Chapter 9 General Discussion
General Discussion

The aim of the research presented in this thesis was to critically describe the current state of economic evaluations of preventive interventions, and to contribute to realistic economic evaluations in this field by performing original research. This thesis is organised into two parts. The first part addresses the current state of economic evaluations of preventive interventions, based on reviews of the literature. In the second part of the thesis, original economic evaluation studies are presented. In this general discussion a summary and synthesis of the main findings are presented. Furthermore, the results are put in a broader perspective and the implications of the studies for policy and practice are discussed. Also, directions for future research in the field of economic evaluation of prevention are provided.

Part one: Current state of economic evaluations in prevention

No relationship between number of economic evaluations and burden of disease

Chapter 2 of this thesis presents an overview of all economic evaluations in the field of prevention published in 2008. In this study, not only the number of economic evaluations conducted, but also the values of the Incremental Cost Effectiveness Ratios (ICERs) reported and the relation with the worldwide burden of disease were investigated. About 80% of all published economic evaluations in the field of prevention presented an ICER below €50,000 per Quality Adjusted Life Year (QALY). The median ICER was €12,500 per QALY. So, the majority of public health interventions seem to be cost effective. However, this result may also be a reflection of selection bias and/or publication bias. When looking at the correlation with the global burden of disease, it appeared that relatively few economic evaluations were reported that were aimed at diseases that cause a large proportion of the total burden of disease worldwide, and vice versa. For instance, relatively little attention is paid to a disease like diabetes mellitus type 2, that is predicted to constitute a large disease burden in the near future. We found relatively many economic evaluations in the field of infectious diseases prevention. A lot of country-specific papers are being published that evaluate new vaccines. Clearly, the interests of pharmaceutical companies act as a
boost here. Finally, many economic evaluations of cancer screening programs were published. This may be due to the high priority of this type of screening in developed countries. There are also many data available from well-organized cancer registries that are suitable to be incorporated into models. The findings in our study on the median ICER and the focus on certain disease areas were confirmed in other recent studies (Bell et al., 2006; Catala-Lopez et al., 2011; Owen, 2011).

**Assumptions regarding adherence to preventive interventions are often too positive**

In chapter 3 we present a review of economic evaluations of two different types of colon cancer screening (colonoscopy and CT-colonography). In these modelling studies we assessed the assumptions regarding test characteristics (sensitivity and specificity) and adherence of the invited population. It appeared that the assumptions regarding test characteristics were in line with those found in the literature. However, this was not the case for adherence. In the studies included in our review, adherence ranged from 50% to 100% in the base case analyses, while international literature indicates adherence rates below 30% for colonoscopy and CT-colonography. Based on these findings, we conclude that researchers seem to systematically overestimate adherence in the base case analysis. However, more than half of those studies used adherence rates in the more empirical range in sensitivity analyses. A consequence of an overestimated adherence is an overestimation of the population health impact of the modelled screening program, and hence a more favourable ICER. The ICER is too favourable, because the estimated net costs of colon cancer screening (investments minus savings – the numerator of the ICER) will be too low and the estimated number of life years gained by screening (the denominator of the ICER) will be too high. In other words, unrealistically high assumptions for adherence results in lower net costs and more health gained, while a more realistic assumption for adherence results in higher net costs and less health gained. Therefore, it can be concluded that using realistic percentages of adherence in the model results in less favourable ICERs. Nevertheless, a higher ICER does not always imply that an intervention is no longer cost-effective. As long as the ICER does not exceed the limits agreed upon, the
intervention remains cost-effective. However, in order to make a useful contribution to decision making, it is of importance to calculate ICERs that correspond with everyday practice.

Costs of adverse events are often excluded in economic evaluations
Chapter 4 describes a review of the cost-effectiveness of screening for abdominal aortic aneurysm (AAA). The economic evaluations that were published based their input parameters on several large trials. In these trials, effects on AAA-related mortality, mortality by other causes, the volume of elective surgery and that of acute surgery were reported. Perioperative mortality was reported in the clinical trials and hence included in the economic evaluations. They did however not report on adverse events and postoperative morbidity following elective surgery. Postoperative morbidity might however be substantial, as indicated by several studies comparing elective surgery by the open surgical technique with the endovascular technique (EVAR), showing that there is a risk of serious complications (5%-10%) with both techniques (Lederle et al., 2009; Prinssen et al., 2004; Greenhalgh et al., 2004). This is in line with observational studies that report morbidity after surgery of an unruptured aneurysm, such as kidney disorders requiring dialysis, stroke, myocardial infarction, leg amputation and COPD (Johnston, 1989; Blankensteijn, 2003). Such an omission will effect the outcome of an economic evaluation. Therefore, model calculations also need to consider the costs as a result of those complications, as well as the impact on quality of life and survival. It is for instance not reasonable to state that the quality of life after surgery is equal to that of other people of the same age and sex (Lee et al., 2002; Legemate & Bossuyt, 2006). So, again too few costs are being considered in the numerator of the ICER and too much effect (overestimation of quality of life) in the denominator, which leads to a more favourable ICER that does not correspond to today’s practice. When these data are included correctly in the model, the question arises whether screening for AAA is still cost-effective (Ehlers et al., 2008). Subsequent to our review, two new economic evaluations were published (Lindholt et
al., 2010; Spronk et al., 2011). One of these studies did incorporate postoperative morbidity in the model (Lindholt, 2010), and concluded that AAA screening indeed was less cost-effective than reported before. Accurate registration of complications is necessary to accurately determine the rate of complications after surgery (Kazmers et al., 1996). Only when all benefits and harms are being considered in the model, an economic evaluation can present a realistic estimation of the cost-effectiveness of preventive interventions.

In the original review we did not concentrate on adherence to screening. However, for this thesis we also investigated the assumptions on adherence in the economic evaluations. That analysis revealed that all economic evaluations of AAA screening used realistic adherence rates, based on findings from clinical trials, ranging from 68% to 80%. The conclusion therefore is that the assumptions regarding adherence were in line with the available literature.

**Economic evaluations from the societal perspective do not include all societal costs**

In chapter 5, we present the results of a review of the different cost parameters included in economic evaluations of preventive interventions for alcohol abuse. We compared these costs to those considered in cost of illness studies regarding alcohol abuse. Although guidelines for economic evaluations state that both direct health care costs and direct and indirect non-health care costs should be incorporated in evaluation studies from a societal perspective, this advice is not always followed. We found in our study that the costs reported in economic evaluations were not in line with the costs considered in cost of illness studies, even when they were conducted from the societal perspective. Economic evaluations generally take fewer cost categories into account than cost of illness studies. The most important difference concerns the costs of traffic accidents related to alcohol, which are rarely considered in economic evaluations. No economic evaluation was found that included all societal costs. A recent publication came to the same conclusion (Barbosa et al., 2010). By not considering all possible costs, the ICER is unfavourably influenced. This could lead to unjust rejection of decision makers to implement such an intervention. The previously
described tendency of economic evaluations showing ICERs that are too favourable is therefore not always true.

Part two: Case studies on the economic evaluation of preventive interventions

Opportunistic screening in primary care and preventive medication with a polypill seems cost-effective

The first modelling study focused on the cost-effectiveness of a polypill in the primary prevention of myocardial infarction and ischemic stroke. A polypill is a combination pill that contains a number of drugs that reduce blood pressure and blood cholesterol levels. We estimated that opportunistic screening based on cardiovascular risk assessment followed by polypill prescription to those with increased risk offers a cost-effective strategy, with ICERs of about €10,000 per QALY gained. The assumptions regarding its effects on blood pressure and cholesterol levels were based on a randomized controlled trial (RCT) (Yusuf et al., 2009). The magnitudes of these effects were confirmed by a recent trial performed in seven countries (PILL Collaborative Group, 2011). Both trials were of short duration, with only 3 months of drug treatment. Therefore we had to make several assumptions for our study. A assumption particularly difficult to estimate appeared to be the response rate of the potential target group to a screening invitation. There are simply no relevant data available on which to base such an estimate, because opportunistic screening and subsequent preventive treatment with a polypill has never been subject of an empirical study. Therefore, we conservatively assumed a response rate to a screening invitation of 50% in the base case analysis. Response rates and adherence to treatment in general are important issues in economic evaluations of preventive interventions, as appeared already from the first part of this thesis. However, in this case our sensitivity analyses showed that response rate and adherence had only a minor influence on the ICERs, while it did have an impact on both projected health benefits and costs.
Resources for COPD should best be allocated to smoking prevention

The second case study is a modelling study on optimal resource allocation to different strategies for COPD prevention and treatment. Resource allocation models give the opportunity to allocate resources as efficiently as possible, taking into account clinical needs, financial constraints, and cost-effectiveness rankings based on optimal values for an entire population (Granata & Hillman, 1998; Earnshaw et al., 2002). As described before, economic evaluations present their results in ICERs, which represents the costs per unit of the outcome measure. However, a low-cost intervention targeting many people may ultimately be more expensive than a costly intervention applied to a limited group of users. Therefore, resource allocation models combine data from economic evaluations of different interventions with demographic and epidemiological data and budget constraints to estimate the optimal allocation of resources over the various interventions (Hoogenveen et al., 2005; Feenstra et al., 2011). The results of this modelling study showed that if the aim of policy makers is to obtain maximal health benefits for a given COPD-budget, all resources should be allocated to smoking cessation.

In this budget allocation study the assumption on adherence to the different interventions was the most uncertain factor. For instance, adherence to a smoking cessation intervention in the general population was based on estimates of Stivoro (Dutch expert centre for tobacco control). Compliance to medication was based on the observed drop-out rates in trials. However, this may be an overestimation because it is known that compliance decreases in long-term interventions and that compliance is generally higher in clinical trials than in daily practice (De Geest & Sabaté, 2003; Sabaté, 2003). So in order to perform a valid budget allocation study, not only effectiveness of interventions but also the long-term willingness to participate in interventions need to be investigated further.
Willingness to participate in a lifestyle intervention does not seem to depend on financial incentives

The third case study concerns the willingness to participate in a hypothetical lifestyle intervention. The results were interesting, in that financial compensation appeared not to lead to an increase in willingness to participate, while mandatory financial contribution led to a decrease in the willingness to participate. All the other attributes of the lifestyle intervention, such as time spent on the intervention, arrangement of the lifestyle intervention program and the kind of physical activity offered were not significantly related to the willingness to participate. Furthermore, our study showed that although respondents graded the hypothetical lifestyle programs as rather good (mean grade of almost 7 out of 10) this positive score was not associated with a higher willingness to participate in the lifestyle intervention. These results show how important it is to gain knowledge of preferences of potential target groups of preventive interventions. Although there are indications that under certain circumstances, such as in work-related life style interventions, a positive financial stimulus may have a positive effect on willingness to participate and adherence (Taitel et al., 2008; Merrill et al., 2011), the influence of financial incentives on changing behaviour generally seems small and only short-term (Kane et al., 2004; Paul-Ebhohimhen & Avenell, 2008; Sutherland et al., 2008; Marteau et al., 2009; Jochelson, 2007). So, the influence of financial incentives needs further study.

Prevention and economic evaluations

Many national and international guidelines have been developed for performing an economic evaluation. Currently those guidelines address economic evaluations in the field of health interventions in general and not prevention in particular. Because of the specific characteristics of preventive interventions, guidelines should be developed for performing economic evaluations specifically directed at the field of prevention. Yet, several researchers have discussed how an economic evaluation in the field of prevention should be performed (Cohen & Patel, 2009; Weatherly et al., 2009; Kelly
Outcomes of public health interventions reach further than health benefits exclusively, and stretch out to other fields. Examples are: improved learning performance (by prevention of binge drinking in youth), improved social well being or participation (by community-based interventions), and less crime (by interventions directed at drug abuse). The classic CEA and CUA do not account for these broader effects, while the costs (or savings) related to these broader outcomes are sometimes included in the numerator of the ICER. In CUAs, the outcome measure is quality adjusted life year (QALY), an outcome measure that includes only health benefits in terms of survival and quality of life. For other effects of an intervention, QALY-like outcome measures could be developed. Weatherly refers to these outcome measures as “criminality QALYs” or “learning performance QALYs” (Weatherly, 2009). It has also been suggested to define productivity benefits (reduced sick leave and improved concentration at work) in PALYs (productivity adjusted life years) (Eysink et al., 2007). Since the objective of an economic evaluation from a societal perspective is to give a complete assessment of all costs and benefits it is more than worthwhile to investigate the applicability of these suggestions. Another option to incorporate the broader effects of prevention is to use a cost benefit analysis (CBA). In a CBA all costs and benefits are expressed in monetary units (Drummond et al., 2005; Weatherly, 2009; Cohen & Patel, 2009; Popovici et al., 2008; Kelly, 2005). CBA has the advantage of increasing comparability of different types of interventions. To acquire complete knowledge, it is important that all costs and benefits are expressed in monetary units, and not to resort to pro memory posts as is often done in societal cost benefit analyses (SCBA). SCBA with multiple pro memory post are uninformative for decision makers (Lubbe & Larsen, 2009; Spreen & Mot, 2008). Several SCBAs with health as one of the outcome measures used a monetary value of the QALY that is higher than the usual value of €20,000 that is often quoted in the Netherlands. In preventive care, a QALY is often valued at €20,000 (van den Berg et al., 2008). However, in a SCBA regarding health in a problem area in The Hague, the QALY was valued at €50,000 (Lubbe & Larsen, 2009) and in a SCBA regarding the banning of smoking in all catering facilities the QALY was even valued at €100,000 (Spreen &
Mot, 2008). So, there is no uniformity in the valuation of the QALY, which limits the usability/comparability of SCBAs. Recent willingness to pay studies indicate values close to the thresholds of €20.000 (Bobinac et al., 2010; Shiroiwa et al., 2009), but also conclude that the value strongly depends on the severity of the condition, age and income. Against this background, a fixed threshold for the QALY seems too rigid. However, as the monetary value of a QALY heavily affects the outcome of a SCBA, more empirical evidence on the value and range of the QALY would be welcome.

**Implications of our findings**

*Guidelines for economic evaluations and publications*

It is impossible to conduct a modelling study without making assumptions. When a modelling study is conducted to support policy making, e.g. to decide whether an intervention should be carried out or not, it is of major importance that the outcomes of the study give a reliable picture of what can be expected from the intervention in a real life setting. To meet those conditions, the assumptions made in the study should represent daily practice as realistically as possible. It should be clear to the user what assumptions were made and on what they were based. When conducting an economic evaluation, more attention should be paid to elements such as adherence, compliance and long term maintenance of the effects (Vos et al., 2010). Furthermore, positive as well as negative effects of an intervention should be addressed. Uncertainty concerning the assumptions in the model should be reflected in the sensitivity analysis (Savelkoul et al., 2011). In an influential guideline for conducting and evaluating economic evaluations, it is stated that the model used should be available to the reviewer (Drummond & Jefferson, 1996). It is recommended to attach the complete model and all input parameters to the publication, as an appendix or as a webpage. This is already standard in other fields of research such as computational neurosciences (http://senselab.med.yale.edu). It will also help users of the information such as policy makers or other researchers, who can make a better assessment or decision the more comparable and complete the information (Severens, 2011; Chaiyakunapruk et al., 2011).
Adherence

As shown in the first part of the thesis, adherence is often overestimated leading to more favourable ICERs. However, in our own studies (second part of the thesis, polypill and COPD studies) it appeared that adherence is a difficult factor to validate. Therefore, it is advisable that research into the preferences of potential participants in new interventions is done in advance of the economic evaluation. This could make assessments of cost-effectiveness prior to implementation more valid and better in line with (future) real life practice. Actual participation should be monitored regularly, which allows for adaptation of the intervention based on empirical data. Based on preferences of potential participants regarding various characteristics, knowledge can be acquired about the level of adherence that can be expected and the conditions for successful implementation of a new intervention. Especially with regard to large investments in public health programs, it is essential to study the preferences of the target group prior to implementation.

Adherence to interventions is an important factor in the economic evaluation of prevention. It often affects the ICER, the budget impact and the health benefits of interventions. Too optimistic assumptions of the willingness to participate may lead to too optimistic conclusions, while with more realistic assumptions the conclusions about the cost-effectiveness would be less favourable. For example in case of colorectal cancer screening, an economic evaluation showed that a decrease in adherence from 100% to a more realistic figure of 30% resulted in a doubling of the ICER, from €21,000 to €42,000. However, the ICER remained below the limit of €45,000, as agreed upon in the USA (Sonnenberg et al., 1999) and it probably had limited influence on the decision of introducing it or not. Sometimes the influence of adherence is smaller, for example in case of our polypill study. The relatively small impact of adherence in this study is related to the fact that this is a one-on-one intervention, for which no major program costs have to be made, as is the case for population-wide screening interventions). Uncertainty in adherence assumptions
should be taken account of by good sensitivity analyses, which are crucial for a proper and useful interpretation of the results of an economic evaluation (Armstrong, 2010; van den Berg et al., 2011).

Implications of this research for policy and science

The results of this thesis imply that policymakers should not blindly rely on the results of an economic evaluation. When they want to use an economic evaluation to support their policy decision they have to realize that ICERs harbour uncertainties (van den Berg et al., 2011). Although most economic evaluations present the sources of the assumptions made, policy makers have to take into account that the effectiveness assumptions are generally based on short-term efficacy studies while in fact reliable insight in long-term effectiveness is needed. Long-term effectiveness refers to the maintenance of effects over time, for example does a vaccine provide with life-long protection or is a booster vaccination needed after some time? Long-term effects depend, amongst others, on compliance; decreased compliance in general negatively affects on effectiveness. In addition, assumptions made about costs may have an influence on the estimated ICER. Therefore, at present it is advisable that policymakers seek assistance from experts in the field of health technology assessment to obtain insight in the uncertainty and sensitivity of these assumptions and to fully understand the meaning and value of the ICER and the decision uncertainty surrounding the different policy alternatives.

Furthermore, this study shows that it is important that economic evaluations include empirical data on the factors that are incorporated in models, such as data on adherence and on negative side effects of the intervention. In order to obtain these data, we suggest using pilot studies that provide these data before widespread implementation of the new intervention is planned. Especially when implementation of large public health interventions is considered, it is necessary to have accurate information. The relevance of this can be illustrated using the case of Chlamydia trachomatis screening in young adults. The results from a relative small trial to determine Chlamydia trachomatis prevalence organised by the Municipal Public
Health Services showed a participation rate of 41% (van Bergen et al., 2005). However, a large pilot implementation project later showed an adherence of only 16% after the first invitation, declining to 9.5% after the third invitation (van den Broek et al., 2012). Following this limited participation of the target group in the intervention, cost-effectiveness could no longer be demonstrated. This example underpins the importance of using real-life data for important parameters of cost-effectiveness models.

For researchers, this may imply that they should restrain themselves and wait with an economic evaluation until accurate empirical data are available. This not only increases the quality of economic evaluations but also its value to policymakers.

**Conclusion**

For a long time, evidence regarding effectiveness was considered a sufficient base for decision making regarding the implementation of preventive interventions. By the end of the last century, cost-effectiveness was added as an important condition for decision making. Economic evaluations assess the cost-effectiveness of interventions. To do so validly, they should be performed with the best available methodology, using empirical data on cost and effects reflecting daily practice. To acquire a higher degree of comparability and transparency, new guidelines should be developed (and then followed) specifically directed at economic evaluations of preventive interventions. Evaluation studies are only useful as input for decision making in policy if the economic evaluations use data that reflect real practice. If this information is not available, the assumptions that are used in the model should be transparently described and sensitivity analyses should be performed. The internet provides ample opportunity to meet these requirements through web based supplements.

It is important to be able to make reliable statements about the cost-effectiveness of a preventive intervention prior to its implementation, to avoid unjust decisions. More behavioral research should be conducted to enhance our knowledge about preferences of participants and how they affect cost-effectiveness of the intervention. It is
advisable to continuously monitor the intervention during its implementation, as it would allow us to adapt the intervention to new insights and needs.

References


van den Berg M, Hamberg-van Reenen HH, Schoemaker CG. Kosteneffectiviteit: euro's per QALY is niet genoeg. Ned Tijdschr Geneeskd, 2011; 155; A3817


