.23	0.599

Figure 4. Meta-analysis weight change (kg) in obese people



A sensitivity analysis was performed on high quality studies (DPP 2005ab, Esposito 2003, Lindstrom 2003, Whelton 1998, Wing 1998). The studies included

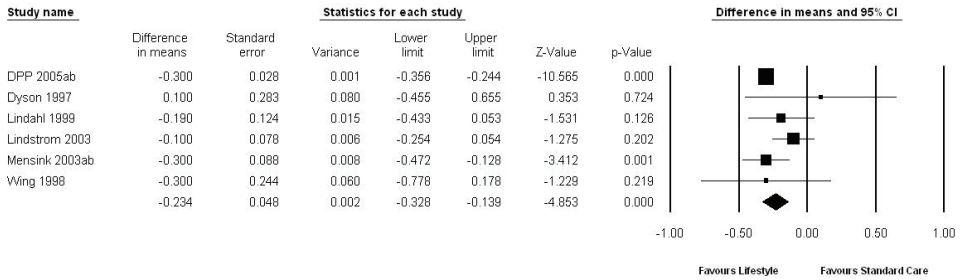
a total of 3023 participants with an average BMI of 33 kg/m^2 and a mean age of 52 years. The results of the sensitivity analysis confirmed the results of the main analysis: compared with standard care, lifestyle intervention reduced significantly body weight and cardiovascular risk factors in obese people with the exception of FPG. The difference in means was -5.1 kg in body weight (p<0.0001), -1.8 kg/m^2 in BMI (p=0.001), -3 mmHg in SBP (p=0.0001), -2 mmHg in DBP (p=0.0001), -1.15 mmol/l in TC (p=0.01), 0.04 mmol/l in HDL (p=0.02), -0.17 in TG (p=0.02), and -0.13 in FPG (p=0.24).

A subgroup analysis was performed in overweight and obese people subjects with impaired glucose tolerance in the prevention of diabetes studies. Nine such studies were identified, in four studies participants had a BMI of < 30 kg/m^2 (Carr 2005, Dyson 1997, Liao 2002, Mensink 2003ab) and in the other five studies participants had a BMI of > 30 kg/m^2 (DPP 2005ab, Lindahl 1999, Lindstrom 2003, Tate 2003, Wing 1998). The studies included a total of 3502 participants, predominantly females, with an average BMI of 33 kg/m^2. Meta-analysis results are presented in Table 6. Compared with standard care, lifestyle intervention reduced significantly body weight and cardiovascular risk factors (see Figure 5) with the exception of LDL and HbA1c.

Table 6. Subgroup analysis in overweight and obese people at risk of diabetes

Outcome	Number studies	Number participants	Difference in means	Standard error	95% CI Lower limit	95% CI Upper limit	p-Value
Weight (kg)	8	3150	-2.93	0.72	-4.35	-1.52	<0.0001
BMI (kg/m^2)	6	2890	-1.29	0.33	-1.94	-0.64	<0.0001
SBP (mmHg)	5	3115	-3.45	0.68	-4.78	-2.13	<0.0001
DBP (mmHg)	5	3115	-1.83	0.34	-2.50	-1.17	<0.0001
TC (mmol/l)	5	867	-0.13	0.06	-0.25	-0.02	0.027
HDL (mmol/l)	5	2842	0.02	0.01	0.00	0.04	0.030
LDL (mmol/l)	3	357	-0.05	0.08	-0.22	0.12	0.555
TG (mmol/l)	6	3028	-0.20	0.07	-0.33	-0.07	0.002
HbA1c (%)	4	682	-0.04	0.09	-0.21	0.14	0.686
FPG (mmol/l)	6	3029	-0.23	0.05	-0.33	-0.14	<0.0001
2h-PG (mmol/l)	3	608	-0.57	0.25	-1.05	-0.09	0.021

Figure 5. Fasting plasma glucose change (mmol/l) in overweight or obese at risk of diabetes



The data input of the meta-analyses performed are provided in the Appendix 2 and the graphical representation of all analyses are presented in Appendix 3.

Discussion

The global rise in obesity prevalence continues to be a threat to people's health. Although health policies have aimed to raise public awareness to prevent obesity, its increasing prevalence implies that successful solutions have not been yet identified. The present systematic review provides new information on the effectiveness of lifestyle interventions by assessing the mid- to long-term effects on weight and cardiovascular risk profile in overweight and obese people. Our findings suggest that - on an average follow-up time of three years - lifestyle interventions reduce significantly body weight and cardiovascular risk factors in overweight and obese people. To estimate whether these results are of clinical relevance we searched the medical literature for studies that have linked the intermediate physiological endpoints i.e. HbA1c, FPG, with hard outcomes such as reduced incidence of diabetes, reduced cardiovascular events and morbidity, and reduced mortality risk.

Several individual studies in obese people designed to prevent diabetes (DPP 2002, Lindstrom 2003) demonstrated that lifestyle intervention was associated to a 58% reduction in the incidence of diabetes in a three years program compared with a control group. A large trial (UKPDS 34) specifically designed to address the hypothesis that glucose lowering therapies may reduce the risk for cardiovascular morbidity or mortality in overweight individuals with type 2 diabetes, demonstrated that relatively small reductions in HbA1c (<1%) were associated with reduced microvascular complications. It has been estimated that each 1% reduction in HbA1c level was associated with a 14% reduction in the incidence of fatal and nonfatal myocardial infarction and 37% reduction in the microvascular complications (Stratton 2000). In our study, lifestyle intervention reduced HbA1c by 0.5% more compared to standard care in overweight people; however, the difference between groups was not significant. This may be partially explained by the small number of studies reporting HbA1c as a reported outcome and, in addition, our main analysis was not stratified for individuals with high risk for developing type 2 diabetes.

Observational analyses using data from a large clinical trial (UKPDS 61) demonstrated that individuals with intermediate FPG values (7.8 to 10 mmol/l) compared with individuals with low FPG values (<7.8 mmol/l) had a significantly lower risk of diabetes related deaths and myocardial infarction. In our analysis, lifestyle intervention reduced significantly FPG in overweight subjects but not in obese subjects. However, when the analysis was performed in obese people at risk of developing diabetes, lifestyle intervention significantly reduced FPG (see Figure 5).

Our analysis showed a significantly greater decrease in waist circumference in the lifestyle intervention group in comparison to standard care. It has been documented that waist circumference is more closely correlated with the volume of visceral adipose tissue than the waist-to-hip ratio or total body fat mass (Despres 1993, Lemieux 1996). It may be hypothesized that a greater loss of visceral adipose tissue, as reflected indirectly by the observed decrease in waist circumference, contributed to the improvement in glycemic status (significant decrease in 2h-PG and FPG) in the lifestyle intervention group. Central obesity, which is measured as increased waist circumference, is also an important component of atherogenic dyslipidemia, which has been identified as predictor of the metabolic syndrome and plays a major role in the pathogenesis of cardiovascular disease (Vinik 2005).

Evidence from epidemiological studies and clinical trials indicate that dyslipidemia is one of the most important modifiable risk factors for coronary heart disease (Ferdinand 2004, Meagher 2004). Dyslipidemia is generally characterized by increased fasting concentrations of TC, LDL and TG, in conjunction with decreased concentrations of HDL (NCEP 2001). Thus, decreasing TC, LDL, TG and increasing HDL represent an important clinical target (Pyorala 1997). For example, analyzing the data from the Scandinavian Simvastatin Survival Study, the authors estimated that each additional 1% reduction in LDL would generate a 1.7% reduction in the risk of major coronary events (Pedersen 1998). In our study lifestyle intervention was associated with a significant decrease in TG, TC and LDL and an increase in HDL in overweight people, as well as with a decrease in TG and TC in obese people. When a subgroup analysis was performed in overweight people with cardiovascular risk factors, the difference in means between lifestyles intervention and standard care was not significantly different with respect to TG and HDL. The observed results pointed in the expected direction but showed no statistically significant difference. A possible explanation for these small changes is the fact that the control group often showed moderate weight loss and TG reductions, minimizing between-group differences.

The association between obesity and cardiovascular disease is well established and up to 60% of overweight and obese patients are hypertensive (Dentali 2005). Cardiovascular complications may, to a large extent, be prevented by lowering blood pressure in patients at risk of developing cardiovascular disease. The Heart Outcome Prevention Study showed that a decrease in systolic blood pressure of 2-3 mmHg in patients with diabetes and one other risk factor for cardiovascular morbidity was associated with a 25% reduction in risk of myocardial infarction, stroke, or cardiovascular death (Gerstein 2002). Another

major trial (UKPDS 38) compared tight with less tight blood pressure control i.e. mean 144/82 mmHg versus 154/87 mmHg. The tight blood pressure control demonstrated considerable benefits reducing heart failure by 56%, stroke by 44% and combined myocardial infarction, sudden death, stroke, and peripheral vascular disease by 34%. According to our analysis, lifestyle intervention reduced significantly systolic and diastolic blood pressure in overweight and obese people.

The present study established the mid- to long-term effectiveness of lifestyle intervention in overweight and obese people by combining the beneficial effects on body weight and cardiovascular risk factors at distal follow-up. The actual lifestyle intervention in all evaluated studies included dietary counseling and physical exercise and lasted from one to six years including an average follow-up time of three years. The question now arise weather the observed beneficial effects of lifestyle interventions are maintained over lifetime. Such information is presently not available. However, the extended follow-up of the Finish Diabetes Prevention Program (Lindstrom 2006) which lasted seven years resulted in sustained lifestyle changes and a reduction in diabetes incidence that was maintained long after the lifestyle counseling had stopped. The study reported a 43% reduction in the relative risk from developing diabetes related to the success in achieving the intervention goals of weight loss, reduced intake of total and saturated fat, increased intake of dietary fibre and increased physical activity. Nevertheless, continued research in overweight and obese people is required to evaluate the effect of lifestyle interventions over lifetime.

Our analysis has several limitations. We may not have identified all relevant literature on this topic given the large applicability of lifestyle interventions in obesity related diseases. However, our search strategies of literature identification was comprehensive, capturing many of the published studies on lifestyle intervention in overweight and obese people independent of the associated co-morbidities. We cannot exclude the possibility of publication bias i.e. small studies with positive results are more likely to be published although the funnel plot suggests no such publication bias. Our results are limited by the poor quality of data reported on the outcome evaluated i.e. BMI, HDL, LDL, HbA1c, as well as the lack of subgroup analysis on different patient population i.e. gender difference, stratification according to different co-morbid conditions. We also observed qualitative and quantitative heterogeneity across studies on sample size, study population, and types of lifestyle intervention.

Questions may arise whether a lifestyle program could accomplish equivalent beneficial effects in different countries and ethnically diverse populations. Socio-cultural factors can influence the efficacy of lifestyle program in different countries, young or older adults, men or women, and the presence of different co-morbidities. Evidence based on the studies selected for the present review strongly suggest that lifestyle intervention programs performed in different countries that target overweight and obese people with different co-morbidities are effective in reducing weight and cardiovascular risk factors. However, further research should adapt lifestyle interventions for the need of each patient population taking into consideration different dietary and physical activity background.

Conclusion

In summary, the results of the present systematic literature review suggest that lifestyle intervention is efficacious in the mid- to long-term prevention and treatment of obesity leading to a significant reduction in weight and cardiovascular risk factors. Lifestyle intervention can be considered an effective prevention tool that can be applied across different disease areas including obesity, diabetes and cardiovascular disease with beneficial effects maintained for more than three years.

Acknowledgements

This study was funded by the Swiss Federal Office of Health.

Appendix 1. Lifestyle intervention components

Author, year	Lifestyle intervention description - Prevention of obesity studies
Anderssen 1996	The dietary counseling focused on a reduction in total caloric intake, increase intake of fish and a reduction in the intake of saturated fat. Physical exercise program consisted in endurance exercise such as aerobic, circuit training, fast walking/jogging, intensity of training being at the level of 60-80% of each participant's heart rate measured by the treadmill test.
Burke 2005	The nutrition component promoted a diet low in fat (<30% energy from total fat; <10% energy from saturated fat), high in fruits and vegetables, low in salt and sugar, and recommended at least four fish meals per week. Participants were encouraged to accumulate at least 30 min of moderate-intensity physical activity on most days and to increase incidental activity. The program encouraged self-directed change in behavior with a focus on barriers to change, costs and benefits of a healthy lifestyle, goal setting and time management.
Carr 2005	Isocaloric diet consisted of <30% of total calories as fat (<7% as saturated fat), 55% as carbohydrate, and the balance as protein, giving <200 mg cholesterol daily. Exercise session consisted in endurance exercise that involved one hour of walking or jogging on a treadmill three times a week, with a goal of exercising at 70% of heart rate reserve.
Dyson 1997	Dietary advice on limiting total fat intake and increasing consumption of unrefined carbohydrate and dietary fiber. The exercise program encouraged continuous rhythmic movements involving the large muscle groups (e.g. swimming, cycling, brisk walking, skipping, jogging and low-impact aerobics). Subjects were seen by a dietitian and fitness instructor every three months and were required to complete food and exercise diaries.
He 2000	Dietary counseling on sodium reduction focused on shopping, cooking, and food selection behaviors aimed at reducing the intake of calories and sodium. Weight loss is encouraged by an increase in the caloric expenditure, primarily by walking at a brisk pace for 45 minutes, 4 to 5 times per week.
Kastarinen 2002	Counseling sessions and behavior modification methods targeting weight reduction, reduction in salt, alcohol and saturated fat consumption, as well as increase in leisure-time physical activity.
Ketola 2001	Individual multifactorial intervention program was tailored for each patient according to the risk factor status and needs of the patients. These include booklets of healthy lifestyle habits, individual dietary counseling by a nurse or dietician, joining a weight reduction group, and group or individual physiotherapy program.

Author, year	Lifestyle intervention description - Prevention of obesity studies
Liao 2002	Dietary prescription included an isocaloric diet comprising 30% of total calories as fat (7% as saturated fat) 55% as carbohydrate, the balance as protein, and <200 mg cholesterol daily. Subjects performed endurance exercise (walk/jog) on a treadmill three times a week for 1 hour at each session. Initially exercise was designed to attain 50% of heart rate reserve. The exercise was gradually increased at 2-week intervals over a period of 3 months until subjects were exercising at a goal of 70% of heart rate reserve.
Mensink 2003	Dietary counseling consisted of carbohydrate intake of at least 55% of total energy intake, total fat intake of <30 to 35% energy intake, saturated fatty acids <10% of energy intake, a cholesterol intake of <33mg/MJ, protein intake of 10-15% of energy intake and an intake of dietary fiber of at least 3 g/MJ. Subjects were encouraged to increase their physical activity (walking, cycling, swimming) to at least 30 minutes of moderate physical activity a day for at least 5 days a week.
Pan 1997	Individual goals were set for total calorie consumption and for daily quantities of cereals, vegetables, meat, milk, and oils. Patients received physician counseling concerning daily food intake. Patients were encouraged to increase the amount of their leisure time physical exercise by at lease 1 unit/day and by 2 units/day if possible for those <50 years of age with no evidence of cardiovascular disease or arthritis. Activities required for one unit of exercise: mild activity 30 minutes (slow walking, shopping, housecleaning), moderate activity 20 minutes (faster walking, cycling, ballroom dancing), and strenuous activity 10 minutes (slow running, climbing stairs, playing volleyball).
Stefanick 1998	Dietary recommendations to use less than 30% total fat, less than 7% saturated fat, and less than 200mg of cholesterol per day. Participants were given goals on engaging in aerobic activity equivalent to at least 16 km of brisk walking or jogging each week.
Trento 2001	Participants were given educational objectives to reach desirable body weight, learn to shop for food, and increase physical activity.
Simkin-Silverman 2003	Counseling was provided on reducing total fat, saturated fat, cholesterol, and integrating alternative lipid lowering dietary strategies (e.g. increasing soy protein, fruits vegetables and fiber) into the participant's meal plan. Participants were asked to increase their physical activity expenditure to the level of 1000 to 1500 kcal/week (e.g brisk walking 10-15 miles).
DPP 2005	The goals of the lifestyle intervention were to achieve and maintain a weight reduction of at least 7% of initial body weight through consumption of a healthy low-calorie, low-fat diet and to engage in physical activity of moderate intensity such as brisk walking for at least 150 min/week.

Author, year	Lifestyle intervention description - Prevention of obesity studies
Esposito 2005	Participants received detailed advice on reducing dietary calories, personal goal setting, and self-monitoring. The mean caloric intake goal was set at 1300 kcal/d for the first year and 1500 kcal/d for the second year. The recommended composition of the dietary regimen was 50% to 60% carbohydrates, 15% to 20% proteins, less than 30% total fat, less than 10% saturated fat, 10% to 15% to monosaturated fat, 5% to 8% polyunsaturated fat, and 18 g of fiber per 1000 kcal. Participants received individual guidance on increasing physical activity, mainly walking, but also with swimming and aerobic ball games.
Harvey-Berino 2006	The weight loss treatment program focused on the modification of eating and exercise habits through the use of behavioral strategies and self-management skills. Subjects were given prescribed menus, grocery lists, and recipes specific to their dietary condition. Calorie goals were formulated to represent a 500 kcal/d restriction from baseline levels.
Jeffery 1995	Standard behavior therapy including instruction on diet, exercise, and behavior modification technique. Dietary goals were assigned at 1000 or 1500 kcal per day depending on initial body weight. Exercise recommendations were to walk or bike 5 days per week, beginning with a weekly goal of 250 kcal per week and gradually increasing to 1000 kcal per week.
Lindahl 1999	The diet recommendation included approximately 20% of energy from fat and with high fibre content. The recommended portion size was calculated to give a daily energy intake of 1800 kcal in men and 1500 kcal in women. Aerobic physical exercise of low to moderate intensity was performed daily for 2.5 hours, e.g. brisk walks, gymnastics, cycling and swimming.
Linström 2003	Lifestyle intervention goals were weight reduction $\geq$ 5%, moderate intensity physical activity $\geq$ 30 min/day, dietary fat < 30% of total energy, saturated fat < 10% of total energy, and fiber $\geq$ 15 g/1000 kcal.
Messier 2004	The goal of the dietary intervention was to produce and maintain an average weight loss of 5% during the 18-month intervention period. Participants were provided with an aerobic exercise prescription that included walking within a heart range of 50-75% of heart rate reserve. Behavior change was facilitated using self-regulatory skills.
Moore 2003	A management model that consisted in training on the clinical benefit of weight loss and effective treatment options, including reduction of dietary energy intake and increased physical activity.
Narayan 1998	The aims of the nutrition intervention were to reduce fat and alcohol and increase fibre intake. The goal of the activity intervention was to increase energy expenditure by 700-1000 kcal per week though physical activity (waking, water aerobics, softball, volleyball).
Sbrocco 1999	Behavioral choice treatment encouraged subjects to eat in moderation, to engage in pleasure activities besides eating such as regular exercise.

Author, year	Lifestyle intervention description - Prevention of obesity studies
Stevens 2001	The dietary intervention focused on reducing caloric intake by decreasing consumption of excess fat, sugar and alcohol. The physical activity goal was to gradually increase activity to 30 to 45 minutes per day, four to five days per week. Exercise intensity was moderate (40% to 55% of heart rate reserve) and consisted in brisk walking. Specific behavior change techniques included self-monitoring (food diaries and graphs of minutes of physical activity per day), setting explicit short term goals and developing specific action plans to achieve those objectives.
Tate 2003	Internet behavioral e-counseling consisted in advice on diet and physical exercise. Subjects were instructed to report calorie and fat intake, exercise energy expenditure and received weekly e-mail behavioral counseling and feedback from a counselor.
Wylie-Rosett 2001	Cognitive behavioral approach to target lifestyle and behavior modifications. The intervention included: a workbook, a computer and staff consultation.
Whelton 1998	Participants were advised on ways to change eating patterns and increase physical activity. The goal for weight loss was achieving and maintaining a weight loss of 4.5 kg or greater.
Wing 1999	Subjects were asked to follow an 800-1000 kcal/day diet, with 20% of calories as fat. Meal plans and shopping lists specific to the calorie and fat goals were distributed weekly. Subjects were encouraged to gradually increase their physical activity to 1500 kcal/week through activities such as brisk walking completed 5 days each week and to monitor their exercise daily throughout the program.
Wolf 2004	Goals of the intervention were modest weight loss (5% of initial weight) and dietary intake as well as increase physical activity.
Yeh 2003	Skill based intervention included counseling on meal planning, dinning out, food label reading, recipe modification, and physical activity.

Appendix 2. Meta-analysis data input

Table 1. Data input from studies in overweight adults

Author, year	Weight (kg)	BMI (kg/m^2)	Waist (cm)	SBP (mmHg)	DBP (mmHg)	TC (mmol/l)	HDL mmol/l)	LDL (mmol/l)	TG (mmol/l)	HbA1c (%)	FPG (mmol/l)	2h-PG (mmol/l)
Andersen 1996	-	LS -2.2 (0.22)	-	L -8.5 (7.1), S -0.5 (11.2)	L -6.8 (5.5) S -0.7 (8.5)	L -0.76 (0.8) S -0.16 (0.6)	L 0.16 (0.2), S -0.02 (1)	-	L -0.72 (1), S 0.17 (0.9)	-	-	-
Burke 2005	L -3.9 (4.8), S -1.4 (5.5)	-	-	S 2 (12.7), S 4 (12.7)	L 0 (8.3), S 2 (8.3)	-	-	-	-	-	-	-
Carr 2005	L -1.8 (2.7), S 0.6 (2.8)	-	L -1.2 (4.9) S 0.8 (5.1)	-		-	-	-	-	-	L -0.04 (0.4), S 0.07 (0.5)	-
Dyson 1997	L -0.4 (5.8), S -0.2 (5.9)	-	-	L -2 (12.7), S 0 (12.7)	L 0 (8.3), S 0.01 (8.3)	L -0.2 (1.1), S -0.2 (1.1)	L -0.01 (0.3), S 0.0 (0.3)	L -0.1 (0.7), S -0.2 (0.7)	L -0.13 (1), S 0.01 (1)	L -0.1 (0.8), S -0.1 (0.8)	L -0.1 (2), S -0.2 (2)	-
He 2000	L 4.9 (7.1), S 4.5 (4.9)	-	-	L 2 (12.4), S 3.7 (13.4)	L -6.5 (7.5), S -5.2 (8.3)	-	-	-	-	-	-	-
Kastarinen 2002	L -1.5 (5.5), S -0.3 (5.8)	-	-	LS -2 (-4.3,0.3)	LS -1.1 (-2.4, 0.2)	L -0.03 (1.1) S 0.07 (1.1)	L 0.1 (0.3), S 0.07 (0.3)	L -0.11 (0.7) S 0.04 (0.7)	L -0.06 (1), S -0.06 (1)	-	-	-
Ketola 2001	L -3 (5.1), S -1.7 (5.4)	-	-	L -5.8 (12.7), S -5.1 (12.8)	L -2 (8.3), S -3.1 (8.3)	L -0.66 (1.1), S -0.4 (1.1)	-	-	-	-	-	-

Author, year	Weight (kg)	BMI (kg/m^2)	Waist (cm)	SBP (mmHg)	DBP (mmHg)	TC (mmol/l)	HDL mmol/l)	LDL (mmol/l)	TG (mmol/l)	HbA1c (%)	FPG (mmol/l)	2h-PG (mmol/l)
Liao 2002	L -1.8 (0.5), S 0.7 (0.6)	L -0.7 (0.2) S 0.2 (0.2)	L -1.2 (0.9) S 0.9 (0.9)	-	-	-	-	-	-	-	-	-
Mensink 2003ab	L -2.4 (4.4), S -0.1 (3.5)	L -0.8 (1.3) S 0 (1.4)	L 1.9 (4.4) S 0.6 (4.2)	-	-	L 0.3 (0.6), S 0.4 (0.7)	L 0.06 (0.2), S 0.05 (0.1)	L 0.32 (0.7), S 0.32 (0.6)	L -0.3 (0.8), S 0.25 (0.8)	L 0 (0.6), S -0.1 (0.7)	L 0.2 (0.6), S 0.5 (0.1)	L -0.9 (2.1), S 0.3 (2.2)
Pan 1997	-	L -1.6 (1.6), S -0.9 (2.9)	-	-	-	-	-	-	-	-	L 1.48 (2), S 2.07 (2)	-
Simkin-Silverman 2003	L -0.1 (5.9), S 2.4 (6.6)	L 0.05 (2), S 0.96 (1.8)	-	L -0.1 (12.7), S 0.2 (12.7)	L 1.5 (8.3), S 2.2 (8.3)	-	L 0.06 (0.3), S 0.08 (0.3)	L 0.09 (0.7), S 0.23 (0.7)	L 0.21 (1), S 0.34 (1)	-	-	-
Stefanick 1998 Female	L -3.1 (3.7), S 0.8 (4.2)	-	-	L -3.1 (8.4), S -2.4 (7.6)	L -2.7 (4.6), S -0.6 (5.9)	L -0.45 (0.6), S -0.03 (0.5)	L -0.03 (0.2), S 0.03 (0.2)	L -0.37 (0.6), S -0.06 (0.4)	L -0.12 (0.6), S 0.02 (0.5)	-	L -0.43 (0.6), S -0.14 (0.8)	L -0.52 (1.6), S -1.8 (1.6)
Stefanick 1998 Male	L -4.2 (4.2), S 0.5 (2.7)	-	-	L -3 (6.8), S 0.3 (7.9)	L -3 (6.6), S 1.8 (6.1)	L -0.53 (0.5), S -0.1 (0.6)	L 0.01 (0.1), S -0.01 (0.1)	L -0.52 (0.5), S -0.12 (0.6)	L -0.08 (0.6), S 0.1 (1)	-	L -0.43 (0.5), S -0.21 (0.6)	L -0.84 (1.7), S -0.32 (1.6)
Trento 2001	L -2.6 (5.2), S -0.9 (5.7)	-	-	L -5.9 (12.7), S -1.9 (12.7)	L -7.1 (8.3), S -6.3 (8.3)	L -0.07 (1.1), S 0.13 (1.1)	L 0.15 (0.3), S 0.05 (0.3)	-	L -0.43 (1), S -0.17 (1)	L -0.3 (0.8), S 1.3 (0.8)	L -0.5 (2), S 0.8 (2)	-

Table 2. Data input from studies in obese adults

Author, Year	Weight (kg)	BMI (kg/m2)	SBP (mmHg)	DBP (mmHg)	TC (mmol/l)	HDL (mmol/l)	LDL (mmol/l)	TG (mmol/l)	HbA1c (%)	FPG (mmol/l)	2h-PG (mmol/l)
DPP 2005ab	L -5.6 (2.8), S -0.1 (2.2)	L -2.4 (2), S -0.15 (2)	L -3.3 (16.4), S -0.6 (16.5)	L -3.8 (9.9), S -1.9 (9.9)	-	L 0.03 (0.3), S 0.001 (0.3)	-	L -0.3 (1), S -0.13 (1)	-	L -0.27 (0.7), S 0.03 (0.7)	-
Esposito 2003	LS -11 (14, -8)	LS -4.2 (-6.4, -2)	-	-	LS -0.1 (-0.31, 0.1)	LS 0.1 (0.05, 0.16)	-	-	-	-	-
Lindhal 1999	L -5.4 (2.8), S -0.5 (1.5)	L -2 (1), S -0.2 (0.5)	L -4.9 (12.1), S 1.3 (9)	L -3.2 (9), S -0.8 (9)	L -0.21 (1.1), S -0.06 (1.1)	-	-	L -0.16 (1.1), S -0.09 (1.1)		L -0.5 (0.6), S -0.31 (1)	L -0.7 (2), S -0.3 (2)
Lindstrom 2003	L -3.5 (5.1), S -0.9 (5.4)	L -1.3 (1.9), S -0.3 (2)	L -5 (14), S -1 (15)	L -5 (9), S -3 (9)	L -0.1 (0.9), S 0.1 (0.8)	L 0.14 (0.2), S 0.11 (0.2)	-	L -0.1 (0.6), S 0 (0.8)	L -0.2 (0.6), S 0 (0.6)	L 0 (0.7), S 0.1 (0.7)	L -0.9 (1.9), S -0.3 (2.2)
Messier 2003	L -5.2 (4.4), S -1.1 (5.6)	-	-	-	-	-	-	-	-	-	-
Moore 2003	LS 1.3 (-1.8, 4.4)	LS 0.1 (-1, 1.1)	-	-	-	-	-	-	-	-	-
Narayan 1998	L 2.5 (6.6), S 0.8 (6.1)	-	L 6 (12.7), S 4.1 (12.7)	L 1 (8.3), S -1 (8.3)	L 0.2 (1.1), S 0.1 (1.1)	-	-	-	-	L 0.1 (1.35), S 0.1 (1.35)	-
Stevens 2001	L -0.2 (5.9), S 1.8 (5.3)	-	L -0.8 (8.7), S 0.6 (8.5)	L -3.2 (6.5), S -2.4 (7)	-	-	-	-	-	-	-
Whelton 1998	L -4.9 (4.1), S -	-	-	-	-	-	-	-	-	-	-

Author, Year	Weight (kg)	BMI (kg/m2)	SBP (mmHg)	DBP (mmHg)	TC (mmol/l)	HDL (mmol/l)	LDL (mmol/l)	TG (mmol/l)	HbA1c (%)	FPG (mmol/l)	2h-PG (mmol/l)
	0.1 (4.2)										
Wing 1998	L -2.5 (8.4), S -0.3 (4.5)	L -0.8 (3), S -0.1 (1.7)	L -4.8 (15), S -1.5 (12)	L -0.2 (10.5), S 2 (8)	L 0.09 (0.7), S 0.18 (0.5)	L 0.02 (0.2), S 0.04 (0.2)	L 0.14 (0.5), S 0.24 (0.7)	L -0.28 (1.3), S 0.52 (1.1)	L 0.04 (1.1), S -0.1 (0.3)	L 0.5 (1.3), S 0.2 (0.4)	-
Wolf 2004	L -2.4 (7.5), S 0.6 (7)	-	-	-	-	-	-	-	-	-	-
Yeh 2003	L -0.6 (0.3), S -1.1 (5.8)	L -0.3 (1.6), S -0.4 (2.3)	-	-	-	-	-	-	-	-	-

*Results are presented for Lifestyle intervention (L) and Standard care group (S) as: mean (standard deviation) for each group or difference in means between groups (standard error) or difference in means between groups (confidence intervals).

Appendix 3. Meta-analyses figures

1. Meta-analysis prevention of obesity

Figure 1. Meta-analysis weight change (kg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Burke 2005	-2.500	0.721	0.520	-3.914	-1.086	-3.465	0.001
Carr 2005	-2.400	0.699	0.489	-3.771	-1.029	-3.431	0.001
Dyson 1997	-0.200	0.835	0.697	-1.836	1.436	-0.240	0.811
He 1999	0.400	1.286	1.655	-2.121	2.921	0.311	0.756
Kastarinen 2002	-1.200	0.423	0.179	-2.028	-0.372	-2.839	0.005
Ketola 2001	-1.300	0.858	0.736	-2.981	0.381	-1.516	0.130
Liao 2002	-2.500	0.145	0.021	-2.784	-2.216	-17.237	0.000
Mensink 2003ab	-2.300	0.842	0.709	-3.951	-0.649	-2.731	0.006
Simkin-Silverman 2003	-2.500	0.556	0.309	-3.590	-1.410	-4.494	0.000
Stefanick 1998_Female	-3.900	0.845	0.715	-5.557	-2.243	-4.614	0.000
Stefanick 1998_Male	-4.700	0.732	0.535	-6.134	-3.266	-6.423	0.000
Trento 2002	-1.700	1.031	1.063	-3.721	0.321	-1.649	0.099
	-2.185	0.316	0.100	-2.805	-1.565	-6.907	0.000

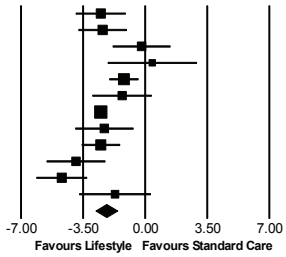


Figure 2. Meta-analysis BMI change (kg/m2) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andersen 1996	-2.160	0.220	0.048	-2.591	-1.729	-9.818	0.000
Liao 2002	-0.900	0.053	0.003	-1.003	-0.797	-17.135	0.000
Mensink 2003ab	-0.800	0.290	0.084	-1.369	-0.231	-2.757	0.006
Pan 1997	-0.700	0.369	0.136	-1.423	0.023	-1.898	0.058
Simkin-Silverman 2003	-0.910	0.168	0.028	-1.240	-0.580	-5.402	0.000
	-1.110	0.229	0.052	-1.559	-0.661	-4.846	0.000

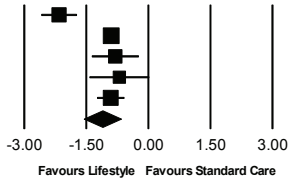


Figure 3. Meta-analysis waist circumference change (cm) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Carr 2005	-2.000	1.272	1.617	-4.493	0.493	-1.573	0.116
Liao 2002	-2.100	0.236	0.056	-2.563	-1.637	-8.885	0.000
Mensink 2003ab	-2.500	0.919	0.844	-4.301	-0.699	-2.721	0.007
	-2.121	0.225	0.051	-2.562	-1.679	-9.415	0.000

Figure 4. Meta-analysis systolic blood pressure change (mmHg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-8.000	2.316	5.364	-12.539	-3.461	-3.454	0.001
Burke 2005	-2.000	1.780	3.167	-5.488	1.488	-1.124	0.261
Dyson 1997	-2.000	1.807	3.266	-5.542	1.542	-1.107	0.268
He 1999	-1.700	2.655	7.048	-6.903	3.503	-0.640	0.522
Kastarinen 2002	-2.000	1.171	1.372	-4.296	0.296	-1.707	0.088
Ketola 2001	-0.700	2.082	4.335	-4.781	3.381	-0.336	0.737
Simkim-Silverman 2003	-0.300	1.126	1.269	-2.508	1.908	-0.266	0.790
Stefanick 1998 _Male	-3.300	1.518	2.305	-6.276	-0.324	-2.173	0.030
Stefanick 1998_Female	-0.700	1.706	2.911	-4.044	2.644	-0.410	0.682
Trento 2002	-4.000	2.400	5.760	-8.704	0.704	-1.667	0.096
	-2.083	0.608	0.370	-3.275	-0.891	-3.426	0.001

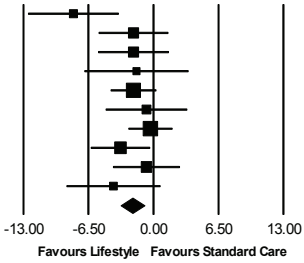


Figure 5. Meta-analysis diastolic blood pressure change (mmHg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-6.100	1.766	3.118	-9.561	-2.639	-3.455	0.001
Burke 2005	-2.000	1.163	1.353	-4.280	0.280	-1.720	0.086
Dyson 1997	-0.010	1.181	1.395	-2.325	2.305	-0.008	0.993
He 1999	-1.300	1.624	2.638	-4.484	1.884	-0.800	0.424
Kastarinen 2002	-1.100	0.662	0.438	-2.398	0.198	-1.661	0.097
Ketola 2001	1.100	1.355	1.837	-1.557	3.757	0.812	0.417
Simkin-Silverman 2003	-0.700	0.736	0.542	-2.143	0.743	-0.951	0.342
Stefanick 1998_ Female	-2.100	1.131	1.280	-4.317	0.117	-1.856	0.063
Stefanick 1998_Male	-4.800	1.312	1.722	-7.372	-2.228	-3.658	0.000
Trento 2002	-0.800	1.569	2.460	-3.874	2.274	-0.510	0.610
	-1.592	0.552	0.304	-2.673	-0.511	-2.886	0.004

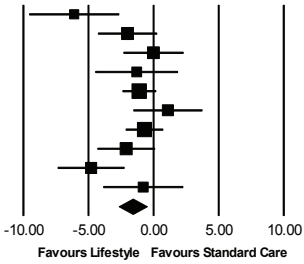


Figure 6. Meta-analysis total cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andersen 1996	-0.600	0.165	0.027	-0.923	-0.277	-3.638	0.000
Dyson 1997	0.000	0.155	0.024	-0.303	0.303	0.000	1.000
Kastarinen 2002	-0.100	0.081	0.007	-0.258	0.058	-1.238	0.216
Ketola 2001	-0.260	0.176	0.031	-0.606	0.086	-1.474	0.140
Mensink 2003	-0.100	0.141	0.020	-0.375	0.175	-0.711	0.477
Stefanick 1998_Female	-0.420	0.112	0.013	-0.639	-0.201	-3.751	0.000
Stefanick 1998_Male	-0.430	0.111	0.012	-0.648	-0.212	-3.860	0.000
Trento 2002	-0.200	0.204	0.042	-0.600	0.200	-0.980	0.327
	-0.262	0.073	0.005	-0.406	-0.119	-3.582	0.000

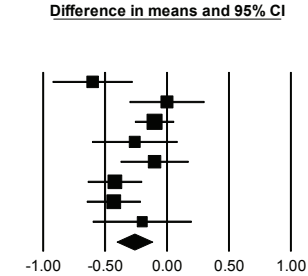


Figure 7. Meta-analysis high density lipoprotein cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	0.180	0.181	0.033	-0.175	0.535	0.995	0.320
Dyson 1997	-0.010	0.042	0.002	-0.091	0.071	-0.241	0.810
Kastarinen 1997	0.030	0.022	0.000	-0.013	0.073	1.383	0.167
Mensink 2003ab	0.010	0.033	0.001	-0.054	0.074	0.304	0.761
Simkin-Silverman 2003	-0.020	0.026	0.001	-0.070	0.030	-0.778	0.437
Stefanick 1998_Female	-0.060	0.035	0.001	-0.129	0.009	-1.706	0.088
Stefanick 1998_Male	0.020	0.026	0.001	-0.031	0.071	0.768	0.443
Trento 2002	0.100	0.055	0.003	-0.007	0.207	1.825	0.068
	0.007	0.014	0.000	-0.022	0.035	0.468	0.640

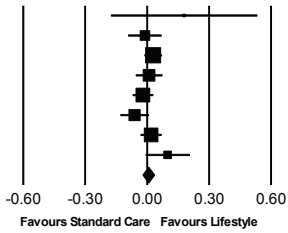


Figure 8. Meta-analysis low density lipoprotein cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Dyson 1997	0.100	0.106	0.011	-0.108	0.308	0.944	0.345
Kastarinen 2002	-0.150	0.055	0.003	-0.258	-0.042	-2.710	0.007
Mensink 2003ab	0.000	0.139	0.019	-0.272	0.272	0.000	1.000
Simkin-Silverman 2003	-0.140	0.066	0.004	-0.269	-0.011	-2.133	0.033
Stefanick 1998_Female	-0.310	0.107	0.012	-0.520	-0.100	-2.888	0.004
Stefanick 1998_Male	-0.400	0.103	0.011	-0.603	-0.197	-3.866	0.000
	-0.155	0.063	0.004	-0.278	-0.032	-2.478	0.013

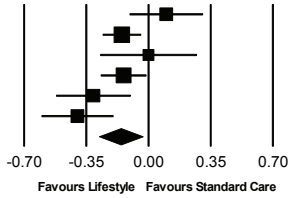


Figure 9. Meta-analysis triglyceride change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-0.890	0.226	0.051	-1.333	-0.447	-3.941	0.000
Dyson 1997	-0.140	0.137	0.019	-0.409	0.129	-1.019	0.308
Kastarinen 2002	0.000	0.072	0.005	-0.141	0.141	0.000	1.000
Mensink 2003ab	-0.550	0.171	0.029	-0.886	-0.214	-3.211	0.001
Simkin-Silverman 2003	-0.130	0.085	0.007	-0.297	0.037	-1.527	0.127
Stefanick 1998_Female	-0.140	0.111	0.012	-0.358	0.078	-1.261	0.207
Stefanick 1998_Male	-0.180	0.164	0.027	-0.501	0.141	-1.100	0.271
Trento 2002	-0.260	0.181	0.033	-0.616	0.096	-1.433	0.152
	-0.230	0.078	0.006	-0.383	-0.076	-2.925	0.003

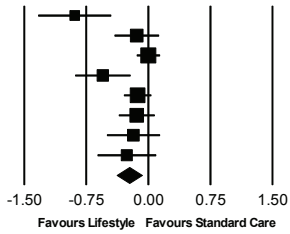


Figure 10. Meta-analysis HbA1c change (%) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Dyson 1997	0.000	0.109	0.012	-0.213	0.213	0.000	1.000
Mensink 2003ab	0.100	0.141	0.020	-0.175	0.375	0.711	0.477
Trento 2002	-1.600	0.144	0.021	-1.882	-1.318	-11.140	0.000
	-0.498	0.521	0.271	-1.519	0.523	-0.956	0.339

Figure 11. Meta-analysis fasting plasma glucose change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Carr 2005	-0.110	0.115	0.013	-0.336	0.116	-0.953	0.341
Dyson 1997	0.100	0.283	0.080	-0.455	0.655	0.353	0.724
Mensink 2003ab	-0.300	0.088	0.008	-0.472	-0.128	-3.412	0.001
Pan 1997	-0.590	0.310	0.096	-1.198	0.018	-1.902	0.057
Stefanick 1998_Female	-0.290	0.155	0.024	-0.593	0.013	-1.876	0.061
Stefanick 1998_Male	-0.220	0.111	0.012	-0.437	-0.003	-1.990	0.047
Trento 2002	-1.300	0.374	0.140	-2.033	-0.567	-3.474	0.001
	-0.278	0.088	0.008	-0.450	-0.107	-3.180	0.001

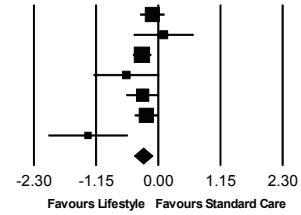


Figure 12. Meta-analysis 2-hour plasma glucose change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	-1.200	0.428	0.183	-2.039	-0.361	-2.804	0.005
Stefanick 1998_Female	-0.340	0.339	0.115	-1.004	0.324	-1.003	0.316
Stefanick 1998_Male	-0.520	0.332	0.111	-1.172	0.132	-1.564	0.118
	-0.627	0.239	0.057	-1.095	-0.160	-2.629	0.009

1.1. Sensitivity analysis prevention of obesity high quality studies

Figure 13. Sensitivity analysis weight change (kg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	-2.300	0.842	0.709	-3.951	-0.649	-2.731	0.006
Simkin-Silverman 2003	-2.500	0.556	0.309	-3.590	-1.410	-4.494	0.000
Stefanick 1998_Female	-3.900	0.845	0.715	-5.557	-2.243	-4.614	0.000
Stefanick 1998_Male	-4.700	0.732	0.535	-6.134	-3.266	-6.423	0.000
Trento 2002	-1.700	1.031	1.063	-3.721	0.321	-1.649	0.099
	-3.072	0.539	0.291	-4.129	-2.015	-5.696	0.000

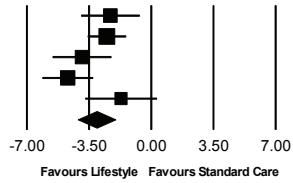


Figure 14. Sensitivity analysis systolic blood pressure change (mmHg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Simkin-Silverman 2003	-0.300	1.126	1.269	-2.508	1.908	-0.266	0.790
Stefanick 1998_Female	-0.700	1.706	2.911	-4.044	2.644	-0.410	0.682
Stefanick 1998_Male	-3.300	1.518	2.305	-6.276	-0.324	-2.173	0.030
Trento 2002	-4.000	2.400	5.760	-8.704	0.704	-1.667	0.096
	-1.623	0.888	0.789	-3.364	0.118	-1.827	0.068

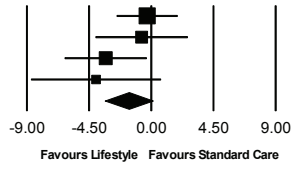


Figure 15. Sensitivity analysis diastolic blood pressure change (mmHg) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Simkin-Silverman 2003	-0.700	0.736	0.542	-2.143	0.743	-0.951	0.342
Stefanick 1998_Female	-2.100	1.131	1.280	-4.317	0.117	-1.856	0.063
Stefanick 1998_Male	-4.800	1.312	1.722	-7.372	-2.228	-3.658	0.000
Trento 2002	-0.800	1.569	2.460	-3.874	2.274	-0.510	0.610
	-2.005	0.925	0.856	-3.818	-0.191	-2.167	0.030

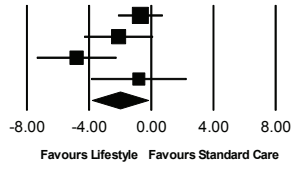


Figure 16. Sensitivity analysis total cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	-0.100	0.141	0.020	-0.375	0.175	-0.711	0.477
Stefanick 1998_Female	-0.420	0.112	0.013	-0.639	-0.201	-3.751	0.000
Stefanick 1998_Male	-0.430	0.111	0.012	-0.648	-0.212	-3.860	0.000
Trento 2002	-0.200	0.204	0.042	-0.600	0.200	-0.980	0.327
	-0.318	0.083	0.007	-0.480	-0.156	-3.852	0.000

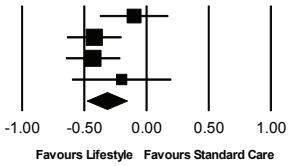


Figure 17. Sensitivity analysis high density lipoprotein cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	0.010	0.033	0.001	-0.054	0.074	0.304	0.761
Simkin-Silverman 2003	-0.020	0.026	0.001	-0.070	0.030	-0.778	0.437
Stefanick 1998_Female	-0.060	0.035	0.001	-0.129	0.009	-1.706	0.088
Stefanick 1998_Male	0.020	0.026	0.001	-0.031	0.071	0.768	0.443
Trento 2002	0.100	0.055	0.003	-0.007	0.207	1.825	0.068
	0.001	0.020	0.000	-0.038	0.040	0.050	0.960

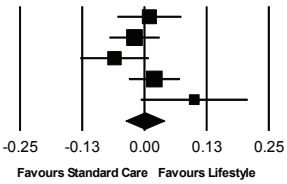


Figure 18. Sensitivity analysis low density lipoprotein cholesterol change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	0.000	0.139	0.019	-0.272	0.272	0.000	1.000
Simkin-Silverman 2003	-0.140	0.066	0.004	-0.269	-0.011	-2.133	0.033
Stefanick 1998_Female	-0.310	0.107	0.012	-0.520	-0.100	-2.888	0.004
Stefanick 1998_Male	-0.400	0.103	0.011	-0.603	-0.197	-3.866	0.000
	-0.219	0.080	0.006	-0.376	-0.062	-2.737	0.006

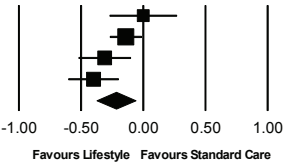


Figure 19. Sensitivity analysis triglyceride change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	-0.550	0.171	0.029	-0.886	-0.214	-3.211	0.001
Simkin-Silverman 2003	-0.130	0.085	0.007	-0.297	0.037	-1.527	0.127
Stefanick 1998_Female	-0.140	0.111	0.012	-0.358	0.078	-1.261	0.207
Stefanick 1998_Male	-0.180	0.164	0.027	-0.501	0.141	-1.100	0.271
Trento 2002	-0.260	0.181	0.033	-0.616	0.096	-1.433	0.152
	-0.210	0.067	0.005	-0.342	-0.078	-3.113	0.002

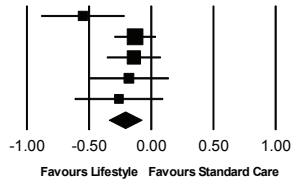


Figure 20. Sensitivity analysis HbA1c change (%) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	0.100	0.141	0.020	-0.175	0.375	0.711	0.477
Trento 2002	-1.600	0.144	0.021	-1.882	-1.318	-11.140	0.000
	-0.750	0.850	0.722	-2.416	0.916	-0.882	0.378

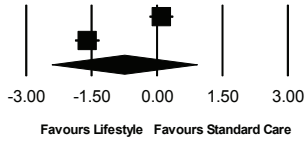
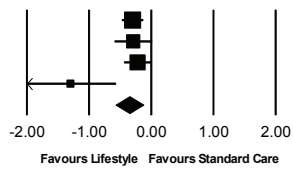


Figure 21. Sensitivity analysis fasting plasma glucose change (mmol/l) in overweight people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Mensink 2003ab	-0.300	0.088	0.008	-0.472	-0.128	-3.412	0.001
Stefanick 1998_Female	-0.290	0.155	0.024	-0.593	0.013	-1.876	0.061
Stefanick 1998_Male	-0.220	0.111	0.012	-0.437	-0.003	-1.990	0.047
Trento 2002	-1.300	0.374	0.140	-2.033	-0.567	-3.474	0.001
	-0.349	0.114	0.013	-0.573	-0.126	-3.063	0.002



1.2. Subgroup analysis overweight people with cardiovascular risk factors

Figure 22. Subgroup analysis weight change (kg) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Burke 2005	-2.500	0.721	0.520	-3.914	-1.086	-3.465	0.001
He 1999	0.400	1.286	1.655	-2.121	2.921	0.311	0.756
Kastarinen 2002	-1.200	0.423	0.179	-2.028	-0.372	-2.839	0.005
Ketola 2001	-1.300	0.858	0.736	-2.981	0.381	-1.516	0.130
Stefanick 1998_Female	-3.900	0.845	0.715	-5.557	-2.243	-4.614	0.000
Stefanick 1998_Male	-4.700	0.732	0.535	-6.134	-3.266	-6.423	0.000
	-2.296	0.703	0.494	-3.674	-0.919	-3.267	0.001

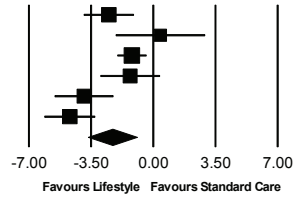


Figure 23. Subgroup analysis systolic blood pressure change (mmHg) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-8.000	2.316	5.364	-12.539	-3.461	-3.454	0.001
Burke 2005	-2.000	1.780	3.167	-5.488	1.488	-1.124	0.261
He 1999	-1.700	2.655	7.048	-6.903	3.503	-0.640	0.522
Kastarinen 2002	-2.000	1.171	1.372	-4.296	0.296	-1.707	0.088
Ketola 2001	-0.700	2.082	4.335	-4.781	3.381	-0.336	0.737
Stefanick 1998_Female	-0.700	1.706	2.911	-4.044	2.644	-0.410	0.682
Stefanick 1998_Male	-3.300	1.518	2.305	-6.276	-0.324	-2.173	0.030
	-2.431	0.778	0.605	-3.955	-0.907	-3.126	0.002

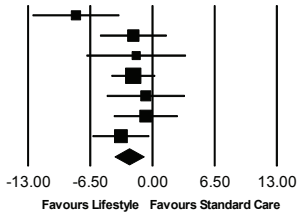


Figure 24. Subgroup analysis diastolic blood pressure change (mmHg) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-6.100	1.766	3.118	-9.561	-2.639	-3.455	0.001
Burke 2005	-2.000	1.163	1.353	-4.280	0.280	-1.720	0.086
He 1999	-1.300	1.624	2.638	-4.484	1.884	-0.800	0.424
Kastarinen 2002	-1.100	0.662	0.438	-2.398	0.198	-1.661	0.097
Ketola 2001	1.100	1.355	1.837	-1.557	3.757	0.812	0.417
Stefanick 1998_Female	-2.100	1.131	1.280	-4.317	0.117	-1.856	0.063
Stefanick 1998_Male	-4.800	1.312	1.722	-7.372	-2.228	-3.658	0.000
	-2.156	0.770	0.593	-3.665	-0.646	-2.798	0.005

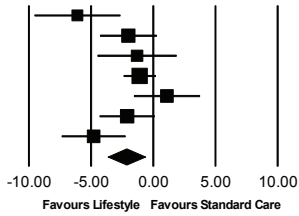


Figure 25. Subgroup analysis total cholesterol change (mmol/l) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andersen 1996	-0.600	0.165	0.027	-0.923	-0.277	-3.638	0.000
Kastarinen 2002	-0.100	0.081	0.007	-0.258	0.058	-1.238	0.216
Ketola 2001	-0.260	0.176	0.031	-0.606	0.086	-1.474	0.140
Stefanick 1998_Female	-0.420	0.112	0.013	-0.639	-0.201	-3.751	0.000
Stefanick 1998_Male	-0.430	0.111	0.012	-0.648	-0.212	-3.860	0.000
	-0.346	0.094	0.009	-0.531	-0.161	-3.661	0.000

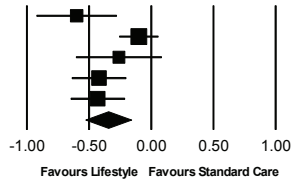


Figure 26. Subgroup analysis triglyceride change (mmol/l) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	-0.890	0.226	0.051	-1.333	-0.447	-3.941	0.000
Kastarinen 2002	0.000	0.072	0.005	-0.141	0.141	0.000	1.000
Stefanick 1998_Female	-0.140	0.111	0.012	-0.358	0.078	-1.261	0.207
Stefanick 1998_Male	-0.180	0.164	0.027	-0.501	0.141	-1.100	0.271
	-0.242	0.141	0.020	-0.519	0.036	-1.709	0.087

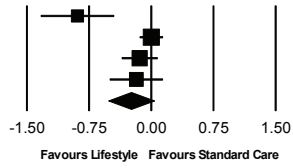


Figure 27. Subgroup analysis high density lipoprotein change (mmol/l) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderssen 1996	0.180	0.181	0.033	-0.175	0.535	0.995	0.320
Kastarinen 1997	0.030	0.022	0.000	-0.013	0.073	1.383	0.167
Stefanick 1998_Female	-0.060	0.035	0.001	-0.129	0.009	-1.706	0.088
Stefanick 1998_Male	0.020	0.026	0.001	-0.031	0.071	0.768	0.443
	0.006	0.024	0.001	-0.041	0.053	0.256	0.798

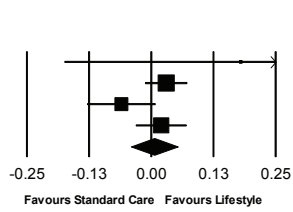
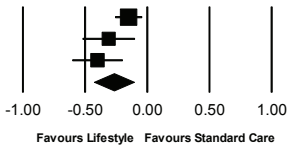


Figure 28. Subgroup analysis low density lipoprotein change (mmol/l) in overweight with cardiovascular risks

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Kastarinen 2002	-0.150	0.055	0.003	-0.258	-0.042	-2.710	0.007
Stefanick 1998_Female	-0.310	0.107	0.012	-0.520	-0.100	-2.888	0.004
Stefanick 1998_Male	-0.400	0.103	0.011	-0.603	-0.197	-3.866	0.000
	-0.267	0.082	0.007	-0.428	-0.106	-3.254	0.001



2. Meta-analysis treatment of obesity

Figure 29. Meta-analysis weight change (kg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-5.500	0.109	0.012	-5.715	-5.285	-50.248	0.000
Esposito 2003	-11.000	1.515	2.295	-13.969	-8.031	-7.261	0.000
Lindahl 1999	-4.900	0.329	0.109	-5.546	-4.254	-14.873	0.000
Lindstrom 2003	-2.600	0.504	0.254	-3.588	-1.612	-5.155	0.000
Messier 2004	-4.100	0.816	0.665	-5.699	-2.501	-5.026	0.000
Moore 2003	1.300	1.578	2.490	-1.793	4.393	0.824	0.410
Narayan 1998	-1.700	1.346	1.812	-4.338	0.938	-1.263	0.207
Stevens 2001	-2.000	0.338	0.114	-2.662	-1.338	-5.919	0.000
Whelton 1998	-4.940	0.780	0.609	-6.469	-3.411	-6.333	0.000
Wing 1998	-2.200	1.706	2.910	-5.543	1.143	-1.290	0.197
Wolf 2004	-3.000	1.210	1.464	-5.371	-0.629	-2.480	0.013
Yeh 2003	0.500	1.550	2.402	-2.538	3.538	0.323	0.747
	-3.487	0.620	0.384	-4.702	-2.272	-5.624	0.000

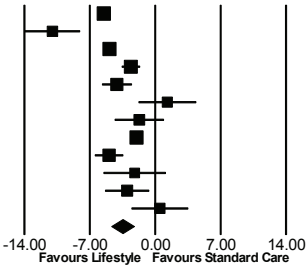


Figure 30. Meta-analysis BMI change (kg/m2) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.270	0.085	0.007	-2.436	-2.104	-26.783	0.000
Esposito 2003	-4.200	1.111	1.234	-6.377	-2.023	-3.781	0.000
Lindahl 1999	-1.800	0.118	0.014	-2.032	-1.568	-15.222	0.000
Lindstrom 2003	-1.000	0.187	0.035	-1.367	-0.633	-5.338	0.000
Moore 2003	0.100	0.534	0.286	-0.948	1.148	0.187	0.852
Wing 1998	-0.700	0.617	0.381	-1.909	0.509	-1.134	0.257
Yeh 2003	0.090	0.749	0.561	-1.378	1.558	0.120	0.904
	-1.329	0.308	0.095	-1.933	-0.724	-4.310	0.000

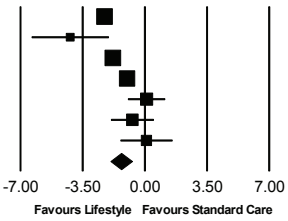


Figure 31. Meta-analysis systolic blood pressure change (mmHg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.700	0.707	0.500	-4.086	-1.314	-3.818	0.000
Lindahl 1999	-6.200	1.564	2.445	-9.265	-3.135	-3.965	0.000
Lindstrom 2003	-4.000	1.270	1.612	-6.488	-1.512	-3.151	0.002
Narayan 1998	1.900	2.677	7.168	-3.348	7.148	0.710	0.478
Stevens 2001	-1.400	0.533	0.284	-2.445	-0.355	-2.625	0.009
Wing 1998	-3.300	3.429	11.759	-10.021	3.421	-0.962	0.336
	-2.781	0.816	0.667	-4.381	-1.180	-3.406	0.001

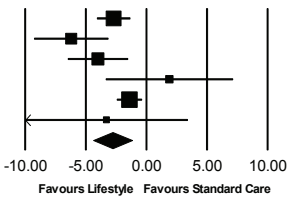


Figure 32. Meta-analysis diastolic blood pressure change (mmHg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-1.940	0.424	0.180	-2.771	-1.109	-4.573	0.000
Lindahl 1999	-2.400	1.320	1.742	-4.987	0.187	-1.818	0.069
Lindstrom 2003	-2.000	0.788	0.621	-3.544	-0.456	-2.538	0.011
Narayan 1998	2.000	1.750	3.062	-1.430	5.430	1.143	0.253
Stevens 2001	-0.800	0.419	0.175	-1.620	0.020	-1.911	0.056
Wing 1998	-2.200	2.357	5.557	-6.820	2.420	-0.933	0.351
	-1.417	0.433	0.188	-2.266	-0.569	-3.273	0.001

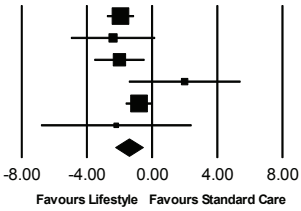


Figure 33. Meta-analysis total cholesterol change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Esposito 2003	-0.100	0.104	0.011	-0.303	0.103	-0.966	0.334
Lindahl 1999	-0.150	0.158	0.025	-0.460	0.160	-0.947	0.344
Lindstrom 2003	-0.200	0.082	0.007	-0.361	-0.039	-2.432	0.015
Narayan 1998	0.100	0.228	0.052	-0.346	0.546	0.439	0.661
Wing 1998	-0.090	0.153	0.023	-0.389	0.209	-0.590	0.555
	-0.136	0.054	0.003	-0.242	-0.031	-2.527	0.011

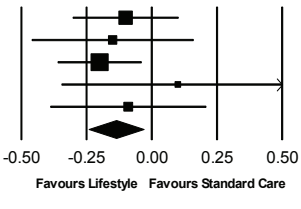


Figure 34. Meta-analysis high density lipoprotein cholesterol change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	0.025	0.012	0.000	0.001	0.049	2.004	0.045
Esposito 2003	0.100	0.027	0.001	0.048	0.152	3.772	0.000
Lindstrom 2003	0.030	0.019	0.000	-0.007	0.067	1.596	0.110
Wing 1998	-0.020	0.057	0.003	-0.131	0.091	-0.352	0.725
	0.040	0.018	0.000	0.004	0.076	2.193	0.028

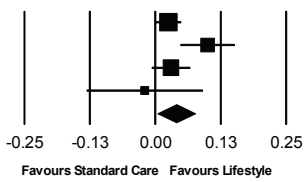


Figure 35. Meta-analysis triglyceride change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.166	0.041	0.002	-0.247	-0.085	-4.019	0.000
Lindahl 1999	-0.070	0.157	0.025	-0.378	0.238	-0.446	0.656
Lindstrom 2003	-0.100	0.067	0.005	-0.232	0.032	-1.484	0.138
Wing 1998	-0.800	0.313	0.098	-1.413	-0.187	-2.560	0.010
	-0.153	0.060	0.004	-0.272	-0.035	-2.540	0.011

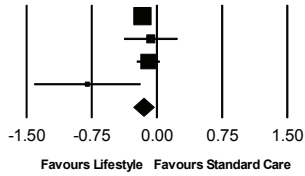


Figure 36. Meta-analysis fasting plasma glucose change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.300	0.028	0.001	-0.356	-0.244	-10.565	0.000
Lindahl 1999	-0.190	0.124	0.015	-0.433	0.053	-1.531	0.126
Lindstrom 2003	-0.100	0.067	0.005	-0.232	0.032	-1.485	0.138
Narayan 1998	0.000	0.285	0.081	-0.558	0.558	0.000	1.000
Wing 1998	0.300	0.244	0.060	-0.178	0.778	1.229	0.219
	-0.146	0.083	0.007	-0.309	0.017	-1.759	0.079

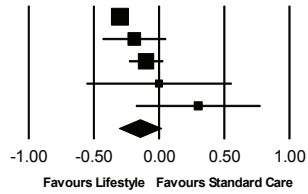


Figure 37. Meta-analysis 2-hour plasma glucose change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Lindahl 1999	-0.380	0.290	0.084	-0.949	0.189	-1.309	0.191
Lindstrom 2003	-0.600	0.183	0.033	-0.958	-0.242	-3.286	0.001
	-0.538	0.155	0.024	-0.841	-0.235	-3.478	0.001

Figure 37a. Meta-analysis HbA1c change (%) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Lindstrom 2003	-0.200	0.058	0.003	-0.313	-0.087	-3.465	0.001
Wing 1998	0.140	0.201	0.040	-0.254	0.534	0.696	0.486
	-0.085	0.161	0.026	-0.400	0.231	-0.525	0.599

2.1. Sensitivity analysis treatment of obesity high quality studies

Figure 38. Sensitivity analysis weight change (kg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-5.500	0.109	0.012	-5.715	-5.285	-50.248	0.000
Esposito 2003	-11.000	1.515	2.295	-13.969	-8.031	-7.261	0.000
Lindstrom 2003	-2.600	0.504	0.254	-3.588	-1.612	-5.155	0.000
Whelton 1998	-4.940	0.780	0.609	-6.469	-3.411	-6.333	0.000
Wing 1998	-2.200	1.706	2.910	-5.543	1.143	-1.290	0.197
	-5.086	0.970	0.942	-6.989	-3.184	-5.241	0.000

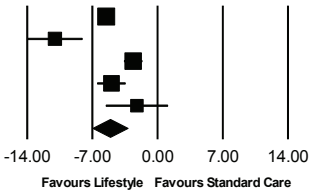


Figure 39. Sensitivity analysis BMI change (kg/m2) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.270	0.085	0.007	-2.436	-2.104	-26.783	0.000
Esposito 2003	-4.200	1.111	1.234	-6.377	-2.023	-3.781	0.000
Lindstrom 2003	-1.000	0.187	0.035	-1.367	-0.633	-5.338	0.000
Wing 1998	-0.700	0.617	0.381	-1.909	0.509	-1.134	0.257
	-1.769	0.521	0.271	-2.790	-0.749	-3.398	0.001

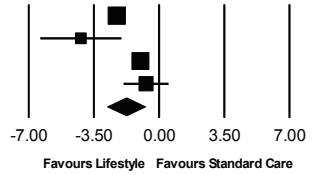


Figure 40. Sensitivity analysis systolic blood pressure change (mmHg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.700	0.707	0.500	-4.086	-1.314	-3.818	0.000
Lindstrom 2003	-4.000	1.270	1.612	-6.488	-1.512	-3.151	0.002
Wing 1998	-3.300	3.429	11.759	-10.021	3.421	-0.962	0.336
	-3.017	0.608	0.370	-4.209	-1.825	-4.962	0.000

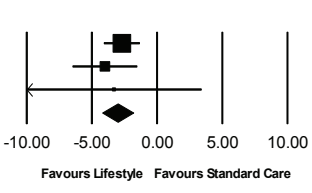


Figure 41. Sensitivity analysis diastolic blood pressure change (mmHg) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-1.940	0.424	0.180	-2.771	-1.109	-4.573	0.000
Lindstrom 2003	-2.000	0.788	0.621	-3.544	-0.456	-2.538	0.011
Wing 1998	-2.200	2.357	5.557	-6.820	2.420	-0.933	0.351
	-1.960	0.369	0.136	-2.683	-1.236	-5.312	0.000

Figure 42. Sensitivity analysis total cholesterol change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Esposito 2003	-0.100	0.104	0.011	-0.303	0.103	-0.966	0.334
Lindstrom 2003	-0.200	0.082	0.007	-0.361	-0.039	-2.432	0.015
Wing 1998	-0.090	0.153	0.023	-0.389	0.209	-0.590	0.555
	-0.151	0.059	0.004	-0.267	-0.034	-2.538	0.011

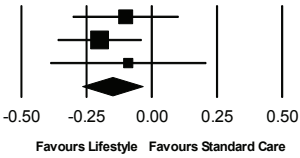


Figure 43. Sensitivity analysis high density lipoprotein cholesterol change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	0.025	0.012	0.000	0.001	0.049	2.004	0.045
Esposito 2003	0.100	0.027	0.001	0.048	0.152	3.772	0.000
Lindstrom 2003	0.030	0.019	0.000	-0.007	0.067	1.596	0.110
Wing 1998	-0.020	0.057	0.003	-0.131	0.091	-0.352	0.725
	0.040	0.018	0.000	0.004	0.076	2.193	0.028

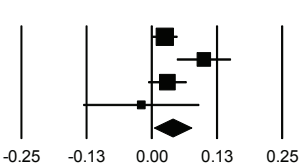
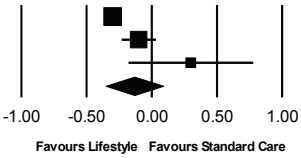


Figure 44. Sensitivity analysis triglyceride change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.166	0.041	0.002	-0.247	-0.085	-4.019	0.000
Lindstrom 2003	-0.100	0.067	0.005	-0.232	0.032	-1.484	0.138
Wing 1998	-0.800	0.313	0.098	-1.413	-0.187	-2.560	0.010
	-0.172	0.075	0.006	-0.319	-0.025	-2.293	0.022

Figure 45. Sensitivity analysis fasting plasma glucose change (mmol/l) in obese people

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.300	0.028	0.001	-0.356	-0.244	-10.565	0.000
Lindstrom 2003	-0.100	0.067	0.005	-0.232	0.032	-1.485	0.138
Wing 1998	0.300	0.244	0.060	-0.178	0.778	1.229	0.219
	-0.132	0.112	0.013	-0.353	0.088	-1.180	0.238



2.2. Subgroup analysis overweight/obese at risk of diabetes

Figure 46. Subgroup analysis weight change (kg) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Carr 2005	-2.400	0.499	0.249	-3.378	-1.422	-4.810	0.000
DPP 2005ab	-5.500	0.109	0.012	-5.715	-5.285	-50.248	0.000
Dyson 1997	-0.200	0.835	0.697	-1.836	1.436	-0.240	0.811
Liao 2002	-2.500	0.145	0.021	-2.784	-2.216	-17.237	0.000
Lindahl 1999	-4.900	0.329	0.109	-5.546	-4.254	-14.873	0.000
Lindstrom 2003	-2.600	0.592	0.351	-3.761	-1.439	-4.390	0.000
Mensink 2003ab	-2.300	0.842	0.709	-3.951	-0.649	-2.731	0.006
Wing 1998	-2.200	1.706	2.910	-5.543	1.143	-1.290	0.197
	-2.933	0.723	0.522	-4.349	-1.516	-4.058	0.000

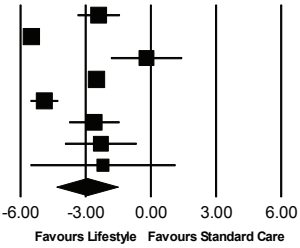


Figure 47. Subgroup analysis BMI change (kg/m2) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.270	0.085	0.007	-2.436	-2.104	-26.783	0.000
Liao 2002	-0.900	0.053	0.003	-1.003	-0.797	-17.135	0.000
Lindahl 1999	-1.800	0.118	0.014	-2.032	-1.568	-15.222	0.000
Lindstrom 2003	-1.000	0.220	0.048	-1.431	-0.569	-4.549	0.000
Mensink 2003ab	-0.800	0.290	0.084	-1.369	-0.231	-2.757	0.006
Wing 1998	-0.700	0.617	0.381	-1.909	0.509	-1.134	0.257
	-1.294	0.332	0.110	-1.944	-0.644	-3.902	0.000

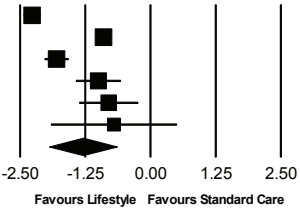


Figure 48. Subgroup analysis systolic blood pressure change (mmHg) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-2.700	0.707	0.500	-4.086	-1.314	-3.818	0.000
Dyson 1997	-2.000	1.736	3.015	-5.403	1.403	-1.152	0.249
Lindahl 1999	-6.200	1.564	2.445	-9.265	-3.135	-3.965	0.000
Lindstrom 2003	-4.000	1.290	1.663	-6.527	-1.473	-3.102	0.002
Wing 1998	-3.300	3.429	11.759	-10.021	3.421	-0.962	0.336
	-3.454	0.677	0.458	-4.781	-2.127	-5.103	0.000

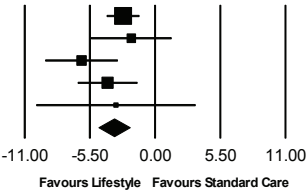


Figure 49. Subgroup analysis diatolic blood pressure change (mmHg) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-1.940	0.424	0.180	-2.771	-1.109	-4.573	0.000
Dyson 1997	-0.100	1.181	1.395	-2.415	2.215	-0.085	0.933
Lindahl 1999	-2.400	1.320	1.742	-4.987	0.187	-1.818	0.069
Lindstrom 2003	-2.000	0.800	0.640	-3.568	-0.432	-2.499	0.012
Wing 1998	-2.200	2.357	5.557	-6.820	2.420	-0.933	0.351
	-1.834	0.341	0.116	-2.502	-1.165	-5.374	0.000

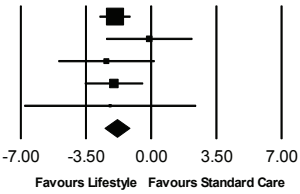


Figure 50. Subgroup analysis total cholesterol change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Dyson 1997	0.000	0.155	0.024	-0.303	0.303	0.000	1.000
Lindahl 1999	-0.150	0.158	0.025	-0.460	0.160	-0.947	0.344
Lindstrom 2003	-0.200	0.094	0.009	-0.385	-0.015	-2.123	0.034
Mensink 2003ab	-0.100	0.141	0.020	-0.375	0.175	-0.711	0.477
Wing 1998	-0.090	0.153	0.023	-0.389	0.209	-0.590	0.555
	-0.130	0.059	0.003	-0.246	-0.015	-2.207	0.027

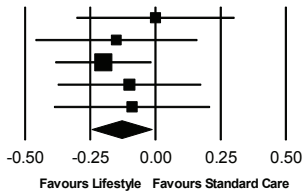


Figure 51. Subgroup analysis high density lipoprotein change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	0.025	0.012	0.000	0.001	0.049	2.004	0.045
Dyson 1997	-0.010	0.042	0.002	-0.091	0.071	-0.241	0.810
Lindstrom 2003	0.030	0.022	0.000	-0.013	0.073	1.380	0.168
Mensink 2003ab	0.010	0.033	0.001	-0.054	0.074	0.304	0.761
Wing 1998	-0.020	0.057	0.003	-0.131	0.091	-0.352	0.725
	0.021	0.010	0.000	0.002	0.041	2.175	0.030

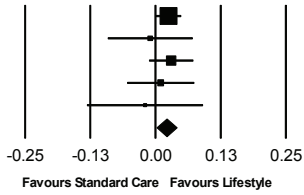


Figure 52. Subgroup analysis low density lipoprotein change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Dyson 1997	0.100	0.106	0.011	-0.108	0.308	0.944	0.345
Mensink 2003ab	-0.150	0.089	0.008	-0.325	0.025	-1.683	0.092
Wing 1998	-0.100	0.157	0.025	-0.407	0.207	-0.638	0.524
	-0.050	0.084	0.007	-0.215	0.116	-0.590	0.555

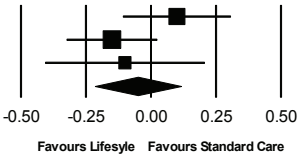


Figure 53. Subgroup analysis triglyceride change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.166	0.041	0.002	-0.247	-0.085	-4.019	0.000
Dyson 1997	-0.140	0.137	0.019	-0.409	0.129	-1.019	0.308
Lindahl 1999	-0.070	0.157	0.025	-0.378	0.238	-0.446	0.656
Lindstrom 2003	-0.100	0.082	0.007	-0.260	0.060	-1.225	0.220
Mensink 2003ab	-0.550	0.171	0.029	-0.886	-0.214	-3.211	0.001
Wing 1998	-0.800	0.313	0.098	-1.413	-0.187	-2.560	0.010
	-0.202	0.067	0.004	-0.333	-0.071	-3.025	0.002

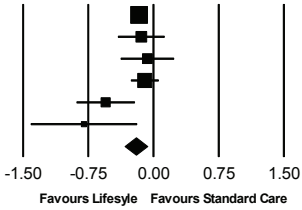


Figure 54. Subgroup analysis HbA1c change (%) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Dyson 1997	0.000	0.109	0.012	-0.213	0.213	0.000	1.000
Lindstrom 2003	-0.200	0.067	0.005	-0.332	-0.068	-2.974	0.003
Mensink 2003ab	0.100	0.141	0.020	-0.175	0.375	0.711	0.477
Wing 1998	0.140	0.201	0.040	-0.254	0.534	0.696	0.486
	-0.035	0.087	0.008	-0.205	0.135	-0.404	0.686

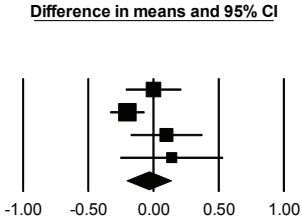


Figure 55. Subgroup analysis fasting plasma glucose change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
DPP 2005ab	-0.300	0.028	0.001	-0.356	-0.244	-10.565	0.000
Dyson 1997	0.100	0.283	0.080	-0.455	0.655	0.353	0.724
Lindahl 1999	-0.190	0.124	0.015	-0.433	0.053	-1.531	0.126
Lindstrom 2003	-0.100	0.078	0.006	-0.254	0.054	-1.275	0.202
Mensink 2003ab	-0.300	0.088	0.008	-0.472	-0.128	-3.412	0.001
Wing 1998	-0.300	0.244	0.060	-0.778	0.178	-1.229	0.219
	-0.234	0.048	0.002	-0.328	-0.139	-4.853	0.000

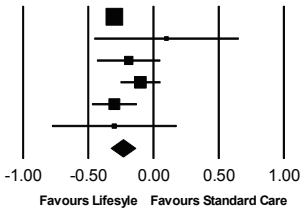
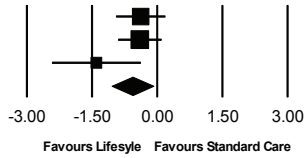


Figure 56. Subgroup analysis 2-hour plasma glucose change (mmol/l) in overweight/obese at risk of diabetes

Study name	Difference in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Lindahl 1999	-0.380	0.290	0.084	-0.949	0.189	-1.309	0.191
Lindstrom 2003	-0.400	0.256	0.065	-0.901	0.101	-1.565	0.118
Mensink 2003ab	-1.400	0.521	0.271	-2.421	-0.379	-2.687	0.007
	-0.567	0.245	0.060	-1.048	-0.087	-2.316	0.021



Chapter 3

Regulatory Requirements: Validation of Electronic Data Capture of Quality of Life Instruments

Summary

Traditionally, health-related quality of life research has predominately been conducted with the use of paper-and-pencil questionnaires. However, as computers become more prevalent, a greater number of studies are utilizing electronic data in clinical trials setting. Regulatory agencies require from health technology manufacturers to provide evidence of the validity of the quality of life instruments in the electronic format as compared to paper version. The aims of the present chapter was to assess the comparability, reliability, and subject acceptability of electronic data capture (EDC) versions of Irritable Bowel Syndrome – Quality of Life (IBS-QOL), EuroQol (EQ-5D) and Work Productivity and Activity Impairment (WPAI:IBS) instruments. Comparability of EDC and paper questionnaires was evaluated in 72 subjects with IBS who completed a baseline EDC or paper questionnaire, a crossover questionnaire 24 hours later, and a retest of the crossover version at 1 week. The EDC version was presented on a hand-held device. Comparability of scores was assessed using paired *t*-test statistics, intraclass correlation coefficients (ICC) and tests for internal consistency (Cronbach's alpha). No significant differences were found between scores obtained by paper questionnaire and EDC at the baseline and crossover assessments. ICCs between baseline and crossover assessments ranged from 0.83 to 0.96 for the IBS-QOL scores, 0.82 to 0.96 for the WPAI:IBS scores, and 0.77 to 0.82 for the EQ-5D. Internal consistency was comparable for the two data collection methods for the IBS-QOL overall score (0.96) and subscales and the EQ-5D Index (0.70 vs. 0.74). Ease of use was comparable for the two modes of administration, but more patients preferred EDC (47.2%) than the paper questionnaire (23.6%). In summary, EDC versions of the IBS-QOL, EQ-5D, and WPAI:IBS are comparable to paper questionnaires in terms of internal consistency, and test-retest reliability, and have a greater patient acceptability.

Introduction

Patient-reported outcomes in clinical studies have generally been obtained by self-administered paper questionnaires. Electronic data capture (EDC) with hand-held or desktop computers is an alternative mode of data collection that offers many potential benefits over paper questionnaires, including customization of questions depending on a prior response, automatic date and time stamping, and immediate data entry that eliminates the possibility of subsequent entry errors. Several recent studies have found that patient-reported outcome data collected with EDC are psychometrically comparable to data collected by the standard paper mode, in terms of validity and reliability, and that EDC has high acceptance, and is generally preferred over the paper mode by the majority of subjects (Bushnell 2003, Drummond 1995, Pouwer 1998, Kleinman 2001, Bliven 2001, Caro 2001).

The validity and acceptability of EDC in studies of patients with irritable bowel syndrome (IBS) have not been investigated. Questionnaires useful for assessing outcomes in IBS studies include validated disease-specific quality of life and work productivity questionnaires, as well as general health questionnaires or utility measures for valuing IBS decrements relative to other diseases. The IBS Quality of Life Questionnaire (IBS-QOL) is an IBS-specific measure with established internal consistency, test-retest reliability and validity (Patrick 1998, Drossman 2000). The Work Productivity and Activity Impairment Questionnaire has been validated for IBS (WPAI:IBS) (Reilly 2004), as well as for other diseases, such as allergies (Reilly 1996), gastroesophageal reflux disease (Wahlqvist 2002) and chronic hand dermatitis (Reilly 2003). The validity and reliability of the EuroQol (EQ-5D) (Kind 1996), a general health measure, has been established in several diseases, including inflammatory bowel disease (Konig 2002), rheumatoid arthritis (Hurst 1997), AIDS (Wu 2002), and IBS (Bushnell 2006). The objectives of this study were to test the equivalence of the EDC versions of the IBS-QOL, WPAI:IBS and EQ-5D with the paper versions of these questionnaires in IBS patients, and to assess respondent acceptability of the EDC format to determine if EDC would be a valid and appropriate methodology for obtaining patient-reported outcomes in IBS studies.

Methods

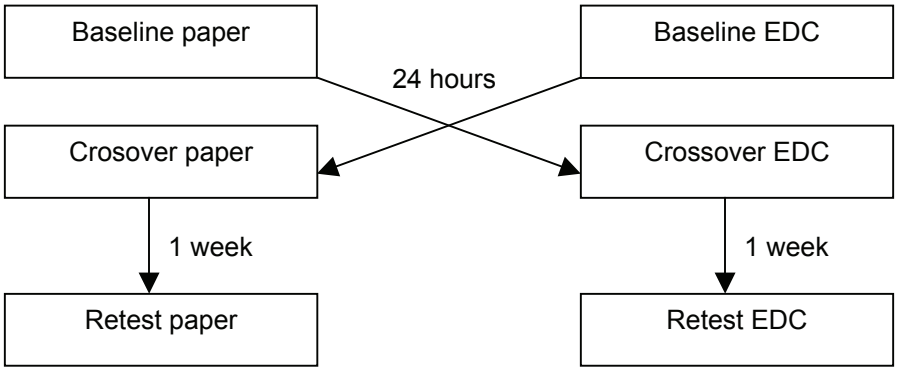
Subject Enrollment and Study Design

This was a randomized crossover study designed to test the effect of mode of questionnaire administration (paper vs. EDC) on patient-reported outcomes. The study was conducted at two USA sites (Seattle, Washington and Rockford, Illinois) with recruitment through general advertisement. Patients aged 18 years or above who met Rome II diagnostic criteria for IBS (Drossman 2000) and signed informed consent were eligible to participate. Patients were randomized in equal numbers to the two sequences of questionnaire administration; efforts were made to recruit patients who were diverse as to gender and type of IBS (constipation predominant, diarrhea predominant and alternating). The study

consisted of the completion of a baseline questionnaire (paper or EDC), a crossover questionnaire within 24 hours, and a retest of the crossover questionnaire seven days later, i.e. a sequence of either paper-EDC-EDC or EDC-paper-paper (See Figure 1).

A hand-held electronic device with entry by stylus was used for EDC. One question was displayed per screen and once a response was entered the next item automatically appeared. There was an option to return to review or change a previous question and all questions required a response. There were minor formatting differences between the paper questionnaire and EDC, e.g. in fonts and bolding, but the questions themselves were essentially identical in both presentations. All paper questionnaires and EDC were completed at the study site (research facility in Seattle and clinical setting in Rockford).

Figure 1. Study design, electronic data capture (EDC)



Additional information was obtained by paper questionnaire throughout the course of the study. Prior to randomization, all qualified patients completed questions about demographics, IBS characteristics and other health information. Following administration of the baseline and crossover questionnaires patients completed questions about the acceptability of the administration mode just completed, and at the crossover assessment, questions about preferences for the mode of administration. At baseline, symptom severity was rated and at the retest visit, patients completed a question assessing the global rating of change in overall quality of life. All patients enrolled completed the study. Patients were compensated for their participation.

Outcomes Questionnaires

The IBS-QOL is a 34-item condition-specific instrument that assesses overall quality of life and 8 domains (dysphoria, interference with activity, body image, health worry, food avoidance, social reactions, sexual and relationships). Each item has a five-point Likert response scale that assesses how much the item

describes the respondent during the past month (not at all, slightly, moderately, quite a bit, and extremely or a great deal). Items scores are summed to derive the overall score and eight subscales; scores are transformed to a 0 to100 scale ranging from 0 (poor quality of life) to 100 (maximum quality of life).

The WPAI:IBS asks questions about the effect of IBS symptoms, for example abdominal discomfort, abdominal pain, bloating, constipation, and diarrhea on ability to work and perform regular daily activities during the past seven days. It consists of six items: employment status; hours missed due to IBS; hours missed for other reasons; hours worked; lost work productivity and daily activity impairment due to IBS. Four scores are calculated: absenteeism (work time missed), presenteeism (impairment while at work), overall work productivity loss (absenteeism + presenteeism) and daily activity impairment. Scores are expressed as percentages, with higher scores indicating more productivity loss.

The EQ-5D consists of a 5-item descriptive system and the EQ Visual Analog Scale (VAS). The descriptive system records the level of self-reported problems on each of the dimensions of the classification (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) on the day the questionnaire is completed. Health states defined by the descriptive system can be converted into a weighted health state index by applying scores elicited from general population samples. Respondents describe their own health status that day using a 20cm VAS, ranging from 0, worst imaginable health state, to 100, best imaginable health state.

Additional Measures

The Short Form 36-Item Health Survey (SF-36) (Ware 1993) is a generic measure of functional status and well-being. It contains 36 questions that measure health across eight dimensions and two summary measures: the physical component (PCS) and the mental component (MCS). Scores within a dimension are reported on a scale from 0 to 100, where higher scores indicate ‘good health’ and 0 indicates ‘poor health’. Items with a recall period refer to the last 4 weeks.

Respondent acceptability of the two modes of administration was assessed with questions regarding ease of reading, ease in changing response and difficulty using the two formats. Responses were scored on 10 cm VAS, from 0, not at all easy (difficult) to 100, extremely easy (difficult). Preference for administration mode was assessed by asking which method was easier to use and which method was preferred. Severity of IBS symptoms during the past 7 days was assessed using a 0- to 10-point numerical scale ranging from no symptoms (0) to very severe symptoms (10). At the final retest visit, patients rated the global change in their overall quality of life as “A lot better”, “Somewhat better”, "About the same", "Somewhat worse”, or "A lot worse".

Statistical Methods

Comparability of the two modes of administration was evaluated by testing differences in scores between the baseline and crossover assessments with

Student's t-test and the intraclass correlation coefficient (ICC). The ICC is the preferred measure of strength of association for determining stability of scores over time because it corrects for lack of independence between measurement intervals (Deyo 1991). The ICC ranges between 0.00 and 1.00, and the minimal acceptable level is 0.70 (SACMOT 2002).

Cronbach's alpha was used to determine if the items within the scaled questionnaires, i.e. IBS-QOL and EQ-5D, had the same degree of association in the two modes of administration at the baseline assessment. A minimum correlation of 0.70 is necessary to support internal consistency and alpha values between 0.85 and 0.95 are preferred (Cronbach 1951). An alpha value of 0.95 has been previously reported for the IBS-QOL total sore (Patrick 1998). Internal reliability does not apply to the IBS:IBS or EQ-5D because these scores are single items.

Test-retest reliability was assessed using the ICC to compare the relationship between the crossover assessment and the retest assessments for the two modes of administration. The test-retest analysis was restricted to subjects rating their overall quality of life "about the same" on the global rating of change item at the retest visit. Previously reported ICC values for IBS-QOL total score was 0.86 (Patrick 1998).

For each mode of administration, the relationship between IBS-QOL, EQ-5D, and WPAI:IBS component scores and the physical and mental summary scores of the SF-36 was assessed using Pearson correlation coefficients, and the magnitude of the coefficients was compared across mode of administration. Discriminant validity of the scores from the two modes of administration (pooled from baseline and crossover for paper and for EDC) was compared relative to IBS symptom severity. An examination of responses to the symptom severity question indicated that they could be categorized as low (0-5), middle (6-7) or high (8-10) to permit a sufficient sample size in each category. Differences in mean scores between severity categories were tested with analysis of variance (ANOVA).

Acceptability of the two modes of administration was evaluated with descriptive statistics. Differences between the ease of administration of the two modes were assessed with Student's paired t-test.

Results for work productivity scores apply only to employed patients and are limited by the small sample size. All analyses were conducted using SPSS (SPSS 1999). A p-value <0.05 was required for significance using two-sided hypothesis tests; no p-value adjustments were made for the analysis of multiple endpoints.

Results

Table 1 shows the characteristics of the population by initial mode of questionnaire administration. Patients had a mean age of 46.2 years and were predominately female (86.1%); 69.4% were currently employed. All IBS types were represented: 25% of patients had IBS with constipation, 33.3% of patients had IBS with diarrhea, and 41.7% had alternating IBS. Compared to patients randomized to the paper-first group, the EDC-first group tended to be younger

(42.4 vs. 48.5 years), female (91.4% vs.81.1%), and less likely to have a college degree (18.9% vs. 34.3%). The EDC-first group was more likely to have IBS with diarrhea (40% vs. 27%) and have a shorter IBS duration (13.0 vs. 16.9 years). Employment rates were comparable.

Table 1 Characteristics of the population by mode of initial administration

	Initial administration, Mean (± s.d.) or percentage		
Characteristic	EDC (n=37)	Paper (n=35)	Total (n=72)
Age (years) [range]	42.4 (13.7) [21-72]	48.5 (14.2) [21-83]	46.2 (13.5) [21-83]
Gender (female, %)	81.1	91.4	86.1
Highest level of school (%)			
High school graduate or less	21.6	28.6	25.0
Some college/2-yr degree	59.5	37.1	48.6
4-year college graduate	18.9	34.3	26.4
Length of time with IBS symptoms (years)	13.0 (9.5)	16.9 (15.2)	14.8 (12.6)
Type of IBS (%)			
Constipation	20.0	29.7	25.0
Diarrhea	40.0	27.0	33.3
Alternating	40.0	43.2	41.7
Currently Employed (%)	70.3	68.6	69.4

EDC, Electronic Data Capture; IBS, Irritable Bowel Syndrome; s.d., standard deviation

Table 2a and 2b show the mean scores for the baseline and crossover assessments by mode of initial questionnaire administration. There were no significant differences between the baseline and crossover scores for subjects in either administration group. ICC statistics were above 0.70 for each IBS-QOL, EQ-5D and WPAI measure for both sequences of administration.

Table 2a. IBS-QOL, EQ-5D and WPAI:IBS scores for baseline and crossover questionnaires by first administration

	Mean (± standard deviation)		P-value[a]	Intraclass correlation coefficient
	Paper	EDC		
Baseline paper questionnaire and crossover electronic data capture (EDC) (n=35)				
IBS-QOL				
Overall	69.3 (18.4)	73.8 (17.6)	0.29	0.96
Dysphoria	71.0 (21.7)	76.9 (20.4)	0.24	0.91
Interference with activity	58.9 (24.9)	63.9 (24.4)	0.39	0.95
Body image	62.4 (22.0)	68.1 (19.9)	0.25	0.96
Health worry	71.2 (22.5)	76.0 (20.6)	0.34	0.83
Food avoidance	62.7 (24.3)	63.5 (23.1)	0.88	0.91
Social reactions	75.3 (21.9)	79.7 (19.7)	0.38	0.91
Sexual	81.9 (20.3)	84.9 (18.0)	0.51	0.93
Relationships	76.2 (21.7)	81.0 (21.5)	0.34	0.93
EQ-5D				
VAS	75.8 (15.4)	75.1 (16.4)	0.85	0.77
Index	0.72 (0.25)	0.72 (0.24)	0.91	0.80
WPAI:IBS				
Absenteeism (n=23)	4.0 (9.3)	3.7 (7.7)	0.52	0.96
Presenteeism (n=23)	27.4 (23.0)	22.6 (20.9)	0.15	0.84
Work Productivity Loss (n=23)	31.4 (28.0)	26.3 (20.9)	0.56	0.90
Daily Activity Impairment (n=35)	33.7 (25.0)	29.7 (24.0)	0.18	0.87

[a]Student's t-test; IBS-QOL, Irritable Bowel Syndrome-Quality of Life; VAS, Viasual Analog Scale; WPAI:IBS, Work Productivity Activity Impairement questionaire – Irritable Bowel Syndrome version

Table 2b. IBS-QOL, EQ-5D and WPAI:IBS scores for baseline and crossover questionnaires by first administration

	Mean (± standard deviation)		P-value[a]	Intraclass correlation coefficient
	Paper	EDC		
Baseline electronic data capture (EDC) and 24-hour crossover paper questionnaire (n = 37)				
IBS-QOL				
Overall	66.1 (20.5)	63.5 (21.0)	0.60	0.96
Dysphoria	66.7 (25.3)	62.9 (26.6)	0.54	0.95
Interference with activity	60.4 (29.2)	56.4 (30.2)	0.57	0.91
Body image	63.8 (22.0)	60.6 (21.3)	0.53	0.93
Health worry	67.0 (25.9)	68.1 (25.9)	0.87	0.91
Food avoidance	58.5 (29.1)	54.9 (29.4)	0.60	0.94
Social reactions	69.3 (26.1)	69.3 (24.1)	1.00	0.91
Sexual and relationships	72.5 (22.8)	72.7 (23.5)	0.97	0.93
	75.0 (22.8)	71.8 (22.5)	0.55	0.90
EQ-5D				
VAS	70.8 (21.0)	64.0 (24.5)	0.21	0.82
Index	0.71 (0.23)	0.63 (0.28)	0.16	0.77
WPAI:IBS				
Absenteeism (n=25)	3.3 (6.4)	3.4 (7.5)	0.78	0.94
Presenteeism (n=25)	30.4 (21.5)	33.6 (23.6)	0.34	0.85
Work Productivity Loss (n=25)	33.7 (24.7)	37.0 (28.2)	0.32	0.90
Daily Activity Impairment (n=25)	40.5 (25.7)	42.0 (23.9)	0.67	0.82

[a]Student's t-test; IBS-QOL, Irritable Bowel Syndrome-Quality of Life; VAS, Viasual Analog Scale; WPAI:IBS, Work Productivity Activity Impairement questionaire – Irritable Bowel Syndrome version

Table 3 shows the results of the internal consistency evaluation of the IBS-QOL and the EQ-5D index at baseline and test-retest reliability of these two measures and the EQ-5D VAS and WPAI:IBS. Internal consistency was comparable for both modes of administration of the IBS-QOL and the EQ-5D; both questionnaires and modes of administration demonstrated internal consistency with alpha values all above 0.70, with the exception of the Relationship domain of the IBS-QOL EDC which was 0.69. Among stable subjects, retest statistics (ICC) were comparable between the EDC and paper versions of the IBS-QOL, with all correlations above 0.88, and for the EDC and paper versions of the EQ-5D, with all correlations above 0.73. The analysis of the reliability of the WPAI:IBS was limited by the small sample of stable employed patients in the paper and EDC groups (n= 15 and 13, respectively). However, all correlations were above 0.75, except for absenteeism, which was 0.68.

Table 3. Internal consistency and test-retest reliability of the questionnaires by mode of administration

	Internal Consistency [a]		Test-Retest Reliability [b]	
	Paper (n=35)	EDC (n=37)	Paper (vs paper) (n=20)	EDC (vs EDC) (n=20)
IBS-QOL				
Overall	0.96	0.96	0.99	0.95
Dysphoria	0.94	0.95	0.99	0.93
Interference with activity	0.82	0.89	0.93	0.96
Body image	0.79	0.72	0.93	0.95
Health worry	0.74	0.79	0.94	0.88
Food avoidance	0.83	0.88	0.95	0.90
Social reactions	0.84	0.80	0.91	0.90
Sexual	0.75	0.77	0.92	0.94
Relationships	0.77	0.69	0.94	0.92
EQ-5D Index	0.74	0.70	0.77	0.75
EQ-5D VAS	NA	NA	0.82	0.73
WPAI:IBS				
Absenteeism	NA	NA	0.68 (n=15)	0.93 (n=13)
Presenteeism	NA	NA	0.75 (n=15)	0.97 (n=13)
Work Productivity Loss	NA	NA	0.84 (n=15)	0.98 (n=13)
Daily Activity Impairment	NA	NA	0.90 (n=20)	0.83 (n=20)

[a] As measured by Chronbach's alpha using baseline administration; [b] As measured by the intraclass correlation coeficient using the crossover and retest assessment at 1 week (includes only those patients reporting no change on the global rating of change at 1 week retest. EDC, Electronic Data Capture; NA, not applicable; VAS, Visial Analog Scale; WPAI:IBS, Work Productivity Activity Impairement questionaire – Irritable Bowel Syndrome version

Table 4 shows the correlation coefficients between the IBS-QOL, EQ-5D and WPAI:IBS component scores and the SF-36 physical and mental summary

scores. The correlation coefficients were generally comparable for the two modes of administration.

Table 4. Concurrence of correlation coefficients between the paper and electronic modes

	SF-36 Physical Component Summary		SF-36 Mental Component Summary	
	Paper (n = 72)	EDC (n = 72)	Paper (n = 72)	EDC (n = 72)
IBS-QOL				
Overall	0.40 [a]	0.36 [a]	0.47 [a]	0.46 [a]
Dysphoria	0.49 [a]	0.45 [a]	0.51 [a]	0.50 [a]
Interference with Activity	0.36 [a]	0.30 [a]	0.37 [a]	0.40 [a]
Body Image	0.21	0.23	0.32 [a]	0.34 [a]
Health Worry	0.32 [a]	0.35 [a]	0.43 [a]	0.44 [a]
Food Avoidance	0.33 [a]	0.24 [a]	0.34 [a]	0.34 [a]
Social Reaction	0.19	0.17	0.29 [a]	0.33 [a]
Sexual	0.22	0.19	0.32 [a]	0.24 [a]
Relationships	0.22	0.26 [a]	0.31 [a]	0.27 [a]
EQ-5D				
VAS	0.55 [a]	0.60 [a]	0.54 [a]	0.45 [a]
Index	0.63 [a]	0.60 [a]	0.32 [a]	0.47 [a]
WPAI:IBS				
Absenteeism[b]	0.04	0.07	-0.33 [a]	-0.41 [a]
Presenteeism[c]	-0.33 [a]	-0.22	-0.35 [a]	-0.46 [a]
Work Productivity Loss[b]	-0.24	-0.17	-0.39 [a]	-0.49 [a]
Daily Activity Impairment	-0.60 [a]	-0.43 [a]	-0.34 [a]	-0.40 [a]

[a] Pearson Correlation is significant at the ≤ 0.05 level (2-tailed) [b] (n=49 paper, n=48 EDC), [c] (n=50 paper, n=48 EDC); EDC, electronic data capture; IBS-QOL, Irritable Bowel Syndrome-Quality of Life; VAS, Viasual Analog Scale; WPAI:IBS, Work Productivity Activity Impairement questionaire – Irritable Bowel Syndrome version

Table 5 shows the mean IBS-QOL, WPAI:IBS, and EQ-5D scores for the two modes of administration (pooled baseline and crossover) by level of disease severity. For the paper version, there were significant differences for all measures by severity category (p values 0.03 to <0.0001), with higher severity scores associated with higher impairment. Each of the summary scores was markedly different at each level of symptom severity, with the exception of the EQ-5D index which had comparable scores for the low and middle severity groups (0.78 and 0.75). For EDC, there were significant differences by level of

severity for the IBS-QOL overall score (p< 0.002) and the WPAI:IBS activity impairment score (p<0.0001). Again, the EQ-5D index scores did not differentiate the low and middle symptom severity groups (0.70 for both). While the trend for the WPAI:IBS overall work productivity loss score indicated that higher symptom severity was associated with higher impairment (21.2%, 37.2% and 40.5% for the low, middle and high severity groups, respectively), the differences were not significant.

Table 5. IBS-QOL, EQ-5D and WPAI:IBS summary scores by symptom severity and mode of questionnaire administration.

IBS symptom severity		Overall IBS-QOL	EQ-5D VAS	Overall work productivity loss	Activity Impairment
Paper Questionnaire (pooled baseline and crossover)					
Low (0-5)	Mean	77.0	0.78	19.9	21.7
	N	24	24	18	24
Middle (6-7)	Mean	67.3	0.75	39.6	40.9
	N	32	32	21	32
High (8-10)	Mean	54.5	0.55	41.5	53.1
	N	16	16	10	16
Total	Mean	67.7	0.72	32.7	37.2
	N	72	72	49	72
P value		P=0.001	P=0.006	P=0.03	P<0.0001
EDC administration (pooled baseline and crossover)					
Low (0-5)	Mean	78.5	0.70	21.2	21.7
	N	24	24	18	24
Middle (6-7)	Mean	67.5	0.70	37.2	38.8
	N	32	32	20	32
High (8-10)	Mean	56.3	0.58	40.5	51.9
	N	16	16	10	16
Total	Mean	68.7	0.67	31.9	36.0
	N	72	72	48	72
P value		P=0.002	P=0.29	P=0.10	P<0.0001

EDC, electronic data capture; IBS-QOL, Irritable Bowel Syndrome-Quality of Life; VAS, Viasual Analog Scale; WPAI:IBS, Work Productivity Activity Impairement questionaire – Irritable Bowel Syndrome version

Table 6 shows ease of use for the two modes of administration by initial mode of administration. Both versions were rated easy to read, regardless of which mode was administered first, with mean scores ranging from 87.9 to 91.8 out of a

possible high score of 100. Patients reported it was significantly easier to go back and change answers on the EDC version than in the paper version, regardless of whether paper or EDC was administered first (p values 0.004 and 0.001), but there were no significant differences in difficulty using the two administration modes.

Table 6. Ease of using the paper questionnaire and electronic data capture (EDC) by mode of first administration.

	Paper first mean (± standard deviation) [Range] n=35		EDC first mean (± standard deviation) [Range]n=37	
	Paper	EDC	Paper	EDC
How easy was the diary to read Not at all easy (0) – Extremely easy (100)	90.7 (8.1) [72 – 100]	87.9 (17.1) [26 – 100]	91.8 (11.8) [47 – 100]	91.2 (11.3) [47 – 100]
How easy was it to go back and change answers? Not at all easy (0) – Extremely easy (100)	77.9 (26.6) [a] [10 – 100]	92.3 (9.1) [64 – 100]	73.1 (35.8) [a] [0 – 100]	93.5 (14.8) [9 – 100]
How difficult was it to use this diary? Not at all difficult (0) – Extremely difficult (100)	10.1 (19.5) [0 – 82]	9.1 (18.2) [0 – 94]	13.5 (26.2) [0 – 100]	5.9 (10.4) [0 –47]

[a] P ≤ 0.004 by Student's paired t-test

Table 7 shows the preference for the two modes of administration. Overall, 47.2% of the patients thought the EDC version was easier to use; 23.6% thought the paper questionnaire was easier to use and 29.2% thought there was no difference between methods. If the patients were to participate in another study, half would prefer EDC, 13.9% would prefer paper questionnaires and 36.1% would have no preference. Missing data was negligible for both modes of administration.

Table 7. Preference for mode of administration by mode of first administration

	Paper first, n = 35	EDC first, n = 37	Total
Method easier to use			
Paper and pencil	5 (14.3%)	12 (32.4%)	17 (23.6%)
Computer	19 (54.3%)	15 (40.5%)	34 (47.2%)
No difference	11 (31.4%)	10 (27.0%)	21 (29.2%)
Preferred method in future study			
Paper and pencil	5 (14.3%)	5 (13.5%)	10 (13.9%)
Computer	18 (51.4%)	18 (48.6%)	36 (50.0%)
No difference	12 (34.3%)	14 (37.8%)	26 (36.1%)

Discussion

Electronic data capture is increasingly being used to collect patient-reported outcomes data in clinical studies. We tested three previously validated questionnaires (IBS-QOL, WPAI:IBS, EQ-5D) to determine if the EDC versions of these questionnaires were comparable to the paper versions and acceptable to subjects, and therefore suitable for use in IBS studies.

We found no significant differences between the scores obtained by paper questionnaire and EDC and no pattern of results emerged that would suggest that one mode of administration was better than the other in terms of its psychometric properties. Scores from both modes of administration had comparable correlations with SF-36 measures of physical and mental well-being. Scores obtained by both paper questionnaire and EDC demonstrated internal consistency, test-retest reliability, and validity, as measured by the relationship of scores to symptom severity. The one exception was the EQ-5D Visual Analog Scale that did not differentiate the low and middle symptom severity groups in either mode of administration. This is not surprising considering that this is a general health measure for a single day and the criterion for discriminant validity was IBS symptom severity for the past 7 days. Acceptability of both modes of administration was high, but patients reported the EDC mode was significantly easier for going back and changing a response, and more patients preferred EDC over the paper questionnaire.

Our subjects were selected to be representative of IBS patients in terms of sex and type of IBS, but our results are limited by the small sample size, particularly among the employed (n = 49) in the investigation of WPAI:IBS work productivity measures. Consequently, although we failed to demonstrate statistically significant differences between the two modes of administration, there may be differences that we were unable to detect. For example, we note that impairement as measured by the IBS-QOL and WPAI:IBS domain scores uniformly decreased from baseline to the 24-hour crossover assessment, regardless of mode of administration, whereas the EQ-5D scores were generally higher when obtained by paper questionaire, regardless of whether paper was

administrated first or second. Investigation of these differences was outside the scope of the planned anaysis and warrant additional investigation.

Another limitation of our findings is that they may not reflect the application of the two modes of administration in the typical clinical setting, in terms of missing information. Because subjects could not skip an entry in EDC and the site coordinators rigorously reviwed missing information from subjects, missing information was not found in either mode of administration. In other settings, paper questionnaires have had higher rates of missing information relative to EDC (Johannes 2000, Ryan 2002, Palermo 2004), so the advantage of EDC in this study may be understated.

Despite the study's limitations, the findings of comparability between the two modes of administration are consistent with the growing body of research in a variety of diseases, populations, and settings that has shown EDC to be comparable to paper questionaires.

Conclusion

Electronic data capture versions of the IBS-QOL, EQ-5D, and WPAI:IBS are generally comparable to paper questionnaires and demonstrate internal consistency, test-retest reliability, and subject acceptability; discriminant validity of the questionnaires by the two modes of administration is comparable.

Acknowledgement

This study was funded by Novartis Pharma AG, Switzerland.

Chapter 4

Decision Analytic Model of Management Strategies for Paget's Disease of Bone

Summary

Bisphosphonates (BP) are the treatment of choice in Paget's disease of bone (PDB). As more potent BP are becoming available, it is important to assess the costs and effects of comparative BP therapy in order determine the optimal way patients with PDB should be managed. The aim of the present chapter is to estimate the cost-effectiveness of a new BP, zoledronic acid 5mg (ZOL), administered as a single 15 minutes intravenous infusion compared to risedronate (RIS) administered orally 30mg/day for 2 months in the treatment of PDB. A Markov model was developed to estimate costs and effects of ZOL versus RIS over a period of two years. The model consists of four half-year cycles and four possible health states, i.e. response, non-response, relapse, no relapse. As measure of effectiveness, time in response was used, with response defined as normalized serum alkaline phosphatase level. Probabilities and time to maintaining response were derived from clinical trial data. Resource use and unit cost estimates for the United Kingdom (UK) were based on databases, published literature and expert panel, and include direct costs of treatment and follow-up. Costs and effects were discounted by 3.5%. The analysis was performed from a National Health System (NHS) perspective and societal perspective in the UK. An extensive probabilistic sensitivity analysis was performed to address uncertainty. This study showed that on average, patients receiving ZOL will maintain response for 5.9 months longer than patients receiving RIS. Additionally, from the societal perspective over two years, the cost savings of ZOL versus RIS was £182 and from the NHS perspective, the cost savings amounted to £195. The key driver of cost-effectiveness in the model was the number of re-treatments required over two years. The average number of treatments per year was estimated at 0.59 in the ZOL group and 0.93 in the RIS group. When taking uncertainty into account, all simulated outcomes indicated that ZOL is both cost-saving and more effective. In conclusion, when compared to the standard therapy RIS, ZOL is the dominant strategy in the management of PDB due to superior effectiveness and lower cost.

Introduction

Paget's disease of bone (PDB) known also as osteitis deformans is the second most common metabolic bone disease in the UK after osteoporosis, affecting up to 5% of those aged 55 years and over. Epidemiological studies undertaken over the last 30 years have suggested high prevalence rates of PDB in Britain, with somewhat lower rates in Australia, North America and parts of Western Europe (van Staa 2002).

PDB is a chronic skeletal disorder characterized by localized areas of increased bone remodeling, bone hypertrophy and abnormal bone structure (Paget 1877). PDB can affect one or more bones within the skeleton. Prevalent signs and symptoms of PDB are skeletal deformity and bone pain. Complications of PDB can involve bone (deformity, pathological fracture, neoplastic degeneration), joints (secondary osteoarthritis), the nervous system (deafness, spinal cord compression), and the vascular system (steal syndrome) (Kanis 1998).

There has been a dramatic change in the therapeutic approach to patients with PDB over the last 40 years. In the 1960s, only symptomatic therapy could be given, with control of pain the main objective. Analgesics and nonsteroidal anti-inflammatory drugs were the most commonly used agents. From 1968 onwards, antiosteoclastic agents became available, including calcitonin, plicamycin (mithramycin) and etidronate, a first-generation bisphosphonate (BP) (Devogelaer 2002). Limitations of these agents, especially potentially deleterious effects on bone mineralization with etidronate, paved the way for increasingly potent second- and third-generation of nitrogen containing bisphosphonates, including clodronate (clodronic acid), pamidronate, alendronate, RIS and recently ZA. With the newer BPs, long-term remission of pagetic lesions, as well as prevention of long-term complications in both symptomatic and asymptomatic patients may be expected (Devogelaer 2002).

Biochemical markers of bone turnover have been used for many years in the diagnosis and monitoring of treatment for PDB (Delmas 1997). All clinical studies of BP in PDB have been short term and have used total serum alkaline phosphatase (total ALP), the key biochemical marker of bone turnover as the primary endpoint. It is believed that suppression of ALP can reduce the risk and prevent progression of complications of the PDB.

As more potent BPs are becoming available, it is important to assess the long-term costs and effects of comparative BP therapy in order determine the optimal treatment for patients with PDB. The increased potency and longer duration of action of newer BP could compensate for their marginally increased cost compared with older BP. This study describes an evaluation of the cost-effectiveness model of two years' treatment with either ZOL or RIS for PDB patients from societal and NHS perspective in the UK.

Methods

A cost-effectiveness analysis (CEA) was performed to compare the costs and effects of ZOL treatment, a single 5 mg infusion over 15 minutes, and RIS treatment, 30 mg/day orally given for two months, in patients with PDB over two

years. The CEA was based on the pooled analysis of the two identical registration clinical trials of ZOL versus RIS in PDB patients, published literature and expert panel (Reid 2005, Hosking 2007). Given the small number of observations beyond two years follow-up, a time horizon of two years was chosen.

The analysis was performed from two perspectives, the societal and NHS perspective in UK. The NHS perspective is usually required by NICE in their technology appraisals (NICE 2004). Additionally, we included the societal perspective, limited to direct costs. Indirect costs due to production losses were not included since the prevalence of PDB increases with age, (Eekhoff 2004) and occur mostly in people over 55 years old. Patients in the two clinical trials had a mean age of 70.8 years in the ZOL group and 70.0 years in the RIS group (see Table 1) (Reid 2005). A discount rate of 3.5% was used for costs and effects in the second year, according to NICE guidelines (NICE 2004).

The clinical studies have been described in detail elsewhere and we summarize the main characteristics relevant to the present study here (Reid 2005). The main patients' characteristics are provided in Table 1.

Table1. Patient characteristics

	ZOL	RIS
Number of patients	176	171
Sex – Male (%)	68.1	67.4
Age (years) Mean (SD)	70.8 (9.82)	70.0 (10.73)
Age – ≥65 years (%)	74.7	73.7
Baseline ALP (U/L) Mean (SD)	427.8 (320.99)	425.3 (312.05)

The primary efficacy variable of the clinical studies was the proportion of patients who achieved therapeutic response. Therapeutic response was defined as a reduction of at least 75% from baseline in total serum alkaline phosphatase (total ALP) excess (difference between measured level and midpoint of the normal range) or normalization of ALP at the end of six months. A secondary efficacy variable was normalization of ALP, i.e. return of ALP to within the normal range. The normal range of ALP for male and female between the age of 20 and 58 years is 31-110 U/L whereas for male and female above 59 years of age the normal range is 35-115 U/L. Note that the ALP-normalized patients form a sub-sample of the patients with a therapeutic response.

Both outcome measures have their own rates for response and relapse, and time-to-response. The results of the pooled studies showed statistically significant superiority of ZOL for the proportion of patients who were therapeutic responders at 6 months (96%) compared to RIS (74%, $p<0.001$). The proportion of patients with ALP normalization was statistically superior with 5 mg ZOL (89%) to that with RIS (58%, $p<0.001$). Additionally, the time to therapeutic response differed significantly, 62.8 days in the ZOL group and 110 days in the RIS group ($p<0.001$). The time to normalization was 92 days in the ZOL group and 136 days in the RIS group [$p < 0.001$] (Reid 2005).

Patients who met the definition of therapeutic response at the end of the core study (at six months), were asked to participate in an extended observation

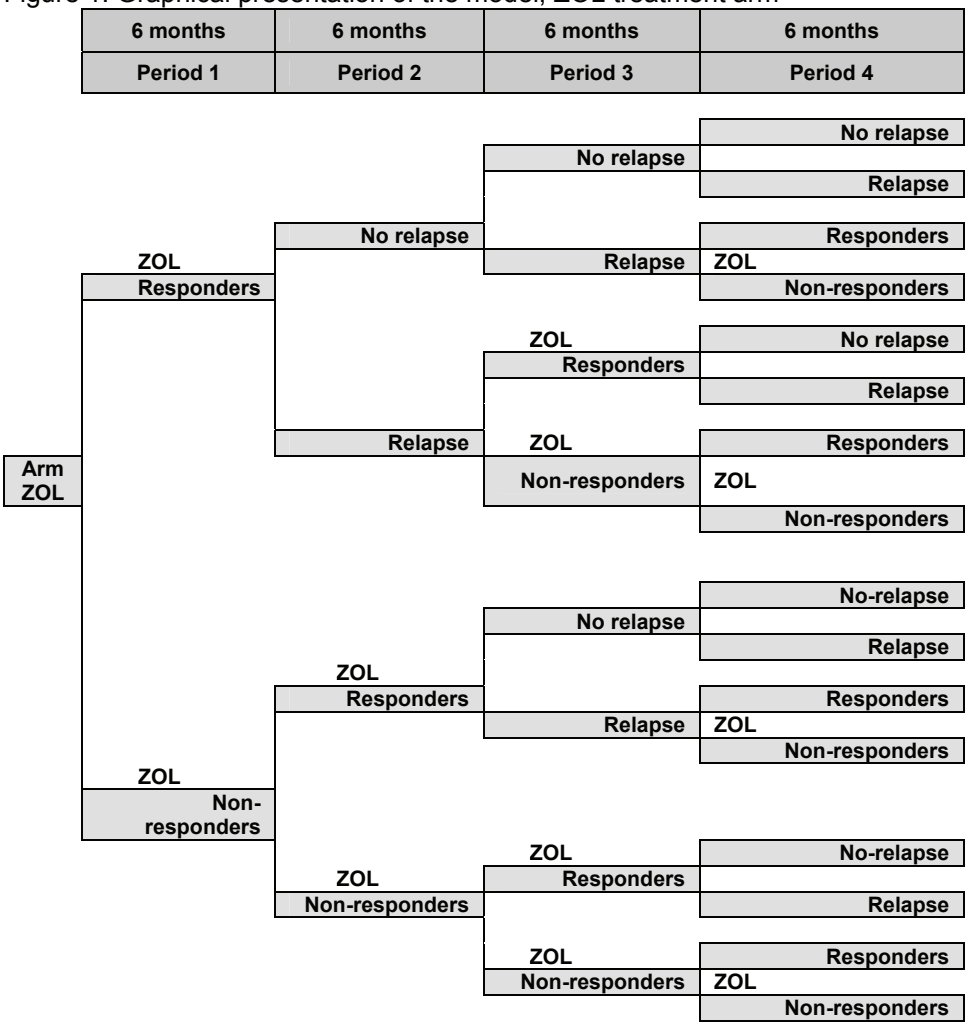
study (Hosking 2007). In this extended observation period patients were followed-up every six months to measure ALP until levels returned to within 20% of baseline or up to the cutoff date of 30-Sep-2004 whichever came first. A partial disease relapse was defined in the clinical protocol as an increase in ALP of at least 50% from the ALP measurement at month 6 and at least 1.25 times the upper normal limit. With up to 18 months follow-up after the end of the core study (24 months overall), 99% of patients in the ZOL group and 70% of patients in the RIS group had had no partial disease relapse. For patients who were normalized at 6 months, these percentages were 100% and 76%, respectively. Table 2 presents the results of the extension trial for patients who were normalized at 6 months.

Table 2. Relapse rate estimates

Follow-up interval (days)	No of events	No of patients at risk (adjusted for censoring)	Probability relapse during 6-month interval
ZOL			
<182	0	141.5	0%
182 – 364	0	132	0%
364 - 546	0	117.5	0%
RIS			
<182	2	89.5	2%
182 – 364	9	84.5	11%
364 - 546	9	68.5	13%

A Markov model was developed to estimate costs and effects of ZOL versus RIS over a period of two years (see Figure 1). The model consists of four half-year cycles and four possible health states: ‘responder’, ‘non-responder’, ‘relapse’ and ‘no relapse’. At the end of every six-month period, ALP measurement is performed and used to classify patients in to one of the four health states. This means that the physician does not know whether a patient is a responder or had a relapse until the end of the six-month period. Response in the CEA was defined as normalization of ALP. The cycle length of the model was chosen to coincide with the frequency of follow-up visits in daily practice.

Figure 1. Graphical presentation of the model, ZOL treatment arm



At the end of period 1, patients in 'non-responder' health status receive another course of treatment with ZOL or RIS, depending on treatment group, whereas patients in 'responder' health status continue to stay in response. After another six months (period 2), a further ALP measurement is performed and patients in 'non-responder' health status are classified again as either a 'responder' or a 'non-responder' to the second course of BP therapy. Based on ALP level at the end of period 2, patients who are in 'responder` status are either classified as 'no-relapse' or 'relapse'. After period 3 and 4, the same algorithm applies; thus, a maximum of four treatment courses can be given.

It was assumed that patients who were in 'non-response' or 'relapse' health status start another course of the same treatment (i.e. no cross-over

occur). These re-treatments did not occur in the clinical studies, as this was not part of the protocol. However, in daily practice, patients whose ALP levels go above the upper limit of the normal range will be re-treated, as will patients who do not respond to the first treatment.

Outcome was expressed as time maintaining response over the 2-year period, with response defined as ALP normalization. To determine time maintaining response we used data from the clinical studies on time to therapeutic response, so in the first period after treatment, time maintaining response is 6 months minus time to response. In a period with no response the time maintaining response is 0, in a period with no relapse this is 6 months, and in the period where relapse occurs, we assumed this takes place (on average) halfway through the cycle, giving a time in response of 3 months.

All estimates of probabilities that were used in the model are presented in Table 3. In the ZOL group, no relapses occurred. Since a true relapse rate of zero seems unlikely, we have assumed a small non-zero percentage, i.e. 1% per 6 months.

Table 3. Model input probabilities plus limits 95% confidence interval

	Description	Point Estimate	95% confidence interval	Distribution (parameters)	Source
Efficacy ZOL	Responders 1st ZOL	89 %	84% - 93%	Beta (156.6, 19.4)	Reid 2005
	Responders subsequent ZOL	71%	57% - 82%	Beta (35.3, 14.7)	Reid 2005
	Relapse period 2	1%	0% - 3%	Beta (1.4, 140.1)	Hosking 2007†
	Relapse period 3	1%	0% - 4%	Beta (1.3, 130.7)	Hosking 2007†
	Relapse period 4	1%	0% - 6%	Beta (1.2, 116.3)	Hosking 2007†
Efficacy RIS	Responders 1st RIS	58%	51% - 65%	Beta (99.2, 71.8)	Reid 2005
	Responders subsequent RIS	46%	35% - 57%	Beta (34.5, 40.5)	Reid 2005
	Relapse period 2	2%	1% - 8%	Beta (1.8, 87.7)	Hosking 2007†
	Relapse period 3	11%	7% - 25%	Beta (9.3, 75.2)	Hosking 2007†
	Relapse period 4	13%	5% - 32%	Beta (8.9, 59.6)	Hosking 2007†
Cost ZOL (£)	Cost responders*	75.52	60.42 - 90.63	Normal (75.52, 7.7)	See tables 4 and 5
	Cost non-responders *	77.52	62.02 - 93.03	Normal (77.52, 7.9)	
	Cost relapse	52.00	41.60 - 62.40	Normal (52.0, 5.3)	
	Cost no relapse	52.00	41.60 - 92.40	Normal (52.0, 5.3)	
Cost RIS (£)	Cost responders **	61.2	48.96 - 73.44	Normal (61.20, 6.2)	See tables 4 and 5
	Cost non-responders **	63.2	50.56 - 75.84	Normal (63.20, 6.4)	
	Cost relapse	46	36.80 - 55.20	Normal (46.0, 4.7)	
	Cost no relapse	46	36.80 - 55.20	Normal (46.0, 4.7)	
Time to response in months	Time to response ZOL (months)	3.02	2.76 - 3.28	Normal (3.02, 0.13)	Reid 2005
	Time to response RIS (months)	4.46	4.15 - 4.76	Normal (4.46, 0.16)	Reid 2005

*Excluding the cost of the drug, ZOL, **Excluding the cost of the drug, RIS, † In the clinical study, the percentage relapse was actually 0%. For the parameters of the beta distribution we used the sample sizes from the study, combined with our assumption of 1% relapse per period.

The clinical study only provided estimates for the rate of response to the first treatment course. The probability of response after an initial non-response was derived from the sub-sample of patients in the trials who had previously used BP since it has been suggested that patients treated for PDB may gradually become resistant to BP therapy (Selby 2002, Johnston 1980). In the trials, the RIS group that had previously used BP had a response rate of 46%, 21% lower than for the whole group. This response rate was assumed to be a good proxy for the probability of response in the period after non-response. The model assumed that the same reduction of 21% would apply to the ZOL group, i.e. a response rate of 71% was used for patients who did not respond to the first cycle of ZOL.

Complications of PDB (e.g. deformity, pathological fracture, deafness, etc) are not included in the model as they were not captured in the clinical trials and precise data could not be taken from literature.

Costs

No cost data was collected in the clinical studies. Instead, resource use has here been estimated based on literature and clinical experts in the UK. The following assumptions were made in the model:

1. PDB in the UK is treated both by specialists and general practitioners (GPs). Given the specificity of the two types of drug administration, we assumed that half of the patients taking oral medication visit a GP and half visit a specialist, whereas all patients that receive an infusion visit a specialist.
2. Oral BPs are poorly absorbed from the gastrointestinal tract causing upper gastrointestinal side effects, such as heartburn and dyspepsia (Kelly 1997). In the model, it is assumed that 15% of patients taking oral BP will manifest upper gastrointestinal side effects (Siris 2003) and are therefore treated with a proton pump inhibitor (PPI), i.e. omeprazole generic, for 2 months.
3. It has been suggested that intravenously administered BP could induce an acute febrile reaction, so called ‘flu like syndrome’ (FLS) (Langston 2004). In the model, it is assumed (based on the trial data) that 11% of the patients receiving intravenously BP manifested an acute febrile reaction (CZOL 2004) and are therefore treated with analgesics, paracetamol, for 3 days.
4. Non-responders to BP treatment are considered more severe cases and are therefore treated with analgesics (ibuprofen) for 1 month on average to relieve the osteoarthritis pain caused by misshapen bones.
5. Concomitant intake of calcium and vitamin D is recommended during BP treatment. In concordance with product labeling, (EMEA 2007) the model assumed 10 days of calcium and vitamin D intake for ZOL and 2 months for RIS.

Resource use was defined for all four health states: ‘response’, ‘non-response’, ‘relapse’, ‘no relapse’, over a period of 6 months. The costs of extra medication, e.g. omeprazole, paracetamol, ibuprofen, calcium and vitamin D are those of generic products in the UK (EMIMS 2005). A summary of resource use for each health state is presented in Table 4.

Table 4. Cost Allocation

Responder		Non-responder		Relapse/No-relapse	
Cost component	Cost (£)	Cost component	Cost (£)	Cost component	Cost (£)
ZOL		ZOL		ZOL	
Drug ZOL	305.6	Drug ZOL	305.6	ALP	7
Nurse time for infusion (30 min.)	17	Nurse time for infusion (30 min.)	17	Visit specialist	29
Medical equipment for infusion	5	Medical equipment for infusion	5	Patient travel expenses to hospital	18†
Calcium Vitamin D (10 days)	1.5	Calcium Vitamin D (10 days)	1.5		
Analgesics FLS (10 days)	0.1	Analgesics FLS (10 days)	0.1		
ALP measurement	7	ALP measurement	7		
Visit specialists	29	Visit specialists	29		
Patient travel expenses to hospital	18†	Patient travel expenses to hospital	18†		
		Analgesics	2		
Total	383.1	Total	385.1	Total	54
RIS		RIS		RIS	
Drug RIS	305.6	Drug RIS	305.6	ALP	7.1
Prescription	5†	Prescription	5†	Visit GP (50%)	14.5
Calcium Vitamin D (2 months)	9	Calcium Vitamin D (2 months)	9	Visit specialist (50%)	14
ALP measurement	7	ALP measurement	7	Patient travel expenses to GP (50%)	3.5†
PPI, omeprazole (1 month)	1.2	PPI, omeprazole (1 month)	1.2		
Visit GP (50%)	14.5	Visit GP (50%)	14.5	Patient travel expenses to hospital (50%)	9†
Visit specialists (50%)	14	Visit specialists (50%)	14		
Patient travel expenses to GP (50%)	3.5†	Patient travel expenses to GP (50%)	3.5†		
Patient travel expenses to hospital (50%)	9†	Patient travel expenses to hospital (50%)	9†		
		Analgesics	2		
Total	368.8	Total	370.8	Total	48

† These items are excluded in the NHS perspective

Estimates of UK unit costs for each health care service were derived from nationally representative data sources (Netten 2005). Costs were calculated to cover medical procedures performed in a two years time frame for patients with PDB and include direct costs of treatment and follow-up. The analysis was performed from societal and NHS perspectives in the UK. When the societal perspective is applied the cost of prescription and of patient travel expenses to hospital and GPs are added to the analysis. A summary of unit costs is provided in Table 5. In Table 4, the resource use is combined with the unit costs to arrive at a cost estimate for each health state.

Table 5. Unit Costs

Cost breakdown	Unit	Cost (£)
Drug Acquisition		
Zoledronic acid	Per cycle	305.62
Risedronate (EMIMS 2005)	Per cycle	305.62
Infusion		
Nurse time (Netten 2005)	Per hour	34
Medical equipment for infusion* (NHS 2004)	Per administration	5
General Practitioner visit (Netten 2005)	Per visit	28
Specialist visit (Netten 2005)	Per visit	29
Analgesics (ibuprofen, paracetamol generic) (EMIMS 2005)	Per month	2
Proton Pump Inhibitor, omeprazole generic (EMIMS 2005)	Per month	8
ALP measurement (NHS 2004)	Per unit	7
Calcium Vitamin D generic (EMIMS 2005)	Per day	0.15
Patient travel expenses to GP (Netten 2005)	Per visist	7
Patient travel expenses to hospital*	Per vistit	18
Prescription oral medication (Netten 2005)	Per prescription	5

* Personal communication and data on file Manchester Royal Infirmary

Sensitivity analysis

To investigate the impact of assumptions made during the analysis and to test the robustness of results given variation in the data input, a probabilistic sensitivity analysis was performed (Briggs 2000). Probabilistic sensitivity analysis involves specifying distributions for input variables in the model allowing the joint effect of input uncertainty to be assessed. Within our model there are two main categories

of variables: model probabilities related to the four health states which have a beta distribution and variables relating to the level of resources consumption with their associated unit cost which have a normal distribution. The parameters of all beta distributions were derived from the point estimate of the probability and the sample size. The standard deviations of the normal distributions (for the cost estimates per health state) were based on the assumption that limits of the 95% confidence interval are the mean ± 20%. This margin for costs reflects the uncertainty about the resource use in each health state. Having specified distributions for the relevant variables, the probabilistic sensitivity analysis was undertaken by randomly sampling from each of the distributions and calculating the expected costs and expected time in response for that combination of input values. This process formed a single replication of the model results, and a total of 2500 replications were performed in order to examine the distribution of the resulting costs and effects. A summary of the point estimates of variables, and their 95% confidence interval used in the probabilistic sensitivity analysis are presented in Table 3.

Results

Table 6 presents the results. From a societal perspective, it was estimated that over two years the costs in the ZOL group were lower (£590) than in the RIS group (£773) leading to savings of £182, after discounting. Also, in the ZOL group the time maintaining response was 5.9 months higher (ZOL 19.5 months, RIS 13.6 months). Thus, ZOL is the dominant treatment. The average (discounted) cost per year was estimated at £295 for ZOL and £386 for RIS, while the average number of treatments per year was estimated at 0.59 in the ZOL group and 0.93 in the RIS group.

When the NHS perspective was taken into consideration, patient travel expenses and prescription costs were excluded. This led to a slightly higher cost savings of £195, after discounting. The average cost per year was now estimated at £260 for ZOL and £357 for RIS.

In both perspectives, costs for ZOL are about 3% higher than RIS in the first 6 months, while they are about 45% lower in the subsequent three periods.

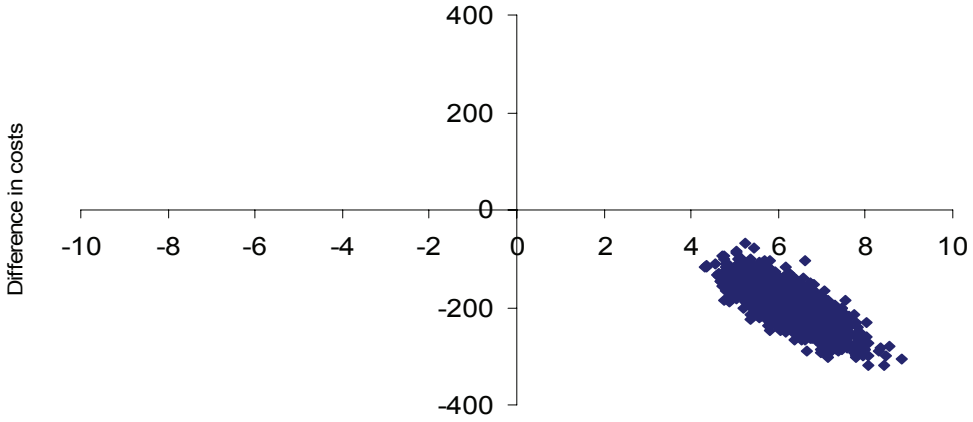
Table 6. Model Outcomes

Time	Costs (£) Societal perspective		Costs (£) NHS perspective		Effects (ALP normalized months)	
	ZOL	RIS	ZOL	RIS	ZOL	RIS
Period 1	381	368	363	350	2.7	0.9
Period 2	88	181	70	167	5.5	3.7
Period 3	66	123	48	109	5.8	4.9
Period 4	59	109	41	95	5.9	4.7
Total (2 years)	594	780	522	721	19.9	13.9
Total discounted	590	773	519	714	19.5	13.6
	Difference £183		Difference £195		Difference 5.9	

Sensitivity analysis

The results of the 2500 replications of the probabilistic sensitivity analysis are presented on the cost-effectiveness plane in Figure 2. The plane is divided into four quadrants indicating four possible situations in relation to the additional costs and additional health outcome effects of ZOL compared to RIS. The graph shows that all the simulated values are in the South-East quadrant, indicating that ZOL is the dominant strategy with both increased time in response and a reduction in costs.

Figure 2. Joint distribution of cost and effects on the cost-effectiveness plane



Discussion

The current study shows that treatment of PDB with ZOL instead of RIS is more effective and cost-saving, even when taking uncertainties into account. These savings are estimated at about £190 per patient over a period of two years, which is a decrease of 24%. The measure of effectiveness used in this study was maintaining ALP normalization. A possible final outcome would be complications of PDB. However, unlike measures such as fracture rate in osteoporosis, the symptoms and complications of PDB are either rare (fracture) or difficult to measure/quantify in randomized controlled trials (warmth, nerve entrapment, osteoarthritis in weight-bearing joints) (Drake 2001). Currently, the aim of PDB therapy is to achieve normalization of biochemical markers of bone turnover, in particular serum alkaline phosphatase, in order to obtain full remission. This objective aims at controlling the activity of the disease and preventing complications (Delmas 1997, Selby 2002, Drake 2000, Jacobs 1999, Rousiere 2003). Therefore, ALP was taken as an adequate and relevant intermediate outcome in this study. However, the complications of PDB may be expected to have a high economic impact and thus further research in PDB is needed, to

explore the long term-effects of BP therapy on complications of PDB. Recently, a large multicentre trial (PRISM) has been set up in the UK to address the effects of the anti-Pagetic treatment on symptoms, complications, as well as ALP level in PDB (PRISM 2005). Until such information is available, using an intermediate measure of effectiveness is necessary.

In the clinical studies, no economic data was gathered. Furthermore, little is known about resource use of patients with PDB. In this study, we relied on literature and expert panels in order to estimate resource use. This is clearly a source of uncertainty, but we expect this uncertainty to be limited, as the management of PDB patients is rather straightforward. By using margins of plus/minus 20% in the sensitivity analysis, we addressed the uncertainty, and it did not have an impact on the outcome of the study.

This study has a limited time frame, which is based on the availability of data. The number of patients observed in the extension trial between after 18 months is rather small, as the follow-up duration was intended to be 18 months. However, based on the small group of patients observed between 18 and 30 months (at 18 months 111 patients at risk in the ZOL group and 55 in the RIS group, at 30 months 52 and 23 at risk, respectively) it seems reasonable to assume that the difference between the relapse rates observed in the first 2 years will continue to increase over time, leading to increased cost savings and time in response associated with ZOL. Thus, the current time frame may be assumed to lead to a conservative estimate of the cost-effectiveness of ZOL.

One major issue with the use of oral BP that was not addressed in this study is compliance with the dosing regimen. Although there are no studies specifically designed to study compliance, several clinical studies have reported data on it (Conte 2004). These show that non-compliance is a serious problem in long-term treatment with oral BP for osteoporosis and bone metastasis from advanced cancer (Conte 2004). It can be assumed that the compliance with oral therapy outside the controlled conditions of a clinical trial is even worse. As ZOL is given once by infusion, this problem occurs only in the RIS group, where patients are required to take the drug orally daily for 2 months. Non-compliance can result in a decreased response in the RIS group. Disregarding this issue means that a conservative estimate of cost-effectiveness has been made.

In conclusion, we found that ZOL is a more cost-effective treatment in management of PDB with superior effectiveness and lower cost, when compared with the oral standard therapy RIS. Thus, ZOL will most likely be the preferred treatment in the long-term management strategies of PDB, both from an economical and a patient convenience point of view.

Acknowledgement

This study was funded by Novartis Pharma AG, Switzerland.

Chapter 5

Probabilistic Cost-Effectiveness Model of Lifestyle Intervention in the Prevention and Treatment of Obesity

Summary

The use of decision-analytic modeling in health technology assessments has increased exponentially in recent years. Mathematical modeling is used widely in economic evaluations of medical interventions. Models represent an important analytic framework to generate estimates of cost-effectiveness based on a synthesis of available data and explicit representation of uncertainty. The aim of the present chapter was to develop a decision analytic model that quantifies the lifetime health and economic consequences of preventing and treating obesity with lifestyle intervention in Switzerland. A Markov model was developed comparing lifestyle intervention and standard care in overweight and obese people. Changes in weight and cardiovascular risk factors over time are modeled from reduction in body mass index (BMI), systolic blood pressure, total cholesterol and high density lipoprotein cholesterol in three-year active treatment period based on data from meta-analysis. A probabilistic sensitivity analysis is performed. Three groups of people are followed in the analysis: overweight subjects (BMI 27 kg/m^2), borderline subjects (BMI 30 kg/m^2) and moderate obese subjects (BMI 33 kg/m^2). The cost-effectiveness of interventions is compared using incremental costs, incremental effects: life-years, quality adjusted life-years, and the cost-effectiveness ratio. Lifestyle intervention results in increased survival duration and quality of life over lifetime. Compared with standard care, the average incremental cost of lifestyle intervention is lower in borderline subjects and moderate obese subjects and higher in overweight subjects. Lifestyle intervention dominates standard care being less costly and more effective in borderline female subjects' aged 35 to 55 years, borderline male subjects' aged 25 to 60 years, moderate obese female subjects' aged 45 years and moderate obese male subjects' aged 55 years. In conclusion, our economic analysis suggests that lifestyle intervention is cost-effective in the long-term prevention and treatment of obesity.

Introduction

Obesity has become one of the most common medical problems due to an explosive increase in the number of people affected by the disease. Obesity is associated with numerous co-morbidities such as cardiovascular disease, diabetes, and hypertension, certain cancers and sleep apnoea (Poirier 2006). Several different treatments are available for the management of obesity: non-pharmacological interventions (e.g. dietary interventions, physical exercise, lifestyle interventions, etc.), pharmacological interventions (e.g. sibutramine, orlistat, etc.), alternative medicine (e.g. acupuncture, hypnotherapy, etc.) and in the case of morbid obesity, surgical treatments (e.g. gastric bypass, gastroplasty, etc). In the latest years a specific interest is focused on prevention strategies that aim to stop the progression of overweight to obesity and the occurrence of obesity complications (World Health Organization 2000).

The economic burden of obesity is substantial. In Switzerland, obesity related expenditures are estimated to a total CHF 2691 million, or almost 2.3-3.5% of total health care spending (Schmid 2005). The top five co-morbidities responsible for the increased costs associated with overweight and obesity are hypertension, hypercholesterolemia, diabetes mellitus, stroke and coronary heart disease (Neilson 2005). Thus, the economic costs of obesity represent a sizeable issue for the Swiss healthcare system. The increasing burden on the budgets of health care providers has resulted in considerable interest in assessing new and existing treatments for their clinical effectiveness and cost-effectiveness. Such an assessment enables decisions to be made on the allocation of limited healthcare resources while ensuring that any additional cost is justified by the additional benefits (Neumann 2005). Among several treatments available, lifestyle interventions have been documented to lead safely to improvements in metabolic abnormalities such as increased body weight, dyslipidemia, elevated blood pressure and glucose control that are linked to the development of obesity, diabetes, and cardiovascular disease (Pritchett 2005). Lifestyle programs are multi-factorial interventions that are designed for each patient or group of patients according to their risk factor status and include promotion of healthy lifestyle habits, dietary counseling, physical exercise training and behavioral change targets.

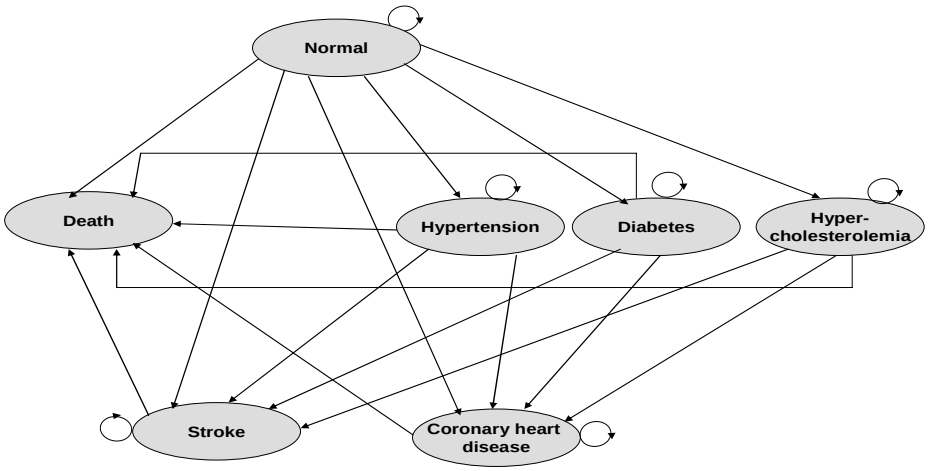
The aim of the present study is to develop a decision analytic model to quantify the lifetime health and economic consequences of preventing and treating obesity with lifestyle intervention in Switzerland.

Methods

A Markov decision model (Sonnenberg 1993) was developed to evaluate the lifetime effect of a three-year lifestyle intervention compared with standard care in overweight and obese people. Seven states are modeled: 'normal', if patients are overweight with body mass index (BMI) 25 to 29.9 kg/m^2 or obese with BMI $\geq$ 30 kg/m^2 and free of complications, 'hypertension' if patients are overweight or obese and developed hypertension (systolic blood pressure greater than 140 mmHg and diastolic blood pressure greater than 90 mmHg, JNC 1997), 'hypercholesterolemia' if patients are overweight or obese and developed

hypercholesterolemia (total serum cholesterol ≥ 6.2 mmol/l, EP 2001), 'diabetes' if patients are overweight or obese and developed type 2 diabetes (fasting glucose of at least 7.8 mmol/l or 2-hour glucose of at least 11.1 mmol/l, Alberti 1998), 'stroke', if patients are overweight or obese and developed stroke, 'coronary heart disease' if patients are overweight or obese and developed coronary heart disease and 'death': patients can die and enter this point at any time in the model. A diagrammatic representation of the Markov model is presented in Fig.1.

Figure 1. Markov model diagram



Following the lifestyle intervention patients enter into the model in the normal health state. The cycle length is one year. At the end of each one-year period, proportions of the cohort can move from one disease state to another or stay in the same disease state. The transition probabilities are based on the disease progression with age, sex, BMI and cycle number. The model is run over a period of 60 years to estimate the lifetime costs and effects of the intervention. Patients entering the model have a minimum age of 25 years. The model runs until subjects reach the age of 85 years considering that the average life expectancy in Switzerland is 77.3 years for men and 82.8 years for women (World Health Organization 2005).

Because of the complexity of the interrelationship between the degree of overweight/obesity and the five obesity-related complications, we made several assumptions. It is assumed that those diseases for which obesity is a risk factor are not interconnected and therefore are not counted as an additive effect on the lifetime health and costs in the model. For simplicity, the possibility of having concomitant diabetes, hypertension and hypercholesterolemia is not incorporated into the model. One reason for excluding the correlation between the existing co-morbidities is the absence of prevalence data for Swiss population. This

assumption is most probably underestimating the burden associated with obesity co-morbidities.

A hypothetical cohort of 10000 subjects overweight or obese receives a lifestyle program or standard care intervention for a period of three years. The components of lifestyle intervention are regular physical activity and healthy eating, including diets rich in fruits and vegetables. In our model, lifestyle intervention is adapted from Finish Diabetes Prevention Study (Lindström 2005) and consist in detailed dietary recommendations: to limit the total intake of fat to less than 30% of energy consumed and of saturated fat to less than 10%, and to increase fibre to at least 15g/1000 kcal as well as advice about specific food types, and asked to undertake moderate exercise for at least 30 minutes per day. Lifestyle intervention group members attended dietitian sessions and supervised exercise sessions during the first three years. For overweight people, standard care consists in no intervention whereas for obese people standard care consists in basic dietary counseling and physical exercise sessions (see cost description section). Treatment effect is modeled as a reduction in BMI, systolic blood pressure, total cholesterol, and high density lipoprotein cholesterol based on data obtained in three-year active treatment. It is assumed that the effect of lifestyle intervention on cardiovascular risk factors and weight loss is maintained up to six years, thereafter subjects start to regain weight linear for a period of four years i.e. after ten years weight loss went back to the initial weight. This assumption is sustained by the extended follow-up of the Finnish Diabetes Prevention Study (Lindstrom 2006) which lasted 7 years and resulted in sustained lifestyle changes that were maintained after the individual lifestyle counselling stopped. The study reported a 43% reduction in the relative risk related to the success in achieving the intervention goals of weight loss, reduced intake of total and saturated fat, increased intake of dietary fibre and increased physical activity. In the cost-effectiveness model, we assume that all patients developing hypertension, diabetes and hypercholesterolemia are diagnosed and treated. We also assumed that patients remain in those states once they have entered, unless they develop cardiovascular disease or die.

Parameter estimates and data sources

We estimate our model using data from a variety of secondary sources, which we will describe in detail. The correlation between BMI and annual risk of developing hypertension, diabetes, and hypercholesterolemia is calculated based on two large American prospective, epidemiological studies: the Nurses Health Study and the Health Professional Follow-up Study, both summarized by Field et al. 2001 (Table 1). Intermediate values of BMI have been interpolated using the polynomial function.

Table 1. Estimated yearly risk of obesity complications

BMI	Normal to Diabetes (%)		Normal to Hypertension (%)		Normal to Hyper-cholesterolemia (%)	
	Male	Female	Male	Female	Male	Female
22	1.16	0.96	1.65	1.75	5.13	6.27
23	1.46	1.28	1.84	1.92	5.32	6.42
24	1.82	1.63	2.03	2.08	5.47	6.52
25	2.24	1.99	2.23	2.23	5.58	6.55
26	2.69	2.37	2.44	2.36	5.66	6.51
27	3.17	2.75	2.63	2.48	5.70	6.42
28	3.67	3.14	2.83	2.59	5.72	6.29
29	4.19	3.52	3.02	2.70	5.72	6.13
30	4.71	3.90	3.19	2.79	5.71	5.97
31	5.22	4.27	3.35	2.87	5.69	5.84
32	5.71	4.62	3.50	2.94	5.66	5.77
33	6.18	4.95	3.62	3.00	5.64	5.80
34	6.61	5.25	3.73	3.06	5.63	5.97
35	7.00	5.52	3.80	3.11	5.63	6.34
36	7.33	5.76	3.85	3.14	5.65	6.95
37	7.61	5.95	3.87	3.18	5.69	7.87
38	7.80	6.10	3.86	3.20	5.76	9.17
39	7.92	6.20	3.80	3.22	5.87	9.17
40	7.95	6.24	3.71	3.23	6.01	9.17

The risk of complications has been adjusted according to age, sex and prevalence of hypertension, diabetes and hypercholesterolemia (Table 2) based on the information provided by the Swiss health survey (SFSO 2006).

Table 2. Yearly probability of developing obesity co-morbidities in Switzerland by age and sex

Age group	Hypertension (%)		Diabetes (%)		Hypercholesterolemia (%)	
	Male	Female	Male	Female	Male	Female
25-34	1.6	0.6	0.7	0.1	0.8	0.5
35-44	3.6	2.2	0.9	0.8	4.4	1.5
45-54	7.7	7.4	2.1	1.9	8.1	5.2
55-64	18.5	18	6.9	3.1	14.4	11.7
65-74	27	26.5	8.1	5.8	16.6	15.3
75+	25	33.2	11.2	8.5	15.3	15.0

The mean BMI by age and gender is taken from World Health Organization country profile (2005) and Monica Project (Mähönen 2000) and presented in Table 3.

Table 3. Mean BMI of the Swiss population by age and gender

Age group	Body Mass Index	
	Male	Female
25-34	24.6	24
35-44	26.2	24.5
45-54	27.7	25.6
55-64	27.8	27.1
65-74	28	27
75+	28	27

The risk of developing coronary heart disease and stroke from 'normal', 'hypertension', 'diabetes', 'hypercholesterolemia' states are based on a risk equation from the Framingham cohort study (Anderson 1990). Risk factors are age, sex, systolic blood pressure, total cholesterol, high density lipoprotein cholesterol, presence of diabetes and smoking status. Data on systolic blood pressure, total cholesterol and high density lipoprotein of the Swiss population are obtained from Monica Project (Mähönen 2000). The mean systolic blood pressure increases with age in both men and women, rising from 127 mmHg in men aged 25-35 years to 145 mmHg in men aged 75 years and over, and from 115 to 144 mmHg in women. The mean blood cholesterol levels increases with age with a slight decrease in the oldest age group. The Framingam equation data input is presented in Table 4. The Framingam equations applied in the model for stroke and coronary heart disease are presented in the Appendix.

Table 4. Framingham equation input

Variable	Unit	Normal	Hypertension	Diabetes	Hyper-cholesterolemia
Age	Years	Range (25-85)	Range (25-85)	Range (25-85)	Range (25-85)
Sex	Male / Female	M / F	M / F	M / F	M / F
SBP	mmHg	Range (115-147)	Range (169-183)	Range (115-147)	Range (115-147)
TC	mmol/l	Range (5.3-6.4)	Range (5.3-6.4)	Range (5.3-6.4)	Range (6.5-8.5)
HDL	mmol/l	Range (1.1-1.4)	Range (1.1-1.4)	Range (1.1-1.4)	Range (0.75-0.9)
Smoking	Yes / No	N	N	N	N
Diabetes	Yes / No	N	N	Y	N
Left ventricular hypertrophy	Yes / No	N	N	N	N

Mortality rates of overweight and obese subjects in the normal health state are assumed to be equivalent to those observed in the general population although there are studies that explored the relationship between BMI and the risk of death (Fontaine 2003, McGee 2005). Obesity and overweight in adults are found to be associated with large decreases in life expectancy and increases in early mortality. However, we decided not to include these increased mortality risks because there is a danger of double counting if the elevated mortality risks are combined with associated complication mortality rates. Age and gender specific mortality data is obtained from Swiss Federal Statistical Office (2006): overall mortality data (Table 5) and disease specific according to International Classification of Diseases (ICD-10 codes): I10-I15 Hypertensive disease (Table 6), I20-I25 Ischemic heart disease (Table 7), I60-I69 Cerebrovascular disease (Table 8) and E10-E14 Diabetes mellitus (Table 9). The yearly probability of diabetes, hypertension, coronary heart disease and stroke are obtained from the actual number of deaths and disease prevalence rates in Switzerland (SFSO 2006).

Table 5. Overall mortality data (excluding ICD-10: I10-I25, I20-I25, I60-I69, E10-E14)

Sex	Age range	Actual population	Actual death	Yearly probability of death (%)
Male	25-34	504703	381	0.08
Female	25-34	510907	190	0.04
Male	35-44	616872	721	0.12
Female	35-44	609610	403	0.07
Male	45-54	512754	1388	0.27
Female	45-54	506602	857	0.17
Male	55-64	421426	2808	0.66
Female	55-64	430819	1741	0.40
Male	65-74	272748	4435	1.61
Female	65-74	327097	2941	0.90
Male	75-84	156349	6676	4.1
Female	75-84	246640	6391	2.56
Male	85+	42600	5080	11.25
Female	85+	103343	9771	9.03

Table 6. Hypertensive disease (Ht) mortality (ICD-10:I10-I15)

Sex	Age range	Actual population	Prev. Ht (%)	Population with Ht	Actual death Ht	Yearly prob. of death Ht (%)
Male	25-34	504703	3.0	15141	0	0.00
Female	25-34	510907	1.7	8685	0	0.00
Male	35-44	616872	5.7	35162	4	0.01
Female	35-44	609610	3.1	18898	0	0.00
Male	45-54	512754	12.4	63582	15	0.02
Female	45-54	506602	10.0	50660	9	0.02
Male	55-64	421426	25.1	105778	51	0.05
Female	55-64	430819	24.5	105551	18	0.02
Male	65-74	272748	37.6	102553	107	0.10
Female	65-74	327097	35.2	115138	67	0.06
Male	75-84	156349	35.0	54722	248	0.45
Female	75-84	246640	43.4	107042	337	0.31
Male	85+	42600	35.0	14910	292	1.94
Female	85+	103343	43.4	44851	885	1.95

Table 7. Diabetes mellitus (Dm) mortality (ICD-10:E10-E14)

Sex	Age range	Actual population	Prev. Dm (%)	Population with Dm	Actual death Dm	Yearly prob. % of death Dm
Male	25-34	504703	0.7	3533	1	0.03
Female	25-34	510907	0.3	1533	1	0.07
Male	35-44	616872	1.4	8636	4	0.05
Female	35-44	609610	1.0	6096	3	0.05
Male	45-54	512754	2.8	14357	28	0.19
Female	45-54	506602	2.6	13172	14	0.11
Male	55-64	421426	8.3	34978	76	0.22
Female	55-64	430819	4.0	17233	25	0.14
Male	65-74	272748	9.5	25911	176	0.68
Female	65-74	327097	6.8	22243	96	0.43
Male	75-84	156349	13.4	20951	292	1.38
Female	75-84	246640	9.5	23431	418	1.77
Male	85+	42600	13.4	5708	971	15.68
Female	85+	103343	9.5	9818	504	5.01

Table 8. Coronary heart disease (CHD) mortality (ICD-10:I20-I25)

Sex	Age range	Actual population	Prev. CHD (%)	Population with CHD	Actual death CHD	Yearly prob. (%) of death CHD
Male	25-34	504703	0.4	2019	11	0.54
Female	25-34	510907	0.3	1533	0	0.00
Male	35-44	616872	0.9	5552	64	1.15
Female	35-44	609610	0.4	2438	16	0.65
Male	45-54	512754	3.5	17946	205	1.14
Female	45-54	506602	2.0	10132	43	0.42
Male	55-64	421426	11.1	46778	459	0.98
Female	55-64	430819	5.9	25418	103	0.40
Male	65-74	272748	21.5	58641	891	1.51
Female	65-74	327097	9.7	31728	376	1.18
Male	75-84	156349	26.4	41276	1781	4.22
Female	75-84	246640	18.4	45382	1398	3.03
Male	85+	42600	26.4	11246	1577	13.10
Female	85+	103343	18.4	19015	3050	14.85

Table 9. Stroke mortality (ICD-10:I60-I69)

Sex	Age range	Actual population	Prev. Stroke (%)	Population with Stroke	Actual death Stroke	Yearly prob. (%) of death Stroke
Male	25-34	504703	0.4	2019	5	0.25
Female	25-34	510907	0.3	1533	5	0.33
Male	35-44	616872	0.3	1851	7	0.38
Female	35-44	609610	0.6	3658	20	0.55
Male	45-54	512754	1.2	6153	42	0.68
Female	45-54	506602	0.9	4559	31	0.68
Male	55-64	421426	2.2	9271	78	0.84
Female	55-64	430819	2.5	10770	61	0.56
Male	65-74	272748	7.6	20729	240	1.15
Female	65-74	327097	5.4	17663	224	1.26
Male	75-84	156349	13.3	20794	731	3.45
Female	75-84	246640	8.9	21951	895	4.00
Male	85+	42600	13.3	5666	651	10.87
Female	85+	103343	8.9	9198	1588	15.89

Data on the effectiveness of lifestyle intervention is obtained using meta-analysis technique (Saint 1999) of the randomized controlled trials performed in overweight and obese people. The aim of the meta-analysis is to combine the long-term effects of lifestyle intervention on weight and cardiovascular risk factors in overweight and obese people from several studies. The lifestyle intervention in all evaluated studies consisted in dietary counseling and physical exercise sessions and lasted from one to six years with an average follow-up time of three years. A summary of the meta-analyses results on the outcomes of interest is presented in Table 10. Effects are combined using a random effects model. The summary outcome measure calculated is the difference in means between lifestyle intervention and standard care. The pooled estimates of effect size are obtained using Comprehensive Meta-Analysis software (CMA 2005).

Table 10. Meta-analysis results

	Outcome	N studies	N	DM	SE	95% CI	p-Value
Overweight	BMI (kg/m^2)	5	926	-1.11	0.23	-1.56, -0.66	<0.0001
	SBP (mmol/l)	9	2239	-2.08	0.61	-3.28, -0.89	0.001
	TC (mmol/l)	7	1516	-0.26	0.07	-0.41, -0.12	<0.0001
	HDL (mmol/l)	7	1875	0.01	0.01	-0.22, 0.04	0.640
Obese	BMI (kg/m^2)	7	3522	-1.33	0.31	-1.93, -0.72	<0.0001
	SBP (mmol/l)	6	4182	-2.78	0.82	-4.38, -1.18	0.001
	TC (mmol/l)	5	893	-0.14	0.05	-0.24, 0.03	0.011
	HDL (mmol/l)	4	2778	0.04	0.02	0.004, 0.08	0.028

N, number; DM, difference in means; SE, standard error; CI, confidence intervals

Utility score represent the strength of patient preferences for their own health on a scale from 0.0 (death) to 1.0 (perfect health). Three published sources of utilities are used: utilities for overweight and obese people (Macran 2004), utilities changes due to decreases in BMI (Hakim 2002) and utilities associated with the complications of obesity (Jia 2005). The utilities associated with overweight and obese people are presented in Table 11 and the utilities associated with obesity complications are presented in Table 12.

Table 11. Utilities for overweight and obese people

Age group	Overweight		Obese	
	Male	Female	Male	Female
25-34	0.92	0.91	0.89	0.88
35-44	0.89	0.86	0.89	0.82
45-54	0.86	0.83	0.84	0.83
55-64	0.81	0.78	0.72	0.74
65-74	0.81	0.76	0.78	0.71
75+	0.77	0.73	0.76	0.68

Table 12. Utilities associated with obesity complications

Health state	Disutility due to co-morbidities	Utility gain for 1 unit BMI reduction
Normal	-	0.017
Hypertension	-0.053	0.015
Diabetes	-0.042	0.029
Hypercholesterolemia	-0.030	0.015
Coronary heart disease	-0.083	0.017
Stroke	-0.080	0.017

The data on resource used by subjects receiving lifestyle intervention or standard care are taken from Finnish Diabetes Prevention Study (Lindström 2005) and adapted for Switzerland. For lifestyle intervention, seven dietitian visits are assumed in the first year, and four visits per year thereafter. Based on the unit cost of the health care calculated for Switzerland, the dietician cost per visit is estimated at CHF 64 (Tarmed 2006). The same price is assumed for physical exercise which is done in group sessions of 20 people for one hour. The group attended four sessions per month in the first year and two sessions per month during the subsequent year. The total estimated costs of lifestyle intervention are CHF 602 per person in the first year and CHF 333 per person per subsequent year. In the standard control group, costs are assumed zero in overweight people. For obese the standard care consists in three dietitian visit in the first year, and one visit per year thereafter and the equivalent of two exercise sessions per month in the first year and one session per month during the subsequent year. The obesity medication costs are not taken into consideration in the standard therapy of obesity given that there are not chronic medications, i.e. the European prescribing guidelines state that the duration of treatment with orlistat should not be longer than two years (EMEA 1998). Thus, a conservative estimate is preferred. The total estimated costs of standard care intervention in obese subjects are CHF 369 per person in the first year and CHF 102 per person per subsequent year.

The costs of obesity complications are taken from published literature and adjusted to 2006 Swiss prices (CHF) using the consumer price index (SNB 2006). This included the average direct and indirect cost of the disease. A summary of the cost components and source of data is provided in Table 13.

Table 13. Cost components

Component	Cost (CHF)[a]	Reference
Lifestyle intervention in overweight/obese[b]	1268	Lindstrom 2005 adapted for Switzerland
Standard intervention in obese[b]	573	Lindstrom 2005 adapted for Switzerland
Hypertension	1653	Schmid 2005
Type 2 diabetes	2890	Schmitt-Koopmann 2004
Hypercholesterolemia	1245	Schmid 2005
Coronary heart disease	6242	Schmid 2005
Stroke	11495	Schmitt-Koopmann 2004

[a] Cost per person per year adjusted for 2006 prices; [b] Cost of three years intervention

Cost-effectiveness analysis

In order to assess the effect of lifestyle interventions in the prevention and treatment of obesity, we defined three groups of people that are followed throughout the analysis: overweight group, borderline group and moderate obese group. The subjects in the overweight group have a BMI of 27 kg/m^2 selected from the overweight range of BMI 25 to 29.9 kg/m^2, the subjects from the moderate obese group have a BMI of 33 kg/m^2 selected from the moderate obesity range BMI 30 to 35 kg/m^2 and the subjects from the borderline group have a BMI of 30 kg/m^2 representing people who are at the upper limit of overweight and lower limit of obesity. For each group of people: overweight, borderline and moderate obese, the results are presented as difference between lifestyle and standard care intervention in mean costs, life-years (LY) and quality adjusted life-years (QALY). A subgroup analysis was performed in overweight subjects, borderline subjects and moderate obese subjects on different age groups (e.g. 25, 35, 45, 55 years) and gender to allow for a direct comparison among patient populations. Half cycle correction is applied to life expectancy calculations assuming that the transition will take place half way through a cycle.

The cost-effectiveness of interventions is compared using the incremental cost-effectiveness ratio, defined as difference in costs (C_L-C_S) divided by difference in effects (E_L-E_S). Given the chronic nature of the diseases incorporated in the model, a lower value of LY compared to QALY is expected. Our model comparatively calculates the incremental cost-effectiveness ratio as cost per LY in a cost-effectiveness analysis and as cost per QALY in a cost-utility analysis. Incremental costs were reported in Swiss Francs (CHF) and Euro (€) to allow for an international comparison (exchange rate 1 CHF = 0.607158 €, May 9, 2007). The analysis is performed from a society perspective. Future costs and effects are presented undiscounted and discounted at 3% rate. The model is developed using Microsoft Excel Software.

Sensitivity analysis

A probabilistic sensitivity analysis (Briggs 2000) is carried out to investigate the robustness of the data input including the baseline risks of transitioning to obesity-complications, the lifestyle and standard care intervention effects, the utility values, the costs of complications, and the costs of interventions. In order to propagate uncertainty in our model, distributions are assigned to all model parameters that are estimated with uncertainty. The distributional forms of the model parameters are: normal distribution for the costs of interventions and cardiovascular risk factors, gamma distribution for the costs of obesity complications and beta distribution for utility scores. Values are drawn at random from the specified distributions using a random number generator for the chosen parameter. Monte Carlo simulation was used to propagate these distributions through the model by recalculating the results over a large number of iterations. The results of running the probabilistic sensitivity analysis by randomly sampling from the parameter distributions are presented on the cost-effectiveness plane (Briggs 1998) and cost-effectiveness acceptability curve (Fenwick 2004).

Results

The average difference in costs and effects between lifestyle intervention and standard care is presented in Table 14. The lifestyle intervention resulted in increased survival and improved quality of life, equivalent to a difference of 0.05 LY and 0.33 QALY per person per year gained over lifetime. Compared with standard care, the average incremental cost of lifestyle intervention is lower in the borderline group (female CHF -281, male CHF -384) and the moderate obese group (female CHF -65, male CHF -176) and higher in the overweight group (female CHF 216, male CHF 124) when probabilistic results are undiscounted. When future costs and effects are discounted at 3% rate, lifestyle intervention has lower costs than standard care in borderline male subjects (CHF -99). In the borderline group, the lifetime incremental cost per LY is estimated in female at CHF -15000 (€ -9107) per LY and male at CHF -13933 (€ -8460) per LY when results are undiscounted. When 3% discount rate is applied, incremental cost-effectiveness ratio is CHF 1700 (€ 1032) per LY in borderline female subjects and CHF -11200 (€ -6800) per LY in borderline male subjects. Lifetime incremental cost per QALY in the borderline group is estimated in female at CHF -969 (€ -588) per QALY and male at CHF -1200 (€ -729) per QALY when results are undiscounted. When discounting is applied, incremental cost-effectiveness ratio is CHF 64 (€ 39) per QALY in borderline female subjects and CHF -354 (€ -215) per QALY in borderline male subjects.

Table 14. Incremental costs and effects per person per year

	Sex	Average [a]	Overweight [b]		Borderline [b]		Moderate Obese [b]	
			0 %	3 %	0 %	3 %	0 %	3 %
Deterministic	F	Cost, CHF	234	510	-161	80	18	207
		Cost, €	142	310	-98	49	11	126
		LY	0.02	0.01	0.02	0.01	0.02	0.01
		QALY	0.27	0.23	0.29	0.25	0.29	0.26
	M	Cost, CHF	156	405	-260	-6	-70	127
		Cost, €	95	246	-158	-4	-43	77
		LY	0.03	0.01	0.03	0.01	0.02	0.01
		QALY	0.29	0.25	0.32	0.28	0.33	0.29
Probabilistic	F	Cost, CHF	216	490	-281	16	-65	145
		Cost, €	131	298	-171	10	-39	88
		LY	0.02	0.01	0.05	0.01	0.02	0.01
		QALY	0.26	0.23	0.29	0.25	0.30	0.26
	M	Cost, CHF	124	384	-384	-99	-176	44
		Cost, €	75	233	-233	-60	-107	27
		LY	0.03	0.01	0.03	0.01	0.03	0.01
		QALY	0.30	0.25	0.32	0.28	0.30	0.29

[a] Difference between lifestyle and standard intervention, average from age 25 to 65 years;
[b] Lifetime results are presented undiscounted (0%) and discounted (3%)

Subgroup analysis is conducted on different age groups (Table 15). The bolded cells show the dominant intervention strategies whose lifestyle intervention has lower costs and a higher health effect compared with standard intervention. The subgroup analysis estimates that lifestyle intervention dominates standard care in overweight female subjects aged 45 years, overweight male subjects aged 25 years, borderline female subjects aged 30 to 65 yeas, overweight male subjects aged 25 to 65, moderate obese female subjects aged 35 to 60 and moderate obese male subjects aged 25 to 60 years, when results are undiscounted. When discounting is applied to future costs and effects, lifestyle intervention dominates standard care in borderline female subjects aged 35 to 55 years, borderline male subjects aged 25 to 60 years, moderate obese female subjects aged 45 years and obese male subjects aged 55 years.

Table 15. Incremental cost-effectiveness ratio per person per year

	Sex	ICER	Age 25[a]		Age 35[a]		Age 45[a]		Age 55[a]	
			0%	3%	0%	3%	0%	3%	0%	3%
Overweight	F	CHF/ QALY	4515	6286	95	2352	**-419**	1131	508	1366
		CHF/ LY	76734	295863	834	45551	**-4311**	23126	7278	33482
	M	CHF/ QALY	**-374**	1854	395	2014	324	1457	237	914
		CHF/ LY	**-2840**	34291	3006	30934	3054	24473	2787	17149
Borderline	F	CHF/ QALY	1534	3023	**-2945**	**-1630**	**-2000**	**-785**	**-898**	**-173**
		CHF/ LY	27510	142619	**-19380**	**-6050**	**-22053**	**-17106**	**-13032**	**-4250**
	M	CHF/ QALY	**-2560**	**-781**	**-283**	**-331**	**-1373**	**-523**	**-1027**	**-508**
		CHF/ LY	**-19496**	**-14886**	**-14196**	**-52927**	**-14158**	**-9595**	**-13282**	**-10417**
Obese	F	CHF/ QALY	1973	3180	**-982**	**-753**	**-1026**	**-88**	**-355**	173
		CHF/ LY	39925	171544	**-11287**	10019	**-14496**	**-2426**	**-6670**	5481
	M	CHF/ QALY	**-1453**	3	395	276	**-741**	9	**-502**	**-69**
		CHF/ LY	**-12657**	58	**-7373**	5580	**-8912**	185	**-8048**	**-1745**

[a] Lifetime probabilistic results are presented undiscounted (0%) and discounted (3%)

To reflect the uncertainty in the estimates, Figure 2 presents a scatter plot of the mean differences in costs and QALY gained between lifestyle and standard care, derived from the Monte Carlo simulation, in borderline people. The x- and y-axis divide the graph into four separate quadrants, which represent the following scenarios for lifestyle versus standard care (clockwise from the top right): (a) more effective and more costly; (b) more effective and less costly; (c) less effective and less costly, and (d) less effective and more costly. The high concentration of points in quadrant (a) and (b) indicate that lifestyle intervention is more effective than standard care. For example in borderline the majority of the replicates lie in the quadrant (b), where lifestyle dominates standard care, 78% simulations of the lifestyle intervention are cost-effective relative to standard care (Figure 2). However, the dispersion of points above and below the x-axis, indicates that there is some uncertainty about whether this gain in QALY is achieved at a lower or higher cost than standard care. If the gain in QALY is achieved at a higher cost, then the critical issue that determines whether lifestyle intervention is cost-effective is how much (if any) the decision maker is willing to pay for an additional unit gain in health outcome.

Figure 2. Scatter plot of the difference in mean costs and effects between lifestyle and standard care intervention

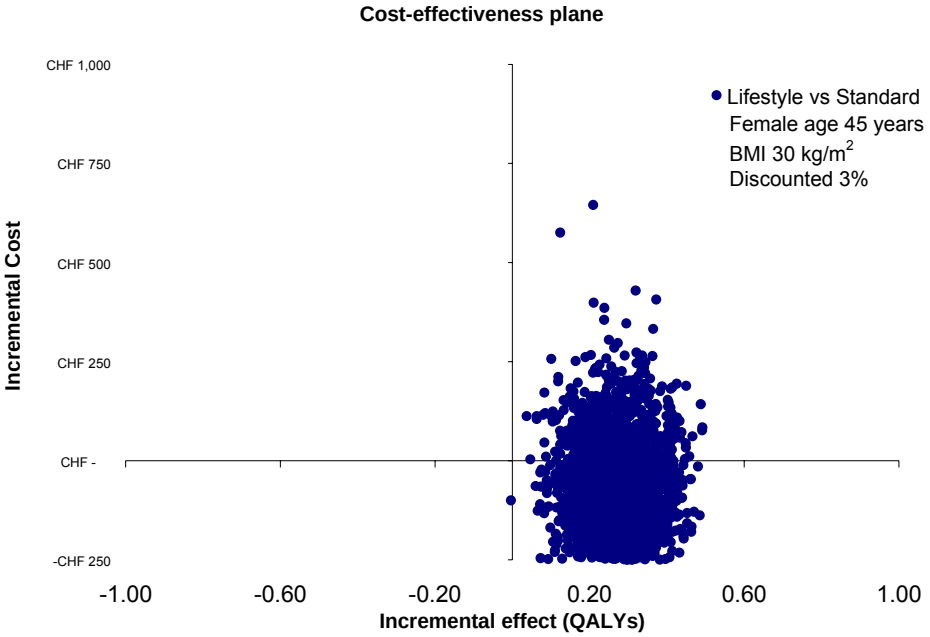


Figure 3 illustrates the cost-effectiveness acceptability curves for lifestyle compared to standard care intervention in borderline group. The curves indicate the probability of lifestyle intervention being more cost-effective than the standard care for a range of potential maximum amounts a decision maker is willing to pay for an additional unit of health gained. We carried out the analysis for a range of values for the society willingness to pay for an additional unit of health gain. When the decision-maker is unwilling to pay anything additional for a health gain the probability that lifestyle intervention is cost-effective is 5% in the overweight group, 78% in the borderline group (Figure 3.) and 47% in the moderate obese group. If the decision-maker is willing to pay CHF 500 per QALY gained, the probability of the lifestyle intervention being cost-effective increases to 12% in overweight subjects, 95% in borderline subjects (Figure 3) and 75 % in moderate obese subjects. Furthermore, if the decision-maker is willing to pay CHF 1000 per QALY gained, the probability of the lifestyle intervention being cost-effective increase further to 35 % in overweight subjects, 99 % in borderline subjects (Figure 3) and 92 % in moderate obese subjects.

Figure 3. Cost-effectiveness acceptability curves lifestyle versus standard care intervention

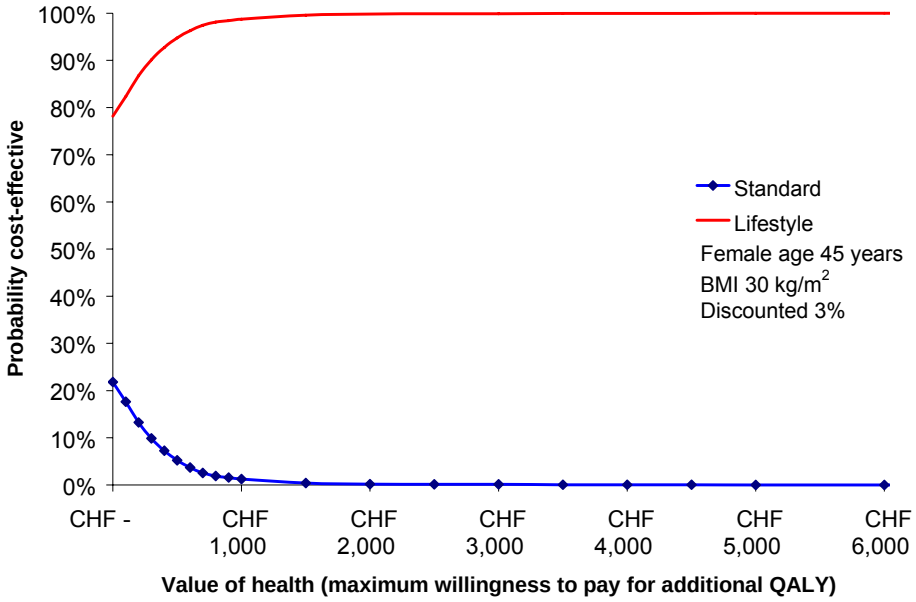


Figure 4 and Figure 5 present the cost-effectiveness acceptability curves in borderline subjects for different age groups in female and male. Overall, the cost-effectiveness for men is greater than for women. Lifestyle intervention is more cost-effective in female aged 35 to 55 years and male aged 25 to 55 years compared to other age groups.

Figure 4. Cost-effectiveness acceptability curves of lifestyle intervention in female

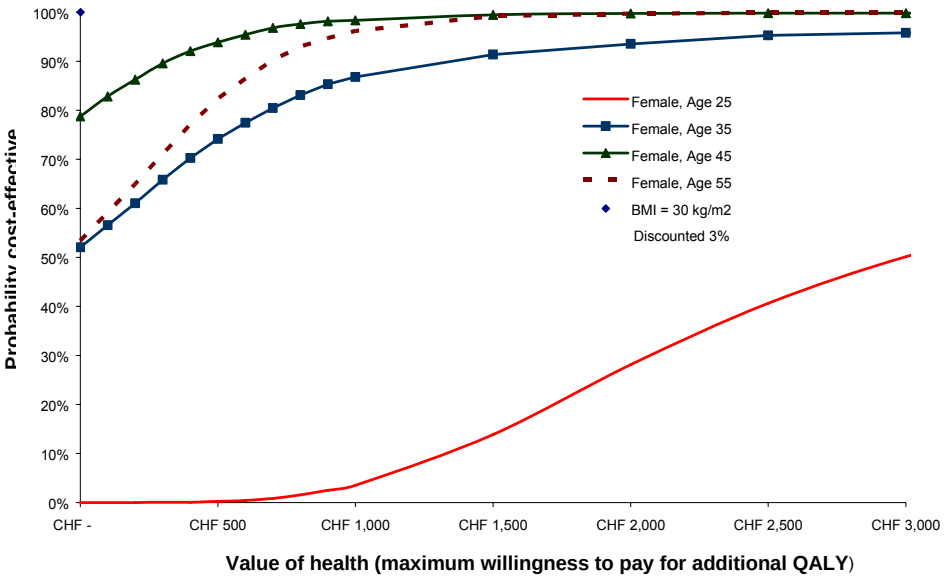
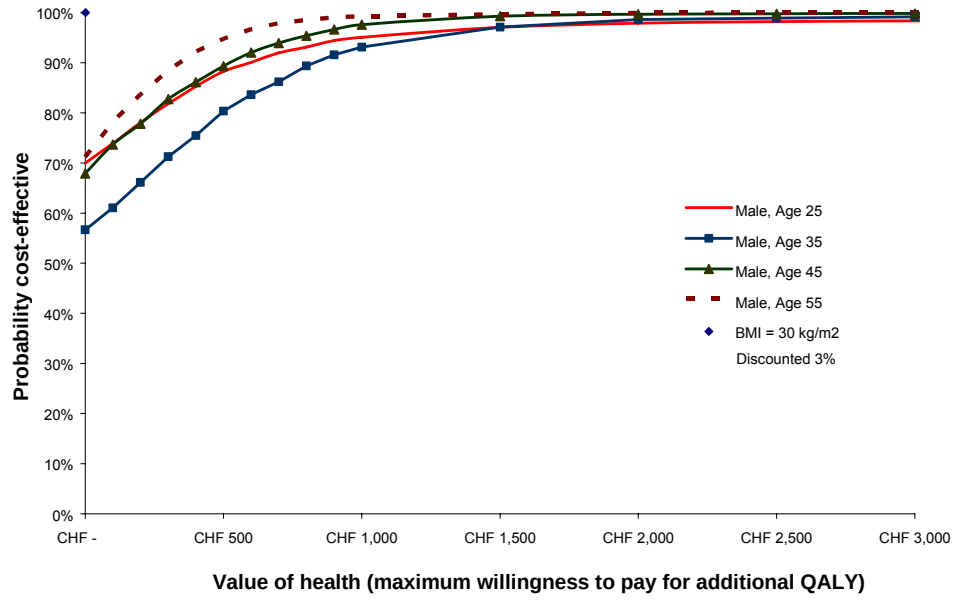


Figure 5. Cost-effectiveness acceptability curves of lifestyle intervention in male



Discussion

We defined based on BMI level three groups of people that are followed thought the study: overweight subjects (BMI 27 kg/m^2), borderline subjects (BMI 30 kg/m^2) and moderate obese subjects (BMI 33 kg/m^2). For the purpose of this study the classification allows us to differentiate prevention from the treatment of obesity and to better understand the importance of the transition from overweight to obesity reflected in borderline group.

Using a lifestyle intervention in overweight and obese people, the present model tries to quantify several elements of benefit associated with weight loss: first, a reduction in the risk of developing diabetes, hypertension, hypercholesterolemia and cardiovascular disease, second, an increase in life expectancy and quality of life, and third, a reduction in treatment costs by reducing the risk of developing obesity complications. In all three groups, independent of age, the lifetime effect of three-year lifestyle intervention resulted in increased survival and an improved quality of life compared with standard care. The difference in cost between lifestyle intervention and standard care is lower in borderline subjects and moderate obese subjects and higher in overweight subjects. In particular, in borderline people lifestyle intervention dominates standard care intervention being less costly and more effective. The majority of the costs savings in borderline group are attributable to a decrease in the risk of developing obesity complications. Despite the lifestyle interventions incurring higher mean costs in the first three years, these additional costs are offset by the reduced probability of developing obesity related complications. The subgroup analyses differentiate the cost-effectiveness for each age group. We observed some important gender differences. Firstly, the incidence of diabetes is higher in men resulting in higher cost saving for men aged 25 to 35 years compared with women in the same age group. Secondly, women have a lower mortality rates and, as a consequence, higher life expectancy compared to men resulting in higher costs associated with obese co-morbidities. Thirdly, investigations have suggested that weight gain during and after menopause may contribute more to cardiovascular disease than actual weight prior to menopause and that weight loss and increase physical activity may mitigate some of the cardiovascular risk factors i.e. high cholesterol, insulin resistance (Carels 2004, Matthews 1989). These observations are consistent with our findings that weight loss with lifestyle intervention is more cost-effective in women aged 45-55 years due to a decrease in the cardiovascular risk factors compared with standard care intervention.

The current analysis suggests that borderline people who are at the upper limit of overweight and lower limit of obesity benefit the most from a lifestyle intervention lasting for three years. The lifetime effect of a lifestyle intervention is reflected in lower costs and higher effects. This may be considered a critical point in the prevention of obesity. It has been suggested that reducing the risk factors that diseases have in common may prove to be an efficient prevention strategy (Epstein 1983). For example, several risk factors such as obesity, physical inactivity, hypertension, hyperglycaemia and hypercholesterolemia predict the development of chronic diseases such as diabetes and cardiovascular disease. Therefore, a successful preventive program based on lifestyle intervention may target several risk factors simultaneously.

One issue of concern in the cost-effectiveness analysis is that societal threshold, the maximum willingness to pay for one unit of health gain, is not known. It has been suggested that incremental costs of less than CHF 24178 (€ 14680) per QALY are considered cost-effective representing strong evidence for adoption (Laupacis 1992); however, the acceptance value is a subject highly debated in the literature. For example, the National Institute for Clinical Excellence in the United Kingdom evaluated orlistat and sibutramine for the treatment of obesity in adults and determined the cost-effectiveness of orlistat at values ranged from CHF 25093 to CHF 110987 per QALY (NICE report 31) and of sibutramine at values ranged from CHF 25334 to CHF 72383 per QALY (NICE report 22). In Switzerland, the cost-effectiveness of orlistat was estimated at CHF 22413 (€ 13 608) per QALY (Ruof 2005). The Swiss authorities decided to reimburse orlistat in the treatment of diabetic patients with BMI $\geq$ 28 kg/m^2, but stipulated that it should be continued beyond six months only in patients who lose $\geq$ 5 kg of their starting weight or achieve a reduction in HbA1c of $\geq$ 0.5 % (Ruof 2005).

To compare the results of our study with other economic analyses performed in overweight and obese people we reviewed the published economic literature on lifestyle interventions. An economic analysis was performed in the American Diabetes Prevention Program (DPP) in obese people with impaired glucose tolerance (Herman 2005). Compared with standard care intervention, lifestyle intervention cost CHF 768 more over a lifetime and produce a gain of 0.57 QALY, resulting in a cost per QALY of approximately CHF 1330 (€ 808). Another modelling study performed in the DPP (Eddy 2005) estimate the 30-year cost per QALY of lifestyle intervention compared with the control group at CHF 74953 (€ 45508). The authors suggest that using lifestyle intervention until after a person develops diabetes would be more cost-effective (CHF 29619 per QALY). In United Kingdom, an economic analysis of lifestyle intervention in obese people at risk of diabetes, estimates the 15-year incremental cost per QALY at CHF 14054 or € 8533 (Avenell 2004). It has been also documented that non-pharmacological treatments that target severe obesity (BMI > 35 kg/m^2) are cost-effective (Tsai 2005). The fact that certain treatments are cost-effective in high risk individuals with severe obesity or associated co-morbidities i.e. diabetes, cardiovascular disease does not answer the question of the cost-effectiveness of the same interventions for lower risk obese or overweight individuals, whose prevention or treatment benefits may also be worth the cost. To answer this question, a better understanding of the implications of weight loss and improvement of risk factors for long-term health outcomes is necessary in overweight and moderate obese people. Policy makers need country specific economic data on overweight and obesity to tackle the growing burden of obesity. Our cost-effectiveness analysis illustrates from a Swiss perspective the lifetime impact of lifestyle intervention in the prevention and treatment of obesity. The obtained results provide an argument for organizations and institutions actively involved in the field of obesity prevention to justify funds for intensified prevention strategies using lifestyle interventions.

Our analysis also has several limitations. The model could be improved by having access to additional Swiss-specific data. So far, epidemiological data

such as the correlation between BMI and the risk of complications, obesity related mortality data and changes in patient utility for weight loss have not been recorded specifically for Switzerland. Further investigations should also take into account other important complications of obesity such as metabolic syndrome, colorectal cancer, gall bladder disease, sleep apnoea and depression. Our model does not take into consideration smoking as a risk factor although it has been documented that obese smokers have a decrease in life expectancy compared with non-smokers (Sempos 1998). Another limitation of our study consists in the estimation of the costs of obesity complications from secondary data sources. In our model, the cost of stroke represents the largest value among all complications. However, this high cost is confirmed by various costs of illness studies performed in Europe (Kolominsky-Rabas 2006, Gerzeli 2005). Overall, our model reflects a conservative approach considering that we do not include obesity medication, all related obesity complications and correlations between diseases incorporated in the model. Although our analysis simplifies the complex reality, it provides a first positive estimate on the lifetime impact of lifestyle intervention on overweight and obese people in Switzerland.

Conclusion

In summary, prevention and treatment of obesity must be a priority for heathcare decision-makers across Europe giving the rapid increase in prevalence. With the aid of a decision analytic model, we synthesised Swiss specific evidence available on the outcomes and costs of lifestyle intervention and standard care in overweight and obese people. Our economic analysis suggests that lifestyle intervention is cost-effective in the long-term prevention and treatment of obesity in Switzerland.

Acknowledgement

This study was funded by the Swiss Federal Office of Health.

Appendix

a) Framingham equation - Coronary Heart Disease

From Anderson et al. (1990), the probability of a new case of CHD at period t is given by

$$CHD(t) = [F(t) – F(t – 1)] / [1 – F(t – 1)]$$

Where

$$F(t) = 1 – \exp(–\exp\{[\ln(t) – \mu(t)] / \sigma(t)\}) \quad \text{(the Weibull function)}$$

$\mu = 15.5305 + 28.4441 \times female – 1.4792 \times \ln[age(t)] – 14.4588 \times \ln[age(t)] \times female + 1.8515 \times \ln[age(t)]2 \times female – 0.9119 \times \ln[sbp(t)] – 0.2767 \times smoker(t) – 0.7181 \times \ln[totalc(t) / HDL (t)] – 0.1759 \times diagnosed\ diabetic – 0.1999 \times diabetic \times female –0.5865 \times LVH(t, gender)$

$\ln(\sigma) = 0.9145 – 0.2784 \times \mu$

Equation input:

- Age in years, input female=1, male=0
- sbp, systolic blood pressure, input range 95-185 mmHg
- Diabetes, 0=no, 1=yes
- Smoking, 0=no, 1=yes
- totalc = total cholesterol level, input range 3.5-10 mmol/l
- HDL, high density lipoprotein cholesterol level, input range 0.65-2.65 mmol/l
- LVH, left ventricular hypertrophy, input 0=no, 1=yes, when information not available LVH=0

b) Framingham equation - Stroke

From Anderson et al. (1990), the probability of a new case of Stroke at period t is given by

$$Prob(S[t]) = [F(t) – F(t - 1)] / [1 – F(t - 1)]$$

Where

$$F(t) = 1 – \exp(–\exp\{[\ln(t) – \mu(t)] / \sigma(t)\}) \text{ (the Weibull function)}$$

$\mu = 26.5116 + 0.2019 \times female – 2.3741 \times \ln[age(t)]– 2.4643 \times \ln[sbp(t)] – 0.3914 smoker(t) – 0.0229 \times \ln[totalc(t) / HDL (t)] – 0.3087 \times diagnosed\ diabetic – 0.2627 \times diabetic \times female – 0.2355 \times LVH$

$\ln(\sigma) = -0.4312$

Equation input:

- Age in years, input female=1, male=0
- sbp, systolic blood pressure, input range 95-185 mmHg
- Diabetes, 0=no, 1=yes
- Smoking, 0=no, 1=yes
- totalc = total cholesterol level, input range 3.5-10 mmol/l
- HDL, high density lipoprotein cholesterol level, input range 0.65-2.65 mmol/l
- LVH, left ventricular hypertrophy, input 0=no, 1=yes, when information not available LVH=0

Chapter 6

Direction of Further Research: Value of Additional Information in Cost-Effectiveness Analysis

Summary

There has been an increasing interest in using value of information analysis in medical decision-making, to quantify the uncertainty in decision-making, to identify the need for further research and as a tool for sensitivity analysis. The aims of the present chapter are to quantify the uncertainty in the cost-effectiveness of lifestyle intervention versus standard care in overweight and obese people and to determine if further research is necessary based on current information. Value of information analysis was applied on a probabilistic cost-effectiveness model to evaluate the uncertainty by calculating the patient expected value of perfect information (EVPI), population EVPI and partial EVPI. The costs were expressed in Swiss Francs (CHF), price year 2006. Results showed that EVPI was higher in overweight than in obese people. The maximum population EVPI was CHF 6.8 million in overweight people and CHF 3.2 million in moderate obese people representing the upper limit on costs associated with decision uncertainty. The partial EVPI estimated a higher uncertainty in the model parameters such as utilities, body mass index, cardiovascular risk factors and systolic blood pressure in overweight and moderate obese subjects. In conclusion, the EVPI analysis indicates that there is some uncertainty regarding the choice between lifestyle intervention and standard care. The parameter EVPI suggests that if further research is commissioned, this should focus on the effectiveness of lifestyle intervention on cardiovascular risk factors and utilities.

Introduction

The increasing prevalence of obesity worldwide underlines the pressing need for finding effective interventions to tackle the disease and its economic consequences. In Switzerland, overweight and obesity account for 37% of the population aged over 15 years (Swiss Federal Statistical Office 2006). Obesity is associated with a high risk of morbidity, mortality as well as reduced life expectancy (Fontaine 2003). The economic burden of obesity is substantial. In Switzerland, obesity related expenditures are estimated to have a cost range approximately between CHF 2691 million and CHF 3229 million, representing 2.3-3.5% of total health care expenditures (Schmid 2005).

The increasing burden on the budget of the Swiss healthcare providers resulted in considerable interest in assessing existing treatments for their clinical effectiveness and cost-effectiveness. A wide variety of treatments for obesity are available including diet, physical exercise, behavioral modification, pharmacological treatment and surgery. Among several treatment options, lifestyle intervention, including dietary counseling and physical exercise, has been documented to lead safely to improvements in metabolic abnormalities such as increased body weight, dyslipidemia, elevated blood pressure, and glucose control that are linked to the development of obesity, diabetes, and cardiovascular disease (Pritchett 2005).

A decision analytic model was developed to assess the cost-effectiveness of lifestyle intervention in overweight and obese people in Switzerland (Galani 2007). Decision analytic models are extensively used in formal decision-making process. One of the requirements for the decision-making is that uncertainty in adopting a decision based on cost-effectiveness must be appropriately characterized and quantified because this affects the value and the interpretation of the model output. Many guidelines for cost-effectiveness analysis recommend probabilistic sensitivity analysis to assess uncertainty associated with the model parameters (Claxton 2002). Value of information analysis has been suggested as a natural methodological extension of the probabilistic sensitivity analysis (Felli 1999). The analysis quantifies the uncertainty by establishing the value of acquiring additional information to inform decision making. The use of value of information analysis has been recently encouraged in decision analytic models (Claxton 2004, Claxton 2001, Sculpher 2005, Fenwick 2005, Philips 2006).

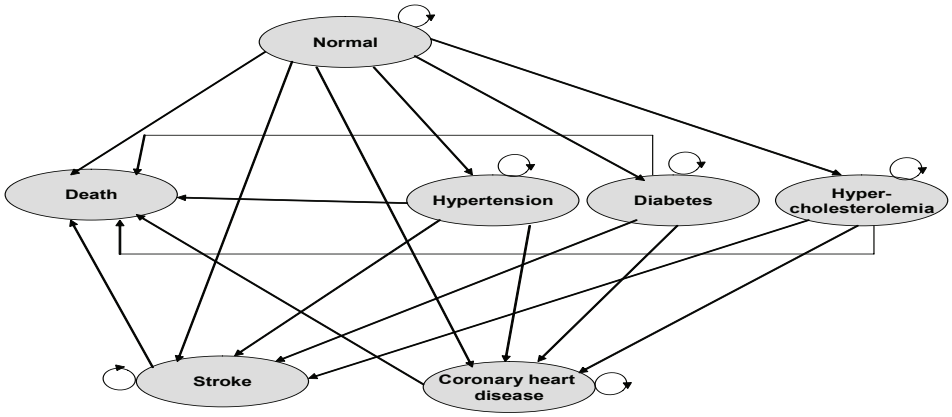
The aim of the present study was to apply the value of information analysis to assess the uncertainty in the cost-effectiveness of lifestyle intervention in overweight and obese people in Switzerland and to determine if further research is necessary based on current information.

Methods

A probabilistic cost-effectiveness model based on Markov process was developed to compare the effect of lifestyle intervention with standard care in overweight and obese subjects. Seven health states were modeled: ‘normal’, if subjects are overweight or obese but free of complications, ‘hypertension’, ‘hypercholesterolemia’, ‘type 2 diabetes’, ‘stroke’, ‘coronary heart disease’ and ‘death’. A representation of the Markov model is presented in Figure 1. Subjects

enter the model in the normal health state. Three characteristics are selected: starting age, starting body mass index (BMI) and gender. Based on BMI subjects are classified as overweight if they have a BMI between 25 kg/m^2 and 29.9 kg/m^2 or obese if they have a BMI of 30 kg/m^2 or over. The cycle length is one year. At the end of each one-year, proportions of subjects can move from one disease state to another or stay in the same disease state.

Figure 1. Model diagram



The transition probabilities between cycles are based on the disease progression with age, gender, BMI and cycle number. We assumed that all subjects developing hypertension, diabetes and hypercholesterolemia are diagnosed and treated. Subjects remain in those states once they have entered, unless they develop cardiovascular disease or die. The possibility of having concomitant diabetes, hypertension and hypercholesterolemia was not incorporated into the model. One reason for excluding the correlation between the existing co-morbidities is the absence of prevalence data for Swiss population. This assumption is most probably underestimating the burden associated with obesity related co-morbidities. Subjects entering the model had a minimum age of 25 years and a maximum of 85 years, therefore, the model runs for a period of 60 years from the age of 25 years, i.e. when a subject enters the model at age of 60 years, the model will run until he/she will have the age of 85 years. The reason for setting the age limit is based on the average life expectancy in Switzerland of 77.3 years for men and 82.8 for women (World Health Organization 2006). We considered that any simulation after the age of 85 years would overestimate the costs and the effects.

A hypothetical cohort of 10000 overweight or obese subjects received a lifestyle intervention or standard care intervention for a period of three years. Lifestyle intervention consisted of regular physical activity and healthy eating, including diets rich in fruits and vegetables. Lifestyle intervention group members

attended dietitian sessions and supervised exercise sessions during the first three years. Standard care in overweight people consisted in no intervention whereas in obese people consisted in basic dietary counseling and physical exercise sessions (see cost description). Treatment effect was modeled as a reduction in BMI, systolic blood pressure, total cholesterol, and high density lipoprotein cholesterol based on data obtained in three-year active treatment. It was assumed that the effect of lifestyle intervention on cardiovascular risk factors and weight loss is maintained up to six years, thereafter subjects start to regain weight linearly for a period of four years, i.e. after ten years weight loss went back to the initial weight. This is in line with the assumption used in the economic evaluations of weight loss medication (Maetzel 2003, Hertzman 2005) and is based on observations from clinical trials (NICE 2007).

We estimated our model using data from a variety of secondary sources. A summary of the data input is presented in Table 1. The correlation between BMI and annual risk of developing hypertension, diabetes, and hypercholesterolemia was calculated based on two large epidemiological studies: the Nurses Health Study and the Health Professional Follow-up Study (Field 2001). Intermediate values of BMI have been interpolated using polynomial function. The risk of complications had been adjusted according to age, gender and prevalence of hypertension, diabetes and hypercholesterolemia based on the information provided by the Swiss health survey (Swiss Federal Statistical Office 2006). The mean BMI by age and gender of the Swiss population was obtained from published literature (World Health Organization 2006, Mähönen 2006).

Table 1. Model data input

Parameter name	Value	Unit	Distribution	Source
BMI overweight [a]	-1.11	kg/m^2	Normal (95%CI -1.56 to -0.66)	meta-analysis
BMI obese [a]	-1.33	kg/m^2	Normal (95%CI -3.28 to -0.89)	meta-analysis
SBP overweight [a]	-2.08	mmHg	Normal (95%CI -3.28 to -0.89)	meta-analysis
SBP obese [a]	-2.78	mmHg	Normal (95%CI -4.38 to -1.18)	meta-analysis
TC overweigh t[a]	-0.26	mmol/l	Normal (95%CI -0.41 to -0.12)	meta-analysis
TC obese [a]	-0.14	mmol/l	Normal (95%CI -0.24 to 0.03)	meta-analysis
HDL overweight [a]	0.01	mmol/l	Normal (95%CI -0.02 to 0.04)	meta-analysis
HDL obese [a]	0.04	mmol/l	Normal (95%CI 0.01 to 0.08)	meta-analysis
Mortality probability	Dependent on age, complication		Not applicable	Swiss Federal Statistical Office
Utilities overweight/ obesity	Dependent on age, BMI		Beta (alpha, beta)	Macran 2004
Disutility due to obesity complications	Dependent on age, BMI, complication		Beta (alpha, beta)	Jia 2005
Utility gain due to 1 unit decrease in BMI	Dependent on age, BMI, complication		Beta (alpha, beta)	Hakim 2002
Cost hypertension	1653	CHF (2006)	Gamma (alpha 262, beta 6)	Schmid 2005

Parameter name	Value	Unit	Distribution	Source
Cost diabetes	2890	CHF (2006)	Gamma (alpha 357 beta 8)	Schmitt-Koopmann 2004
Cost hypercholesterolemia	1245	CHF (2006)	Gamma (alpha 149, beta 8)	Schmid 2005
Cost stroke	11459	CHF (2006)	Gamma (alpha 81 beta 142)	Schmitt-Koopmann 2004
Cost coronary heart disease	6054	CHF (2006)	Gamma (alpha 67, beta 94)	Schmid 2005
Cost lifestyle overweight/obesity year 1	602	CHF (2006)	Normal (95%CI 700 to 500)	Lindstrom 2005, adapted for CH
Cost lifestyle overweight /obesity year 2 and year 3	333	CHF (2006)	Normal (95%CI 400 to 200)	Lindstrom 2005, adapted for CH
Cost standard care obesity year 1	269	CHF (2006)	Normal (95%CI 400 to 200)	Lindstrom 2005, adapted for CH
Cost standard care obesity year 2 and year 3	102	CHF (2006)	Normal (95%CI 150 to 90)	Lindstrom 2005, adapted for CH

[a] Incremental effect (Lifestyle intervention - Standard Care), BMI, body mass index, SBP, systolic blood pressure, TC, total cholesterol, HDL, high density lipoprotein cholesterol, CHF, Swiss Francs (year 2006), CH, Switzerland

The risk of developing coronary heart disease and stroke from 'normal', 'hypertension', 'diabetes', 'hypercholesterolemia' health states were based on a risk equation from the Framingham cohort study (Anderson 2001). The risk factors were age, gender, systolic blood pressure, total cholesterol, high density lipoprotein cholesterol, presence of diabetes and smoking status. Data on systolic blood pressure, total cholesterol and high density lipoprotein of the Swiss population were obtained from literature (Mähönen 2006). The mean systolic blood pressure increased with age in both men and women, rising from 127 mmHg in men aged 25-35 years to 145 mmHg in men aged 75 years and over, and from 115 mmHg to 144 mmHg in women. The mean blood cholesterol levels increased with age with a slight decrease in the oldest age group.

Mortality rates of overweight and obese subjects in the normal health state were assumed to be equivalent to those observed in the general population although there are studies that explored the relationship between BMI and the risk of death (Fontaine 2003, McGee 2005). Obesity and overweight in adults are found to be associated with large decreases in life expectancy and increases in early mortality. However, we decided not to include these increased mortality risks because there is a danger of double counting if the elevated mortality risks are combined with associated complication mortality rates. Age and gender specific mortality data were obtained from Swiss Federal Statistical Office: overall mortality data and disease specific according to International Classification of Diseases (ICD-10 codes): I10-I15 Hypertensive disease, I20-I25 Ischemic heart disease, I60-I69 Cerebrovascular disease and E10-E14 Diabetes mellitus. The yearly probability of developing diabetes, hypertension, coronary heart disease and stroke were obtained from the actual number of deaths and disease prevalence rates in Switzerland (Swiss Federal Statistical Office 2006).

Data on the effectiveness of lifestyle intervention was obtained using meta-analysis technique (Saint 1999) of the randomized controlled trials performed in overweight and obese people. The meta-analysis combined the long-term effects of lifestyle intervention on weight and cardiovascular risk factors in overweight and obese people from several studies. The lifestyle intervention in all evaluated studies consisted in dietary counseling and physical exercise sessions and lasted from one to six years with an average follow-up time of three years. The summary outcome measure calculated was the difference in means between lifestyle intervention and standard care. Effects were combined using a random effects model (Table 1).

Utility score represent the strength of patient preferences for their own health on a scale from 0.0 (death) to 1.0 (perfect health). Three published sources of utilities were used: utilities for overweight and obese people (Macran 2004), utilities changes due to decreases in BMI (Hakim 2002) and utilities associated with the complications of obesity (Jia 2005).

The data on resource used by patients receiving lifestyle intervention or standard care were obtained from the Finnish Diabetes Prevention Study (Lindström 2005) and adapted for Switzerland. In the lifestyle intervention group seven dietitian visits were assumed in the first year, and four visits per year thereafter. Based on the unit cost of the health care calculated for Switzerland, the dietician cost per visit was estimated at CHF 64 (TARMED 2006). The same

price was assumed for physical exercise which was done in group sessions of 20 people for one hour. The group attended four sessions per month in the first year and two sessions per month during the subsequent year. The total estimated cost of lifestyle intervention was CHF 602 per person in the first year and CHF 333 per person per subsequent year. In the standard intervention group, costs were assumed zero in overweight people. For obese subjects the standard care intervention consisted in three dietitian visits in the first year, and one visit per year thereafter and the equivalent of two exercise sessions per month in the first year and one session per month during the subsequent year. The obesity medication costs were not taken into consideration in the standard therapy of obesity. Thus, a conservative estimate was preferred. The total estimated cost of standard care intervention in obese subjects was CHF 269 per person in the first year and CHF 102 per person per subsequent year.

The costs of obesity complications were obtained from published literature and adjusted to 2006 Swiss Francs (CHF) prices using the consumer price index (Swiss National Bank 2006). A top down method using a prevalence approach has been used to estimate the direct and indirect costs of obesity complications.

We evaluated the cost-effectiveness of lifestyle intervention versus standard care intervention using a cost utility analysis. The cost-effectiveness of interventions was compared using the incremental costs (CHF), the incremental effects (quality adjusted life-years, QALY) and the cost-effectiveness ratio (CHF/QALY). In order to assess the effect of lifestyle intervention in overweight and obese subjects, we defined two groups of people that were followed throughout the analysis: overweight subjects (BMI 28 kg/m^2) and moderate obese subjects (BMI 33 kg/m^2). A subgroup analysis was performed in male and female subjects aged 30, 40, 50, and 60 years.

A probabilistic sensitivity analysis was performed (Briggs 2000). To reflect the uncertainty in the model parameters they were incorporated in the model as probability distribution. The cardiovascular risk factors data input were characterized as normally distributed with standard deviations based on the meta-analysis results. Utilities score were characterized as beta distributions for two reasons: beta distribution takes values between 0 and 1, and it is a continuous distribution, which is a desirable property for representing uncertainty. Gamma distribution, which is constrained on the interval 0 to positive infinity, was used to characterize the costs of obesity complications. Gamma distribution can be highly skewed to reflect the skew often found in cost data. We characterized the distribution of the costs of interventions as normal with standard deviations equal to the standard error, because it cannot take values less than zero and it is positively skewed. Monte Carlo simulation was used to propagate these distributions through the model by recalculating the results over a large number of simulations. The results of running the probabilistic sensitivity analysis by randomly sampling from the parameter distributions are presented on the cost-effectiveness acceptability curve (Fenwick 2004).

The cost-effectiveness analysis adopted the society perspective. The model was developed using Microsoft Excel.

Value of information analysis

Value of information analysis was undertaken for the cost-effectiveness model by calculating the patient expected value of perfect information (EVPI) (Claxton 2001), population EVPI (Sculpher 2005) and the partial EVPI associated with model parameters (Brennan 2007). The output from the simulations was used to estimate the EVPI. The EVPI for an individual patient was calculated as the difference between the expected value of the decision made with perfect information and the decision made on the basis of existing evidence (Sculpher 2005). The population EVPI was obtained using patient EVPI applied to an estimated annual incidence of overweight and obesity in Switzerland (Swiss Federal Statistical Office 2006). A discount rate of 3% was applied to the population size in the EVPI calculation. It was assumed that the information on interventions would be valuable for 10 years. We used a conservative assumption considering that in 10 years time, advancements in technology could influence the development of new interventions able to better tackle obesity and its complications. To determine which parameters have the greatest value of information and require further research, we looked on following parameters of the cost-effectiveness model: cardiovascular risks (body mass index, systolic blood pressure, total cholesterol, high density lipoprotein), utilities (overweight and obese patient preferences), cost of interventions (lifestyle intervention, standard care intervention), cost of complications (diabetes, hypertension, hypercholesterolemia, coronary heart disease, stroke). The partial EVPI calculated the value of information of the remaining parameters of the model if we assumed perfect information for the parameter of interest. The partial EVPI for a parameter or group of parameters was the difference between the expected value of the decision made with perfect information and the decision made on the basis of existing evidence (Sculpher 2005).

Results

Cost-effectiveness results

Table 2 presents the results of the cost-effectiveness analysis based on current evidence. The lifestyle intervention resulted in increased quality of life in overweight and obese subjects when results were undiscounted and discounted at 3% rate. In overweight and obese people, the difference in quality of life between lifestyle intervention and standard care ranged from 0.19 to 0.41 QALY (undiscounted) and from 0.16 to 0.37 QALY (discounted) per person per year gained over lifetime, depending on gender and age group. Compared with standard care intervention, the average incremental cost of lifestyle intervention was higher in overweight than in obese subjects, ranging from CHF 510 to CHF 704 in overweight female subjects and from CHF 402 to CHF 434 in overweight male subjects, when results were discounted. When results were undiscounted, the lifestyle intervention dominate standard care being less costly and more effective in obese female subjects aged 40 to 60 years and obese male subjects aged 30 to 60 years.

Table 2. Probabilistic cost-effectiveness results

		Female						Male					
	Age	Undiscounted			Discounted 3%			Undiscounted			Discounted 3%		
		Costs[1]	QALY[1]	ICER	Costs[1]	QALY[1]	ICER	Costs[1]	QALY[1]	ICER	Costs[1]	QALY[1]	ICER
Overweight	30	427	0.19	2266	704	0.16	4358	104	0.24	437	434	0.20	2189
	40	47	0.24	195	423	0.21	2037	164	0.26	626	438	0.22	1959
	50	89	0.30	298	388	0.26	1516	187	0.30	616	408	0.26	1547
	60	334	0.35	956	510	0.31	1660	263	0.37	707	402	0.32	1237
Obese	30	102	0.19	539	310	0.16	1922	-146	0.24	D[2]	105	0.21	508
	40	-151	0.25	D[2]	100	0.21	469	-128	0.29	D[2]	98	0.25	400
	50	-169	0.33	D[2]	55	0.28	195	-179	0.34	D[2]	40	0.30	133
	60	-43	0.38	D[2]	116	0.34	342	-124	0.41	D[2]	27	0.37	73

[1]Difference between lifestyle intervention and standard care; QALY, quality adjusted life years; ICER, incremental cost-effectiveness ratio; Costs are expressed in Swiss Francs (CHF) year 2006; [2]D, lifestyle intervention dominate standard care intervention being more effective and less costly

Figure 2 presents the probability that lifestyle intervention is cost-effective for different threshold values in overweight people. Lifestyle intervention had a higher probability of being cost-effective in male compared to female subjects from the same age group. If the decision-maker is willing to pay CHF 2000 for a unit of health gain, lifestyle intervention will have a probability of being cost-effective ranging from 48% to 91%, depending on gender and age group with the exception of overweight female and male subjects aged 30 years.

Figure 2. Cost-effectiveness acceptability curves of lifestyle intervention in overweight subjects (discounted 3%)

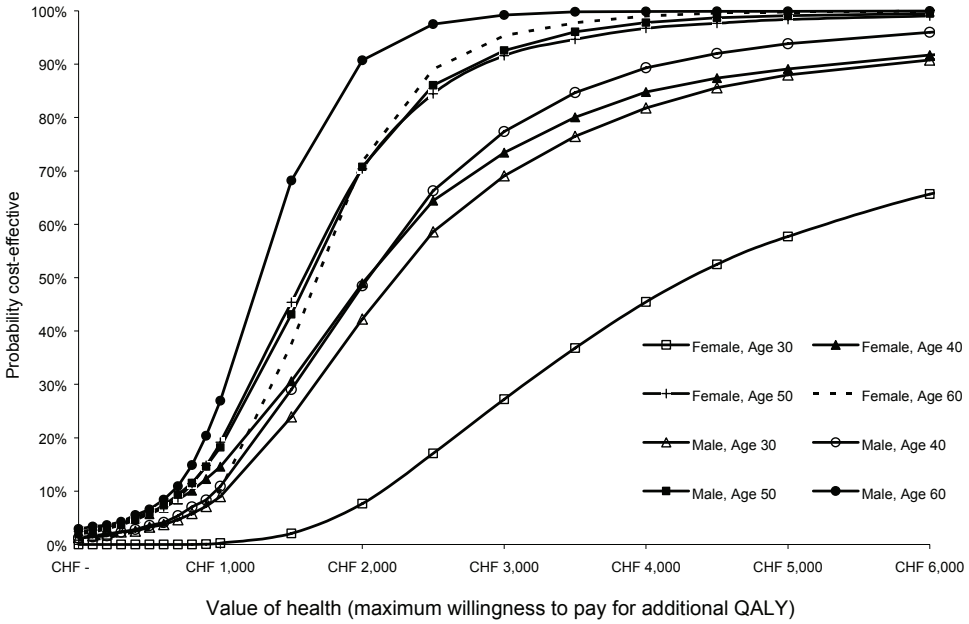
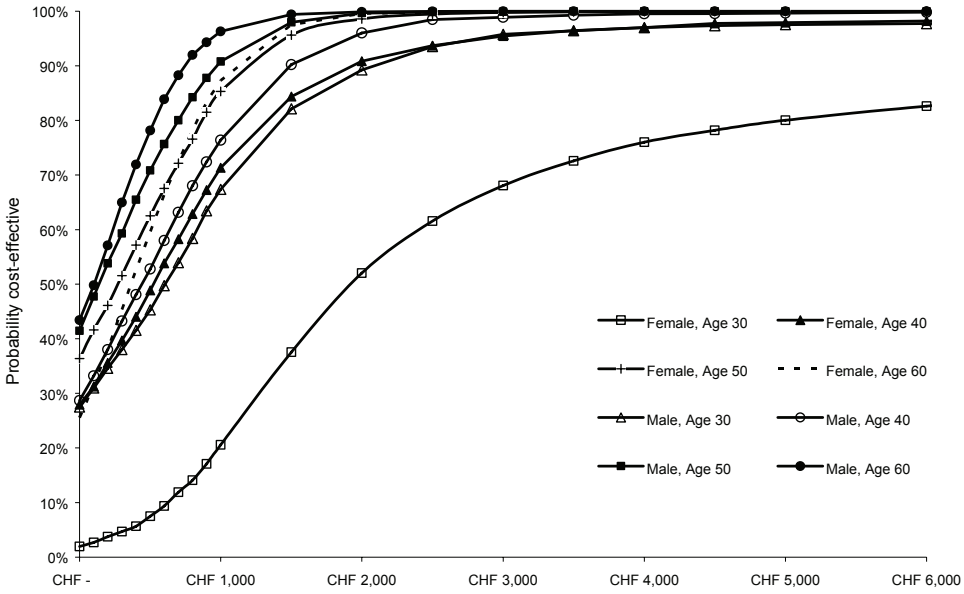


Figure 3 presents the probability that lifestyle intervention is cost-effective for different threshold values in moderate obese people. Lifestyle intervention had a higher probability of being cost-effective in moderate obese subjects compared to overweight subjects from the same age group. Within moderate obese group, lifestyle intervention had a higher probability of being cost-effective in male compared with female subjects from the same age group. If the decision-maker is willing to pay CHF 2000 for a unit of health gain, lifestyle intervention will have a probability of being cost-effective in moderate obese subjects ranging from 52% to 100% depending on gender and age group.

Figure 3. Cost-effectiveness acceptability curves of lifestyle intervention in obese subjects (discounted 3%)



Value of information results

Table 3 presents the patient EVPI for overweight and obese people. The patient EVPI reached a maximum at the incremental cost-effectiveness ratio i.e. when we are most uncertain about the decision based on current information. A higher uncertainty was observed in overweight female subjects aged 30 to 40 years and in moderate obese female subjects aged 30 years compared to other age groups.

Table 3. Patient expected value of perfect information (EVPI)

		Age	Undiscounted	Discounted 3%
			EVPI (CHF)	EVPI (CHF)
Overweight	Female	30	133	198
		40	101	103
		50	83	76
		60	68	48
	Male	30	87	83
		40	87	89
		50	79	67
		60	72	36
Obese	Female	30	80	100
		40	37	73
		50	32	78
		60	69	64
	Male	30	38	78
		40	40	63
		50	30	72
		60	42	77

CHF, Swiss Francs (year 2006)

Table 4 presents the populations EVPI for overweight and obese people at alternative thresholds values. When results were undiscounted, the population EVPI values ranged from zero to CHF 4.1 million in overweight people and from zero to CHF 2.2 million in moderate obese people, depending on age, gender and threshold value. When results were discounted at 3% rate, the population EVPI ranged from zero to CHF 6.8 million in overweight people and from zero to CHF 3.2 million in moderate obese people, depending on age, gender and threshold value.

Table 4. Populations expected value of perfect information

	Age		Yearly Incidence	Threshold value							
				CHF 0		CHF 1000		CHF 2000		CHF 5000	
				UD	D	UD	D	UD	D	UD	D
Overweight	Female	30	3992	151672	0	1036025	3763	4086515	349605	2964772	6785783
		40	4306	3215684	103257	1263160	775092	498117	3820791	202449	930752
		50	5206	2143027	177949	873539	1364801	110280	1708165	2254	113810
		60	2242	129103	9788	1345412	222124	86162	670135	0	5620
	Male	30	6391	2931215	83508	2497611	689783	900063	4239067	361361	2116957
		40	3513	1104473	20491	1724479	408452	465928	2721024	73454	351013
		50	5719	1254367	122617	1817459	1142732	204300	1833086	11229	37324
		60	662	84714	20710	178891	203992	6951	46084	0	8
Obese	Female	30	1825	679929	16727	911252	354991	538373	1618853	496053	995249
		40	1325	409488	409662	61116	427004	22427	144496	20462	65683
		50	2269	748764	1123705	38155	252393	4051	31075	0	1672
		60	3512	2168665	953722	58058	318340	2519	12670	0	0
	Male	30	2020	648960	701914	83085	711594	28475	231894	22479	87940
		40	2786	1201050	1014432	93891	786962	14872	152226	2165	26294
		50	3297	1414504	3199596	48073	318386	5563	22172	0	0
		60	990	259161	408086	2574	16539	263	83	0	0

Results are presented undiscounted (UD) and discounted (D) at 3% rate; CHF, Swiss Francs year 2006

A graphical representation of the population EVPI in overweight people is presented for female subjects (Figure 4), and male subjects (Figure 5). The population EVPI in overweight female subjects was higher than in overweight male subjects reflecting a higher uncertainty. The subgroup analysis in overweight people estimated a higher population EVPI in male and female subjects aged 30 years compared to other age groups.

Figure 4. Population expected value of perfect information (EVPI) in overweight female subjects aged 30 to 60 years (discounted 3%)

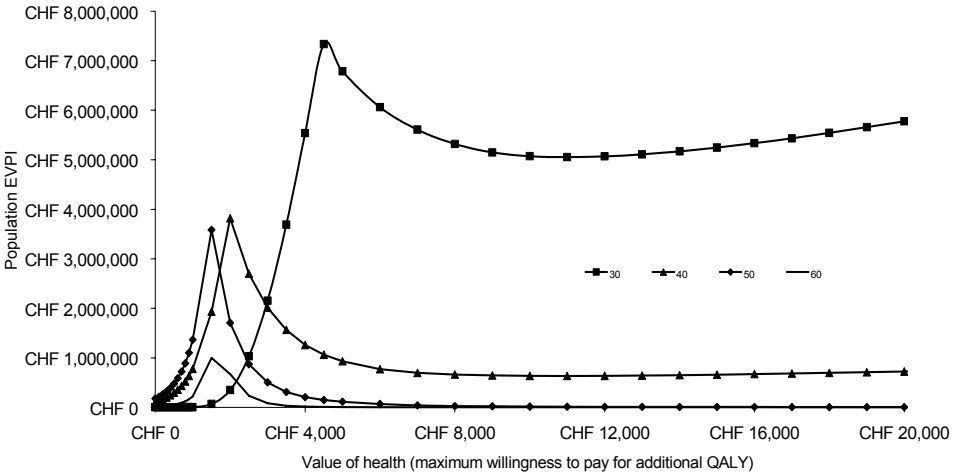
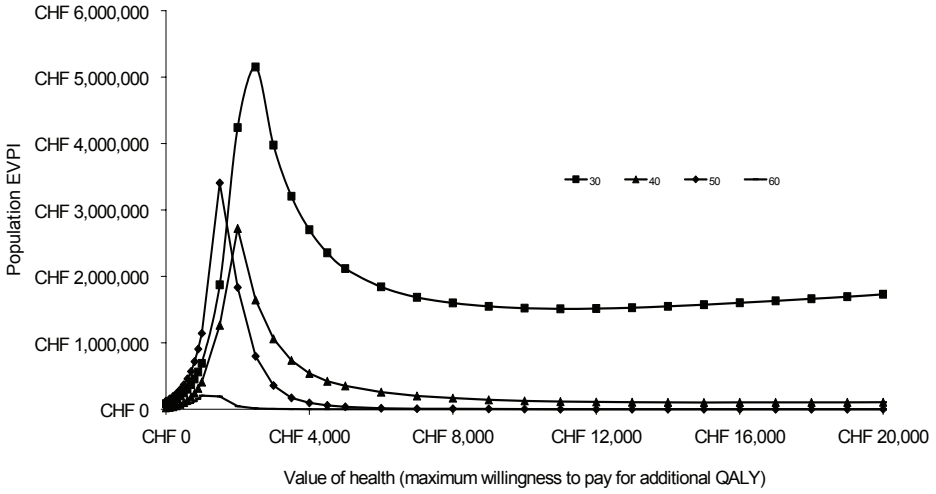


Figure 5. Population expected value of perfect information (EVPI) in overweight male subjects aged 30 to 60 years (discounted 3%)



We examined the partial EVPI for five groups of parameters (cardiovascular risks, utilities, cost of lifestyle intervention, cost of standard care intervention, cost of complications, all costs) and nine individual parameters (BMI, systolic blood pressure, total cholesterol, high density lipoprotein cholesterol, cost of hypertension, cost of diabetes, cost of hypercholesterolemia, cost of stroke, cost of coronary heart disease). Overall, partial EVPI associated with the group of parameters and the individual parameters was higher in overweight subjects than in moderate obese subjects depending on age and gender, when results were discounted. In overweight female subjects, the partial EVPI with the highest uncertainty was observed in the utilities of subjects aged 30 years (CHF 4.7 million, Figure 6), in the BMI of subjects aged 40 years (CHF 1.3 million), and in the cost of stroke of subjects aged 30 years (CHF 1.1 million). In overweight male subjects, the partial EVPI with the highest uncertainty was observed in the utilities of subjects aged 30 years (CHF 2.4 million) and in the BMI of subjects aged 50 years (CHF 1.5 million). In moderate obese female subjects, the partial EVPI with the highest uncertainty was observed in the utilities of subjects aged 30 years (CHF 1.3 million) and in the BMI of subjects aged 60 years (CHF 1.3 million, Figure 7). In moderate obese male subjects, the partial EVPI with the highest uncertainty was observed in the cardiovascular risk factors of subjects aged 50 years (CHF 3 million), in the BMI of subjects aged 50 years (2.3 million) and in the systolic blood pressure of subjects aged 50 years (CHF 1.7 million).

Figure 6. Partial expected value of perfect information (EVPI) in overweight female subjects aged 30 to 60 years (discounted 3%)

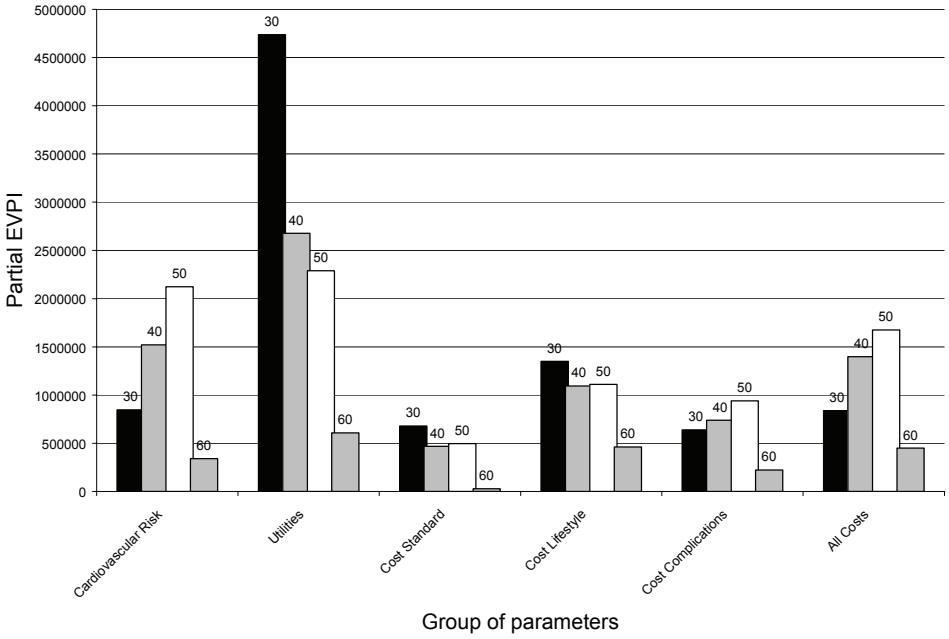
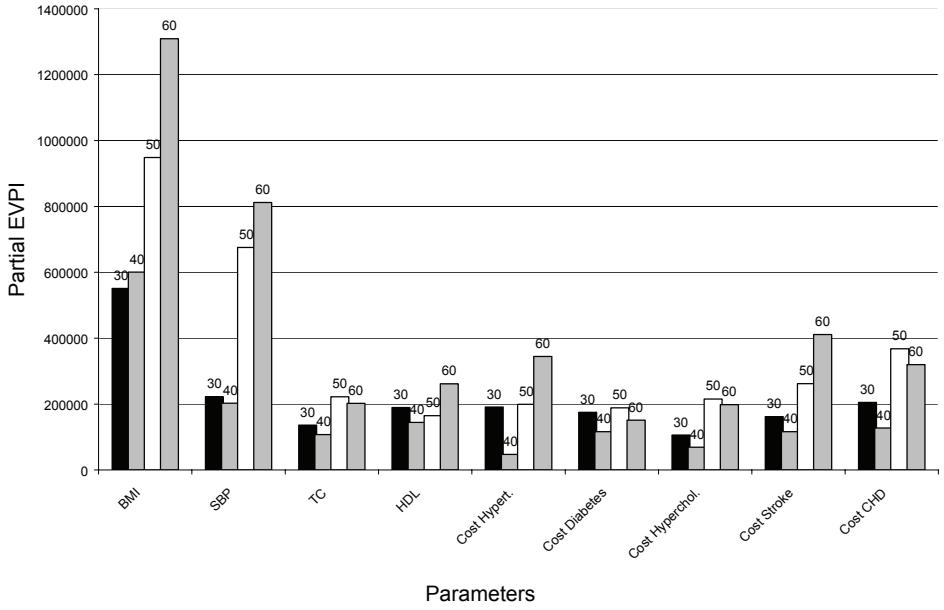


Figure 7. Partial expected value of perfect information (EVPI) in moderate obese female subjects aged 30 to 60 years (discounted 3%)



Discussion

Our study demonstrated the application of value of information analysis to quantify the uncertainty. We tried to address the following issues: to determine if the selection of lifestyle intervention is optimal based on the current information available, to assess if it is worth collecting additional information to inform this decision in the future and to determine which parameters in the obesity cost-effectiveness model have the greatest value of information and require further research.

The cost-effectiveness results demonstrated that, based on existing evidence, the lifestyle intervention can be regarded as cost-effective only in certain situations depending on gender, age group, and threshold value. When no discount was applied, the lifestyle intervention dominated standard care in moderate obese people being less costly and more effective. When a discount of 3% was applied, lifestyle intervention was cost-effective at an incremental cost per QALY ranging from 1237 to 4358 CHF/QALY in overweight people and from 73 to 1922 CHF/QALY in moderate obese people, depending on age and gender. A recent economic review of non-pharmacological weight loss treatments found that if weight loss, relative to the observed trend, remains constant for 5 years post-intervention before returning to baseline, the cost per QALY in the best performing non-pharmacological studies ranges from 429 to 24566 CHF/QALY (NICE 2007). Our cost-effectiveness model assumed that lifestyle intervention

effect on weight and cardiovascular risk factors lasts for 6 years based on the results obtained from meta-analysis of three years randomized clinical trials in overweight and obese people. This assumption is in line with the extended follow-up of the Finnish Diabetes Prevention Study which lasted 7 years and resulted in sustained lifestyle changes and a reduction of diabetes incidence, which was maintained after the individual lifestyle counseling stopped (Lindstrom 2006). The study reported a 43% reduction in the relative risk related to the success in achieving the intervention goals of weight loss, reduced intake of total and saturated fat, increased intake of dietary fibre and increased physical activity.

The value of information analysis places a limit on returns to further research. If the costs of the research exceed the EVPI, then the proposed research is not cost-effective (Claxton 2001). In our analysis, the maximum population EVPI was CHF 6.8 million in overweight subjects and CHF 3.2 million in moderate obese subjects. These values represent an upper limit on the costs associated with the decision uncertainty. Therefore, costs associated to proposed future research should not exceed this amount if the research would be considered cost-effective.

The partial EVPI with the highest uncertainty was observed in the utilities of overweight male and female subjects aged 30 years and moderate obese female subjects aged 30 years. One possible explanation could be that utilities used in the model were obtained from the published literature (Macran 2004). Further investigations are necessary to evaluate the overweight and obese patient's preferences on weight loss treatments in Switzerland. Further research is needed to calculate the expected benefits and the cost of sample information in order to determine the optimal study design such as follow-up time, sample size and patient allocation.

A higher partial EVPI was observed in the BMI of overweight female subjects aged 40 years, overweight male subjects aged 50 years; moderate obese female subjects aged 60 years and moderate obese male subjects aged 50 years. A higher uncertainty was observed also in moderate obese male subjects aged 50 in parameters such as cardiovascular risk factors and systolic blood pressure. Therefore, further research on the cardiovascular risk factors in overweight and obese people in Switzerland is recommended.

Based on the results of our study we identified four factors that influence the uncertainty and implicit the EVPI. The first factor was the level of the maximum willingness to pay for an additional unit of health gain with lifestyle intervention. The EVPI depended on the value of the maximum acceptable threshold. This was due to the interaction between the maximum acceptable threshold and the uncertainty surrounding the decision. When the maximum acceptable threshold was close to the incremental cost-effectiveness ratio then the uncertainty surrounding the decision was maximized. The second factor was the uncertainty surrounding the decision to adopt the lifestyle intervention. The uncertainty surrounding this decision was an important element in the calculation of EVPI, with the EVPI increasing with increased uncertainty. Therefore, when the uncertainty surrounding a decision was low, the EVPI was negligible, for example in obese male subjects aged 40 years the EVPI was CHF 63 when results were discounted. The third factor was the size of the eligible population. The size of the

population eligible for treatment had a direct impact on the population level estimates of the EVPI. Where the population was large, the scaled up population values were larger. For example, the population values of the EVPI were larger for overweight male subjects aged 50 years than for overweight male subjects aged 60 years due to an increase incidence of male overweight subjects aged 50 years. The fourth factor that influenced the uncertainty and EVPI in our model was represented by patient characteristics such as age, gender and BMI. Our analysis suggests that the allocation of funds between lifestyle intervention and standard care in the prevention and treatment of obesity and future research will depend crucially upon these four factors.

Our study results are limited by several factors related to the structure of the cost-effectiveness model and the EVPI analysis. The model could be improved by having access to additional Swiss-specific data. So far, epidemiological data such as the correlation between BMI and the risk of complications, obesity related mortality data and changes in patient utility have not been recorded specifically for Switzerland. Further investigations should also take into account other important complications of obesity such as metabolic syndrome, colorectal cancer, gall bladder disease, sleep apnea and depression. Another limitation of our study consists in the estimation of the costs of obesity complications from secondary data sources.

Both probabilistic sensitivity analysis and parameter EVPI were computed using Monte Carlo methods. Unfortunately, to obtain these measures accurately, very large numbers of model evaluations are needed, potentially millions. For computationally expensive computer models, evaluating these measures may then require lengthy computing times. Our model used for the probabilistic sensitivity analysis a number of 3000 simulations and for the partial EVPI a number of 50 simulations. To assess the validity of the study results, we performed additional analyses for different age groups in male and female using a number of 5000 simulations for the probabilistic sensitivity analysis and 250 for the partial EVPI. The new results of the parameter EVPI showed the existence of a slight over-estimating bias in using small numbers of Monte Carlo samples. However, the overall trend of the parameters presented in the results section was maintained. Further theoretical investigation of Monte Carlo bias in the context of parameter EVPI would be useful.

Conclusion

In summary, we applied the value of information analysis to evaluate the uncertainty in the cost-effectiveness of lifestyle interventions in overweight and obese people in Switzerland. The value of information analysis indicates that there is some uncertainty regarding the choice between lifestyle intervention and standard care intervention. The extent of the uncertainty depends on the maximum acceptable threshold, the uncertainty surrounding the decision to adopt the lifestyle intervention, the size of the eligible population and patient characteristics. The parameter EVPI suggests that if further research is commissioned, this should focus on the effectiveness of lifestyle intervention on the cardiovascular risk factors and quality of life of the overweight and obese people in Switzerland.

Acknowledgement
This study was funded by the Swiss Federal Office of Health.

Chapter 7

Health Policy Implications: Decision-Makers' Attitude Towards Economic Evaluations of Medical Technologies

Summary

Increasing costs have generated concern among governments and healthcare providers who have realized the need for cost containment measures and more efficient resource utilization. Health economics is one potential source of information that can make healthcare more efficient. We conducted a literature review to summarize published literature on self-reported attitudes of healthcare decision-makers towards economic evaluations of medical technologies and to determine the extent to which economic evaluations are used in health policy decisions. Fifty-five articles investigated the use of economic evaluations on three levels of decision-making: central, local and physician level. Results indicate the use of economic evaluation information increased from limited/minor to moderate use. The influence of economic evaluations increased with the level of centralization of healthcare system. Barriers to use health economics research varied across levels and included health economics research-related barriers such as timely availability, lack of credibility, insufficient methodological quality and decision-context-related barriers including limited decision-makers' knowledge, inflexibility in healthcare budgets and variability among healthcare organizations. For consistent policy-making it is important that similar recommendations for cost-effective interventions and programs are developed at all levels and that implementation is promoted by incorporating the appropriate incentives in healthcare provision.

Introduction

Healthcare costs have increased dramatically in recent decades in all industrialized countries. Increasing costs have generated concern among governments and healthcare providers who have realized the need for cost containment measures and more efficient resource utilization. Health economics is one potential source of information that can make healthcare more efficient. Health economics has been described in various ways, but most commonly as a set of analytical techniques to assist healthcare decision-making to promote efficiency and equity (Shiell 2002). Publications in economics journals in the area of health economics increased almost 350% from 1990 to 2000 (Rubin 2003). Economic evaluation of medical technologies, an integral part of health economics, is used to help decision-makers when addressing problems arising due to the scarcity issue. Economic evaluations can play an important role in different types of decision-making, such as defining the basic health benefit package (Schreyögg 2005), setting the price of a new technology (García-Alonso 2008), reimbursement decisions (Cohen 2007), formulary decisions (Weart 2007) and individual patient care (Chauhan 2008). The perspectives, techniques, decision constraints and available information differ across countries and often across decision–makers.

The aim of the economic evaluations of healthcare programs is to serve as an aid to decisions that affect policy-making. If economic evaluations of healthcare programs were to have no impact on decision about allocation of resources to healthcare programs, carrying out such evaluations is a meaningless activity. Thus, it is crucial to explore the link between economic evaluations of medical technologies and their use in policy decision-making. We conducted a literature review to summarize and synthesize published literature on self-reported attitudes of healthcare decision-makers towards economic evaluations of medical technologies. The aims of this literature review was to determine the extent to which economic evaluations are used in health policy decision-making, and to consider factors associated with the utilization of such research findings.

Methods

A systematic literature review of published English language studies was conducted using MEDLINE, EMBASE and HEED from January 1995 to December 2007. We used a full set of Medical Subject Headings terms in the literature search. The following terms were used individually or in association: ‘health economics’, ‘survey’ ‘decision-making’, ‘health policy’, ‘research utilization’, ‘priority setting’ ‘healthcare rationing’, ‘drug cost’, ‘medical technology’, ‘formulary inclusion’ ‘pharmacy administration’, ‘pharmacies and therapeutics committees’. Cross-references were checked.

This review was restricted to empirical research i.e. surveys investigating the attitudes toward economic evaluations among decision-makers and actual use pattern. In order to focus on the analyses of actual decision-making process and to gain the best insight into the decision-making process as a whole, we included only qualitative studies.

The term medical technology refers to pharmaceuticals, medical devices and health interventions such as surgical procedures. Economic evaluation studies were meant to include all formal studies comparing the costs and consequences of relevant alternatives using indicators like cost-effectiveness, cost-utility, cost-minimization and cost-benefit.

Health economics evaluations may be used by a wide range of health care decision-makers. Politicians, public officials, insurers, health service administrators, and clinicians are among the most important users of such data. When considering the relevance of health economics studies, we assessed the impact on three levels of decision-makers: central, local and physician level. The central level includes decision-makers with a national or regional healthcare perspective responsible for the availability of effective and affordable healthcare programs for the whole population. This includes decisions about inclusion of drugs on national formularies, national guidance on the use of medical technologies, or development of national policies or programs (e.g. screening, immunization). The local level includes decisions on the development of local programs, the purchase of equipment, and the inclusion of drugs in hospital, practice, or health plan formularies. The physician level covers decisions of individual health care professionals at the patient level. A reviewer abstracted the relevant study characteristics using a standardized template. Results are presented as percentage (%) of characteristics and reported issues from evaluated studies. The estimated relative importance of the decision-making criteria in the studies is based on our interpretation of the study results.

Results

Included studies and population

Using combined search strategies, we identified 3150 abstracts, of which 2757 were excluded on the basis of titles. The remaining 393 abstracts were reviewed. Of these, 195 did not include empirical research, 94 were excluded based on the study population, 13 were systematic reviews, 36 were health policy discussions and the remaining 55 articles were included in our review. Of 55 studies evaluated, 45 studies evaluated only one level of healthcare decision-maker, 6 studies evaluated two levels of decision-makers and 4 studies evaluated three levels of decision-makers. Table 1 provides a summary of selected studies. Of studies evaluated, 21 studies investigated healthcare decision-making at central level, 33 at local level and 15 at physician level.

Table 1. Selected studies by level of decision-making

Central level			Local level			Physician level		
Studies*	N*	RR* (%)	Studies*	N*	RR* (%)	Studies*	N*	RR* (%)
Anell 2000	210	69	Cox 2000	16	100	Buusman 2007	15	63
Bloom 2004	104	90	DeRoeck 2004	91	100	Duthie 1999	11	100
Bryan 2007	28	100	Drummond 1997	4	58	Erkan 2002	375	38
DeRoeck 2004	165	100	Duthie 1999	11	100	Fattore 2006	183	47
Drummond 1997	66	66	Evans 2000	41	100	Ginsburg 2000	512	52
Duthie 1999	12	100	Fattore 2006	191	47	Hasle-Pham 2005	44	44
Hoffman 2000	347	65	Fijn 1999	38	70	Hoffman 2000	347	65
Hoffman 2002	12	100	Grabowski 1997	5	100	Jacoby 2003	56	43
Iglesias 2005	NR	-	Greenberg 2005	61	46	Jansson 2006	738	49
Luce 1995	24	100	Grizzle 2000	31	100	Kangis 1996	60	100
Martin 2001	11	73	Hasle-Pham 2005	94	44	Prosser 2005	15	100
Nuijten 2003	NR	-	Hoffman 2000	347	65	Schumock 2004	75	100
PausJenssen 2003	7	100	Jenkings 2004	6	100	Ubel 2003	560	65
Prosser 2005	15	100	Luce 1995	24	100	Wu 2004	24	44
Ross 1995	34	100	Luce 1996	51	82	Zwart-van Rijkom 2000	11	100
Singer 2000	32	100	Lyles 1997	51	100			
Vuorenkoski 2003	18	100	Martin 2003	17	100			
Weatherly 2002	68	67	Marx 2007	42	34			

Central level			Local level			Physician level		
Studies*	N*	RR* (%)	Studies*	N*	RR* (%)	Studies*	N*	RR* (%)
Williams 2007a	30	100	McDonald 2001	21	81			
West 2002	NR	-	Motheral 2000	241	8			
Zwart-van Rijkom 2000	12	100	Nichol 2007	20	10			
			Odedina 2002	212	95			
			Prosser 2005	15	100			
			Schumock 2004	75	100			
			Shalansky 2003	164	82			
			Sloan 1997	103	65			
			Spath 2003	NR	-			
			Steiner 1996	231	41			
			Tan 2007	84	61			
			Tordoff 2006	24	83			
			Walley 1997	178	65			
			Williams 2007b	31	100			
			Zwart-van Rijkom 2000	11	100			

*Studies, first author, publication year and reference; N, number of responders; RR, response rate; NR, not reported

The number of empirical research investigating the use of economic evaluations in healthcare decision-making process increased substantially in the last 13 years in response to the growing number of health economics literature. Figure 1 presents the cumulative number of studies evaluating the decision-maker's attitude towards health economics research.

Figure 1. Cumulative number of studies evaluating the use of economic evaluations in decision-making process.

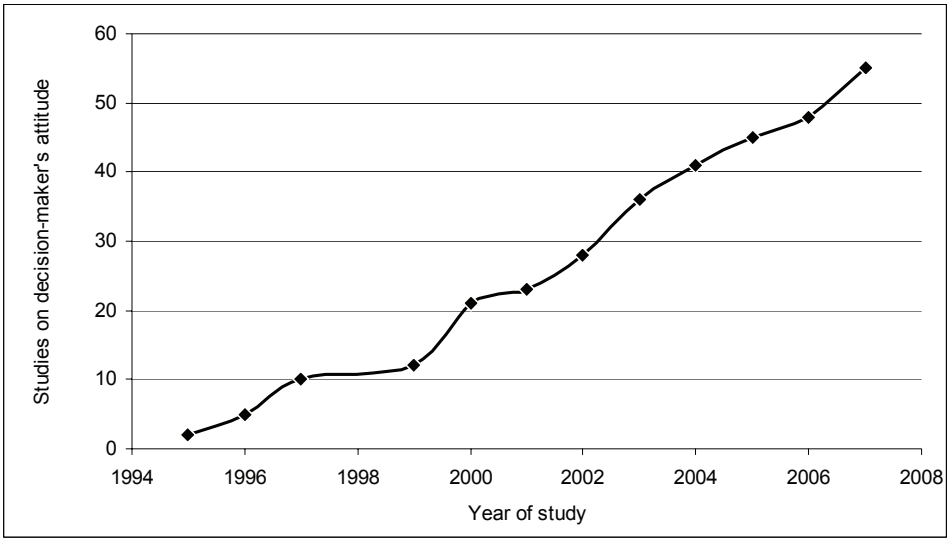


Table 2 presents the study research population according to the level of decision-maker. The study population includes the decision-makers interviewed in the selected studies. As expected there is an increasing number of responders with decreasing level of decision-maker i.e. the highest number of responders is on physician level followed by the local level and the central level.

Table 2. Research population by level of decision-makers

Central level			Local level			Physician level		
Research population	NS*	N*	Research population	NS*	N*	Research population	NS*	N*
Health authorities	13	749	Healthcare managers	4	660	Physicians (not specified)	9	2530
Central formulary committee members	2	217	Pharmacy and therapeutics committee	8	599	GPs*	5	121
Third-party payers	2	128	Medical and pharmaceutical advisers	4	392	Specialists	1	375
National decision-makers (e.g. NICE)	2	58	Pharmacy decision makers	8	378			
Policy advisers	2	43	Drug formulary decision- making	3	296			
			Health plans	3	296			
			Hospital directors	2	164			
			Managed care organization	3	122			

*NS, number of studies; N, number of responders; GPs, general practitioners

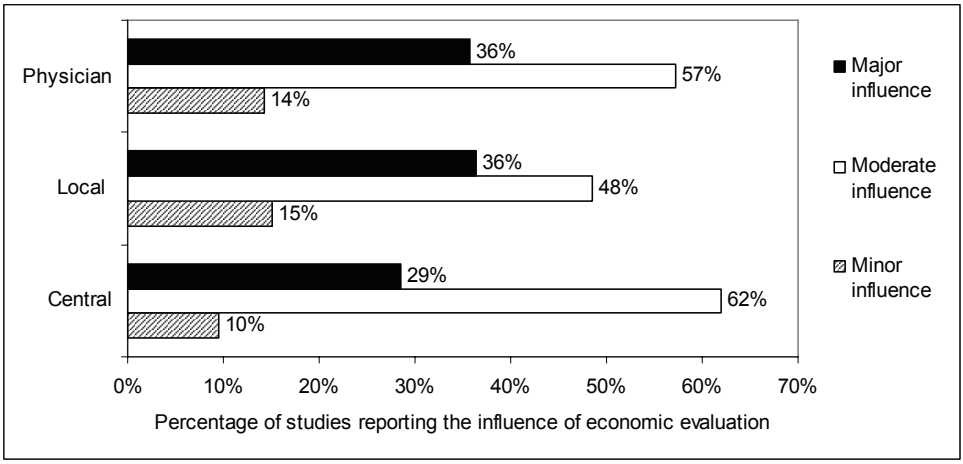
The most investigated countries were United States (US) 18 studies, United Kingdom (UK) 17 studies, Canada 6 studies and the Netherlands 4 studies. For most countries, for example Germany, France, Italy and Spain, they were only one or two empirical studies. The most applied research method was face to face interviews (42%) followed by mail questionnaire (40%), telephone interview (17%) and web questionnaire (2%). Some of the studies combined different methods, for example mail survey and interviews for investigating committees. The majority of studies concentrated on pharmaceutical policy (60%), followed by policy regarding other health care interventions (17%) or both.

Influence of economic evaluations on decision-making process

The type of economic analysis used was mostly cost-effectiveness analysis (51%), followed with considerable distance by cost-minimization (16%), cost-benefit analysis (13%) and cost-utility analysis (5%). Fifteen percent of the studies did not specify the type of analysis used by healthcare decision-makers. Cost-effectiveness analysis was performed mostly on central level, whereas cost-minimization analysis was used mostly at local and physician level where the price of the medical technology was sometime rated as one of the most important aspects. The sources of economic evaluations used by decision-makers were published literature 25%, evaluations performed by government organizations (e.g. National Institute for Clinical Excellence NICE) 24%, evaluations performed by the industry 22%, expert opinion 12%, clinical guidelines 9% and in-house data 7%.

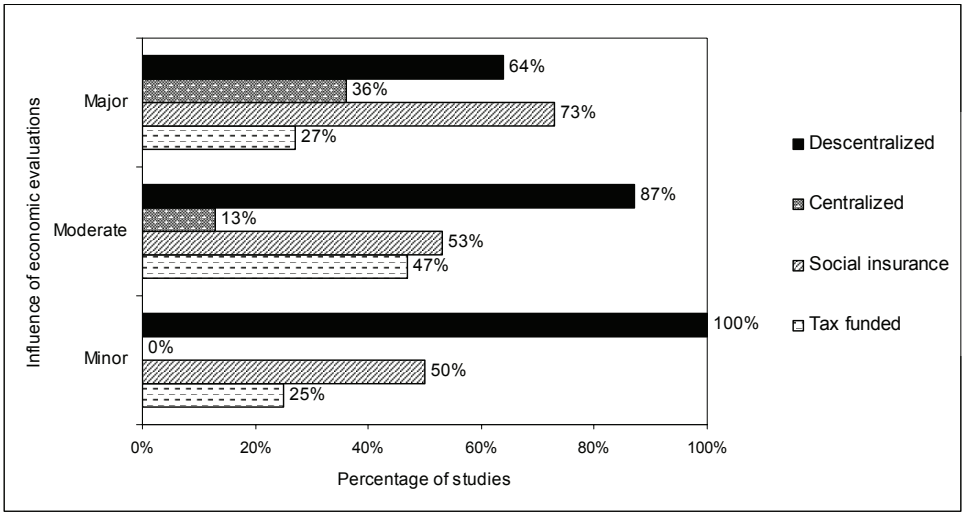
Figure 2 presents the number of studies reporting the influence of economic evaluations on healthcare decision-makers. Overall, the participants interviewed considered economic evaluations useful to inform policy decisions. The impact on policy is reported moderate in majority of the studies. However, an increasing trend in using the economic evaluation information is observed on all levels of decision-making: 36% on central and local level and 29% on physician level. In addition to economic evaluation arguments other factors appear to influence healthcare decision-making as well. Clinical aspects such as efficacy and effectiveness, and safety data were still considered the most influential arguments. Different aspects characterized each decision level. At central level there was substantial influence of regulatory and political arguments. At local level, there was influenced by economics, especially the drug acquisition cost. The physician level was dominated by the patient, disease, and the administrative burden.

Figure 2. Studies reporting the influence of economic evaluations on health care decision-making



The way in which health economics research evidence translates into policy and practice tends to differ across levels of decision-makers. We investigated the extent to which information on economic evaluations of medical technologies is used by national and regional authorities according to the type of healthcare system (tax funded or social insurance) and the degree of regionalization (centralized or decentralized healthcare system). From more than 24 countries investigated, 54% have a social security system and 46% have a tax funded healthcare system. Of the evaluated countries 71% have a decentralized health care system. Figure 3 presents the perceived influence of economic evaluations of medical technologies on decision-making process according to different structure of the healthcare system and funding mechanism. We observed that the influence of economic evaluations increased with the level of centralization of healthcare system. In other words, the higher the level of decentralization the lower the level of impact of economic evaluations is observed.

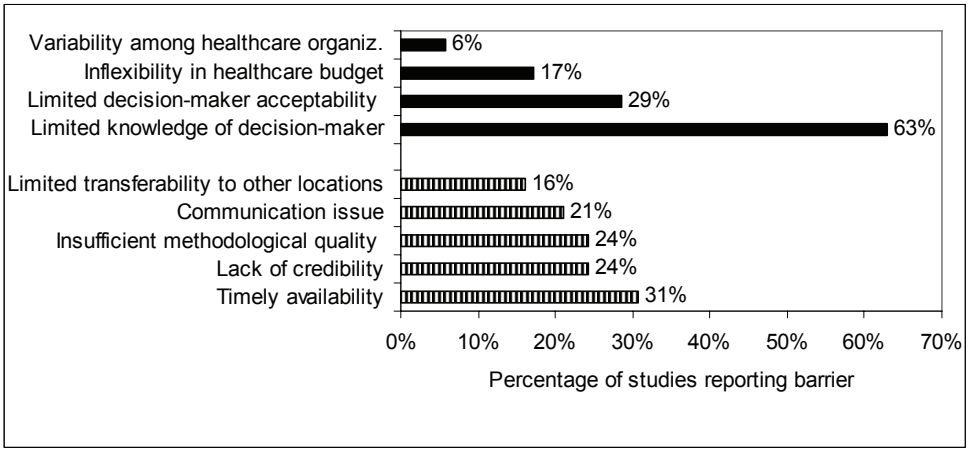
Figure 3. Influence of economic evaluations by healthcare system and funding



Barriers and incentives to use economic research in decision-making

The majority of the studies reported barriers to the use of health economics data. There is a certain pattern of answers which can be found in all studies that reported barriers. Two major types of barriers have been identified: health economics research-related barriers and decision-context-related barriers. Figure 4 presents the barriers in using the economic research and their perceived influence as reported by healthcare decision-makers. Health economics research-related barriers included timely availability (the data is usually not available when decisions have to be made), lack of credibility (studies are often sponsored by pharmaceutical industry), insufficient methodological quality (studies used too many questionable assumptions and poor quality data), communication issue (lack of transparency in reporting, limited peer-reviewed publications) and limited transferability of economic evaluations to other locations. The most important decision-related-context barriers included limited knowledge of decision-makers in health economics research field, limited decision-makers acceptability of economic evidence and inflexibility in healthcare budgets i.e. difficulty in moving resources from one budget to another.

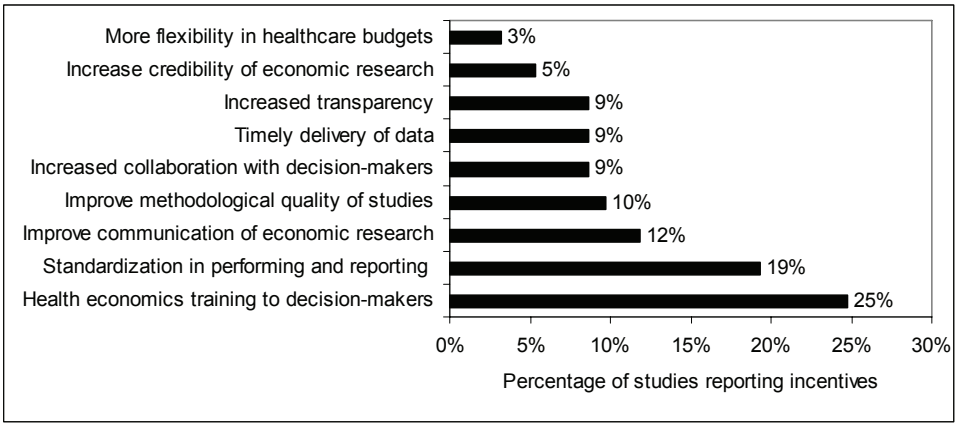
Figure 4. Barriers to use economic evaluations in decision-making process



Another barrier to the acceptance of the economic evaluations results is the so called 'silo mentality'. Health care accounting system cannot recognize how increased expenditure in one area reduces spending in another area. For example, decision-makers who have the responsibility for the hospital budget may be reluctant to approve the use of a costly new medical technology, even if it reduces overall costs to the payer, the patient or even society as a whole.

Several studies suggested ways to overcome barriers in using economic research in policy decisions. Figure 5 presents the reported incentives to use economic evaluations data. The majority of studies suggested that considerable effort must be directed at educating decision-makers at all levels about the techniques of health economics and the relevance of this information to their organization or practice. This in turn would contribute to a wider use of health economics research. The integration of cost-effectiveness information in clinical guidelines was seen as an important step to the wider application of research economics research findings. The EUROMET survey (Hoffman 2000), which investigated the impact of health economics studies on decision-making in nine European countries, suggested that the biggest barriers to use health economics findings were institutional barriers, including the difficulties in transferring budgets. In studies that investigated the impact of economic research at physician level some specific barriers were identified: difficulties in substituting one drug for another, fear of creating a credibility gap in the doctor patient relationship, difficulties in denying patients drugs they ask for and direct-to-consumer advertising about drugs and treatment in the US.

Figure 5. Incentives to use economic evaluations in decision-making process



Discussion

Demonstration of meaningful clinical and economic value of medical technologies is becoming essential when negotiating access to funding. Examples of the factors contributing to the pressures faced when making resource allocation decisions include an ageing population, greater pressure on budgets, a rising incidence of disease, and changes in regulatory requirements. Irrespective of the factors contributing to attitude changes, the influence of economic evaluations information in healthcare decision-making remains critical. Even though the evaluation procedures will varies between countries, a more formalized approach to health technology assessment is becoming norm in most countries (Jönsson 2008). The present study may thus be seen as further contributing to understanding of the 'reasons behind allocation of healthcare resources'. In addition and sometime in contrast to previous reviews on decision-makers' attitudes (Kernick 2000, Drummond 2003, van Velden 2005), our research suggests the use of economic evaluation information in healthcare decision-makers increased from limited/minor to moderate use (Figure 2). Moreover, an increasing trend in using economic research in health policy decision-making is observed at all level of decision-makers. This is not surprising giving the increased interest of government agencies and local decision-makers to determine value-for-money of medical technologies around the world.

We observed that differences in healthcare systems/funding structures across countries can lead to variations in decision-maker's reaction to economic evaluations information and requirements to support its dissemination. It is not surprising that the level of economic influence increased with the level of centralization of health care system. For example, in the tax funded health care system such as UK, the NICE incorporates health economics information systematically in its technology appraisals and disseminates these by means of leaflets, monographs, databases and web sites to all purchasers and providers in the UK. Recent research in the UK confirms that economic analysis is highly integrated into NICE's technology appraisal process (Buxton 2006). Opposed to

this is the decentralized nature of the organization and funding of the healthcare system such as Sweden, Italy, and Spain. Regionalized councils and municipalities, where the health economics expertise is scarcer, are responsible for many aspects of health and social care, and sometimes regions even compete with each other in the uptake of new medical technologies without consideration of the economic evidence.

At the central level, some countries have gone beyond merely developing academic guidelines for health economics evaluations. They now use such data as an element in reimbursement decisions and/or clinical recommendations i.e. Australia, Canada, UK. Whereas socialized medicine prevails in Australia, Canada and Europe, the private system dominates the US healthcare system. Notably, 30% of the surveys evaluated were conducted in the US. Federal and state governments and their agencies are minimally involved in regulating medical technologies expenditures in the US, therefore economic evaluations information have limited influence on decision-making process. Instead, local decision-makers such as health maintenance organizations, pharmacy benefit management companies are increasingly interested in the potential of economic evaluations especially when making reimbursement decisions. Recently, a panel composed of medical and pharmacy directors of public and private health plans developed a strategic plan for incorporating cost-effectiveness analysis into US health policy decisions (Neumann 2008). The plan has three long-term goals: increasing use of cost-effectiveness analysis by the Centers for Medicare and Medicaid Services, creating infrastructure to support research and integrating cost-effectiveness analysis into other public and private initiatives. Whether this plan will successfully incorporate economic analysis in decision-making process at central and local level in US needs to be seen. However, such initiatives should be encouraged.

At the physician level, clinicians recognize that the resources to treat all patients to the best of their abilities cannot be made available, and, therefore, there is a need in clinical decision-making to consider the costs and well as the effectiveness of treatments. This is illustrated by the increasing number of appraisals conducted by clinicians appearing in medical journals and by the evidence that general practitioners believe cost should be considered when choosing treatments for patients (Ryan 1990). There are some possible forms of action which may be taken to try to make clinicians consider the cost-effectiveness of the various alternatives available some of them already in place in some countries. Firstly, it may be possible to provide regulations forcing clinicians to consider costs as well the benefit of treatment decisions. Alternatively, it may be possible to provide guidelines to enforce the effectiveness and cost-effectiveness of the existing healthcare technologies. The degree to which such incentives and regulations are introduced may determine the level of influence of economic evaluations. Practice guidelines can influence the individual treatment decisions of clinicians and local healthcare providers, but economics based recommendations still constitute a minority in these practice guidelines in almost all countries evaluated. Differences among countries can be noted in this respect as well. In the UK, NICE provides national guidance also through clinical guidelines, which are often also based on cost-effectiveness information. In

France, priorities for introducing clinical practice guidelines are mostly determined by the practitioners' need (bottom-up demand), although policy-makers may also identify issues for which they would like to promote clinical guidelines (Orvain 2004). In the Netherlands, a special program was funded to develop 31 practice guidelines incorporating cost-effectiveness information (Niessen 2007), but after termination of this program there was no more systematic attention for economic arguments in guideline development.

Our study has several limitations. The heterogeneous mix of the studies on different decision-makers and across different geographical regions may have limited the generalizability of the study results due to complexity of the environment in many countries. Some of the surveys included in the review are quite dated at this point and may not be very relevant anymore given all of the changes in the healthcare environment since then. However, to have a better understanding of the impact of economic studies on healthcare decision-makers, those studies have been included. We may have not identified all relevant literature on this topic given the large implications in the international healthcare policy; however, an attempt to approach this complex issue was made by systematically assessing the literature.

Self-reported attitudes of decision-makers towards economic research can be a useful and relatively rapid means of gaining the perspective and insights of key stakeholders who make or influence policy decisions. There are however, a number of limitations to such surveys. These include the difficulty in ensuring that all key decision-makers are included and the difficulty in obtaining sufficient feedback on the findings and conclusions from responders for validation. Like all surveys, they also capture a moment in time. This especially can be limiting with policy-maker surveys, since changes in governments can result in changes in key decision-makers and in program priorities. Given these limitations, policy-makers surveys are most appropriate as means of initially identifying key issues such as factors influencing the use of economic evaluations, barriers and incentives to its use.

In conclusion, we observed different ways of incorporating economic evidence at the central, local and physician level. For consistent policy-making it is important that similar recommendations for cost-effective interventions and programs are developed at all levels and that implementation is promoted by incorporating the appropriate incentives in health care provision, in educating healthcare providers to understand the economics based guidance and to enhance the basis for implementation by involving them in developing the guidance. Our results indicate a trend in higher impact of economic arguments in all levels of decision-making. May be the best way forward is to focus on the cooperation between different disciplines in providing the best scientific evidence and produce guidance which is balanced in terms of integrating different societal values and is therefore widely accepted. In state run systems like the UK this has been achieved at the central level by NICE. In social insurance and more decentralized systems, however, the development and implementation of practice guidelines which are set up by multidisciplinary teams as well and systematically include cost-effectiveness information may proof the best way of overcoming the gap between the availability and use of health economics evidence.

Acknowledgement
This study did not receive funding.

Chapter 8

Discussion

This chapter discusses the main conclusions of this thesis. All separate chapters ended with detailed conclusions and discussions of the results. We will not repeat these here. This final chapter brings together the various conclusions along the lines of the challenges presented in the introduction. In addition, this chapter will identify and discuss areas for further research.

Methodological and practical challenges of health technology assessments

The use of meta-analysis in systematic literature review

Systematic review and meta-analytical methods are already common approaches to the assessment of health technology and increasing adoption of such approaches may be foreseen in response to increasingly wide emphasis on evidence-based approaches to medicine and healthcare. Chapter 2 presented a systematic literature review with meta-analysis to assess the long-term effectiveness of lifestyle interventions in the prevention and treatment of obesity. Despite its widespread use, meta-analysis continues to be a controversial issue. An important consideration is the possibility of heterogeneity between study

outcome estimates. It has been argued that producing an overall combined estimate for heterogeneous studies is wrong and leads to a result which is misleading, and impossible to interpret. A much used quote is that it is equivalent to: 'combining apples and oranges and the occasional lemon' (Rhee 2007). However, there are no clear guidelines outlining how variable study results have to be before it is deemed invalid to combine them. In chapter 2 we addressed the issue of heterogeneity by taking into consideration the following factors: we determined the study question of the evaluated studies, evaluated the similarities or dissimilarities in the study design, and looked if the heterogeneity of the outcomes can be explained. However, no consensus has been reached concerning the best strategy for dealing with heterogeneity; currently a large degree of subjectivity is required on the part of the reviewer. In chapter 2 we investigated the heterogeneity using the following techniques: funnel plot for publication bias, random effect model, subgroup analysis, sensitivity analysis, and quality assessment scores.

Methodologies for critical appraisal of the research evidence, including ways of assessing the quality of the primary studies have been developed. These encompass the relatively simple fixed effect approaches, through random effects models, to more sophisticated hierarchical modeling (Branscum 2007). The more complex methods were largely devised to deal with heterogeneous outcomes, systematic variation between studies, and the need to incorporate a fuller set of components of variability into the model. Several of these methods have come under criticism. The fixed effect model assumes no heterogeneity between the study results i.e. the studies are all estimating one single true value underlying all the study results (Hasselblad 1995). Hence, all observed variation in the treatment effects between the studies is considered due to sampling error alone. The random effects method (DerSimonian 1986) incorporates an assumption that the different studies are estimating different, yet related, treatment effects. The random effects method is based on the inverse variance approach, making an adjustment to the study weights according to the extent of variation, or heterogeneity, among the varying treatment effects. Mengersen et al (2005) carried out a meta-analysis of the effect of passive smoking on lung cancer, and investigated the difference in results when using different methods. They state that different conclusions may have been drawn if only fixed or random effect methods had been used.

Some authors (Covey 2007) investigated the effects of presenting treatment benefits in different formats on the decisions of both patients and health professionals. Three formats were investigated: relative risk reductions, absolute risk reductions, and number needed to treat or screen. The meta-analysis showed that treatments were evaluated more favorably when the relative risk format was used rather than the absolute risk or number needed to treat format. In chapter 2 we measured the outcome as the 'difference in means' that is a standard statistic that measures the absolute difference between the mean values in the two groups in a clinical trial. It estimates the amount by which the treatment changes the outcome on average. It can be used as a summary statistic in meta-analysis when outcome measurements in all trials are made on the same scale.

The robustness of our meta-analysis was investigated through sensitivity analysis by incorporating the quality of the evaluated studies. Our sensitivity analysis did not change the results of the main analysis; therefore it strengthens the confidence that can be placed in these results. However, there are several practical issues to consider when assessing the quality of studies. The first issue is whether to blind the assessors to aspects of the randomized controlled trials. Previous research (Jadad 1996) investigated the effects of blinding, and found evidence to suggest that blinded assessment produced significantly lower and more consistent scores than open assessment. However, lifestyle interventions cannot be blinded in randomized controlled trials. To address this issue in chapter 2, we made an adjustment in the 5-point Jadad score to use an appropriate quality check for open assessment in randomized controlled trials of lifestyle intervention. Another issue is that some large and complex trials report the details of study methodology in separate earlier publications. Some authors (Detsky 1992) argue that looking at this material would probably increase quality score of the trial above the score it would achieve when considering it in isolation.

It has long been accepted that research with statistically significant results is more likely to be submitted and published than work with null or non-significant results (Easterbrook 1991), which leads to a preponderance of false-positive results in the literature (Begg 1989). The implications of this for meta-analysis are that, even if all published studies have been identified, these may be only a subset of the studies actually carried out. Since positive results are more likely to be published than negative ones, combining only the published studies (Begg 1989) uncritically may lead to an over optimistic conclusion. Publication bias has long been recognized as a problem in this regard since it means that the likelihood of finding studies is related to the results of those studies (Dickersin 1992a). One way to investigate whether a review is subject to publication bias is to prepare a ‘funnel plot’ and examine this for signs of asymmetry. Effect estimates from small studies will therefore scatter more widely at the bottom of the graph, with the spread narrowing among larger studies. In chapter 2 we investigated the effect of publication bias by plotting the outcome measure, in this case body weight, against the standard error. Here the treatment effect did not lead to funnel plot asymmetry. However, it is worth mentioning that although funnel plots may alert reviewers to a problem which needs considering, they do not provide a solution to this problem. The only satisfactory way to address publication bias and the inadequate quality of individual trials is through prospective registration of trials (Dickersin 1992b) and improvements in the quality of the conduct, analysis and reporting of studies, meta-analyses and systematic reviews (Begg 1996, Moher 1998).

However, despite these drawbacks, meta-analysis has several advantages. If it is well conducted, meta-analysis allows for a more objective appraisal of the evidence, which may lead to resolution of uncertainty and disagreement. Meta-analysis may reduce the probability of false negative results and thus prevent undue delays in the introduction of effective treatments into clinical practice. Meta-analysis of a large number of individual studies or of individual patient data allows testing of a priory hypothesis regarding treatment effects in subgroups of patients. In conclusion, meta-analysis is a potentially powerful technique to

systematically review, analyze, and synthesize the body of research on a specific medical intervention.

Measurement of patient reported outcomes

Chapter 3 demonstrated the validity and subject acceptability of the EDC version of the Irritable Bowel Syndrome-Quality of Life measure (IBS-QOL), EuroQoL (EQ-5D), and the Work Productivity and Activity Impairment Questionnaire for Irritable Bowel Syndrome (WPAI:IBS) as compared to the existing paper versions. Traditionally, health-related quality of life data have been collected through patient self-report questionnaires printed on paper forms. Recently, the electronic methods of data collection are increasingly used in clinical trials. Although a concern exists that patients, particularly elderly, may be resistant to using new technology, several studies in different disease areas have suggested that EDC methods were preferred over traditional paper-and-pen methods by elderly volunteers and patients (Yarnold 1996). Surveys in patients with diabetes, gastrointestinal and psychiatric disease found that interactive computer programs were well accepted by the patients and provided reliable information (Pouwer 1998). The development of automated computer systems is therefore a promising approach to collection of health-related quality of life data in busy clinical practices. However, before a new method of administration of existing validated questionnaires can be recommended for wider use, the method should be evaluated for its effect on the reliability and validity of the instrument and for its effect on patient responses. This type of study represents a challenge for producers of medical technologies as they are including this evidence as part of their submission for market approval. In particular, the challenges related to the electronic validation study presented in chapter 3 include the following aspects: small sample size (72 patients), recruitment problems, e-diary programming, and data analysis issues. The study tried to estimate a balance between the following impacts: the minimum sample needed to effectively run some of the statistical tests is usually 50 but depends on the number of items in measure; the types of tests needing to be run (e.g. a larger sample would be needed if we are factoring subscales), the target population and the condition, and resource availability (e.g. needs of the study design, logistical limitation). The study encountered recruitment problems due to the fact that IBS is a disease that occurs predominantly in female, therefore an equal number of male patients was challenging. Another challenge was related to programming of the e-diary: some questions had a long text and therefore were divided into two screen shots, the Visual Analog Score of the EQ-5D could not be programmed on the vertical axis but horizontally, the skip button function was not included so that the patients would not have the option to skip a question. The later would have had implications for the missing data which is crucial for a small sample size study. All these programming issues posed a danger of EDC not being equivalent to the paper version format. However, such decisions had to be made in order to ensure the completeness of the study. As the evaluation of PRO endpoints in medical interventions continues to grow, there is an increasing interest in using these outcomes for labeling and promotional purposes. It is vital that these claims be based upon sound scientific evidence.

The use of decision analytic models

The use of decision-analytic modeling for HTA has increased exponentially in recent years. The process of decision analytic modeling is now seen as central to the process of HTA (Philips 2004). Usually their objective is to obtain a clear understanding of the relationship between incremental cost and effect in order to assess relative cost-effectiveness and to determine which interventions should be adopted given existing information. Decision analytic models provide a key role in translating the uncertainty associated with parameters (e.g. the cost of a particular adverse event, the quality of life impact of a condition or the relative risk reduction associated with a specific intervention) into the uncertainty associated with making a decision regarding the use of particular technologies (Claxton 2002). They are also a valuable tool for quantifying the implications (in terms of resource costs and health gain forgone) of that decision uncertainty which, in the form of expected value of perfect information, can be used to set priorities for future research (Claxton 2001).

Chapter 4 and 5 presented two examples of decision analytic models developed to inform decision-making process. The two cost-effectiveness models resulted in the dominance of the new medical intervention over existing treatment and are therefore in contrast with the more common outcome, that one intervention is both more effective and more costly allowing a less straightforward recommendation.

Given that decision models are being used more widely as part of HTA, it is important that the rigor of such studies is constantly enhanced. There are, however, a large number of specific issues for methods in decision modeling such as identification of parameter estimates from literature, bias in parameter estimates and extrapolation of treatment beyond the duration observed in clinical studies.

Decision models usually include a large range of different types of parameters, such as treatment effects, baseline event rates which may relate to the natural history of the condition, quality of life effects, health state values or utilities, resource use and unit costs. It is very rarely the case that one source of information will provide data for all these parameter estimates. The process of searching for information to populate each parameter in a decision model is extremely resource intensive. Decision models often have to be developed rapidly to inform particular decisions at a certain point in time. We address this issue in chapter 5 by performing a systematic literature review to search for data input. We focused our attention on parameters that we expected to be more sensitive and likely to have the largest influence on the results of the model such as body weight, cardiovascular risk factors, and cost of medical interventions. As part of the process of literature searching for parameter estimation, we also identified information that contributed to the development of the structure of the model.

Another issue in decision modeling relates to bias in parameter estimates. Despite an emphasis in the methods literature on the need to derive treatment effect estimates from randomized controlled studies, which, in principle, exhibit absence of selection bias, this may not be possible for all decision problems. In the absence of trial data, there is little guidance on how to adjust estimates of

treatment effect from observational studies where there are risks of selection bias. However, to avoid this issue, both models presented in chapter 4 and 5 used parameter estimates from randomized controlled studies.

One big challenge that regulatory agencies have to face in terms of the uncertainty surrounding existing evidence relates to costs and outcomes which have not been observed directly in trials. A feature of many trials is their short-term follow-up. This is particularly true of Phase III regulatory trials where there is a strong need to satisfy the licensing authorities and hence to get the product to market as swiftly as possible. For those interventions between which costs and benefits are likely to differ over an extensive time period, there will inevitably be a mismatch between trial follow-up and the appropriate time horizon of the cost-effectiveness analysis as was the case in chapter 4. This required the decision model to estimate the costs and health outcomes beyond the trial, together with the uncertainty associated with the extrapolation.

In conclusion, decision analytic models represent an explicit way to synthesize evidence currently available on the outcomes and costs of alternative (mutually exclusive) medical interventions. Given that the evidence base associated with new medical interventions will always have weakness and limitations; the adoption decision requires an analytical framework which is explicit in its handling of uncertainty.

Characterizing uncertainty in decision analytic models

Uncertainty is ubiquitous in all HTA decision models. Within modeling, researchers have tended to distinguish between parameter, structural and methodological uncertainty (Akehurst 2000). It has been suggested that the limitations of the chosen model structure should be acknowledged (Ramsey 1999) and a sensitivity analysis using alternative model structures should be performed (ISPOR 2003). In chapter 4 and 5 the structure of our models was dictated by the available data input, therefore the structural uncertainties could not be properly addressed. No proven methods exist to evaluate the structural uncertainties except to compute cost-effectiveness estimates for each alternative structural assumption and to examine the appropriateness of the results (Hay 1999). We addressed the methodological uncertainty in chapter 4 and 5 by using alternative discounting assumptions in line with the recommended way of addressing this issue (Halpern 1998).

The principal focus of uncertainty in literature has been on parameter uncertainty. Some suggest that parameters with the greatest level of uncertainty, such as those derived from expert opinion or those with greatest influence on the model outcomes, should be subjected to sensitivity analysis (Halpern 1998, Soto 2002) whereas others suggest that it should be the key clinical variables and the main cost drivers (Nuijiten 1998, Ramsey 1999). Nevertheless, the parameters and the value chosen for the sensitivity analysis should be presented and justified (ISPOR 2003). There is disagreement about the most appropriate approach to exploring parameter uncertainty. Parameter uncertainty can be addressed by univariate, multivariate or probabilistic sensitivity analysis, analysis of extremes, joint confidence intervals, bootstrap techniques or Monte Carlo simulation (Hay 1999). To address parameter uncertainty, chapter 4 and 5 used probability

distribution and Monte Carlo simulation rather than a point estimate. One reason for performing probabilistic sensitivity analysis in chapter 5 was to determine whether it would be worthwhile to seek better data for future decisions. Recommendations for the conduct or future design of research to obtain further data can be based on informal interpretation of the implications of sensitivity analysis or formal value of information techniques.

There has been an increasing interest in using expected value of information theory in medical decision-making, to identify the need for further research to reduce uncertainty in decision and as a tool for sensitivity analysis (Claxton 2001). Chapter 6 highlights the importance of identifying, quantifying, and incorporating parameter uncertainty in decision models. In addition to the precision of the data, the quality of the evidence available on a particular part of the model may be limited. In our model, this was seen in the link between weight loss and cardiovascular risk factors.

Chapter 6 demonstrated that Bayesian decision theory and value-of-information analysis is a valuable and practical framework within which two conceptually separate decisions problems were addressed: the selection of the lifestyle intervention strategy given existing information, and identification of the value of further information collection to inform this choice in the future. However, by observing an expected value of perfect information greater than the cost of additional research provides only the necessary but not sufficient condition for deciding to acquire more experimental information i.e. conducting a clinical trial. For a full analysis it is necessary to estimate the benefits of sampling, or the expected value of sample information for the patient population, and the cost of sample information, including the additional treatment and reporting cost. The expected value of sample information was not studied in our research.

Given that reimbursement authorities call for additional research on a new technology, even when these authorities approve the use of the therapy in question, data on value of information is expected to be very valuable. The formal use of value of information techniques will likely be increased over time in line with the recently growth in conditional reimbursement.

The use of economic evaluations in decision-making

Chapter 7 looked at the way health economics information is used in healthcare decision-making process. It evaluated the influence of economic research at the national, local and physician levels. Our review indicates that there is a considerable interest amongst policy decision-makers in health economics research, but that only moderate use was found at central, local and physician level. The influence of economic evaluations depends on several considerations, including the information needs of decision-makers, transparency of the economic evaluation and subsequent decision-making, mechanisms for enforcing the decision, and processes for monitoring and reappraising the evidence (Hutton 2006, Zentner 2005). For example, the gap between the long-term perspective of technology assessments and the short-term perspectives of policy-makers can limit the usefulness of recommendations (Neumann 2004). Moreover, broader health system characteristics, such as decentralized management, inadequate public resources or ‘silo’ budgeting, as well as existing incentives for

manufacturers and academics to deliver research that is interesting rather than practical and focused, may also prevent the best use of economic evaluation (Rutten 2005; OECD 2003).

The ability of HTA to maximize health for a given budget is difficult to assess in practice. Obstacles to effective assessment include the many other factors influencing policy and practice decisions, and the long-term nature of some of the effects of HTA such as changes in expectations and behavior patterns of users (Hailey 2007). If recommendations are to be implemented and the technologies taken up, there must be a clear and well communicated decision-making process in place. A lack of a defined process can create doubts over the legitimacy of decisions and therefore be less likely to have the support of stakeholders. It is difficult to incorporate evidence into poor-defined decision-making processes because the producers of evidence will be less likely to deliver timely and relevant advice. Part of instituting a clear decision-making process involves identifying an assessment framework that aligns incentives with evidence and health system objectives.

HTA has become an important mechanism for supporting priority-setting and decision-making. In particular, the growth of HTA reflects the demand for well founded information to support evidence-based decisions on the adoption and provision of health technologies. While there is general consensus that HTA provides value, this brief has highlighted a number of issues that can affect – positively and negatively – the effectiveness and impact of HTA. Successful implementation of HTA can be facilitated if there are appropriate policy instruments and regulatory levels available; a prior commitment by decision-makers to use assessment reports in decision-making process; available resources to implement decisions; stakeholder involvement; and transparency in both assessment and decision-making process. It nevertheless remains a challenge and one of the least developed areas of the overall HTA process.

Areas for further research

Many technical and methodological hurdles remain, and they need further investigation and research. They include the ability of summary measures to capture other benefits important to patients and the public; the generalisability of studies beyond a particular setting or country; the inability to account for the opportunity costs of expensive, new technologies; and the comparability between health state elicitation instruments.

HTA assessments are only helpful if they are used to support decision-making. Relevant stakeholders should be involved in order to facilitate the acceptance and implementation of decisions. There must be a transparent and well-communicated decision-making process to give legitimacy to subsequent recommendations. The availability of relevant policy instruments and collaboration between national and international HTA bodies also facilitate effective and efficient implementation. Existing evidence shows that stakeholder involvement can lead to greater transparency, relevance and acceptance of decisions, but little attention has been paid to how they are involved in the assessment process and how and when their perspectives are considered. More studies should be supported on the role and influence of various stakeholders, especially patients.

International collaboration between across HTA bodies can help facilitate methods development and more efficient assessment process, thereby improving the impact of HTA. An example is the European Network for Health Technology Assessment (EUnetHTA 2007), which was developed in response to an EU request for a formal, sustainable European network. International partnerships and networks are necessary to improve coordination, reduce duplication of effort, develop practical tools for HTA and improve the transfer of HTA into policy.

There is a lack of understanding about the 'real world' impact of HTA, not only on decision-making, but also on health outcomes, care delivery, healthcare costs, and research innovation. Several challenging questions remain regarding the circumstances surrounding the practical use of economic evidence in decision-making and priority-setting (When is it specifically used? How are criteria applied in practice and how are they weighted against the broad spectrum of decision factors? For a given disease area or public health problem, has HTA appropriately and accurately identified interventions that have led to improved health outcomes? Has the use of HTA led to better managed healthcare budgets or a decrease in healthcare costs? Does HTA provide sufficient incentives to facilitate innovative research and development? Has this 'fourth hurdle' in the reimbursement process prevented manufacturers from investing in new and innovative therapies? How can HTA be more broadly applied?). More focused research in these areas is needed. Greater efforts should be made to set up a formal evaluation component in the HTA process. Only by securing a better understanding of the decision-making process and the practical application of HTA can the impact of economic evaluation be enhanced.

There is limited information on the use of HTA for identifying areas of de-investment. More research is needed to identify ineffective and obsolete technologies and interventions. Reassessment after a technology has been used in practice is also an important mechanism in ensuring effective implementation. Regular review and re-evaluation are crucial in ensuring the availability of cost-effective and value-added medical technologies. While significant advances have been made on assessment methodologies, there is limited knowledge of how non-quantifiable factors are considered in the HTA process; this is especially true of equity concerns. Further exploration is needed to find out how such issues are taken into account in both assessment and subsequent decision-making, in order to address the social implications and constraints of efficient and equitable health care.

There is also a lack of research on the systematic assessment of public health interventions, especially those focused on prevention. Until now HTA has focused particularly on pharmaceuticals. In this thesis we give examples of how other medical interventions such as lifestyle intervention could be assessed. The application of the principles and methods of economic evaluations to preventive measures should be further explored, in an attempt to move towards a more evidence-based approach to important population health issues such as obesity or smoking. Given the limited evidence on the economic evaluation of public health interventions, more research should be funded to identify what assessments have been done so far, and what they revealed.

There is an increasing need to better account for uncertainty in assessment and decision processes, such as conditional approvals to make new technologies available while gathering additional data to address areas of uncertainty. Conditional approvals provide for the later collection of real-world data and reduce the potential opportunity costs of making inappropriate or inaccurate decisions. This thesis gave examples of how to methodologically address the uncertainty in the economic evaluations of medical technologies. However, further research is needed in conditional approvals using real word data to assess the uncertainty surrounding new and emerging technologies.

In conclusion, HTA offers extensive opportunities to support governments and other stakeholders, although issues remain concerning its use in, and impact on, healthcare policy and decision-making. Many of these have been highlighted in this thesis. The role of HTA in decision-making has grown substantially, but the need and demand for policy-makers to employ and translate evidence-based decisions into direct effects on health care costs and outcomes will probably increase in the next years.

References

Anell A, Persson U. Reimbursement and clinical guidance for pharmaceuticals in Sweden: do health-economic evaluations support decision making? Eur J Health Econ. 2005 Sep;6(3):274-9

Acquadro C, Berzon R, Dubois D et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. Value Health 2003; 6(5):522-531.

Akehurst R, Anderson P, Brazier J, Brennan A, Briggs A, Buxton M, et al. Decision analytic modelling in the economic evaluation of health technologies – a consensus statement. Pharmacoeconomics 2000;17:443–4.

Alberti KG, Zimmet PK. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1. Diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 1998, 15(7):539-53

Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991, 12 (A Pt 2) :293-8

Anderssen SA, Hjermann I, Urdal P, Torjesen PA, Holme I. Improved carbohydrate metabolism after physical training and dietary intervention in individuals with 'atherotrombotic syndrome'. The OSLO Diet and Exercise Study (ODES). A randomized trial. J Intern Med. 1996, 240(4):203-9

Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, Smith WC, Jung RT, Campbell MK, Grant AM. Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health improvement. Health Technol Assess. 2004, 8(21): iii-iv, 1-182

Bailar JC 3rd. The promise and problems of meta-analysis. N Engl J Med. 1997 Aug 21;337(8):559-61.

Banta D. The development of health technology assessment. Health Policy. 2003 Feb;63(2):121-32

Battista RN, Hodge MJ. The evolving paradigm of health technology assessment: reflections for the millennium. CMAJ. 1999 May 18;160(10):1464-7

Begg CB, Berlin JA. Publication bias and dissemination of clinical research. J Natl Cancer Inst 1989;81:107–15.

Begg CB, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996;276:637-639.

Berg M, van der Grinten T, Klazinga N. Technology assessment, priority setting, and appropriate care in Dutch health care. Int J Technol Assess Health Care. 2004 Winter;20(1):35-43

Berman NG, Parker RA. Meta-analysis: neither quick nor easy. BMC Med Res Methodol. 2002 Aug 9;2:10

Bliven BD, Kaufman SE, Spertus JA. Electronic collection of health-related quality of life data: validity, time benefits, and patient preference. Qual Life Res 2001;10(1):15-22.

Bloom BS. Use of formal benefit/cost evaluations in health system decision making. Am J Manag Care 2004, 10(5):329-35

Branscum AJ, Hanson TE. Bayesian Nonparametric Meta-Analysis Using Polya Tree Mixture Models. Biometrics. 2007 Dec 6. [Epub ahead of print]

Brennan A, Kharroubi S, O'hagan A, Chilcott J. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. Med Decis Making. 2007 Jul-Aug;27(4):448-70

Briggs, A., Handling uncertainty in economic evaluation and presenting the results, in economic evaluation in health care. Merging theory with practice, M. Drummond and A. McGuire, Editors. 2001, Oxford University Press: Oxford. p. 172-214.

Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000; 17(5):479-500

Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions Health Technol Assess 1999;3(2)

Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ 1998, 7(8):723-40

Bryan S, Williams I, Mciver S. Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. Health Econ 2007; 16:179-193

Bushnell DM, Martin ML, Parasuraman B. Electronic versus paper questionnaires: A further comparison in persons with asthma. Journal of Asthma 2003, 40(7):751-762

Bushnell DM, Martin ML, Ricci JF, Bracco A. Performance of the EQ-5D in Patients with Irritable Bowel Syndrome. Value in Health 2006; 9(2)

Burke V, Beilin LJ, Cutt HE, Mansour J, Wilson A, Mori TA (2005). Effects of a lifestyle programme on ambulatory blood pressure and drug dosage in treated hypertensive patients: a randomized controlled trial. J Hypertens. 23(6):1241-9

Buusman A, Andersen M, Merrild C, Elverdam B. Factors influencing GPs' choice between drugs in a therapeutic drug group. A qualitative study. Scand J Prim Health Care. 2007;25(4):208-13.

Buxton MJ. Economic evaluation and decision making in the UK. Pharmacoeconomics. 2006;24(11):1133-42

Carels RA, Darby LA, Cacciapaglia HM, Douglass OM. Reducing cardiovascular risk factors in postmenopausal women trough a lifestyle change intervention. Journal of Women's Health 2004, 13(4):412-26

Caro JJ Sr, Caro I, Caro J, Wouters F, Juniper EF. Does electronic implementation of questionnaires used in asthma alter responses compared to paper implementation? Qual Life Res 2001;10(8):683-91.

Carr DB, Utzschneider KM, Boyko EJ, Asberry PJ, Hull RL, Kodama K, Callahan HS, Matthys CC, Leonetti DL, Schwartz RS, Kahn SE, Fujimoto WY. A reduced-fat diet and aerobic exercise in Japanese Americans with impaired glucose tolerance decreases intra-abdominal fat and improves insulin sensitivity but not β-cell function. Diabetes 2005, 54(2): 340-47

Chauhan D, Mason A. Factors affecting the uptake of new medicines in secondary care - a literature review. J Clin Pharm Ther. 2008;33(4):339-48

Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. Lancet 2002; 360:711–15.

Claxton K, Neuman PJ, Araki S, Weinstein MC. Bayesian value of information analysis. An application to a policy model of Alzheimer disease. International Journal of Technology Assessment in Health Care 2001, 17(1):38-55

Claxton K, Fenwick E, Palmer S, Sculpter M, Abrahams K, Sutton A. Building a reference case for Bayesian applications to health economics and outcome research, Centre for Health Economics 2004, Technical Paper Series 35, The University of York

CMA, Comprehensive Meta Analysis 2005 Version 2.2.023. (www.Meta-Analysis.com)

Clinical. Clinical Overview CZOL446K (zoledronic acid) in Paget's disease of the bone (osteitis deformans); Clinical Study Report Study No: ZOL446K230; Clinical Study Report Study No: CZOL446H2305. Novartis, confidential 2004., Contract No.: Document Number|.

The Cochrane Collaboration. The Cochrane Manual Issue 4, 2008.

Cohen J, Stolk E, Niezen M. The increasingly complex fourth hurdle for pharmaceuticals. Pharmacoeconomics. 2007;25(9):727-34

Conte P, Guarneri V. Safety of intravenous and oral bisphosphonates and compliance with dosing regimens. Oncologist. 2004;9 Suppl 4:28-37.

Cookson R, Maynard A. Health technology assessment in Europe. Improving clarity and performance. Int J Technol Assess Health Care. 2000 Spring;16(2):639-50

Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. Med Decis Making. 2007 Sep-Oct;27(5):638-54. Epub 2007

Cronbach LF. Coefficient alpha and the internal structure of tests. Psychometricka 1951:16;297-334.

Cox E, Motheral B, Griffis D. Relevance of pharmacoeconomics and health outcomes information to health care decision-makers in the United States. Value Health 2000, 3(2):162

Davis P, Howden-Chapman P. Translating research findings into health policy. Soc Sci Med. 1996 Sep;43(5):865-72

Delmas, P. D., and Meunier, P. J. The management of Paget's disease of bone. N Engl J Med 1997; 336:558-66

Dentali F, Sharma AM, Douketis JD. Management of hypertension in overweight and obese patients: a practical guide for clinicians. Curr Hypertens Rep. 2005, 7(5):330-6

DeRoeck D. The importance of engaging policy-makers at the outset to guide research on and introduction of vaccines: the use of policy-maker surveys. J Health Popul Nutr. 2004;22(3):322-30

DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986 Sep;7(3):177-88.

Despres J. Abdominal obesity is an important component of insulin-resistance syndrome. Nutrition 1993, 9(5):452-459

Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbé KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. J Clin Epidemiol. 1992 Mar;45(3):255-65

Devogelaer, J. P. Modern therapy for Paget's disease of bone: focus on bisphosphonates. Treat Endocrinol 2002;1:241-57

Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Controlled Clin Trials 1991:12:142s-158s.

Diabetes Prevention Program (DPP) Research Group. Impact of lifestyle and metformin therapy on cardiovascular disease risk factors in the Diabetes Prevention Program. Diabetes Care 2005a, 28(4):888-894

Diabetes Prevention Program (DPP) Research Group. Intensive lifestyle intervention or metformin on inflammation and coagulation in participants with impaired glucose tolerance. Diadetes 2005b, 54:1566-1572

Dickersin 1992a. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards. JAMA 1992; 263:374-8.

Dickersin 1992b. Dickersin K. Keeping posted. Why register clinical trials? - revisited. Controlled Clin Trials 1992; 13:170-7.

Draborg E, Gyrd-Hansen D, Poulsen PB, Horder M. International comparison of the definition and the practical application of health technology assessment. Int J Technol Assess Health Care. 2005 Winter;21(1):89-95

Drake WM, Kendler DL, Brown JP. Consensus statement on the modern therapy of Paget's disease of bone from a Western Osteoporosis Alliance symposium. Biannual Foothills Meeting on Osteoporosis, Calgary, Alberta, Canada, September 9-10, 2000. Clin Ther. 2001 Apr;23(4):620-6.

Drossman DA, Patrick DL, Whitehead WE, Toner BB, Diamant NE, Hu Y, Jia H, Bangdiwala SI. Further Validation of the IBS-QOL: A disease-specific quality-of-life questionnaire. Am J Gastroenterol 2000; 95(4):999-1007.

Drossman et al. Rome II the functional gastrointestinal disorders. Second edition. Degnon Associates, McLean, VA, 2000.

Drummond HE, Ghosh S, Ferguson A, Brackenridge D, Tiplady B. Electronic quality of life questionnaires: a comparison of pen-based electronic questionnaires with conventional paper in a gastrointestinal study. Qual Life Res 1995;4(1):21-6.

Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: evidence from the UK. Soc Sci Med 1997, 45(4):583-595

Drummond M, Weatherly H. Implementing the findings of health technology assessments. If the CAT got out of the bag, can the TAIL wag the dog? Int J Technol Assess Health Care. 2000 Winter;16(1):1-12

Drummond SPECTRUM International trends in the use of health technology assessment. Decision Resources December 2007

Drummond MF, Schwartz JS, Jönsson B, Luce BR, Neumann PJ, Siebert U, Sullivan SD. Key principles for the improved conduct of health technology assessments for resource allocation decisions. Int J Technol Assess Health Care. 2008 Summer;24(3):244-58.

Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, Barbieri M; ISPOR Task Force. Use of pharmacoeconomics information--report of the ISPOR Task Force on use of pharmacoeconomic/health economic information in health-care decision making. Value Health. 2003;6(4):407-16.

Duthie T, Trueman P, Cancellor J, Diez L. Research into the use of health economics in decision making in the United Kingdom- Phase II. Is health economics 'for good or evil'? Health Policy 1999, 46(2):143-57

Dyson PA, Hammersley MS, Moris RJ, Holman RR, Turner RC. The fasting hypercalcemia study: II Randomized controlled trial of reinforced healthy-living advice in subjects with increased but not diabetic fasting plasma glucose. Metabolism 1997, 46 (12 Suppl.1):50-5

Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet1991;337:867–72.

Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at risk for diabetes. Ann Intern Med. 2005, 143(4):251-64

Eekhoff ME, van der Klift M, Kroon HM, Cooper C, Hofman A, Pols HA, Papapoulos SE. Paget's disease of bone in The Netherlands: a population-based radiological and biochemical survey--the Rotterdam Study. J Bone Miner Res. 2004 Apr;19(4):566-70. Epub 2004 Jan 26

EMEA. Summary of product characteristics Aclasta. 2007 [updated 2007; cited 2008 11 July]; Available from: http://www.emea.europa.eu/humandocs/Humans/EPAR/aclasta/aclasta.htm

EMEA, European Agency for the Evaluation of Medicinal Products 2005. Reflection paper on the regulatory guidance for the use of health-related quality of life measures in the evaluation of medicinal products. EMEA 2005,http://www.emea.eu.int/

EMEA, European Agency for the Evaluation of Medicinal Products. Committee for proprietary medicinal products European Public Assessment Report (EPAR): Xenical. London; EMEA 1998: p.1-39

EMIMS, Public price. Source www.emims.net. Accessed 21 April 2005 [updated Accessed 21 April 2005; cited]; Available from.

EP, Expert Panel on detection, evaluation, and treatment of high cholesterol in adults. Executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation and treatment of high blood cholesterol in adults (adult treatment panel III). JAMA 2001, 285(19):2486-97

Epstein FH, Holland WW. Prevention of chronic diseases in the community – one disease versus multiple disease strategies. Int J Epidemiol. 1983, 12(2):135-7

Erkan D, Yazici Y, Harrison MJ, Paget SA. Physician treatment preferences in rheumatoid artritis of differing disease severity and activity: the impact of cost on first-line therapy. Arthritis Rheum 2002; 47(3):285-90

Esposito K, Pontillo A, Di Paolo C, Giugliano G, Masella M, Marfella R, Giugliano D. Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women. JAMA 2003, 289(14):1799-804

Evans C, Dukes EM, Crawford B. The role of pharmacoeconomic information in the formulary decision-making process. JMCP 2000; 6(2), 108-21

Fattore G, Tobirca A. Economic evaluation in health care: the point of view of informed physicians. Value Health 2006, 9(3):157-67

Felli C, Hazen GB. A Bayesian approach to sensitivity analysis. Health Econ 1999;8:263-8

Fenwick, E., K. Claxton, and M. Sculpher, Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001,10(8): 779-87

Fenwick E, O'Brien B, Briggs A. Cost-effectiveness acceptability curve – facts, fallacies and frequently asked questions. Health Econ. 2004, 13(5):405-15

Fenwick E, Claxton K, Sculpter M. The value of implementation and the value of information: combined and uneven development. Centre for Health Economics Research paper 5, September 2005, University of York

Ferdinand KC.The importance of aggressive lipid management in patients at risk: evidence from recent clinical trials. Clin Cardiol 2004, 27(6 Suppl 3):III12-5

Field AE, Coakley EH, Must A, Spadano JL, Laird N, Dietz WH, Rimm E, Colditz GA. Impact of overweight on the risk of developing common chronic diseases during a 10-year period. Arch Intern Med.2001, 161(13):1581-6

Fijn R, Brouwers JR, Knaap RJ, De Jong-Van Den Berg LT. Drug and Therapeutics (D & T) committees in Dutch hospitals: a nation-wide survey of structure, activities, and drug selection procedures. Br J Clin Pharmacol. 1999;48(2):239-46.

Fontaine KR, Redden DT, Wang C, Westfall AO, Allison DB. Years of life lost due to obesity. JAMA. 2003, 289 (2):187-193

Food and Drug Administration- Guidance for Industry Patient Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2006 US Department of Health and Human Services, Food and Drug Administration Accessed: 7-13-2006. Available at: http://www.fda.gov/cber/gdlns/prolbl.pdf.

Gabrowski H, Mullins CD. Pharmacy benefit management, cost-effectiveness analysis and drug formulary decisions. Soc Sci Med 1997; 45(4), 535-544

Galani C, Schneider H, Rutten FF. Modelling the lifetime costs and health effects of lifestyle intervention in the prevention and treatment of obesity in Switzerland. Int J Public Health. 2007;52(6):372-82.

Galani C, Schneider H. Prevention and treatment of obesity with lifestyle interventions: review and meta-analysis. Int J Public Health. 2007;52(6):348-59.

García-Alonso MD, García-Mariñoso B. The strategic interaction between firms and formulary committees: effects on the prices of new drugs. J Health Econ. 2008;27(2):377-404.

Gerstein HC. Reduction of cardiovascular events and microvascular complications in diabetes with ACE inhibitor treatment: HOPE and MICRO-HOPE. Diabetes Metab Res Rev. 2002, 18 Suppl 3:S82-5

Gerzeli S, Trricone R, Zolo P, Colangelo I, Busca MR, Gandolfo C. (2005) The economic burden of stroke in Italy. The EcLIPSE study: economic longitudinal incidence-based project for stroke evaluation. Neurol Sci 26(2):72-80

Ginsburg ME, Kravitz RL, Sandberg WA. A survey of physician attitude concerning cost-effectiveness in patient care. West J Med 2000;173:390-394

Goddard M, Hauck K, Smith PC. Priority setting in health - a political economy perspective. Health Econ Policy Law. 2006 Jan;1(Pt 1):79-90

Goodman CS. Healthcare technology assessment: methods, framework, and role in policy making. Am J Manag Care. 1998 Sep 25;4 Spec No:SP200-14; quiz SP215-6

Greenberg D, Peterburg Y, Vekstein D, Pliskin JS. Decisions to adopt new technologies at the hospital level: insights from Israeli medical centers. Int J Technol Assess Health Care 2005; 21(2), 219-227

Grizzle A, Motheral B, Garrity B, Cox E, Amstrong E. A qualitative assessment of managed care decision-makers's views and use of pharmacoeconomic information. Value Health 2000; 3(2),162-3

Hailey DM. Health technology assessment in Canada: diversity and evolution. Med J Aust. 2007 Sep 3;187(5):286-8

Halpern MT, McKenna M, Hutton J. Modeling in economic evaluation: an unavoidable fact of life. Health Econ 1998;7:741–2.

Hakim Z, Wolf A, Garrison LP. Estimating the effect of changes in body mass index on health state preferences. Pharmacoeconomics 2002, 20 (6):393-404

Harvey-Berino J, Pintauro S, Buzzell P, Gold EC. Effect of internet support on the long-term maintenance of weight loss. Obes Res 2004, 12(2):320-9

Hasle-Pham E, Arnould B, Späth HM, Follet A, Duru G, Marquis P. Role of clinical, patient-reported outcome and medico-economic studies in the public hospital drug formulary decision-making: results of a European survey. Health Policy 2005; 71:205-212

Hasselblad V, Mosteller F, Littenberg B, Chalmers TC, Hunink MG, Turner JA, Morton SC, Diehr P, Wong JB, Powe NR. A survey of current problems in meta-analysis. Discussion from the Agency for Health Care Policy and Research inter-PORT Work Group on Literature Review/Meta-Analysis. Med Care. 1995 Feb;33(2):202-20

Hay J, Jackson J. Panel 2: methodological issues in conducting pharmacoeconomic evaluations – modeling studies. Value Health 1999;2:78–81.

He J, Whelton PK, Appel LJ, Charleston J, Klag MJ. Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension. Hypertension 2000, 35(2):544-9

Herman WH, Hoerger TJ, Brandle M, Hicks K, Sorensen S, Zhang P, Hamman RF, Ackermann RT, Engelgau MM, Ratner RE. The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. Ann Intern Med 2005, 142 (5):323-32

Hoffman C, Schulenburg JMG. The influence of economic evaluation studies on decision making A European survey. Health Policy 2000; 52, 179-192

Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF. Do health-care decision makers find economic evaluations useful? The findings of focus group research in UK health authorities. Value Health 2002, 5(2):71-78

Hosking D, Lyles K, Brown JP, Fraser WD, Miller P, Curiel MD, Devogelaer JP, Hooper M, Su G, Zelenakas K, Pak J, Fashola T, Saidi Y, Eriksen EF, Reid IR. Long-term control of bone turnover in Paget's disease with zoledronic acid and risedronate. J Bone Miner Res. 2007 Jan;22(1):142-8

Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997;36(5):551-9.

Hutton J, McGrath C, Frybourg JM, Tremblay M, Bramley-Harker E, Henshall C. Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). Int J Technol Assess Health Care. 2006;22(1):10-8.

Iglesias CP, Drummond MF, Rovira J. Health-care decision-making process in Latin America: problems and prospects for the use of economic evaluation. Int J Technol Assess Health Care 2005; 21(1), 1-14

ISPOR Task Force. Principles of good practice for decision analytic modeling in health-care evaluation. Value Health 2003;6:9–17.

Jacobs JW, Huisman AM, van Paassen HC, Bijlsma JW. [Paget's disease of the bones: diagnosis and treatment]. Ned Tijdschr Geneeskd. 1999 Apr 3;143(14):719-25.

Jacoby A, Smith M, Eccles M. A qualitative study to explore influences on general practitioners' decisions to prescribe new drugs. Br J Gen Pract. 2003;53(487):120-5.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds JM, Gavaghan DJ, McQuay DM. Assessing the quality of reports on randomized clinical trials: Is blinding necessary? Controlled Clin Trials 1996, 17:1-12

Jansson S, Anell A. The impact of decentralized drug-budgets in Sweden - survey of physicians' attitudes towards costs and cost-effectiveness Health Policy 2006, 299-311

Jeffery RW, Wing RR. Long-term effects of interventions for weight loss using food provision and monetary incentives. J Consult Clin Psychol 1995, 63(5):793-6

Jenkings KN, Barber N. What constitutes evidence in hospital new drug decision making. Social Science Medicine 2004; 58:1757-66

Jia H, Lubetkin EI. The impact of obesity on health-related quality-of-life in the general adult US population. J Public Health 2005, 27(2):156-64

Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure. The sixth report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC VI). Arch Intern Med 1997, 157:2413-2446

Johannes C, Woods J, Crawford S, et al. Electronic versus paper instruments for daily data collection. Ann Epidemiol 2000;10:457

Johannes CB, Crawford SL, Woods J, Goldstein RB, Tran D, Mehrotra S, Johnson KB, Santoro N. An electronic menstrual cycle calendar: comparison of data quality with a paper version. Menopause 2000; 7:200-208.

Johnston CC, Jr., Khairi MR, Meunier PJ. Use of etidronate (EHDP) in Paget's disease of bone. Arthritis Rheum. 1980 Oct;23(10):1172-6.

Jonsson E, Banta HD, Henshall C, Sampietro-Colom L. Summary report of the ECHTA/ECAHI project. European Collaboration for Health Technology Assessment/Assessment of Health Interventions. Int J Technol Assess Health Care. 2002 Spring;18(2):218-37.

Jönsson B. Health technology assessment: Regulators or payers-Who will take the lead? Clin Ther. 2008;30(5):960-3

Kanis, J. Pathophysiology and treatment of Paget's disease of bone. London: Martin Duritz; 1998.

Kangis P, van der Geer L. Pharmaco-economic information and its effect on prescriptions J Manag Med 1996; 10(5):66-74

Kastarinen M, Puska MP, Korhonen MH, Mustonen JN, Salomaa VV, Sundvall JE, Tuomilehto JO, Uusitupa MI, Nissinen AM. LIEHEF Study Group. Non-pharmacological treatment of hypertension in primary health care: a 2-year open randomized controlled trial of lifestyle intervention against hypertension in eastern Finland. J Hypertens 2002, 20(12):2505-12

Kelly R, Taggart H. Incidence of gastrointestinal side effects due to alendronate is high in clinical practice. BMJ. 1997 Nov 8;315(7117):1235.

Kernick DP. The impact of health economics on healthcare delivery. A primary care perspective. Pharmacoeconomics. 2000;18(4):311-5.

Ketola E, Makela M, Klockars M. Individualized multifactorial lifestyle intervention trial for high-risk cardiovascular patients in primary care. Br J Gen Pract 2001, 51(465):291-4

Kind P. The EuroQol instrument: an index of health-related quality of life. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd edition. Philadelphia (PA): Lippincott-Raven Publishers, 1996:191-201.

Kleinman L, Leidy NK, Crawley J, Bonomi A, Schoenfeld P. A comparative trial of paper-and-pencil versus computer administration of the Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire. Med Care 2001;39(2):181-9.

Kolominsky-Rabas PL, Heuschmann PU, Marschall D, Emmert M, Baltzer N, Neundörfer B et al. (2006). Lifetyme cost of ischemic stroke in Germany: results and national projections from a population-based stroke registry. The Erlangen Stroke Project. Stroke, 37: 1179-1183

Konig HH, Ulshofer A, Gregor M, von Tirpitz C, Reinshagen M, Adler G, Leidl R. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. Eur J Gastroenterol Hepatol 2002 ;14(11):1205-15.

Langston AL, Ralston SH. Management of Paget's disease of bone. Rheumatology (Oxford). 2004 Aug;43(8):955-9.

Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. CMAJ 1992, 146(4):473-81

Lemieux S, Prud'homme D, Bouchard C, Tremblay A, Despres JP. A single threshold value of waist girth identifies normal-weight and overweight subjects with excess visceral adipose tissue. Am J Clin Nutr 1996, 64(5):685-93

Liao D, Asberry PJ, Shofer JB, Callahan H, Matthyas C, Boyko EJ, Leonetti D, Kahn SE, Austin M, Newell L, Schwartz RS, Fujimoto WY. Improvement of BMI, body composition, body fat distribution with lifestyle modification in Japanese Americans with impaired glucose tolerance. Diabetes Care 2002, 25(9):1504-10

Lindahl B, Nilsson TK, Jansson JH, Asplund K, Hallmans G. Improved fibrinolysis by intense lifestyle intervention. A randomized trial in subjects with impaired glucose tolerance. J Intern Med 1999, 246(1):105-12

Lindstrom J, Louheranta A, Mannelin M, Rastas M, Salminen V, Eriksson JG, Uusitupa M, Tuomilehto J. The Finnish Diabetes Prevention Study (DPS). Lifestyle intervention and 3-year results on diet and physical activity. Diabetes Care 2003, 26(12):3230-6.

Lindström J, Peltonen M, Toumilehtp J. Lifestyle strategies for weight control: experience from the Finnish Diabetes Prevention Study Proc Nutr Soc 2005, 64(1):81-8

Lindström J, Ilanne-Parikka P, Peltonen M, Aunola S, Eriksson JG, Hemio K, Hamalainen H et al. Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finish Diabetes Prevention Study. Lancet 2006, 368(9548):1673-9

Luce BR, Brown RE. The use of technology assessment by hospitals, health maintenance organizations, and third-party payers in the United States. Int J Technol Assess Health Care 1995; 11(1), 79-92

Luce BR, Lyles A, Rentz AM. The view from managed care pharmacy. Health Affairs 1996; 15(4), 168-76

Lyles A, Luce BR, Rentz AM. Managed care pharmacy, socioeconomic assessments and drug adoption decisions Soc Sci Med 1997; 45(4), 511-21

Macran S. The relationship between body mass index and health related quality of life. Discussion paper 190, 2004 The University of York, Centre for Health Economics

Maetzel A, Ruof J, Covington M, Wolf A. Economic evaluation of orlistat in overweight and obese patients with type 2 diabetes mellitus. Pharmacoeconomics. 2003;21(7):501-12

Matthews KA, Meilahn E, Kuller LH, Kelsey SF, Caggiula AW, Wing RR. Menopause and risk factors for coronary heart disease. N Engl J Med 1998, 321(10):641-6

Martin DK, Pater JL, Singer PA. Priority-setting decisions for new cancer drugs: a quantitative case study. The Lancet 2001; 358(17), 1676-81

Martin DK, Hollenberg D, MacRae S, Madden S, Singer P. Priority setting in a hospital drug formulary: a qualitative case study and evaluation. Health Policy 2003; 66:295-303

Marx S, Wong P, Priory S. Great research or lack of training. ISPOR Connections 2007, 13(2):19-20

Mähönen M, Rajakangas A-M, Kuulasmaa K, Tunstall-Pedoe H. (2000) WHO MONICA coronary event registration data book 1980-1995 Accessed on September 2006 at www.who.com/publications

McDonald R, Burrill P, Walley T Managing the entry of new medicines in the National Health Service: health authority experiences and prospects for primary care groups and trusts. Health Soc Care Community. 2001;9(6):341-7

McGee DL. Body mass index and mortality: a meta-analysis based on person-level data from twenty six observational studies. Ann Epidemiol 2005, 15(2):87-97

Meagher EA. Addressing cardiovascular disease in women: focus on dyslipidemia. J Am Board Fam Pract 2004, 17(6):424-37

Mengersen KL, Tweedie RL, Biggerstaff BJ. The impact of method choice in meta-analysis. Aust J Stats 1995;37:19–44.

Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Servick MA, Ettinger WH Jr, Pahor M, Williamson JD. Exercise and dietary weight loss in overweight and obese older adults with knee osteoartritis: The arthritis, diet, and activity promotion trial. Arthritis Rheum 2004, 50 (5):1501-10

Mensink M, Feskens EJ, Saris WH, de Bruin TW, Blaak EE. Study on lifestyle-intervention and impaired glucose tolerance Maastricht (SLIM): preliminary results after one year. Int J Obes Relat Metab Disord 2003a, 27(3):377-84

Mensink M, Blaak EE, Corpeleijin E, Saris WH, de Bruin TW, Feskens EJ. Lifestyle intervention according to general recommendations improves glucose tolerance. Obesity Research 2003b, 11(12):1588-96

Michaels JA Improving NICE's social value judgments. BMJ. 2006 Jan 7;332(7532):48-50.

Neumann J. Using cost-effectiveness analysis to improve health care: opportunities and barriers. Oxford University Press: New York

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609-13.

Moore H, Summerbell CD, Greenwood DC, Tovey P, Griffiths J, Henderson M, Hesketh K, Woolgar S, Adamson AJ. Improving management of obesity in primary care: cluster randomized trial. BMJ 2003, 8;327(7423):1085

Motheral B, Grizzle A, Amstrong E, Cox E, Bataoel J, Bennett D, Shahriar J. A national survey of decision-makers in managed care on their views and use of pharmacoeconomic information. Value Health 2000, 3(2):163

Narayan KM, Hoskin M, Kozak D, Kriska AM, Hanson RL, Pettitt DJ, Nagi DK, Bennett PH, Knowler WC. Randomized clinical trial of lifestyle interventions in Prima Indians: a pilot study. Diabet Med 1998, 15(1): 66-72

National Cholesterol Education Program. Third report of the expert panel on detection, evaluation and treatment of high blood cholesterol in adults. NIH Publication no. 01-3670, 2001National Institute of Health, Bethesda, MD

Netten A, Lesley C. Unit costs of Health and Social Care 2004. wwwPSSRUacuk. Accessed 21 April 2005.

Neumann PJ. Using Cost-Effectiveness Analysis to Improve Health Care: Opportunities and Barriers. Oxford University Press. 2005

Neumann PJ, Palmer JA, Daniels N, Quigley K, Gold MR, Chao S; Panel on Integrating Cost-Effectiveness Considerations into Health Policy Decisions. A strategic plan for integrating cost-effectiveness analysis into the US healthcare system. Am J Manag Care. 2008 Apr;14(4):185-8

NICE, National Institute for Clinical Excellence 2004 Guide to the Methods of Technology Appraisal

NICE, National Institute for Clinical Excellence. Guidance on the use of orlistat for the treatment of obesity in adults. London. Report no.22, 2004

NICE, National Institute for Clinical Excellence. Guidance on the use of sibutramine for the treatment of obesity in adults. London. Report no.31, 2004

Nichol MB, Knight TK, Epstein J, Honda DH, Tretiak R. Opinions regarding the academy of managed care pharmacy dossier submission guidelines: results of a small survey of Managed care organizations and pharmaceutical manufactures. JMCP 2007; 13(4),360-70

Niessen LW, Grijseels E, Koopmanschap M, Rutten F; Dutch Ministry of Health. Economic analysis for clinical practice--the case of 31 national consensus guidelines in the Netherlands. J Eval Clin Pract. 2007;13(1):68-78

NHS Reference Costs 2004. www.dh.gov.uk. Accessed 21 April 2005 [updated Accessed 21 April 2005; cited 21 April 2005]; Available from: www.dh.gov.uk.

Niessen LW, Grijseels E, Koopmanschap M, Rutten F, Dutch Ministry of Health. Economic analysis for clinical practice - the case of 31 national consensus guidelines in the Netherlands. Journal of Evaluation in Clinical Practice 2007, 13(1): 68-78

Nuijten MJC, Szende A, Kosa J, Mogyorosy Z, Kramberger B, Nemecek K et al. Health care reform in six Central European countries. A focus on health economic requirements in the drug pricing and reimbursement process. Eur J Health Econom 2003; 4(4), 286-91

Nuijten MJC, Pronk MH, Brorens MJA, Hekster YA, Lockefeer JHM, de Smet P, et al. Reporting format for economic evaluation: Part II: Focus on modelling studies. Pharmacoeconomics 1998; 14:259–68.

OECD 2003. Survey of pharmacoeconomic assessment in eleven countries. OECD working papers no 4. OECD.Paris

Odedina FT, Sullivan J, Nash R, Clemmons CD. Use of pharmacoeconomic data in making hospital formulary decisions. Am J Health-Syst Pharm 2002; 59,1441-44

Orvain J, Xerri B, Matillon Y. Overview of health technology assessment in France. Int J Technol Assess Health Care. 2004;20(1):25-34

Paget, J. On a form of chronic inflamation of bones (Osteitis deformans). Medico-Chirurgical Transactions of London 1877; 60:37-63

Palermo TM, Valezuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability. Pain 2004;107:213-19

Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, Hu ZX, Lin J, Xiao JZ, Cao HB, Liu PA, Jiang XG, Jiang YY, Wang JP, Zheng H, Zhang H, Bennett PH, Howard BV. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and diabetes study. Diabetes Care 1997, 20(4):537-44

Patrick DL, Drossman DA, Frederick IO, Dicesare J, Puder KL. Quality of life in persons with irritable bowel syndrome: Development and validation of a new measure. Digestive Diseases and Sciences 1998;43(2):400-411.

PausJenssen AM, Singer PA, Detsky AS. Ontario's formulary Committee How recommendations are made. Pharmacoeconomics 2003; 21(4):285-294

Pedersen TR, Olsson AG, Faergeman O, Kjekshus J, Wedel H, Berg K, Wilhelmsen L, Haghfelt T, Thorgeirsson G, Pyorala K, Mittinen T, Christophersen B, Tobert JA, Musliner TA, Cook TJ. Lipoprotein changes and reduction in the incidence of major coronary heart disease events in the Scandinavian Simvastatin Survival Study (4S). Atheroscler 1998, Suppl. 5(3):99-106

Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess 2004;8(36).

Philips Z, Bojke L, Sculpter M, Claxton K, Golder S. Good clinical practice guidelines for decision analytic modelling in health technology assessment: a review and consolidation of quality assessment. PharmacoEconomics 2006, 24(4):355-371

Philips Z, Claxton KP, Palmer S, Bojke L, Sculpter MJ. Priority setting for research in health care: an application of value of information analysis to glycoprotein IIb/IIa antagonists in non-DT elevation acute coronary syndrome. International Journal of Technology Assessment in Health Care 2006, 22:3,379-387

Poirier P, Giles TD, Bray GA, Hong Y, Stern JS, Pi-Sunyer FX, Eckel RH. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect on weight loss. Artherioscler Thromb Vasc Biol. 2006, 26(5):968-76

Pouwer F, Snoek FJ, van der Ploeg HM, et al: A comparison of the standard and the computerized versions of theWell-Being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). Qual Life Res 1998, 7:33-38

PRISM, Randomised trial of intensive versus symptomatic management of Paget's Disease (the PRISM trial I. http://www.abdn.ac.uk/hsru/hta/prism.shtml. Accessed 19 December 2005.

Pritchett AM, Foreyt JP, Mann DL. Treatment of the metabolic syndrome: the impact of lifestyle modification. Curr Atheroscler Rep. 2005, 7(2):95-102

Prosser H, Walley T. A qualitative study of GPs' and PCO stakeholders' views on the importance and influence of cost on prescribing. Soc Sci Med. 2005;60(6):1335-46

Pyorala K, Pedersen TR, Kjekshus J, Faergeman O, Olsson AG, Thorgeirsson G. Cholesterol lowering with simvastatin improves prognosis of diabetic patients with coronary heart disease. A subgroup analysis of
Scandinavian Simvastatin Survival Study (4S) Group. Diabetes Care 1997, 20(4):614-20

Ramsey SD, Sullivan SD. Weighing the economic evidence: guidelines for critical assessment of cost-effectiveness analyses. J Am Board Fam Pract 1999;12:477–85.

Reid IR, Miller P, Lyles K, Fraser W, Brown JP, Saidi Y, Mesenbrink P, Su G, Pak J, Zelenakas K, Luchi M, Richardson P, Hosking D. Comparison of a single infusion of zoledronic acid with risedronate for Paget's disease. N Engl J Med. 2005 Sep 1;353(9):898-908.

Reilly MC, Lavin PT, Kahler KH, Pariser DM. Validation of the Dermatology Life Quality Index (DLQI) and the Work Productivity and Activity Impairment: Chronic Hand Dermatitis Questionnaire (WPAI-ChHD) in Chronic Hand Dermatitis (ChHD). J Am Academy Dermatol.2003; 48:128-30.

Reilly MC, Bracco A, Ricci J-F, Santoro J, Stevens T. The validity and accuracy of the Work Productivity and Activity Impairment questionnaire – irritable bowel syndrome version (WPAI:IBS). Aliment Pharmacol Ther 2004; 20: 1-9.

Reilly MC, Tanner A, Meltzer EO. Work, classroom and activity impairment instruments: validation studies in allergic rhinitis. Clinical Drug Investigations 1996; 11(5):278-288.

Revicki DA, Osoba D, Fairclough D, Barofsky I, Berzon R, Leidy NK, Rothman M. Recommendations of health-related quality of life research to support labeling and promotional claims in United States. Qual Life Res 2000, 9:887-900

Rhee JS. Comparing apples to oranges in meta-analysis studies. Arch Facial Plast Surg. 2007 Mar-Apr;9(2):139-40

Ross J. The use of economic evaluation in health care: Australian's decision makers perceptions. Health Policy 1995, 31(2):103-10

Rousiere M, Michou L, Cornelis F, Orcel P. Paget's disease of bone. Best Pract Res Clin Rheumatol. 2003 Dec;17(6):1019-41.

Rubin RM, Chang CF. A bibliometric analysis of health economics articles in the economics literature: 1991-2000. Health Econ. 2003;12(5):403-14.

Ruof J, Golay A, Berne C, Collin C, Lentz J, Maetzel A. Orlistat in responding obese type 2 diabetic patients: meta-analysis findings and cost-effectivenesss as rationales for reimbursement in Sweden and Switzerland Int J Obes 2005, 29(5):517-23

Rutten F, Brouwer W, Niessen L. Practice guidelines based on clinical and economic evidence; indispensable tools in future market oriented health care. European Journal of Health Economics 2005, 6: 91-93

Ryan JM, Corry JR, Attewell R, Smithson MJ. A comparison of an electronic version of SF-36 general health questionnaire to standard paper version. Qual Life Res 2002;11:19-26

SACMOT, Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002;11:193–205.

Saint S, Vesntra DL, Sullivan SD. The use of meta-analysis in cost-effectiveness analysis Pharmacoeconomics 1999, 15(1):1-8

Sbrocco T, Nedegaard RC, Stone JM, Lewis EL. Behavioural choice treatment promotes continuing weight loss: preliminary results of a cognitive-behavioral decision-based treatment for obesity. J Consult Clin Psychol 1999, 67(2):260-6

Schmid A, Schneider H, Golay A, Keller U. Economic burden of obesity and its comorbidities in Switzerland. Soz.-Präventivmed. 2005, 50(2):87-94

Schmitt-Koopmann I, Schwenkglenks M, Spinas GA, Szucs TD. Direct medical costs of type 2 diabetes and its complications in Switzerland. Eur J Public Health. 2004, 14(1):3-9

Schreyögg J, Stargardt T, Velasco-Garrido M, Busse R. Defining the "Health Benefit Basket" in nine European countries. Evidence from the European Union Health BASKET Project. Eur J Health Econ. 2005;Suppl:2-10

Schumock GT, Walton SM, Park HY, Nutescu EA, Backburn JC, Finley JM, Lewis RK. Factors that influence prescribing decisions. Ann Pharmacother 2004; 38(4):557-62

Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty – When is there sufficient evidence? Value in Health 2005, 8(4), 2005

Selby PL, Davie MW, Ralston SH, Stone MD. Guidelines on the management of Paget's disease of bone. Bone. 2002 Sep;31(3):366-73.

Sempos CT, Durazo-Arvizu R, McGee DL, Cooper RS, Prewitt TE. The influence of cigarette smoking on the association between body weight and mortality. The Framingham Heart Study revised. Ann Epidemiol. 1998, 8(5):289-300

SFSO, Swiss Federal Statistical Office (2006): Swiss Health Survey (Bundesamt für Statistik, Schweizerische Gesundheitsbefragung 2002)

Shalansky SJ, Virk R, Ackman M, Jackevicius C, Kertland H, Tsuyuki R, Humphries K. Access to new cardiovascular therapies in Canadian hospoitals: a national survey of the formulary process. Can J Cardiol 2003; 19(2):173-9

Shiell A, Donaldson C, Mitton C, Currie G. Health economic evaluation. J Epidemiol Community Health. 2002;56(2):85-8.

Simkin-Silverman LR, Wing RR, Boraz MA, Kuller LH. Lifestyle intervention can prevent weight gain during menopause: results from a 5-year randomised clinical trial. Ann Behav Med 2003, 26(3):212-20

Singer PA, Martin DK, Giacomini M, Purdy L. Priority setting for new technologies in medicine: qualitative case study. BMJ 2000;321:1316-8

Siris E, Roodman G. Paget's disease of bone. Section 5, Chapter 82.: American Society for Bone and Mineral Research; 2003 .

Sloan FA, Whetten-Goldstein K, Wilson A. Hospital pharmacy decisions, cost containment, and the use of cost-effectiveness analysis. Soc Sci Med 1997; 45(4), 523-533

SNB. Swiss National Bank 2006, Monthly Statistical Bulletin, Accessed September 2006 at www.snb.ch

Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Dec Making 1993, 13(4):322-38

Soto J. Health economic evaluations using decision analytic modeling. Principles and practices – utilization of a checklist to their development and appraisal. Int J Technol Assess Health Care 2002; 18:94–111.

Spath HM, Charavel M, Morelle M, Carrere MO. A quantitative approach to teh use of economic data in teh selection of medicines for hospital formularies: a French survey. Pharm World Sci 2003; 25(6): 269-275

SPSS, Inc. Statistical Package for the Social Sciences® Base 10.0 for Windows User's Guide. USA, Chicago, Ill.: SPSS, Inc., 1999.

Stefanick ML, Mackey S, Sheehan M, Ellsworth N, Haskell WL, Wood PD. Effects of diet and exercise in men and postmenopausal women with low levels of HDL cholesterol and with high levels of LDL cholesterol. N Engl J Med 1998, 339(1):12-20

Steiner CA, Powe NR, Anderson GF, Das A. The review process used by US health care plans to evaluate new medical technology for coverage. J Gen Intern Med 1996; 11(5):294-302

Stevens VJ, Obarzanek E, Cook NR, Lee IM, Appel LJ, Smith WD, Milas NC, Mattfeldt-Beman M, Belden L, Bragg C, Millstone M, Raczynski J, Brewer A, Singh B, Cohen J. Trials for the hypertension prevention research group. Long-term weight loss and changes in blood pressure: results from the Trials of Hypertension Prevention, Phase II. Ann Intern Med 2001, 134(1):1-11

Tan EL, Day RO, Brien JE. Prioritizing drug and therapeutics committee (DTC) decisions: a national survey. Pharm World Sco 2007; 29:90-96

TARMED 2006: Tarif code 00.0510 Accessed September 2006 at Tarif-Browser: www.tarmed.ch/tarif/

Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes. JAMA 2003, 289(14):1833-6

Tordoff JM, Murphy JE, Norris PT, Reith DM. Use of centrally developed pharmacoeconomic assessments for local formulary decisions. Am J Health-Syst Pharm 2006; 63(1), 1613-18

Trento M, Passera P, Tomalino M, Bajardi M, Pomero F, Allione A, Vaccari P, Molinatti GM, Porta M. Group visits improve metabolic control in type 2 diabetes: a 2-year follow-up Diabetes Care 2001, 24(6):995-1000

Tsai AG, Glick HA, Shera D, Stern L, Samaha FF. Cost-effectiveness of a low-carbohydrate and a standard diet in severe obesity. Obes Res 2005, 13(10):1834-40

Tunis SR. Why Medicare has not established criteria for coverage decisions. N Engl J Med. 2004 May 20;350(21):2196-8.

Ubel PA, Jepson C, Baron J, Hershey JC, Asch DA. The influence of cost-effectiveness information on physicians' cancer screening recommendations. Social Science Medicine 2003; 56, 1727-3

UKPDS 34, UK Prospective Diabetes Study Group: Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes. Lancet 1998, 352(9131):854-65,

UKPDS 38, UK Prospective Diabetes Study Group: Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes. BMJ 1998, 317(7160):703-13

UKPDS 61, UK Prospective Diabetes Study Group: Are lower fasting plasma glucose levels at diagnosis of type 2 diabetes associated with improved outcomes? Diabetes Care 2002, 25(8):1410-7

van Staa, T. P., Selby, P., Leufkens, H. G., Lyles, K., Sprafka, J. M., and Cooper, C. Incidence and natural history of Paget's disease of bone in England and Wales. J Bone Miner Res 2002; 17:465-71; 2002

Van Velden ME, Severens JL, Novak A. Economic evaluations of healthcare programmes and decision making: the influence of economic evaluations on different healthcare decision-making levels. Pharmacoeconomics 2005, 23(11):1075-82

Velikova G, Wright EP, Smith AB, Cull A, Gould A, Forman D, Perren T, Stead M, Brown J, Selby PJ. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999, 17:998-1007.

Vinik AI. The metabolic basis of atherogenic dylipidemia. Clinical Cornerstone. 2005, 7(2-3):27-35

Vuorenkoski L, Toiviainen H, Hemminki E. Drug reimbursement in Finland-a case of explicit prioritizing in special categories. Health Policy. 2003;66(2):169-77

Wahlqvist P, Carlsson J, Stalhammar NO, Wiklund I. Validity of a Work Productivity and Activity Impairment questionnaire for patients with symptoms of gastro-esophageal reflux disease (WPAI-GERD)- results from a cross-sectional study. Value in Health 2002; 5:106-113.

Walley T, Barton S, Cooke J, Drummond M. Economic evaluations of drug therapy: attitudes of primary care prescribing advisers in Great Britain. Health Policy. 1997;41(1):61-72

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.

Weart W, Bauman GR. The case for a value-based formulary: striving for total lowest net cost. Manag Care Interface. 2007;20(4):42-7

Whelton PK, Appel LJ, Espeland MA, Applegate WB, Ettinger WH Jr, Kostis JB, Kumanyika S, Lacy CR, Johnson KC, Folmar S, Cutler JA. Sodium reduction and weight loss in the treatment of hypertension in older persons: a randomised controlled trial of nonpharmacologic interventions in the elderly (TONE). TONE collaborative research group. JAMA 1998, 279(11):839-46

Weatherly H, Drummond M, Smith D. Using evidence in the development of local health policies. Some evidence from the United Kningdom. Int J Technol Assess Health Care 2002; 18(4):771-81

West R, Borden EK, Collet JP, Rawson NS, Tonks RS,. 'Cost-effectiveness' estimates result in flawed decision-making in listing drugs for reimbursement. Can J Public Health 2002; 93(6):421-5

Wild D, Grove A, Martin M et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005; 8(2):94- 104.

Williams I, Bryan S, McIver S. How should cost-effectiveness analysis be used in health technology coverage decisions? Evidence from the National Institute for Health and Clinical Excellence approach. J Health Serv Res Policy. 2007;12(2):73-9.

Williams IP, Bryan S. Cost-effectiveness analysis and formulary decision making in England: findings from research. Soc Sci Med. 2007;65(10):2116-29.

Wing RR, Venditti E, Jkicic JM, Polley BA, Lang W. Lifestyle intervention in overweight individuals with a family history of diabetes. Diabetes Care 1998, 21(3):350-9

Wolf AM, Conaway MR, Crowther JQ, Hazen KY, Nadler JL, Oneida B, Bovbjerg VE. Translating lifestyle intervention to practice in obese patients with type 2 diabetes. Improving control with activity and nutrition (ICAN) study Diabetes Care 2004, 27(7):1570-6

World Health Organization Obesity: preventing and managing the global epidemic. Geneva, WHO Technical Report Series, No. 894, 2000 Accessed on September 2006 at http://www.who.int/nutrition/publications/obesity/en/index.html

World Health Organization, European region, country profile Switzerland, InfoBase ref: 101551a3, 2005 Accessed September 2006 at www.who.com

Wu AW, Jacobson KL, Frick KD, Clark R, Revicki DA, Freedberg KA, Scott-Lennox J, Feinberg J. Validity and responsiveness of the Euroqol as a measure of health-related quality of life in people enrolled in an AIDS clinical trial. Qual Life Res 2002;11(3):273-82.

Wu O, Knill-Jones R, Willson P, Craig N. The impact of economic information on medical decision making in primary care. Journal Evaluation Clinical Practice 2004, 10(3), 407-11

Wylie-Rosett J, Swencionis C, Ginsberg M, Cimino C, Wassertheil-Smoller S, Caban A, Segal-Isaacson CJ, Martin T, Lewis J. Computerized weight loss intervention optimises staff time: the clinical and costs results of a controlled clinical trial conducted in a managed care setting. J Am Diet Assoc 2001, 101(10):1155-62

Yarnold PR, Stewart MJ, Stille FC, et al: Assessing functional status of elderly adults via microcomputer. Percept Mot Skills 1996, 82:689- 690

Yeh MC, Rodriguez E, Nawaz H, Gonzalez M, Nakamoto D, Katz DL.Technical skills for weight loss: 2-y follow-up results of a randomised trial. Int J Obes Relat Metab Disord 2003, 27(12):1500-6

Zentner A, Valasco-Garrido M, Busse R. Methods for the comparative evaluation of pharmaceuticals. GMS Health Technol Assess 2005, 1:Doc09

Zwart-van Rijkom JE, Leufkens HGM, Busschbach JJV, Broekmans AW, Rutten FFH. Differences in attitudes, knowledge and use of economic evaluations in decision-making in the Netherlands. The Dutch results from the EUROMET project. Pharmacoeconomics 2000; 18(2): 149-160

Summary

Health technology assessment (HTA) originated from the spread of costly medical equipment and growing concerns over the ability and willingness of taxpayers and health insurers to pay for them. HTA include a range of medical interventions, including drugs, medical devices, medical and surgical procedures, and the organizational and support systems used in care provision. It involves evaluating an intervention through the production, synthesis, and/or systematic review of a range of scientific and non-scientific evidence. The evidence typically considered includes safety, efficacy, cost, and cost-effectiveness, as well as the social, organizational, legal, and ethical implications. The main aim of HTA is to provide a range of stakeholders, typically those involved in funding, planning, purchasing, and investing in healthcare, with accessible, useable, and evidence-based information that will guide decisions about technology and the efficient allocation of resources. It has been called 'the bridge between evidence and policy making', because it provides information for health care decision-makers at macro-, meso-, and micro-levels. In particular, the increased use of medical technologies has encouraged decision-makers to rely on HTA to help determine the reimbursement status and pricing of interventions. HTA also contributes in many ways to the knowledge base for improving the quality of care, especially in supporting the development of clinical practice guidelines and health service standards.

The present thesis addresses some of the methodological challenges of the health technology assessment and evaluates the impact of economic evaluations in healthcare decision-making process.

The use of meta-analysis in systematic literature review

Systematic reviews have a central role in evidence based-medicine. The quantitative systematic review, also known as meta-analysis provides a logical structure for quantifying the existing evidence. Meta-analysis offers the opportunity to critically evaluate and statistically combine results of comparable studies or trials. The aim is to get a consistent estimation of the global effect of a procedure on a specified outcome by increasing the number of observations and statistical power. This thesis presents the results of a systematic literature review with meta-analysis on the effectiveness of lifestyle intervention in the prevention and treatment of obesity. The global rise in obesity prevalence continues to be a threat to people's

health. This review is part of a HTA in the prevention and treatment of obesity performed for Swiss Ministry of Health to support the decision to allocate funds in the prevention of obesity programs in Switzerland in 2007. The systematic review provides new information on the effectiveness of lifestyle interventions by assessing the mid- to long-term effects on weight and cardiovascular risk profile in overweight and obese people. An important consideration when performing meta-analyses is the possibility of heterogeneity between studies outcome estimates. It has been argued that producing an overall combined estimate for heterogeneous studies is wrong and leads to a result which is misleading, and impossible to interpret. However, there are no clear guidelines outlining how variable study results have to be before it is deemed invalid to combine them. We address this issue of heterogeneity by taking into consideration the following factors: we determined the study question of the evaluated studies, evaluated the similarities or dissimilarities in the study design, and looked if the heterogeneity of the outcomes can be explained. However, no consensus has been reached concerning the best strategy for dealing with heterogeneity; currently a large degree of subjectivity is required on the part of the reviewer. In this thesis we investigated the heterogeneity using the following techniques: funnel plot for publication bias, random effect model, subgroup analysis, sensitivity analysis, and quality assessment scores. We concluded that meta-analysis is a potentially powerful technique to systematically review, analyze, and synthesize the body of research on a specific medical intervention.

Measurement of patient reported outcomes

Patient reported outcomes provide the patient perspective on the effectiveness of treatment. Patients, clinicians, pharmaceutical industry, decision-makers, payers and regulatory authorities acknowledge the need to understand the impact of symptoms and diseases on patients' lives and to evaluate how treatment affects patient functioning and well-being as a criterion for licensing new medications and for policy decisions. Traditionally, patient reported outcome data in clinical trials have been collected through patient self-report questionnaires printed on paper forms. Recently, the electronic methods of data collection are increasingly used in clinical trials. Although a concern exists that patients, particularly elderly, may be resistant to using new technology, several studies in different disease areas have suggested that EDC methods were preferred over traditional paper-and-pen methods. For patient reported outcome endpoint data to be accepted as evidence of treatment effectiveness there must be evidence documenting the instrument's conceptual framework, content validity, and psychometric qualities, including reliability, validity and responsiveness.

This thesis presents the data collected from a study that establishes the validity of the electronic versions of three quality of life measures in comparison to the existing paper versions. A selective 5-HT4 receptor partial agonist was developed for the treatment of functional gastrointestinal dysmotility disorders by Novartis Pharma. A multinational clinical trial has been completed using electronic data collection. In order to evaluate changes in patient reported outcomes, standard instruments were included to assess change in patients' quality of life (IBS-QOL), productivity (WPAI:IBS-C), and utility (EQ-5D). For health technology

manufacturers it is important to provide evidence of the validity of the quality of life instruments in the electronic format as compared to paper version. This type of study represents a challenge for producers of medical technology which have to include this evidence as part of their submission for market approval. In particular, the challenges encountered by electronic validation study presented in this thesis included the following aspects: small sample size (72 patients), recruitment problems, e-diary programming, and data analysis issues. All these programming issues posed a danger of the electronic data capture not being equivalent to the paper version format. However, the study proved to be successful. It is vital that patient reported outcome claims be based upon sound scientific evidence.

The use of decision analytic models

Mathematical modeling is used widely in economic evaluations of medical interventions. Health economics models represent an important analytic framework to generate estimates of cost-effectiveness. The purpose of modeling is to structure evidence on clinical and economic outcomes in a form that can help to inform decisions about clinical practices and healthcare resource allocations. Models synthesize evidence on health consequences and costs from many different sources, including data from clinical trials, observational studies, insurance claim databases, case registries, public health statistics, and preference surveys. The use of decision-analytic modeling for health technology assessment has increased exponentially in recent years.

Given that decision models are being used more widely as part of HTA, it is important that the rigor of such studies is constantly enhanced. There are, however, a large number of specific issues for methods in decision modeling such as identification of parameter estimates from literature, bias in parameter estimates and extrapolation of treatment beyond the duration observed in clinical studies. We addressed these issues in two decision analytic models developed to inform decision-making process. One model presents the development and results of a cost-effectiveness model of zoledronic acid versus risedronate in Paget's disease of bone. Zoledronic acid is a third generation of nitrogen containing bisphosphonate seeking approval for Paget's disease of bone indication. For registration purposes two six-month randomized clinical trials were performed in order to evaluate the safety and efficacy of intravenous zoledronic acid for the treatment of Paget's disease of bone using oral risendronate as a comparator. The economic model, based on the two clinical trials, goes beyond six-month and evaluates the cost-effectiveness of zoledronic acid versus standard therapy, risedronate, for the first two years of market access of zoledronic acid in the United Kingdom.

The second model presents the development and results of an economic analysis that evaluates the lifetime effects of three-year lifestyle intervention in the prevention and treatment of obesity. The model estimates the cost-effectiveness of lifestyle intervention versus standard treatment in overweight and obese people in Switzerland. The model is a key component of the HTA that was carried out for lifestyle intervention in obesity area. Probabilistic sensitivity analysis was undertaken to establish the uncertainty associated with the decision to adopt the lifestyle intervention program in the prevention and treatment of obesity.

We concluded that decision analytic models represent an explicit way to synthesize evidence currently available on the outcomes and costs of alternative medical interventions. Given that the evidence base associated with new medical interventions will always have weakness and limitations; the adoption decision requires an analytical framework which is explicit in its handling of uncertainty.

Characterizing uncertainty in decision analytic models

Most review agencies conduct or require sensitivity analyses on all variables that could potentially influence the overall results of the economic evaluation. This is because of the uncertainty inherent in conducting economic evaluations, specifically over the value of particular estimates and their relative effect on costs and benefits. The stipulation for sensitivity analyses comes from the need to test or verify the robustness of the findings. Traditionally, the uncertainty has been examined using sensitivity analysis. In the recent years there has been considerable emphasis on the development of appropriate statistical methods for handling uncertainty in economic evaluations of medical interventions, with a tendency to move from univariate sensitivity analysis towards probabilistic descriptions of uncertainty e.g. cost-effectiveness planes, cost-effectiveness acceptability curves and distributions of incremental net benefit. Value of information analysis is a natural methodological extension of Bayesian decision theory that quantifies the existing level of uncertainty and estimates the impact on the expected net benefit of alternative decision options through obtaining perfect information on model parameters. We addressed the uncertainty in this thesis by performing probabilistic sensitivity analysis and by evaluating the value of additional information. We used the value of information analysis on a probabilistic cost-effectiveness model of lifestyle intervention in overweight and obese people to evaluate the uncertainty. Our analysis quantified the uncertainty surrounding the decision to adopt lifestyle intervention.

Given that reimbursement authorities call for additional research on a new technology—even when these authorities approve the use of the therapy in question—data on value of information is expected to be very valuable. The formal use of value of information techniques will likely be increased over time in line with the recently growth in conditional reimbursement.

The use of economic evaluations in decision-making

The importance of health economics research utilization in policy-making and of understanding the mechanisms involved is increasingly recognized. The existence of relevant research, though necessary, is not sufficient. Evidence-based policy is difficult to achieve and it is widely agreed that health policies do not reflect research evidence to the extent that in theory they could. In this thesis we assessed the use of research evidence relating to economic analyses in healthcare decision-making. We conducted a literature review to summarize and synthesize published literature on self-reported attitudes of healthcare decision-makers towards economic evaluations of medical technologies. The aims of this literature review was to determine the extent to which economic evaluations are used in health policy decision-making, and to consider factors associated with the utilization of

such research findings. Examination of the policy-making process confirms it to be complex, with many genuine obstacles to evidence-based policy-making at the same time as there are factors that could increase research utilization.

The place of HTA in the decision-making process can affect the extent to which evidence is used to inform policy and priority-setting. Countries often disagree on the use of HTA recommendations. Some support recommendations on the grounds that experts are the best people to provide them, while others prefer decision-makers to make recommendations in the light of political context and other country-specific circumstances. However, decision-makers may not have the technical expertise to understand the methodological strengths and weaknesses of an assessment. Improvements are still needed, but much has been done by assessment bodies to enhance the accessibility and usability of HTA among different audiences such as policy-makers, health professionals, and general public. Although different decision structures provide policy-makers with a wide range of discretion, not employing HTA evidence may lead to inefficient, ineffective, and inequitable healthcare.

Samenvatting

Health technology assessment (HTA) is een beoordelingsmethode voor gezondheidstechnologie die is ontstaan vanuit de bredere beschikbaarheid van kostbare medische apparatuur en een toenemende bezorgdheid over het vermogen en de bereidheid van belastingbetalers en ziektekostenverzekeraars om daarvoor te betalen. Deze methode richt zich op uiteenlopende medische interventies, zoals geneesmiddelen, medische apparatuur, geneeskundige en chirurgische procedures en de organisatorische en ondersteunende systemen die bij het bieden van zorg worden gebruikt. Een onderdeel ervan is de evaluatie van een interventie door verschillende wetenschappelijke en niet-wetenschappelijke empirische gegevens te produceren, te combineren en/of systematisch te beoordelen. Kenmerkende aspecten die hierbij worden bekeken zijn veiligheid, werkzaamheid, kosten en kosteneffectiviteit, alsmede sociale, organisatorische, juridische en ethische implicaties. Het hoofddoel van HTA is het aan de verschillende belanghebbenden, met name aan degenen die in de gezondheidszorg betrokken zijn bij de financiering, planning, inkoop en investeringen, verschaffen van toegankelijke, bruikbare en empirisch onderbouwde informatie als een richtlijn bij de besluitvorming over technologie en een doelmatige toewijzing van middelen. Deze beoordelingsmethode wordt wel ‘de brug tussen empirie en beleid’ genoemd, omdat ze informatie levert aan de besluitvormers in de gezondheidszorg op macro-, meso- en microniveau. Met name het toegenomen gebruik van medische technologie heeft ertoe geleid dat de beslissers op HTA vertrouwen bij het bepalen van de vergoedingsstatus en de prijs van interventies. HTA draagt ook op allerlei manieren bij aan kennis waarmee de kwaliteit van de zorg kan worden verbeterd, met name als hulpmiddel bij het ontwikkelen van richtlijnen voor de klinische praktijk en van standaarden in de gezondheidszorg.

Het gebruik van meta-analyse in systematisch literatuuronderzoek

Systematisch literatuuronderzoek speelt een centrale rol bij empirisch onderbouwde geneeskunde. Kwantitatief systematisch literatuuronderzoek, ook wel meta-analyse genoemd, biedt een logische structuur om het bestaande empirische

materiaal te kwantificeren. Meta-analyse geeft de gelegenheid om de resultaten van vergelijkbare onderzoeken of experimenten kritisch te beoordelen en statistisch te combineren. Het doel hiervan is een consistente schatting te krijgen van het algemene effect van een procedure op een bepaald resultaat door het aantal waarnemingen en de statistische kracht te verhogen. Dit proefschrift presenteert de resultaten van een systematisch literatuuronderzoek waarbij een meta-analyse is uitgevoerd naar de doeltreffendheid van leefstijlinterventie bij de preventie en behandeling van obesitas. De algemene toename van obesitas is nog altijd een bedreiging voor de volksgezondheid. Dit onderzoek maakt deel uit van een HTA van de preventie en behandeling van obesitas die voor het Zwitserse Ministerie van volksgezondheid werd uitgevoerd als ondersteuning van de beslissing om in 2007 programma's voor de preventie van obesitas te financieren. Het systematische onderzoek levert nieuwe informatie over de doeltreffendheid van leefstijlinterventies door een schatting te geven van de middellange- en langetermijneffecten op het gewicht en het cardiovasculaire risicoprofiel van mensen met overgewicht en obesitas. Een belangrijk aspect bij het uitvoeren van een meta-analyse is de mogelijke heterogeniteit van de schattingen op grond van de onderzoeksresultaten. Men heeft betoogd dat het bepalen van een overkoepelende gecombineerde schatting bij heterogene onderzoeken verkeerd is en tot een misleidend resultaat leidt dat onmogelijk geïnterpreteerd kan worden. Er bestaan echter geen duidelijke richtlijnen die aangeven hoe sterk onderzoeksresultaten moeten verschillen, wil het combineren ervan niet langer geldig worden geacht. Op het punt van de heterogeniteit zijn de volgende factoren bekeken: de onderzoeksvraag van de geëvalueerde onderzoeken is bepaald, de overeenkomsten en verschillen in de onderzoeksopzet zijn geïnventariseerd, en nagegaan is of de heterogeniteit in de uitkomsten verklaarbaar was. Er bestaat echter geen overeenstemming over wat de beste strategie is voor de aanpak van heterogeniteit; momenteel is nog een grote mate van subjectiviteit van de onderzoeker vereist. In dit proefschrift is de heterogeniteit onderzocht met behulp van de volgende technieken: een funnel plot voor de publicatiebias, een random effect model, subgroepanalyse, gevoeligheidsanalyse en quality assessment-scores. De conclusie is dat meta-analyse in potentie een krachtige techniek is voor het systematisch onderzoeken, analyseren en samenvatten van bestaand onderzoek naar een specifieke medische interventie.

Meting van door de patiënt gerapporteerde resultaten

Door de patiënt gerapporteerde resultaten geven het perspectief van de patiënt aan op de doeltreffendheid van de behandeling. Patiënten, clinici, de farmaceutische industrie, beleidsvormers, financiers en regelgevende instanties erkennen dat er als criterium voor het toelaten van nieuwe medicijnen en voor beleidsbeslissingen behoefte is aan inzicht in de gevolgen van symptomen en ziektes voor het leven van patiënten en aan een evaluatie van de invloed van de behandeling op het functioneren en het welzijn van de patiënt. Traditioneel werden deze gegevens over het door de patiënt ervaren resultaat verzameld aan de hand van door de patiënt zelf ingevulde vragenlijsten op papier. Sinds kort worden bij klinische experimenten steeds vaker elektronische vormen van dataverzameling gebruikt. Weliswaar is er enige bezorgdheid dat met name oudere patiënten zich misschien verzetten tegen het gebruik van deze nieuwe techniek, maar uit een aantal onderzoeken op het gebied van verschillende ziektes blijkt dat elektronische dataverzamelingsmethodes de voorkeur hebben boven de traditionele methode met pen en papier. Door de patiënt gerapporteerde resultaten gelden pas als empirisch bewijs voor de doeltreffendheid van de behandeling als van het meetinstrument het begripsmatige kader, de validiteit van de inhoud en de psychometrische eigenschappen zijn vastgelegd, met inbegrip van de betrouwbaarheid, validiteit en gevoeligheid.

Dit proefschrift presenteert gegevens die zijn verzameld in een onderzoek waarin de validiteit van de elektronische versie van drie metingen van de kwaliteit van leven wordt vergeleken met de bestaande papieren versies. Het onderzoek betrof een selectieve 5-HT4-receptor partiële agonist, die was ontwikkeld voor de behandeling van functionele gastro-intestinale dysmotiliteitsstoornissen. Hierbij werd een multinationaal klinisch onderzoek afgesloten met behulp van elektronische dataverzameling. Om verschillen in door de patiënt gerapporteerde resultaten te evalueren werden standaardinstrumenten opgenomen waarmee verandering wordt gemeten in de kwaliteit van leven (IBS-QOL), de productiviteit (WPAI:IBS-C) en de utiliteit (EQ-5D) van de patiënt. Voor fabrikanten van gezondheidstechnologie is het belangrijk om empirisch te bewijzen dat de meetinstrumenten voor de kwaliteit van leven in elektronische vorm even valide zijn als de papieren versie. Dergelijk onderzoek is een uitdaging voor de producenten van medische technologie, want dit empirische bewijs is onderdeel van de aanvraag voor toestemming om hun product om op de markt te brengen. In dit proefschrift komen met name de volgende aspecten van elektronisch validatieonderzoek aan de orde: de kleine steekproefomvang (72 patiënten), wervingsproblemen, het programmeren van elektronische dagboeken en zaken rond de gegevensanalyse. Alle punten rond het programmeren hielden het risico in dat de elektronische methode van gegevensverzameling niet gelijkwaardig was met de papieren versie. Het onderzoek bleek echter succesvol. Het is van vitaal belang dat uitspraken over door de patiënt gerapporteerde resultaten op degelijk wetenschappelijk empirisch bewijs zijn gebaseerd.

Het gebruik van besliskundige modellen

Bij een economische evaluatie van medische interventies worden vaak wiskundige modellen gebruikt. Het gebruik van modellen in gezondheidseconomie vormt een belangrijk analytisch instrument om schattingen van de kosteneffectiviteit te verkrijgen. Het doel van modellen is dat ze de klinische en economische empirische gegevens zodanig structureren dat dit bijdraagt aan de onderbouwing van beslissingen in de klinische praktijk en aan het toewijzen van middelen in de gezondheidszorg. Modellen vatten empirische gegevens over de gezondheidseffecten en kosten samen vanuit veel verschillende bronnen, zoals gegevens uit klinische proeven, observatieonderzoeken, databanken met verzekeringsclaims, gevalsbeschrijvingen, gezondheidsstatistieken en preferentieonderzoeken. Het gebruik van besliskundige modellen bij de beoordeling van gezondheidstechnologie is de laatste jaren exponentieel gestegen.

Omdat beslissingsmodellen steeds vaker als onderdeel van HTA worden gebruikt, is het belangrijk dat de validiteit van dergelijk onderzoek wordt vast gesteld. Er zijn allerlei methodologische punten die specifiek zijn voor het opstellen van beslissingsmodellen, zoals het correct afleiden van parameterschattingen uit de literatuur, het vermijden van vertekening in de schattingen van parameters en een verantwoordede extrapolatie naar een behandeling die langer duurt dan de periode die in de klinische proeven is geobserveerd. Op deze punten wordt ingegaan voor twee besliskundige modellen die werden ontwikkeld om het besluitvormingsproces te onderbouwen. Het eerste model geeft de ontwikkeling en resultaten weer van een model voor de kosteneffectiviteit van zoledroninezuur versus risedronaat bij de botziekte van Paget. Zoledroninezuur is een derde generatie stikstof bevattend bisfosfonaat, waarvoor toestemming werd gevraagd voor gebruik bij de botziekte van Paget. Voor deze registratie werden twee gerandomiseerde klinische tests van zes maanden uitgevoerd om de veiligheid en werkzaamheid te evalueren van intraveneuze toediening van zoledroninezuur bij de behandeling van de botziekte van Paget, met als vergelijking orale toediening van risedronaat. Het economische model dat op deze twee klinische proeven is gebaseerd, gaat verder dan zes maanden en evalueert de kosteneffectiviteit van zoledroninezuur versus de standaardbehandeling met risedronaat voor de eerste twee jaren waarin zoledroninezuur in het Verenigd Koninkrijk op de markt is.

Het tweede model geeft de ontwikkeling en resultaten weer van een economisch ramingsmodel voor de invloed op de levensduur van drie jaar leefstijlinterventie versus de standaardbehandeling van mensen met overgewicht en obesitas in Zwitserland. Dit model was een sleutelonderdeel van de HTA die werd uitgevoerd voor leefstijlinterventie op het gebied van obesitas. Met behulp van probabilistische gevoeligheidsanalyse werd de onzekerheid bepaald bij de beslissing om het leefstijlinterventieprogramma in te voeren ter voorkoming en behandeling van obesitas.

De conclusie was dat besliskundige modellen een expliciete methode zijn om de huidige beschikbare gegevens over de resultaten en kosten van alternatieve medische interventies samen te vatten. Omdat de grondslag van empirische gegevens in verband met nieuwe medische interventies altijd zwakke punten en beperkingen zal hebben, vereist dit de toepassing in een analytisch kader waarin de aanpak van onzekerheid expliciet is aangegeven.

Aangeven van onzekerheid in besliskundige modellen

In de meeste onderzoeksmethodes worden gevoeligheidsanalyses uitgevoerd of vereist voor alle variabelen die mogelijkerwijs van invloed zijn op de overkoepelende resultaten van de economische raming. Dat is een gevolg van de onzekerheid die inherent is aan economische ramingen, met name voor de waarde van bepaalde schattingen en hun relatieve invloed op kosten en baten. Deze voorwaarde van gevoeligheidsanalyse komt voort uit de noodzaak om de robuustheid van de uitkomsten te toetsen of te verifiëren. Traditioneel wordt de onzekerheid onderzocht met behulp van gevoeligheidsanalyse. De laatste jaren is er een grote nadruk geweest op het ontwikkelen van geschikte statistische methodes om de onzekerheid aan te pakken bij economische ramingen voor medische interventies, waarbij de tendens is om van univariate gevoeligheidsanalyses over te stappen op probabilistische beschrijvingen van de onzekerheid, zoals kosteneffectiviteitsvlakken, aanvaardbaarheidskrommen voor de kosteneffectiviteit en verdelingen van de marginale netto opbrengst. De analyse van informatiewaarde is een logische uitbreiding van de Bayesiaanse beslissingstheorie, die het bestaande niveau van onzekerheid kwantificeert en de invloed ervan op de verwachte netto opbrengst van mogelijke alternatieve beslissingen schat door de waarde van volmaakte informatie over de parameters van het model te beschouwen. In dit proefschrift is de onzekerheid onderzocht met een probabilistische gevoeligheidsanalyse en door de waarde van extra informatie in te schatten. De analyse van informatiewaarde is toegepast op een probabilistisch kosteneffectiviteitsmodel voor leefstijlinterventie bij mensen met overgewicht en obesitas. Deze analyse kwantificeert de onzekerheid rond de beslissing om over te gaan tot leefstijlinterventie.

Omdat vergoedingsinstanties bij een nieuwe technologie om aanvullend onderzoek vragen – zelfs als die instanties het toepassen van de betreffende therapie goedkeuren – is te verwachten dat gegevens over de informatiewaarde heel waardevol zijn. Het formele gebruik van technieken om de informatiewaarde te bepalen zal waarschijnlijk in de loop van de tijd toenemen, in het verlengde van de recente groei van voorwaardelijke vergoedingen

Het gebruik van economische ramingen bij besluitvorming

Dat het belangrijk is om van gezondheidseconomisch onderzoek gebruik te maken bij beleidsbeslissingen en om inzicht te krijgen in de betrokken mechanismes, wordt in toenemende mate onderkend. Het bestaan van relevant onderzoek is weliswaar noodzakelijk, maar niet voldoende. Empirisch onderbouwd beleid is moeilijk te verwezenlijken en er is brede overeenstemming dat de beleidsmaatregelen op het terrein van de gezondheid de onderzoeksbevindingen niet zo goed weerspiegelen als in theorie mogelijk zou zijn. In dit proefschrift is het gebruik beoordeeld van aan economische analyses gerelateerde onderzoeksresultaten bij de besluitvorming in de gezondheidszorg. Voor dit proefschrift is een literatuuronderzoek uitgevoerd waarin een overzicht en een samenvatting werd gegeven van de gepubliceerde literatuur over de zelf-gerapporteerde houding van beleidsmakers in de gezondheidszorg tegenover economische ramingen voor medische technologie. Het doel van dit literatuuronderzoek was te bepalen in hoeverre economische

ramingen worden gebruikt bij beleidsmatige besluitvorming in de gezondheidszorg en de factoren te bekijken die meespelen bij het gebruik maken van dergelijke onderzoeksresultaten. Onderzoek van het beleidsvormingsproces bevestigt dat dit complex is, met veel reële hindernissen voor empirisch onderbouwd beleid, terwijl er tegelijkertijd ook factoren zijn die het gebruik maken van onderzoek kunnen versterken.

De plaats van HTA in het besluitvormingsproces kan van invloed zijn op de mate waarin empirisch materiaal wordt gebruikt als onderbouwing voor het beleid en voor het stellen van prioriteiten. Landen zijn het vaak niet eens over het toepassen van de aanbevelingen van HTA. Sommigen steunen deze omdat deskundigen de meest aangewezen personen worden geacht om aanbevelingen te doen, terwijl anderen er de voorkeur aan geven dat besluitvormers aanbevelingen doen in het licht van de politieke context en andere voor dat land specifieke omstandigheden. De besluitvormers hebben echter misschien onvoldoende technische expertise om de methodologische sterke en zwakke punten van een beoordeling te begrijpen. Verbeteringen zijn nog steeds noodzakelijk, maar er is door beoordelingsinstanties al veel aan gedaan om de toegankelijkheid en bruikbaarheid van HTA te verbeteren voor de verschillende doelgroepen, zoals beleidsmakers, gezondheidsprofessionals en het algemene publiek. Hoewel uiteenlopende beslissingsstructuren beleidsmakers van een breed scala aan beoordelingsmogelijkheden voorzien, kan het niet of verkeerd gebruiken van de resultaten van HTA leiden tot ondoelmatige, ondoeltreffende en onrechtvaardige gezondheidszorg.

List of abbreviations

ALP	Alkalin phosphatase
BMI	Body mass index
BP	Bisphosphonates
CEA	Cost-effectiveness analysis
CH	Switzerland
CHF	Swiss Francs
CMA	Comprehensive meta-analysis
DBP	Diastolic blood pressure
EDC	Electronic data capture
EMEA	European Agency for the Evaluation of Medicinal Products
EVPI	Expected value of perfect information
EQ-5D	EuroQol (standardized instrument for use as a measure of health outcome)
FPG	Fasting plasma glucose
FLS	Flu like syndrome
FDA	Food and Drug Administration
GP	General practitioner
HbA1c	Heaemoglobin A1c
HDL	High density lipoprotein cholesterol
HTA	Health technology assessment
IBS	Irritable bowel syndrome
IBS-QOL	Irritable bowel syndrome quality of life questionnaire
ICC	Intraclass correlation coefficients
ICD	International Classification of Disease

LY	Life-years
LDL	Low density lipoprotein cholesterol
MCS	Mental component score
NHS	National Health System
NICE	National Institute of Clinical Excellence
PDB	Paget's disease of bone
PCS	Physical component score
PPI	Proton pump inhibitor
PRO	Patient reported outcome
RIS	Risedronate
SAP	Serum alkaline phosphatase
SBP	Systolic blood pressure
TC	Total cholesterol
TG	Triglyceride
VAS	Visual analog scale
ZOL	Zoledronic acid
WHO	World Health Organization
WPAI:IBS	Work productivity and activity impairment irritable bowel syndrome questionnaire
QALY	Quality adjusted life years
2h-PG	Two hours plasma glucose

Portfolio Summary

Summary of PhD training and teaching activities

Name PhD student: Carmen Galani Erasmus MC Department: Health Policy & Management Research School: NIHES	PhD period: 2004-2009 Promotor: Prof. Dr. Frans Rutten Copromotor: Dr. Maiwenn Al

1. PhD training	Year	Workload (Hours/ECTS)
General academic skills		
- Biomedical English Writing and Communication	2005 - 2008	
- Prepared abstracts and manuscripts for submission to international congresses and peer-reviewed journals: Value in Health, International Journal of Public Health, Current Medical Research and Opinion	2007, 2008	20
- Participated as reviewer of medical journals Journal of American Medical Association, Value in Health, Journal of Occupational and Environmental, Medicine, European Journal of Health Economics	2007, 2008	10
- Presented research findings at international congresses	2005, 2007	2
Research skills		
- Statistics		
- Attended courses at ISPOR Annual International Meeting 2005, Washington, US		
▪ 'Bayesian Analysis – Overview'	2005	5
▪ 'Bayesian Analysis – Applications'	2005	5
- Completed TreeAge modelling course	2005	8
- Methodology		
- Attended courses at ISPOR Annual European Congress 2004, Hamburg, Germany	2004	

▪ 'European Databases and Retrospective Databases Analysis' ▪ 'The Analysis and Interpretation of Quality of Life and Patient Reported Outcome' - Meta-analysis : Theory and Applications course	2004 2004 2005	5 5 10
In-depth courses (e.g. Research school, Medical Training) - Advanced Health Economics Modelling Course, A. Briggs, Oxford University - Issues in Public Health course	 2005 2004	 10 10
Presentations - ISPOR Annual International Meeting 2005, Washington, US - ISPOR Annual European Meeting 2007, Dublin, Ireland - 12th United European Gastroenterology Week 2004, Prague, Czech Republic	 2005 2007 2004	 2 2 2
International conferences - ISPOR 2004, 2005, 2006, 2007, 2008 - UEGW 2004		 8 2
Seminars and workshops -		
Didactic skills -		
Other -		
2. Teaching activities		
	Year	Workload (Hours/ECTS)
Lecturing -		
Supervising practicals and excursions -		
Supervising Master's theses -		
Other -		
Total		106 hours

Acknowledgements

It is a pleasure for me to thank to all people who made this thesis possible. First of all, I would like to thank my promoter Frans Rutten and copromoter Maiwenn Al for their inspiring and encouraging way to guide me to a deeper understanding of health economics and for their useful comments during the whole work with this dissertation.

I gratefully acknowledge the funding received to pursue a PhD from Novartis Pharma Ag. Thanks to Jean-Francois Ricci for his encouragement and supervisory role and also Jens Grueger who made possible to obtain the grant.

I would like to express my gratitude to Heinz Schneider and Rainer Rohrbacher, whose support and expertise have influenced me greatly. They offered me a position as a health economics consultant; they taught me to think from an economics perspective and motivated me to deepen my knowledge in methodological research. I would particularly like to thank my colleague and friend, Elisabeth Brock, who has encouraged me through this entire process.

I wish to thank my parents, Smaranda and Gheorghe, for their constant support and encouragement in all my professional endeavours. Finally, I wish to express my gratitude to my husband Andrea and our daughter Stephanie for their continuous moral support.