National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Modelling Future Mortality in Ontario: Extension of the PREVENT Model and Development of an Ontario Database

by

Margaret Elizabeth Herbert

Thesis submitted to the School of Graduate Studies and Research in partial fulfilment of the requirements for the M.Sc. degree in Epidemiology

University of Ottawa

December 1994

ISBN 0-612-00583-6

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

PREVENT is a user-friendly, interactive computer program developed in the Netherlands which integrates the effects of demographic structure, risk factor prevalence, relative risks and mortality rates to estimate the impact of changes in risk factor prevalence on mortality. Using relative risks from large international studies, the program uses Population Attributable Risk to model future mortality with and without simulated interventions on 6 risk factors affecting 3 diseases.

A new database for PREVENT has been developed with the following Ontario data: population structure, all-cause and disease-specific mortality, life expectancy, and birth projections. The current prevalence of smoking, obesity, alcohol and seat belt use is estimated from the Ontario Health Survey; the prevalence of hypertension and hypercholesterolemia from provincial heart health surveys. Future changes in risk factor prevalences for the simulation period are extrapolated from past trends.

Over its maximum 50 year simulation period, PREVENT estimates that a 50% increase in the Ontario's youngest smoking cohort would result in an excess 1574 deaths per year after 50 years. A 10% reduction in both hypertension and high cholesterol over 5 years would result in a peak reduction of 3279 deaths per year due to Ischemic Heart Disease (IHD). A 95% or more use of seat belts achieved in one year would prevent only 57 deaths in the first year but over the 50 year simulation would accrue over 2000 Actual Years of Life Gained. A 50% decrease in alcohol consumption over 10 years is modelled to impact on four causes of death, including

IHD with its high mortality. In spite of this very large intervention, at the end of the 50 year simulation period the improvement in life expectancy at birth is just 0.5 years. A 50% decrease on obesity among older women results in less than 100 breast cancer deaths prevented per year.

Sensitivity testing shows that the mortality varies directly with both the prevalence and relative risks. The shape of the mortality prevention curve is affected by two factors. In the first twenty years the latency and lag times and the spread in implementation of the intervention dictate the pattern. In later years mortality is affected by the changing demographic structure, age-specific mortality rates and future trends in risk factor prevalence that are in the model.

Historical testing of the risk factor smoking, using a 1966 Ontario database, shows good agreement between PREVENT estimates and observed total mortality. The agreement between predicted and observed mortality is not as close for lung cancer, IHD and COLD.

ACKNOWLEDGEMENTS

There were many people whose suggestions, advice and support have helped me in this project. I would like to specifically acknowledge and thank several. My family supported my decision to undertake an academic program and have continued to support me in this research project.

This research was undertaken as a project of the Community Health Research Unit (CHRU). I am grateful for their financial support in the form of a student assistanceship. It is through another CHRU project that I was able to obtain detailed data from the Ontario Health Survey. Richard Aylesworth and Jack Boulet were very helpful in extracting that data in the format required.

Louise Gunning-Schepers was most gracious in our correspondence by giving me detailed information concerning PREVENT and by supporting the development of the Ontario database.

I especially thank Dr. Robert Spasoff for his patience, consistent support and practical guidance in supervising this endeavour.

Table of Contents

List of Tables

List of Figures

1.0 INTRODUCTION

1.1 Public Health Policy and Planning

Increasingly scarce health resources demand emphasis on programs with maximum health benefit for the population. Among other things this emphasizes the need to identify and address important health "targets" and to choose cost-effective programs. For some time both the federal and provincial health policy papers and goal statements have advocated change in this direction [1 2 3]. These policy papers also indicate that programs of disease prevention will be an important avenue to achieving these objectives.

The emphasis on maximization of benefit has resulted in corresponding interest in the evaluation of health outcomes and produced a demand for predictive information on which to base program selection and funding decisions. With increasing competition for limited resources, advocates of disease prevention are now expected to support their proposals with statements of expected health benefits.

Programs for disease prevention are at a competitive disadvantage compared to programs of disease treatment for the following reasons: the benefits are usually realised only in the long term, many people are required to take action that will benefit only a few[4] (paradox of prevention), they involve non-occurrence of diseases so that there is no change in health status for easy identification of a positive outcome, and major multifactorial intervention trials[5,6] have had disappointing results in terms of overall mortality. If proposals for disease prevention are to be successful in attracting funding they must apply what is known about risk factor modification in

the population and realistically estimate expected health benefits. The epidemiologic methods required to estimate future disease incidence based on past data are still in development.

Chronic disease epidemiology focuses on identification and quantification of risks and causes. Factors with significant association to a disease incidence are labelled as hazardous (or protective). The confidence of such labelling is enhanced if there is a plausible biological explanation, consistent and strong association, and a gradient of higher risk associated with larger exposure. For important chronic diseases such as ischemic heart disease (IHD) there has been extensive investigation of possible risk factors. The Framingham Study of cardiovascular disease [7] and the British Physicians' Study of lung cancer [8] are classic examples of investigations which have contributed fundamental epidemiologic data on risk factors and their relationship to non-infectious diseases. The model for lung cancer illustrates the simple case of one risk factor emerging as the overwhelming determinant of the disease. The case of heart disease is more complex, with at least three major risk factors (hypertension, smoking and hypercholesterolemia) as determinants.

With identification of the determinants of heart disease came intervention trials, notably MRFIT, the WHO collaborative trial and the North Karelia Project [4,5,9]. While all of these studies showed conclusively that risk factors could be changed, all of them had disappointing results in terms of health benefits (particularly prevented mortality) compared to reference groups. At the same time, studies such as Framingham and the Alameda county project [6,10], were demonstrating a significant relationship

between health status and lifestyle (risk factor behaviour). Even though underlying biological mechanisms linking behaviour and disease are not always well defined, the results of such studies have had major impact on the public's concern about health and may have resulted in the gradual improvements seen in prevalence of some risk behaviours.

Unfortunately, these and similar studies have had limited impact on resource allocation to preventive projects. One problem with preventive proposals is the inability to predict accurate outcomes. If increased support is to be gained for disease prevention and risk factor intervention, it will be due to realistic expectations of outcomes that are estimated in advance using sound epidemiological techniques [11]. In other words, reliable population health models will have to be developed, accepted and used.

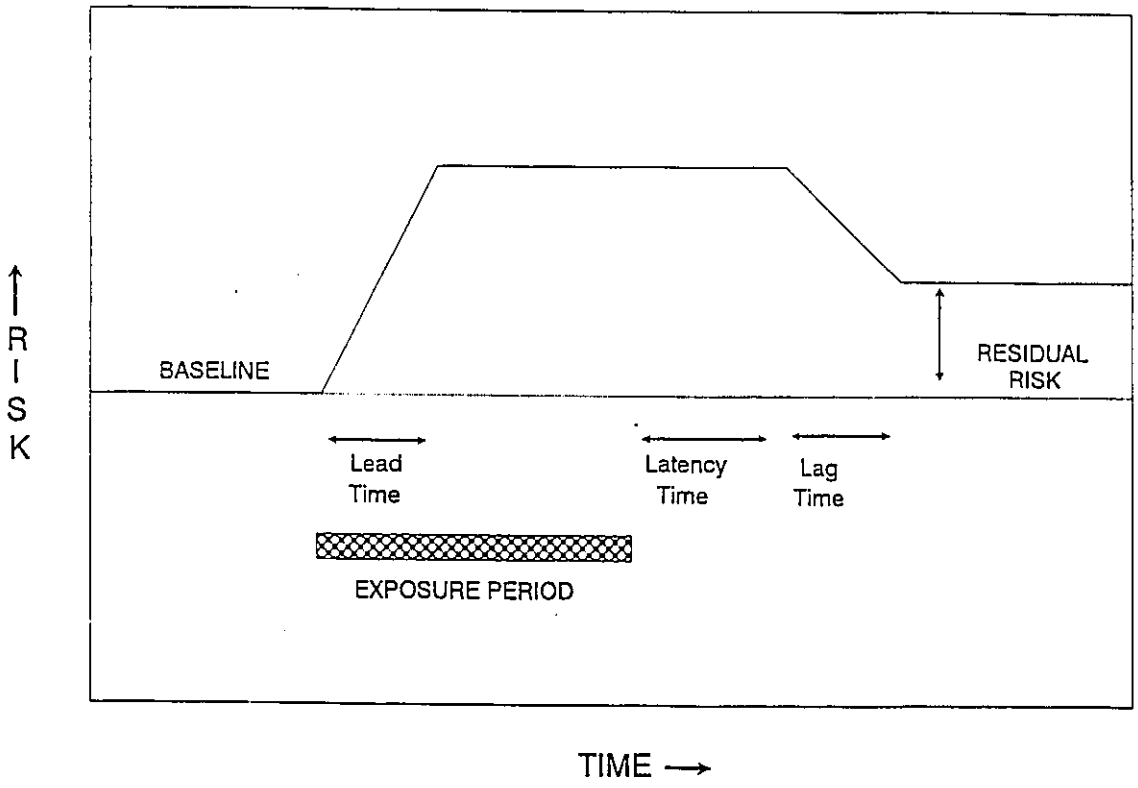
The PREVENT program [12], developed by Dr. Louise Gunning-Schepers in the Netherlands, uses information for specific populations combined with epidemiologic measures of population risk to estimate the future influence of hypothetical interventions within the specified population. The PREVENT program is the basis of my thesis work and I have focused on developing a database for the population of Ontario.

1.2 Epidemiologic Techniques to Estimate Interventions

In order to obtain realistic estimates of the benefits of preventive interventions, models should include the following characteristics. A multifactorial facility is

required whereby one risk factor may impact on several diseases and one disease may be influenced by several risk factors. Time dimensions over which risk magnitude increases upon exposure and reduces upon cessation are needed: specifically, the lead time (period between beginning of exposure and maximum risk), the latency (period between cessation of exposure and initial reduction in risk), and the lag time (period after cessation of exposure over which risk gradually declines to a residual level). These time dimensions are presented in Figure 1. Changing trends in the population

Figure 1 Time Dimensions and Individual Risk Patterns Associated with Exposure



prevalence of risk factors over the simulation period need to be considered. A demographic element which adjusts for changes in population structure over the estimation period is needed. Ideally, outcomes related to both morbidity and mortality should be generated. Finally, cost estimates of health services and appraisal of economic impact are useful.

In order to estimate the impact of an intervention it is first necessary to identify:

i) modifiable risk factor(s) for one or more diseases to intervene upon

ii) a specific target population

iii) meaningful outcome measures

The identification of the risk factor(s), the target disease(s) and the target population is the first step in estimation of intervention outcome.

Next is the choice of methods to estimate the impact of changes in risk factor prevalence (resulting from intervention) and the choice of outcome measure(s). Disease incidence and disease-specific mortality can be determined from a population's morbidity and mortality data. Other health outcome measures such as Potential Years of Life Lost (PYLL) or Gained (PYLG) and life expectancy may be derived from them. Hypothetically, either morbidity or mortality might be used to estimate future disease levels after an intervention. In Ontario the mortality data are far more accurate, in spite of errors of misclassification on death certificates [13]. With the exception of data from cancer registries, morbidity data do not allow an accurate measure of chronic disease incidence. There are two main reasons: hospital data will

only identify incident cases of sufficient severity to warrant hospitalization, and hospital and OHIP data will not identify incident versus follow up cases. Mortality is the only primary outcome that remains feasible for Ontario data.

Another element necessary in the methodology is the identification and quantification of risk. The epidemiologic measures of risk are generally referred to as Relative Risk or more specifically Incidence Density Ratio, Cumulative Incidence Ratio, and Odds Ratio. Causative factors of a number of diseases have been identified and risks quantified by investigators worldwide. Among these are important chronic diseases, ischemic heart disease , chronic obstructive lung disease (COLD) and cirrhosis as well as several malignancies, including lung cancer[14]. Many risk factors identified for these diseases are modifiable. Literature reports from large, well controlled studies show reasonable agreement on the overall magnitude of risk for many of these risk factor-disease pairs. Unfortunately, there has been very little research into comparative risk across age and gender subgroups or exposure levels.

In estimates of population health, relative risks derived from reliable research are assumed to be generalizable to other populations. From the public health perspective, the relative risk alone, no matter how large, is of limited importance. The impact of the risk factor and disease on the population are reflected by a combination of the relative risk, the prevalence of the risk factor in the population and the mortality rate for the disease. An epidemiologic measure of the importance of a risk factor in the population under study is required. The Etiologic Fraction or Attributable Fraction (Population) is such a measure[15], reflecting both magnitude of

the relative risk and prevalence of the risk factor in the population.

$$AF(P) = \frac{p(RR - 1)}{p(RR - 1) + 1}$$

where AF(P) is Attributable Fraction (Population)

RR is Relative Risk

p is proportion of population exposed to the risk

The Population Attributable Risk (Percent) is defined as the attributable fraction in the population, expressed as a percentage[16].

Over the last twenty years investigators have considered the possibility of combining some measure of Population Attributable Risk with other measures to estimate impact of risk factor changes on health outcomes in the population.

In 1977, Sturmans et al.[17] recognized that Attributable Risk Percent can be used to estimate reduction in incidence as a consequence of intervention. They calculated the impact of hypertension, hypercholesterolemia and smoking, both marginal and jointly, on the incidence of coronary heart disease. Calculations of joint impact assumed a multiplicative interaction of risks. They used risk data from the Framingham Study and projected outcomes of possible interventions for the male population of that study.

In 1979, Ouellet et al[18] estimated the percentage of premature mortality and the potential years of life lost due to smoking and alcohol use in the Canadian population. The calculations were based on the Attributable Fraction using fixed relative risks and the distribution of the risk factors in the population. The authors identified their lack of adjustment for lead time (period between first exposure and development of full risk) and the possible interaction between risk factors as important limitations. They

concluded that the available data were insufficient to take these factors into account at the time.

Walter[14] suggested that attributable risk calculations might serve as an important element in the choice of preventive programs. Morgenstern and Bursic[19] proposed that epidemiologic research on risk factors be used to estimate the "Potential Impact" of risk factor intervention. They define the "Potential Impact Fraction" (PIF) as the proportional reduction in the total number of new cases of a disease resulting from a specific change in distribution of a risk factor in the population at risk. This measure is closely related to the attributable fraction.

$$PIF = \frac{(p - p')(RR - 1)}{p\,(RR - 1) + 1}$$

where PIF is Potential Impact Fraction
RR is Relative Risk
p is proportion exposed before intervention
p' is proportion exposed after intervention

The Potential Impact Fraction is used in the PREVENT model.

Health Risk or Health Hazard Appraisal[20] is a related application, focused on the individual. Most health risk appraisal methods are computer-based analyses of the cumulative risks experienced by an individual expressed as a probability of disease occurrence or a life expectancy. The purpose of these appraisal methods is generally for health education on an individual level and as an incentive for the individual to modify risk behaviours.

1.3 Computer Models and Population Health

The magnitude and complexity of calculation required for multifactorial estimation of individual and population health outcomes led to computer applications.

Table 1 Attributes of Population Health Models

MODEL	PRE-VENT	CAN*TROL	SAMMEC II	ARDI	CHOLES TEROL	CHD	PAR UBC	CRISP-ERS	GoM/ MANTON	POHEM
METHODS										
PAR	✓	✓	✓				✓			
Monte Carlo		or						✓	✓	✓
Other		✓		✓	✓	✓				
TIMES										
Latency	✓	✓			✓	✓	No	✓		✓
Lag	✓						No	✓		✓
Lead		✓					No	✓		✓
MULTIPLE DISEASES	✓	✓	✓	✓	No	No	✓	No	✓	✓
MULTIPLE RISKS	✓	✓	No	No	✓	✓	✓	✓		✓
COST/BENEFIT ANALYSIS		✓	✓	✓	No	✓	No			✓
AUTOMATIC TREND ADJUSTMENT										
Population	✓	✓	No	No		✓	No	✓	✓	No
Risk Factor	✓		No	No		No	No			✓
OUTCOME MEASURES										
Morbidity										
Mortality	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PYL(G/L)	✓	✓	✓	✓	✓	✓		✓	✓	✓
Others	✓	✓	✓	✓				✓		✓
		✓	✓	✓				✓	✓	✓
OPERATES ON PERSONAL COMPUTER	✓	✓	✓	✓		No	✓	No	No	✓
MODELS INTERVENTIONS	✓	✓	No	No	No	✓	No	✓	No	No
MODEL FLEXIBILITY										
Intervention Types	✓	✓				✓		✓		
Intervention Periods	✓	✓						✓		
Populations	✓	✓	✓	✓		✓		✓		
MODELS FORWARD ≥ 10 YEARS	✓	✓	No	No	✓	✓	No	✓	✓	✓
COHORT OPTION	✓									
SENSITIVITY ANALYSIS REPORTED	✓	✓						✓		
HISTORICAL TESTING REPORTED	✓				✓					

Blank boxes indicate insufficient information or not applicable to model.

Computer capacity allows researchers to develop complex and integrated health

models. This section reviews some of these models and Table 1 presents a comparison.

1.3.1 CAN*TROL Model

One of the most comprehensive population health models is CAN*TROL[21], a computer model for designing national cancer control strategies. CAN*TROL was developed at the Center for Health Policy Research and Education in North Carolina and has been used to model cancer outcomes for several national populations and many subpopulations. CAN*TROL contains five interactive components: a population structure model which uses a specified current population and models its structure into the future, a cancer incidence model which estimates the number of cases each year based on relative risks and risk factor prevalence, a screening and detection model which estimates the proportion of cancer found at particular stages, a treatment and support model which calculates survival of patients treated at various stages, and an 'other cause' mortality model which calculates the number of deaths from other causes. CAN*TROL has been used to simulate combinations of prevention, screening, treatment and support on future outcomes of cancer incidence, mortality, quality of life and cost. The model appears to use either a population attributable risk or linear regression method to estimate risk factor impact in the cancer incidence model component. The model has many advantageous features: it models multiple cancer types and multiple risk factors, includes lead times, automatically estimates changes in population structure in the future, includes a wide variety of outcome measures, can be adjusted to model many populations, contains

estimates of cost effectiveness, and can be run on a personal computer. One advantage that CAN*TROL has over other population models such as PREVENT is the ability to model disease incidence. This advantage is due directly to accurate incidence data available from cancer registries.

1.3.2 SAMMECII and ARDI Models

Two population health models dealing with single health risks have also been developed in the United States. SAMMECII[22] (Smoking Attributable Mortality, Morbidity and Economic Cost) was developed for the Office on Smoking and Health, Public Health Service, to model the impact of smoking on a comprehensive list of diseases. ARDI[23] (Alcohol-Related Disease Impact) was developed for the Center for Disease Control to model the impact of alcohol consumption on more than 30 diseases and causes of death. Both programs are available for personal computers and were designed to allow individual states to calculate outcomes and help them plan health policy initiatives. SAMMECII uses estimates of relative risks and risk factor prevalence to calculate population attributable risks. ARDI takes published estimates of alcohol attributable fraction and uses them directly. The programs estimate outcomes for mortality, morbidity and costs. ARDI also estimates non-health sector costs such as destruction of property and loss of productivity due to alcohol. Both programs have been designed for use by regional health workers. They operate on personal computers, require less than 1 Meg RAM memory, are menu driven, and have data files arranged in an interconnected series of easily edited spread sheets. Their primary purpose is to model the existing impact of the risk factor rather than to

model interventions. These appear to be static models for populations in the year of input, and do not simulate future outcomes.

1.3.3 Coronary Health Policy Model

There have also been computer models simulating the effect of multiple risk factors on one disease. The Coronary Heart Policy Model was developed at the Institute for Health Research and the Department of Health Policy and Management at Harvard School of Public Health[24]. It forecasts the future incidence, mortality and cost of Coronary Heart Disease (CHD) for the U.S. population as related to four risk factors; smoking, blood pressure, cholesterol level and weight. The model contains three submodels. The first is a Demographic-Epidemiologic submodel, which simulates the distribution of risk factors and the incidence of CHD in a demographically evolving population. This submodel stratifies the population by sex, one year age categories, dichotomous smoking status and three categories each of blood pressure, total cholesterol and weight. It uses risk factor prevalence from national surveys and relative risks from the Framingham study[25] in a logistic risk function to estimate incidence. Second is the Bridge submodel, which simulates the outcomes of incident cases in the first 30 days after the initial coronary event. The probability of death (in first 30 days) is estimated for this subpopulation stratified by sex, age, and incident event. Cost estimates are also made in this submodel. The surviving cases enter the Disease History submodel. The disease subpopulation is stratified by sex, age and 12 disease status categories, and estimates mortality outcomes to age 85. The following are the overall outcomes available from the model: the numbers and rates of CHD

and other causes of death, number of persons surviving to age 85 with and without CHD, the numbers and rates of CHD events, the number and prevalence of persons with CHD, and the resource costs of CHD. The major limitations of the model are: a mainframe computer is required for calculations; the assumption of independent distribution of risk factors has been made; the assumption of unchanged risk factor prevalence over the simulation period has been made; the input data for CHD incidence and case fatalities may not be accurate and/or nationally representative; and validity has yet to be established by accurate prediction of future outcomes. Initial predictions from the model show that although age-specific incidence rates will decline by 10% in the U.S. during the period form 1980 to 2010, the absolute number of people experiencing CHD will increase. This effect is due almost entirely to aging of the "baby-boomers" in the population.

1.3.4 Cholesterol Model

Another computer model estimating the effect of a single risk factor, total serum cholesterol level, on one disease, coronary heart disease (CHD) was developed for the Canadian population by Grover et al.[26] at the Montreal General Hospital and McGill University. Their purpose was to estimate the effects of life long risk modification on life expectancy and other outcomes. There are two submodels: the primary events submodel and the secondary events submodel. The primary events submodel estimates annual CHD events, CHD mortality and other cause mortality based on multivariate logistic regression risk coefficients from the Framingham study. The calculation varies with age, sex, diastolic blood pressure, serum cholesterol level,

the presence of left ventricular hypertrophy, the presence of glucose intolerance, and cigarette smoking. All-cause mortality is based on 1986 Canadian life tables adjusted for risks associated with smoking, glucose intolerance, and diastolic blood pressure as estimated in the Canada Health Survey. The probability of other cause mortality is calculated as the difference between all-cause and CHD mortality. The secondary events submodel estimates the risk of death within 12 months of a coronary event, based on Framingham data. The model does not generate estimates for the population as a whole but rather for individuals and subgroups with specific risk profiles, rather like health risk appraisal models. The program was evaluated using patient profiles from three major CHD primary prevention trials and accurately forecast the results observed over 5 to 10 years of follow up. The forecasted lifetime benefits of cholesterol reduction varied greatly depending on baseline levels, other risks, age and sex. Greatest benefits were forecast for cholesterol risk modification in young men with very high cholesterol levels. Benefits forecast for women were generally lower than for men with similar profiles. The model assumes that there is no increased risk of other cause mortality associated with reduction of cholesterol levels. Some of the limitations of this model are: the inability to model at a large population level, the focus on a single risk factor, and inability to adjust for future trends in factors such as smoking over the simulation.

1.3.5 PARUBC Model

There is another Canadian computer program known as PARUBC developed by Milsum and Jones from the Department of Health Care and Epidemiology at the

University of British Columbia[27]. This program uses personal computer spreadsheet software to determine the population attributable risks (PAR) of 8 modifiable risk factors impacting on 14 causes of death. The program uses the relative risks and population prevalence data from the Health Hazard Appraisal Implementation Project[28], as well as population structure by sex and three age categories, and finally mortality rates. The program also includes 13 non-modifiable precursors concerning family history, and clinical or environmental risks. The exposure to all risk factors and precursors has been dichotomized for simplicity. The input allows flexibility of population entered and computes outcomes of total and modifiable PAR, attributable mortality rates, and attributable mortality given population size. Graphic output is also included. The program makes the following assumptions: a multiplicative model of risk interaction, independent distribution of risk factors, and dichotomous exposure status. The program is limited by the following: no morbidity outcomes, no accounting for previous exposures (eg ex-smokers), a static model which estimates only for the year of input and does not simulate future outcomes. Initial runs with the program have estimated that the causes of death for which modifiable risk factors are most important are myocardial infarction and motor vehicle crashes with attributable fractions of 67% to 91%.

1.3.6 CRISPERS Model

A more computationally complex model has been developed by the Micropopulation Simulation Resource Center, Health Computer Sciences, at the University of Minnesota. The Chronic Disease Risk Intervention Simulation

Programs for Epidemiologic Research Studies, or CRISPERS[29,30], is a computer model for chronic disease. It uses generalized Monte Carlo techniques to simulate morbidity and mortality from a chronic disease, within a simulated population. Using a specified starting population, composed of a set of synthetically generated individual health profiles, the program simulates the future health development of that entire population. In CRISPERS, the health events of individuals are generated by a Monte Carlo technique based on assigned risk probabilities. The probabilities of various events for each individual are expressed as a cumulative distribution from 0.0 to 1.0. A random number is assigned for each individual-event possibility. The "event" occurs for individuals if the random number falls within a set range (determined by the overall event probability). Health outcomes, in terms of morbidity and mortality, are estimated by summing the individual health status of all individuals in the population. To date the model appears to have been used exclusively to model coronary heart disease. CRISPERS has the following advantages: it models the impact of multiple risks, it simulates a latency period between disease incidence and mortality, it adjusts for future demographic changes in the population, it provides outcome measures related to both morbidity and mortality, and it can model many years into the future. CRISPERS is limited by its ability to model only one disease at a time and by its complexity (which suggests mainframe computer capacity and expert assistance to operate). Although CRISPERS was not originally designed to model interventions, a prototype sub-program, named CRISPERT[31], has been developed to support intervention modelling within the main program.

1.3.7 Grade of Membership Model

Another model using microsimulation techniques had been developed at the Center for Demographic Studies at Duke University. Published research studies by Manton and Stallard [32,33] indicate that this program is used to model health outcomes in elderly populations. In this model profiles are established for individuals at age 65 and the program models subsequent survival using all-cause mortality rates. At each age where the individual is alive, he(she) is assigned an overall level of functional status represented by multiple Grade of Membership (GoM) scores. Scores are drawn at random based on predetermined probabilities. This method attempts to model the multidimensional nature of disability in the elderly and to reflect this in measures of Disability Free Live Expectancy (DFLE). Rather than defining a discrete amount of DFLE for individuals this GoM method allows "fuzzy" state membership (or weighted average membership) over a set of disability states. This process of generating individual profiles is repeated until a sufficiently large sample (population) of synthetic profiles has been generated. Outcome measures related to life expectancy are produced by determining averages for the life profiles of the entire sample. This program models functional status among elderly populations well. It is unique among the other models reviewed in its use of overall mortality data and focus on older age groups. Its overall complexity suggests that mainframe computer capacity is necessary and that its use is limited to experts. The research studies reviewed indicate that the program is not designed to assess specific risks of disability or to model intervention on them.

1.3.8 POHEM Model

One of the most intriguing and innovative approaches to population health modelling is the POHEM[34] (Population Health Model) project undertaken by Statistics Canada and the Canadian Institute for Advanced Research. This model uses information from three domains: a population of individuals and their characteristics; environment (physicochemical, socio-cultural, economic and health); and "health-affecting" interventions. The main groups of variables are: socioeconomic (SES) (education, labour, income, marital status and fertility); risks (genetic, physical lifestyle, environmental); diseases; functional status; costs (for health services) and health (as a summary status score). POHEM is based on microsimulation modelling, that is, simulation at the level of the individual. Based on a set of starting parameters, an entire population of individuals is computer generated and followed forward in time. Life events occur for each individual over time, based on their characteristics (age, sex, history) and risk status. These events are determined by a set of conditional probabilities and are "assigned" to individuals using Monte Carlo (chance) techniques. These probabilities (with appropriate stratification) are estimated using Canadian statistical and survey data on mortality, morbidity, risk prevalence and cost. Population health outcomes may be obtained by summation over the simulated individual files at any time point in the simulation. POHEM has many advantages over other models of population health[35]. Among the advantages are: it models multiple disease and risks; incorporates time dimensions such as lead and lag times; reports multiple outcomes (mortality, morbidity, cost and many more); simulates well into the future, and has flexibility of input variables to allow simulation of different

populations and profiles. The major disadvantage of the program is that its complexity restricts its use to experts only. To date its main application has been to estimate breast cancer outcomes. One of the greatest limitations to more extensive application of POHEM is the lack of the integrated health data required as input.

A major limitation to all these models, as identified and commented on by most authors, is lack of appropriate and population-specific input data.

2.0 THE PREVENT MODEL

One of the more comprehensive, yet user friendly, computer models developed for estimating population health is PREVENT[12]. PREVENT was developed in the Netherlands by Louise Gunning-Schepers and Jan Barendregt, with an original database reflecting the Dutch population in 1985. PREVENT's major function is the modelling of risk factor interventions. PREVENT operates on a personal computer, as a menu-driven computer program which allows the user to specify the intervention, simulation period and outcome format. A separate program controls modification of the input data (database). The user can interactively: build new databases, add or remove risk factors and causes of death, increase or reduce stratification (by age, exposure level, etc.) to match available data, modify trends and change prevalence or risk data. The difficulty in developing a new database relates not to the program but to limitations in the input data available. Because many of the data are either not available or not in compatible formats, developing a database requires generalization, interpolation, extrapolation, many assumptions, and critical appraisal and decisions on data sources.

The original goal of the PREVENT project was to devise an instrument for health policy makers that would assist them in formulating realistic targets and setting priorities. There were three specific objectives:

- To adapt existing epidemiologic techniques to develop a methodology which would estimate the effects of multiple risk factors on many diseases. Results would be presented both as proportional changes in disease incidence and as changes in absolute measures of health.

II Using existing demographic and epidemiologic data to apply this methodology to the Netherlands, to see if data from a variety of international sources could be successfully transferred to a single population.

III To use the methodology and data to develop an instrument of use to policy makers and which would ultimately be used by them.

The PREVENT model has incorporated three important features. First the program is able to model the effect of multiple risk factors on a single disease. It is also able to model the effect on multiple diseases from a single risk factor. To achieve this two assumptions have been made: that a multiplicative model of risk factor interaction holds for calculation of relative risks and related proportional variables (potential and trend impact fractions); and that risk factor exposures are independent of each other. A third assumption, that relative risks for individuals of a specific gender, age group and exposure level are uniform within that category, is also inherent in the model.

Second, PREVENT incorporates two time dimensions; latency, the period between cessation of exposure and the initial reduction of risk, and lag time, the period over which risk diminishes to a minimum residual level after exposure ceases. Lead time, the period between first exposure and development of maximum risk, is not included. These time dimensions are graphically presented in Figure 1.
The basic assumptions necessary for use of these time dimensions are that lag time, latency period and case fatality rates are constant over the entire simulation period of the model. An additional assumption is that mortality rates are stable and change

only in response to changes in risk factor prevalence.

Third, the program adjusts for changes caused by the demographic evolution of the population and natural trends in risk factor exposure over the simulation period. This means that the program recognizes the increasing mortality rates of an aging population. Similarly, if smoking is gradually decreasing, the program simulates a continuation in this natural trend into the future and adjusts outcomes appropriately.

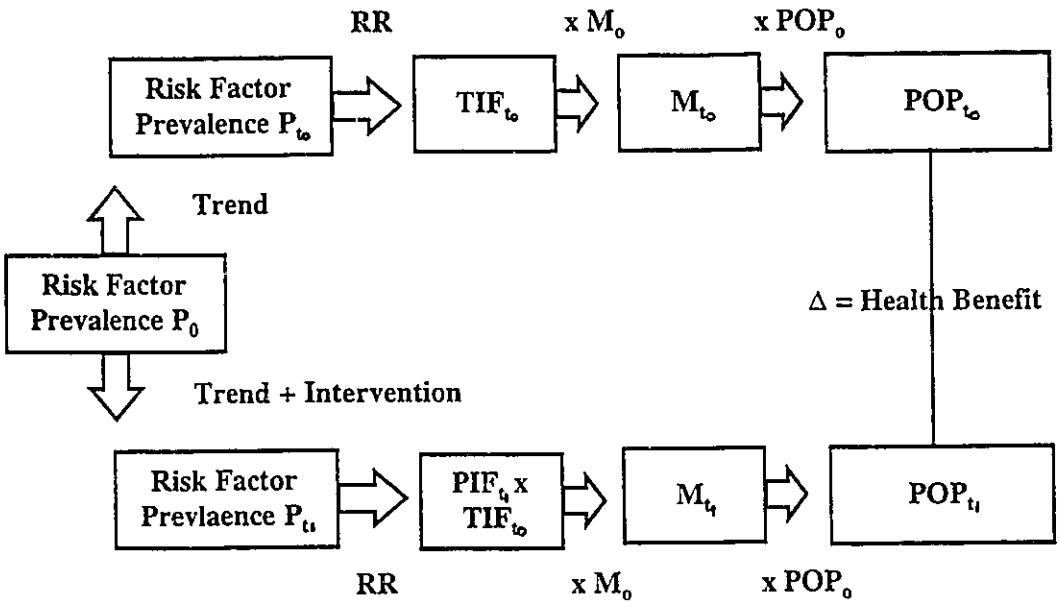
This thesis describes the development of an Ontario database for PREVENT. A full description of the program, its database and its capacities will be presented in later sections. The development of this database for Ontario allows researchers and planners to model the impact of risk intervention for this province or for any municipality in Ontario, with only minor assumptions and minimal modification of the population structure file.

2.1 The Computer Model

The calculations in PREVENT are based on Impact Fraction, a generalization of Population Attributable Risk. A simplified flow chart of the steps in PREVENT is presented in Figure 2. This flow chart shows the calculations for the first year in the simulation of a single risk factor impacting on two diseases. The model itself is not complex; rather the complication arises because of adjustments necessary for latency and lag times, as well as the number of iterations required to manage stratification by gender, age group, risk factor and disease, exposure category, intervention and each year of the simulation. Through all intermediate steps in the model, PREVENT uses proportions (of the population) in its calculations. In the final step these proportions

are multiplied by actual population numbers to obtain the absolute values (e.g., absolute number of deaths) required for some of the output options.

Figure 2 Simplified Flow Chart of PREVENT Program



Influences on risk factor prevalence that occurred in the past, whether due to past interventions or evolution in social behaviour, are considered separately from modelled interventions and are referred to as autonomous trends. The PREVENT program calculates an Impact Fraction attributable to them and refers to it as a Trend Impact Fraction (TIF).

PREVENT considers the difference in proportion of the population exposed with and without a modelled intervention, and calculates the Potential Impact Fraction (PIF) for that intervention.

In this text a simplified explanation for the five steps presented graphically in Figure 2, will be presented. For more detailed calculations, specifying levels of stratification, adjustment for latency and lag times, and definitions please refer to Appendix A.

All projections are made using an arbitrarily selected base year. As a first step the risk factor prevalence, P_0, in the base year is determined (from input data).

In the second step, the risk factor prevalence is estimated for the year, t, after the base year. The risk factor prevalence, P_{t0}, is the estimated population prevalence for the year t, adjusted for change in prevalence due to autonomous trends without intervention. The second risk factor prevalence, P_{t1}, is adjusted both for trend and intervention.

In step three, the Relative Risks (RRs) associated with each disease (A,B, etc.), enter the model with the risk factor prevalences, and the Impact Fractions for both trend and intervention are calculated. These trend (TIF) and potential (PIF) impact fractions are multiplied to represent the total effect of intervention.

In step four, the proportion of disease prevented for the trend alone and for trend plus intervention are multiplied by the disease-specific mortality rates of the base year, and mortality rates for year t are generated. Two sets of mortality rates are calculated, representing the disease and risk factor-specific mortality rates at time t both for trend and intervention.

In step five, these mortality rates are multiplied by the population in the previous year and the population remaining at time t is determined by difference. The projected health benefit of the intervention is the difference between the population remaining at time t after adjustment for trend alone, and the population remaining after adjustment for trend plus intervention.

2.2 Risk Factors and Diseases

Table 2 Risk Factors and Related Causes of Death in the Ontario Database for PREVENT

Causes of Death	Risk Factors
Lung Cancer	Smoking
COLD	Smoking
IHD	Smoking
	Hypertension
	High cholesterol
	Alcohol
CVA	Hypertension
Cirrhosis	Alcohol
Accidental Fall	Alcohol
MVC	Alcohol
	Seat Belt Use
Breast Cancer	Obesity

The Ontario database models six risk factors as they impact on eight causes of death, as presented in Table 2. The Dutch database contained five of the above risk factors and the same related causes of death. For the Ontario database I have added seat belt use as a risk for motor vehicle crash (MVC). The Ontario database includes only modifiable risk factors that impact on noncommunicable diseases.

3.0 OBJECTIVES OF THIS PROJECT

There are three objectives for this project:

I To develop an Ontario database for PREVENT.

II To extend the PREVENT model by adding new risk factor-disease combinations to the model.

III To test the validity of the model by determining its sensitivity to variation in the input data and whether it can accurately predict actual disease mortality from historical prevalence data on one of the risk factors (smoking).

4.0 METHODS

The following sections systematically deal with the methodologic issues encountered in developing the Ontario database. The sections are organized using a standard set of subheadings.

The files contained in the Ontario Database for PREVENT are very lengthy and are not included in hard copy with this thesis. A copy of the data set of 49 files, in ASCII format, on diskette, is available from the Margaret Herbert and the Department of Epidemiology and Community Medicine, University of Ottawa.

4.1 Base Year

Data Required All data in a PREVENT database are anchored to a base year, from which all simulations start. Ideally, the choice of base year should be dictated by the research questions of the individual users. Population, mortality, prevalence and trend data for the population of interest should then be entered for that base year. Because this process is quite complex, the Ontario database developed for this thesis is intended for general use. The most recent year for which dependable data are available was chosen as base year.

Data Sources Statistics Canada produces very detailed provincial population and mortality data on a yearly basis. The Ontario Health Survey (1990), conveniently contains data on the prevalence of four of six risk factors in the model.

Problems Availability of prevalence data turned out to be the limiting concern. Risk factor prevalence data are usually obtained by population survey methods, and in Canada, such surveys are only conducted periodically.

Resolution 1990 was chosen as the base year.

4.2 Population Structure

Data Required A PREVENT database focuses on a specific population. The database includes a data set giving the specific population structure for the base year in one-year age categories, by sex.

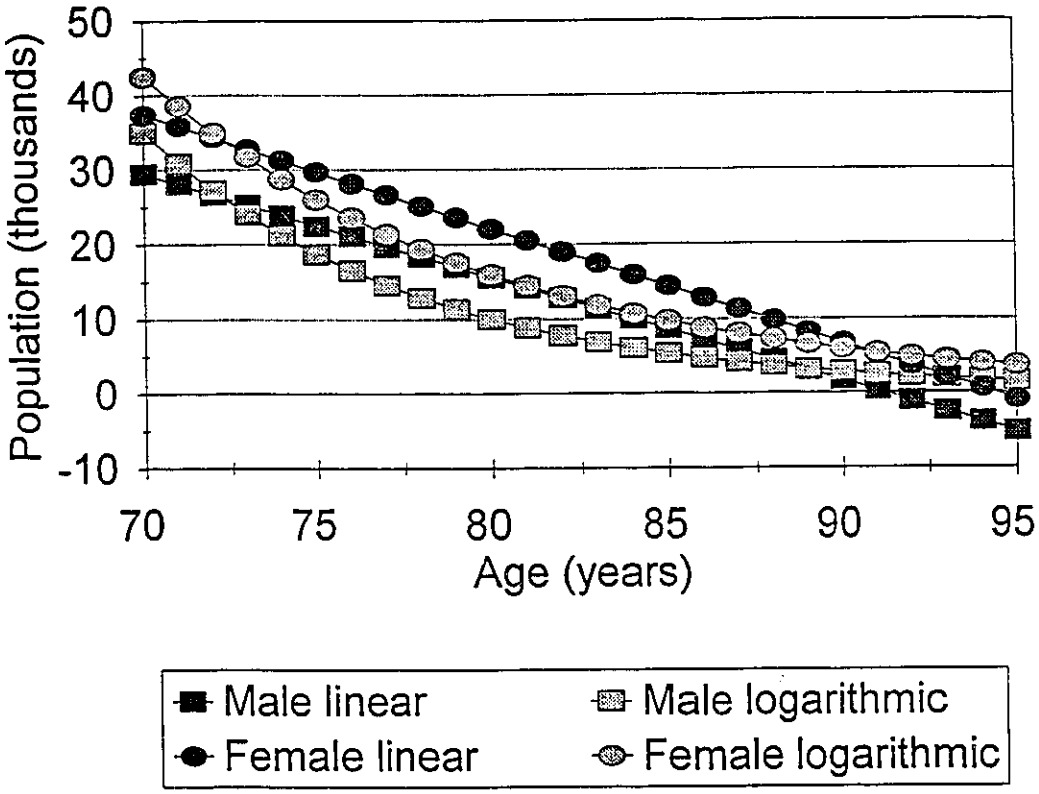
Data Sources For the Ontario database this information was obtained from a publication of Statistics Canada[36]. This source provides one year age categories, by sex, up to age 69, and then in five year categories up to age 90.

Problems PREVENT requires one year age entries up to age 95.

Discussion Some form of data interpolation was required for ages 70 to 89, and extrapolation was needed for ages 90 to 95. Both linear and logarithmic scales were used to plot the data points available for ages 70 or more, and lines were fitted using regression techniques. The linear scale was rejected because the very steep decline in population it modelled resulted in a zero population before age 95. Both estimations are presented in Figure 3.

Next, the one year population counts for ages 70 to 89 were estimated by interpolation (line fitting by linear regression on a logarithmic scale), and population counts for ages 90 to 95 were estimated using extrapolation by the same method. The values estimated in this manner appeared to be reasonable upon first inspection, but a problem emerged when these estimates were used along with the number of deaths (published data) to obtain all-cause mortality. The all cause mortality rates obtained indicated a reduction in all-cause mortality in very old age for both males and

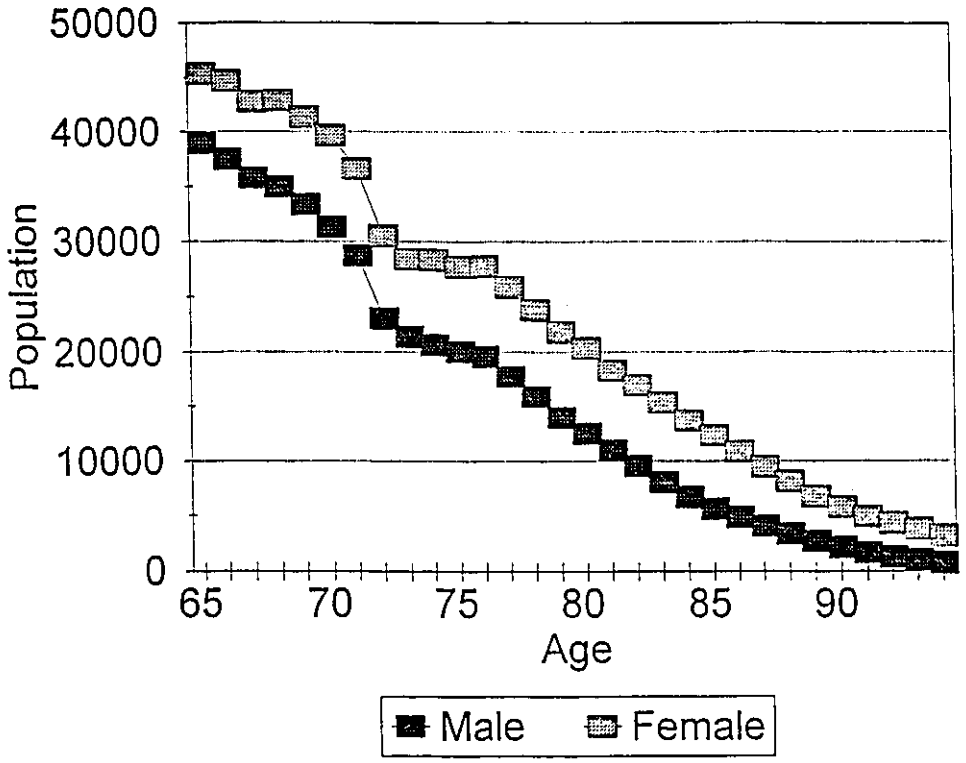
Figure 3 Extrapolation of 1990 Ontario Population Structure in Oldest Ages



females. This was clearly incorrect and indicates that the logarithmic estimate over-estimates population in the oldest age categories.

Resolution Population structure is published in one year age categories for males and females to age 89 in 1991, the census year. Comparison with published 1990 statistics over the age range 55 to 69 (by single years) show a fairly consistent ratio: 0.946 for males and 0.962 for females. These ratios were then applied to estimate 1990 population for ages 70 to 89. The population rates for ages 90 to 95 were estimated by continuation of the declining population curve. The resulting estimates

Figure 4 1990 Ontario Population Data Used in PREVENT



are presented graphically in Figure 4. The all-cause mortality rates calculated from these estimates show a consistent and increasing all-cause mortality rates for older ages.

4.3 Life Expectancy

Data Required The database for PREVENT includes life expectancy tables for males and females in the population under study for the base year. These are the only data sets for which PREVENT internally generates information. The data

modification program PREVDATA asks for the life expectancy of 95 year old males and females in the population for the base year. From these data, together with mortality data, life expectancies for all ages are generated for the base year, using the life table method.

Data Sources Statistics Canada publishes life table data on a periodic basis. The latest published life tables[37] for Ontario, for the years 1985-87, were used. This publication lists age-specific life expectancies for males and females in Ontario up to age 85.

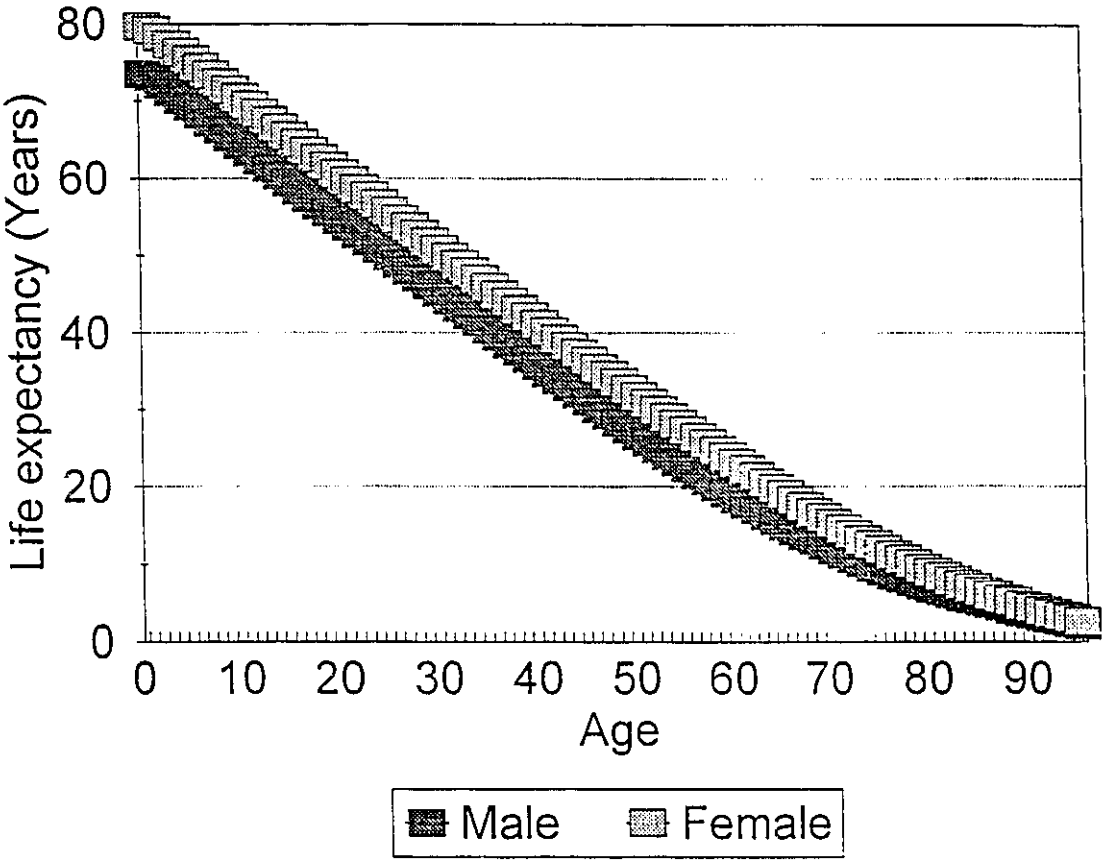
Problems The life expectancy for males and females aged 95 were not directly available.

Discussion Using available published data, the life expectancies were extrapolated (manual curve fitting) to age 95 for both males (2.10 years) and females (2.65 years). The life expectancies and extrapolations are presented in Figure 5. The life expectancies at birth generated by the program, 73.84 years for males and 79.64 years for females, agree well with the Statistics Canada published data, 73.49 years for males and 79.73 years for females.

Resolution A life expectancy of 2.10 years for 95 year old males and 2.65 for 95 year old years females was entered into the database.

Remaining Limitations The Statistics Canada data that I worked from were for 1985-87. If trends in life expectancy have not changed, a very slight increase might be expected for the year 1990 for most ages. Life expectancy at birth generated from my extrapolation shows a slight increase over the 1987 published figure for males and shows a slight decrease for females.

Figure 5 1990 Ontario Life Expectancy Data Used in PREVENT



4.4 Birth Projections

Data Required The PREVENT database requires birth projections for the base year and the following 50 years, for both males and females.

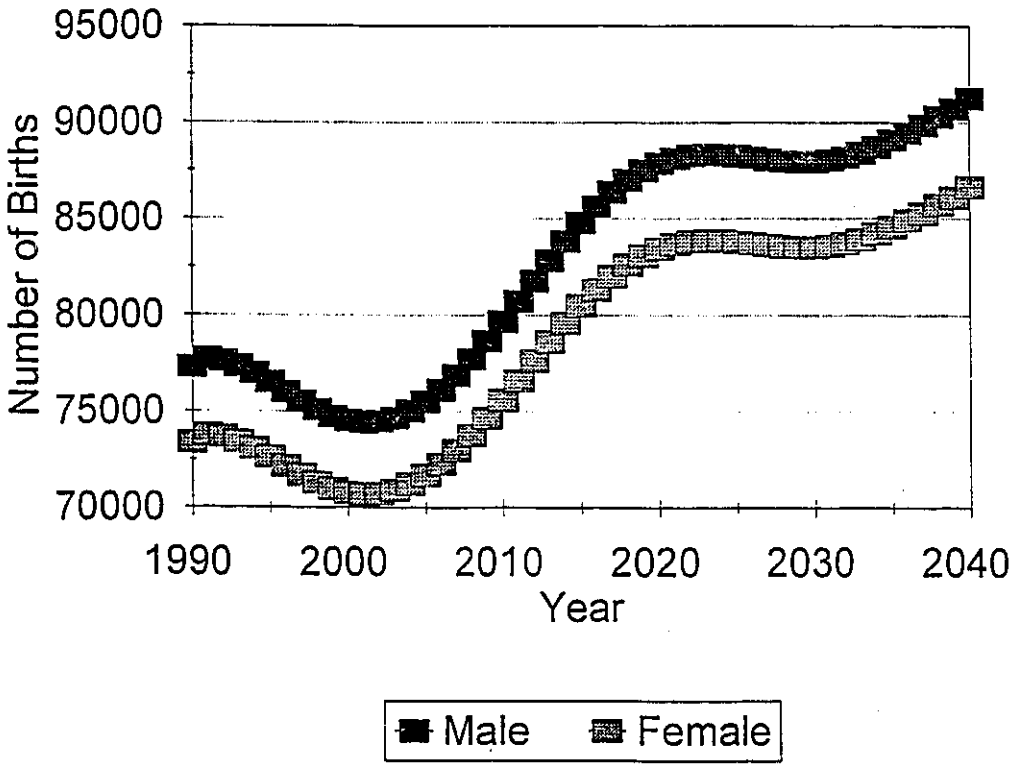
Data Sources Statistics Canada publishes birth projections for Canada and the provinces. Their most recent publication[38] lists total births projected for Ontario for 1990 to 2010. Upon inquiry, I was able to obtain unpublished projections[39] for Ontario to the year 2035.

Problems There were two problems with the data available: total births

unstratified by sex were listed; and birth estimates for 2036 to 2040 were not given.

Discussion To estimate the birth projections by sex the proportions of male and female births were needed. Using published data on Ontario Births[40] for 1980 to 1989, the proportion of male and female births was determined for each year. The proportions for the ten years were averaged, at 0.513 for males and 0.487 for females.

Figure 6 Ontario Birth Projections used in PREVENT



Total birth projections for Ontario for 1990 to 2035 are presented in Figure 6. There is no consistent trend to allow extrapolation over a large group of data points.

Starting at the last trough (minimum) in the data set, at 2029, birth projections were extrapolated, by manual curve fitting, to 2040. Line fitting using both linear and logarithmic scales was considered and rejected due to non-linearity in the seven data points.

Resolution Visual extrapolation of the terminal portion of the birth projection curve produced an estimate of 177,900 births in Ontario in 2040. The calculated proportions for male (0.513) and female (0.487) births were multiplied by the total birth projections to estimate birth projections, by sex, for each year from 1990 to 2040.

Remaining Limitations As PREVENT simulates outcomes, it recalculates the population for each year in the simulation period. This recalculation is based on population structure, deaths and births. PREVENT is unable to make adjustments for migration. Estimates of annual net migration for Ontario for 1990 to 2035 range from 73,852 to 104,123 people[38].

4.5 All-Cause Mortality Rates

Data Required The PREVENT database requires all-cause mortality rates (deaths per 100,000 population) for the base year, in one year age categories (to age 95), by sex.

Data Sources The numbers of deaths, by single year of age (to 95) and sex, for Ontario, in 1990 are published by Statistics Canada[41].

Problem The death rates rather than the actual number of deaths are needed for the database.

Resolution The number of deaths, for each single year of age and sex, were divided by the corresponding population number (by single year and sex, as determined in Section 4.2 above), and then multiplied by 100,000.

4.6 Cause-Specific Mortality Rates

Data Required For each cause of death in the PREVENT database, a data set is needed for cause-specific mortality rates for the base year, in five year age categories (to age 95+), by sex.

Data Sources Statistics Canada publishes total deaths for all ages, by individual International Classification of Disease (ICD) code, on an annual basis for the provinces. Death rates are available by five year age group only at a national level. Neither of these formats was directly useful.

Table 3 ICD and LCDC Codes for Causes of Death in the Ontario Database

Cause of Death	ICD Code	LCDC Code
Ischemic Heart Disease	410-412	117
Lung Cancer	162	46
COLD and Asthma	491-493,496	132
Cerebrovascular Accident	430-438	118
Breast Cancer	174-175	64
Cirrhosis	571	148
Accidental Fall	E880-888	159
Motor Vehicle Crash	E810-819	157

The Laboratory Centre for Disease Control (LCDC) has developed a computer program which uses highly stratified mortality data to present mortality information in a variety of formats. Upon request, I was able to obtain printouts of 1990 mortality rates for Ontario, by five year age groups (to age 85+) and sex, for each cause of

death in the Ontario database[42]: ischemic heart disease, lung cancer, chronic obstructive lung disease, cerebrovascular accident, breast cancer, cirrhosis, accidental fall and motor vehicle crash. These causes of death, with their LCDC and corresponding ICD codes are presented in Table 3.

Problems The available data present the cause-specific mortality rate for the oldest age category as "85 to 85+". The PREVENT database requires data to 95+ years.

Discussion For some causes of death, such as cirrhosis, lung cancer and motor vehicle crash, mortality rates stabilize or decline somewhat for older age groups. All other causes of death in this database demonstrate escalating rates in the oldest age categories.

The cause-specific mortality data available for the oldest age category represent pooled rates for ages 85 and older. There is no consistent mathematical model (e.g. linear regression using linear or logarithmic scale) by which to estimate separate rates for 80-84, 85-89, 90-94, and 95+ years.

Resolution For each cause of death, the data available for the oldest age groups (age 50 and older) were plotted, checked for magnitudes and trends against age-specific mortality rates for Canada[43], and extrapolated to estimate mortality rates for age groups 85-89, 90-94, and 95+ years. Various methods of extrapolation were considered and one general method was used. For each cause of death, the proportion of all-cause mortality in age categories above 50 years was calculated. For the oldest age categories each cause of death demonstrated one of three possible patterns: increasing proportion of all-cause mortality, a stable proportion of all-cause

mortality, or a decreasing proportion of all-cause mortality. For each cause of death, mortality rates for age groups 85-89, 90-94 and 95+ were estimated by choosing rates representing a continuation in the aging trend for proportion of all-cause mortality.

The following sections present a short discussion of the cause-specific mortality rates as obtained from LCDC, the problems encountered and the data chosen for use in the Ontario database. The structured sections on data required, data sources, and problems, presented above (Section 4.36), are applicable to all-causes of death in the model so that only a "Discussion" section of the problems particular to each cause of death is presented.

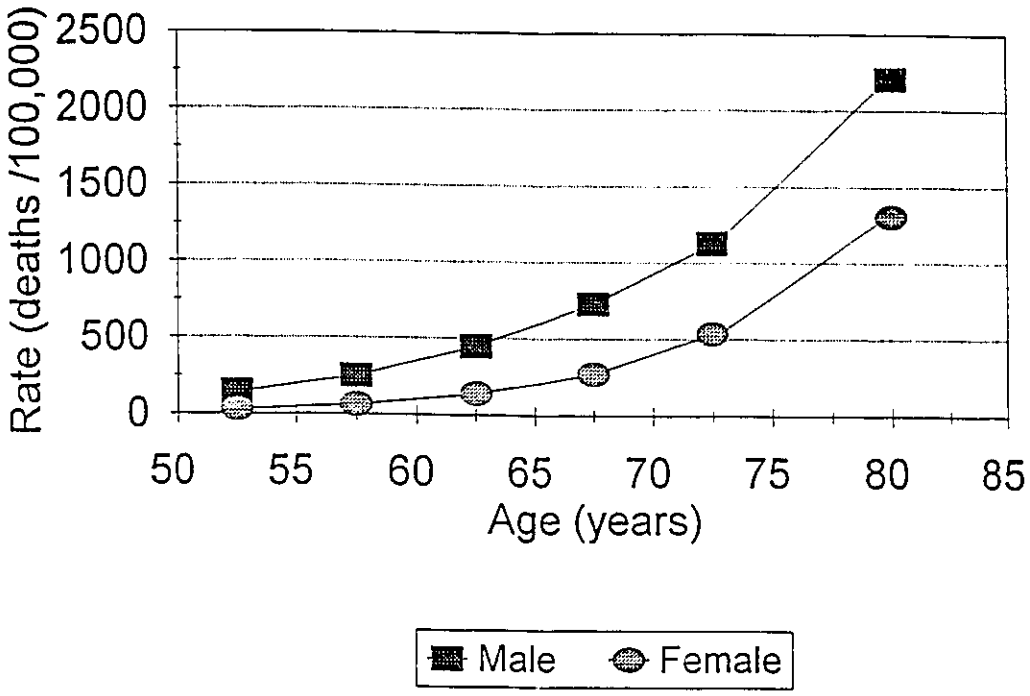
4.6.1 Ischemic Heart Disease Mortality

Discussion The IHD mortality rates for age groups above 50 years are presented in Figure 7. As plotted, these data are clearly not linear.

When the mortality rates were plotted on a logarithmic scale the data appeared to be very linear with a correlation of 0.999 for males and 0.998 for females. Although this extrapolation method seems ideal, and the correlation between observed points and predicted line is excellent, it does not work epidemiologically. The mortality rates for 95 year olds predicted by this method, exceed 90% of all-cause mortality rates for both males and females.

As an epidemiology-based alternative to mathematical modelling, patterns in the percentage of all-cause mortality rates were examined, Figure 8. IHD accounts for approximately one quarter of all-cause mortality in males and approximately 20

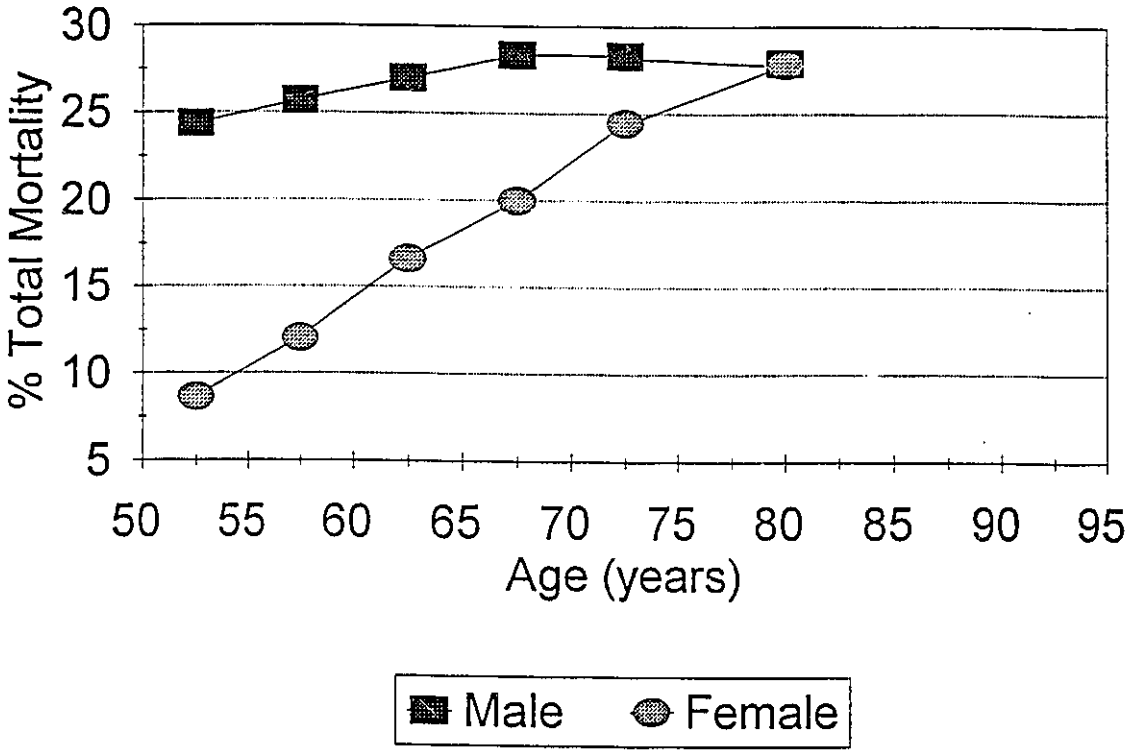
Figure 7 IHD Mortality Rates For Ages >50 Years



percent in females older than 65. The trend appears to be changing very little for males and increasing at a rate of at least 5% over 5 years for females as age increases.

Resolution The IHD mortality rates for ages 85 to 89, 90 to 94 and 95+ were estimated as a fixed proportion, 27.9%, of all-cause mortality for males. The IHD mortality rates for females for ages 85 to 89, 90 to 94 and 95+ were estimated as 5% greater (as a percent of all-cause mortality) for each advancing 5 year age group.

Figure 8 IHD Mortality Rate as a Percentage of All-Cause Mortality Rates



4.6.2 Cerebrovascular Accident

Discussion Cerebrovascular Accident accounts for approximately 3% of all-cause mortality at age 55 and increases to at least 6% by age 75. In males the rate of increase is at least 1.5% over 5 years and in females the rate of increase is at least 2.0% over 5 years.

Resolution The cerebrovascular accident mortality rates for males ages 80 to 84, 90 to 94, and 95+ were estimated as 1.5% greater (as a percent of all-cause mortality) for each advancing 5 year age group. For females ages 80 to 84, 90-94, and 95+

cerebrovascular accident rates were estimated as 2.0% greater (as a percent of all-cause mortality) for each advancing 5 year age group.

4.6.3 Chronic Obstructive Lung Disease

Discussion COLD accounts for approximately 1% of all-cause mortality at age 55 and increases to at least 5% by age 75. In males the percentage of all-cause mortality increases consistently by approximately 1% over five years. In females the percentage of all-cause mortality increases to age 70 and then starts to decline at a rate of 0.5% over five years. This inflection point and decline could be an artifact, or it could be a cohort effect. Women currently over 70 years of age were adults before smoking became popular for women (in the 1950s). This cohort of women has, by far, the greatest proportion of nonsmokers.

Resolution The COLD mortality rates for males ages 80 to 84, 90-94 and 95+ were estimated as 1.0% greater (as a percent of all-cause mortality) for each advancing 5 year age group. For females ages 80 to 84, 90-94, and 95+ COLD rates were estimated as 0.5% less (as a percent of all-cause mortality) for each advancing 5 year age group.

4.6.4 Lung Cancer Mortality

Discussion Lung cancer accounts for approximately 13% of all-cause mortality for those 65 years of age and declines as age increases for both sexes. The rate of decline is approximately 2% over 5 years for the oldest age groups.

Resolution The lung cancer mortality rates for ages 85 to 89, 90 to 94, and 95+

were estimated as 2% less (as a percent of all-cause mortality) for each advancing five year age group for both men and women. When the overall percent of all-cause mortality in women dropped below 2%, the percentages were estimated at half that of the previous age group.

4.6.5 Breast Cancer Mortality

Discussion Breast cancer accounts for approximately 16% of all-cause mortality for women 55 years of age, and declines to 6% by age 75. The rate of decline is approximately 1.0% over 5 years for the oldest age groups. The percent of all-cause mortality in males is negligible (less than 0.1%) for all age groups.

Resolution The female breast cancer mortality rates for ages 85 to 89, 90 to 94, and 95+ were estimated as 1% less (as a percentage of all-cause mortality) for each increasing five year age group. When the overall percentage dropped below 2%, the percentages adopted were half those of the previous group. The breast cancer mortality rates for males aged 85 to 89, 90 to 94, and 95+ were estimated as a fixed proportion, 0.044%, of all-cause mortality.

4.6.6 Cirrhosis Mortality

Discussion Cirrhosis accounts for approximately 5% of all-cause mortality in males and 2.5% in females at age 55. These percentages decrease to 1.5% for men and 1% for women by age 75. The rates of decline are between .8 and 1.5% over five years. At age 75 the overall percentage of all-cause mortality is already very low, approximately 1% for both men and women.

Resolution The cirrhosis mortality rates for males and females aged 85 to 89, 90 to 94, and 95+ were estimated as half the percentage of all-cause mortality of the previous group. This modelled a declining trend without allowing mortality rates to reach zero.

4.6.7 Accidental Fall Mortality

Discussion Accidental falls account for less than 1% of all-cause mortality for those 65 years of age, but are known to be an important cause of death in the very old [44]. Only over age 70 does the percentage of all-cause mortality seem to increase steadily. There are not enough data points to accurately predict the trend.

Resolution The accidental fall mortality rates for ages 85 to 89, 90 to 94, and 95+ were estimated as a 1% increase in percentage of all-cause mortality for each advancing five year age group.

4.6.8 Motor Vehicle Crash Mortality

Discussion Motor vehicle crashes account for approximately 1.5% of all-cause mortality for those 55 years of age and declines steadily as age increases. The percentage of all-cause mortality is below 0.5% for both sexes by age 75.

Resolution The motor vehicle crash mortality rates for ages 85 to 89, 90 to 94, and 95+ were estimated as half the percentage of all-cause mortality of the previous group. This models a declining trend without allowing mortality rates to reach zero.

4.7 Risk Data

Whenever possible, the relative risks used in the original Dutch database for PREVENT were also used in the Ontario database. This was possible for the risk factors whose prevalence was obtained from the Ontario Health Survey (smoking, alcohol use and obesity) because I was able to obtain prevalence data specific to the stratification cut points used in the Dutch database. The cut points in the prevalence data used for hypertension also matched the Dutch database and the hypertension risks from that database were also used. Since these relative risks were, in effect, adopted from previous work, they will not be presented here. The international epidemiologic studies which were reviewed in the choice of relative risks and short discussions of the issues in appropriate choice of these relative risks are presented in Appendix B.

The cut points for disease severity of hypercholesterolemia did not match. It was necessary to review sources of risk data for high cholesterol and IHD and choose relative risks appropriate to the severity cut points used in the available prevalence data. The general considerations in choice of relative risks in PREVENT are presented below.

Data Required For each risk factor and related cause of death, levels of stratification must be established and consistently adhered to in several of the data sets. These stratification factors are: sex, age groups (rounded to the nearest 5 years) for which different levels of risk apply, and the number of exposure categories for which different risks apply. PREVENT requires separate estimates of relative risk for each sex, age category and exposure level chosen. It also requires input of residual risks for each of these strata and estimates of latency and lag times. A

residual risk is the minimum relative risk which is experienced after the exposure ceases. For some risk factors and diseases (e.g., smoking and lung cancer) relative risk is thought to decrease gradually over a long period (defined by the Lag time in PREVENT) before the minimum residual risk is achieved.

Problems Once the age group and exposure categories are chosen, data on relative risks, residual risks and risk factor prevalence must be obtained to match those categories.

Discussion The relative risk data are usually obtained from published reports of large prospective studies. Ideally, risk estimates should be available for single-year age groups, both sexes and a wide range of exposure categories. Risk estimates are generally reported in age group and exposure categories convenient to the investigators; raw data are rarely available for restratification. Often selection criteria specified in original protocols for these studies have predetermined the levels of exposure reported and narrowed the age range. Some of these studies are exclusive to convenient subpopulations (e.g., physicians, nurses, communities) that do not reflect a general population distribution with respect to sex, age, socio-economic status, or ethnicity. Most of these studies have not presented information on residual risks in their published results.

Resolution In developing the Ontario database, I reviewed and chose to keep the relative risks, residual risks, latency and lag times used in the Dutch database whenever possible.

4.7.1 Hypercholesterolemia as a Risk for Ischemic Heart Disease

Table 4 Mortality Ratios for Ischemic Heart Disease Associated with Elevated Total Cholesterol Levels

Study	Age	Total Cholesterol Category (mmol/L)	Risk Ratio Male	Risk Ratio Female
Framingham Study	35-44	≥ 6.9 versus < 5.7	5.5	5.0
	45-54	"	2.4	1.5
	55-64	"	1.7	1.3
	all	"	2.5	1.5
Pooling Project	40-64	≥ 6.2 versus < 5.7	2.0	
Whitehall Study	40-62	≥ 6.1 versus < 4.2	2.0	
Western Collaborative Study	39-49	≥ 6.8 versus < 6.8	2.9	
	50-59		1.7	
Stockholm Prospective Study	35-39	≥ 7.3 versus < 7.3	2.1	
Evans County Study	15-75+	≥ 6.8 versus < 5.7	1.3	
MRFIT Intervention Trial	35-39	< 4.7	1.00	
		4.7-5.4	1.14	
		5.4-5.8	2.32	
		5.9-6.2	2.44	
		>6.2	7.75	
	40-44	< 4.7	1.00	
		4.7-5.4	1.79	
		5.4-5.8	2.63	
		5.9-6.2	3.69	
		>6.2	5.79	
	45-49	< 4.7	1.00	
		4.7-5.4	1.26	
		5.4-5.8	1.79	
		5.9-6.2	2.36	
		>6.2	3.88	
	50-54	< 4.7	1.00	
		4.7-5.4	1.28	
		5.4-5.8	1.43	
		5.9-6.2	1.96	
		>6.2	2.93	
	55-59	< 4.7	1.00	
		4.7-5.4	1.20	
		5.4-5.8	1.74	
		5.9-6.2	1.93	
		>6.2	2.40	

* Reference 12

Data Sources Table 4 presents mortality ratios for IHD deaths, associated with hypercholesterolemia, from seven studies[12]. Unfortunately, there is no information

available on residual risk stratified by sex, age group, or exposure level.

Table 5 Relative Risks and Residual Risks for IHD Associated with Hypercholesterolemia

Exposure	Hypercholesterolemia	
	mild	severe
35-44 years		
Male	2.77	6.77
Female	2.77	6.77
45-54 years		
Male	1.89	3.41
Female	1.89	3.41
55-64 years		
Male	1.84	2.40
Female	1.89	2.40
≥ 65 years		
Male	1.84	2.40
Female	1.84	2.40
Latency	0	
Lag time	3	
Residual Risk		
35-44		
Male	1.0	3.0
Female	1.0	3.0
45-54		
Male	1.0	1.5
Female	1.0	1.5
55-64		
Male	1.0	1.5
Female	1.0	1.0
≥ 65		
Male	1.0	1.5
Female	1.0	1.0

Discussion In all the studies reviewed, total cholesterol levels are the basis of risk ratios presented. The risk ratios from these studies show quite a wide range of values. While some of this variation is due to differences in populations and age groups studied, the main source of variation is difference in cut points. Those studies that choose lower cholesterol ranges as their reference (minimum) category, have larger risk ratios for higher cholesterol categories. While the Framingham Study was

chosen as the main source of relative risks in the Dutch database, the exposure cut points do not agree well with those available for the Ontario prevalence data (see section 4.8.3). The cut points and risk ratios adapted from the MRFIT study (Table 4) have a better match of reference category for use with Ontario data.

Overall, in the studies reviewed, risk of IHD increases with increasing cholesterol level. The relative risk is highest for the youngest age groups studied, and diminishes rapidly in older age groups. Only the Framingham study reports risk ratios for women separately. Over all age groups the risk for women appears to be less than that for men. It is interesting to note that although high cholesterol levels are common in elderly women, their associated risk of IHD is slight.

It should be noted here that some researchers have found very low total cholesterol levels to be associated with increased non-IHD mortality (notably lung cancer, respiratory disease, digestive disease, trauma and residual deaths). The mortality rates and cholesterol levels follow a U-shaped relationship. A recent review paper[45] finds the evidence to be inconclusive and the authors recommend further research. If these findings are substantiated and the non-IHD risks quantified, future PREVENT databases could be adjusted to incorporate these effects.

None of the studies reported the magnitude of risk reduction after control of cholesterol levels (by dietary intervention or drug therapy); nor do they indicate latency or lag periods. Since these data are required in the PREVENT database, unsubstantiated estimates will be necessary.

Resolution The relative risks of IHD associated with hypercholesterolemia, that were chosen for the Ontario database PREVENT are presented in Table 5. The cut

points chosen reflect available prevalence data. The reference category (relative risk 1.0) is less than 5.2 mmol/L total cholesterol; the mild category is 5.2 to 6.1 mmol/L; and the severe category is 6.2 mmol/L or more. The relative risks are stratified by sex, two exposure categories (mild and severe), and four age groups (35-44 years, 45-54, 55-64 and $\geq$ 65 years). The relative risks from the MRFIT study were used with some adjustment for collapsing age groups and exposure categories. Despite evidence of lower risk for women in other studies, women were assigned the same values as men, due to lack of available estimates for the exposure categories used.

In the choice of residual risks, latency and lag times, I have followed the estimates used in the Dutch database.

4.8 Prevalence

Ontario data for prevalence of several risk factors considered by PREVENT are available from the Ontario Health survey of 1990. These include smoking, alcohol consumption and obesity. In the period during which I was gathering data for the Ontario database, researchers in the Department of Epidemiology and Community Medicine at the University of Ottawa who were associated with the Community Health Research Unit were working directly with the raw data from the Ontario Health Survey. They were able to provide me with Ontario-specific estimates of population prevalence for the exact exposure cut points that I required for the database.

Other risk factors such as hypertension are not as readily available but were

extrapolated from published summaries of heart health surveys recently conducted in other provinces.

In all cases, the exposure data are estimated for the year 1990 and are stratified by sex, age group and exposure categories. These stratification factors must match the relative risk data, while age should generally be in multiples of 5 years to match mortality data. Generally I have tried to use the same stratification used in the Dutch data set.

4.8.1 Prevalence of Smoking

Data Required Prevalence of smoking is needed for males and females in the age categories 20-24 years, 25-34 years, 35-49 years, 50-64 years and 65+ years. Prevalence data are also needed for three exposure levels; 1-12, 13-22, and 23+ cigarettes per day. The prevalence of ex-smokers from each of these age, sex and exposure categories is also needed. The reference category, never smokers, are not listed in the data but calculated in the program.

Data Sources This information was abstracted directly from the Ontario Health Survey data.

Resolution The prevalence data used are presented in Table 6.

4.8.2 Prevalence of Hypertension

Data Required In the Dutch database prevalence of uncontrolled hypertension is available for males and females in the age categories, 35-44 years, 45-49 years, 50-54 years, 55-59 years and 60+ years. Prevalence data are also needed for two exposure

Table 6 Smoking Prevalence in Ontario, 1990

Exposure	Sex	Age Categories				
		20-24	25-34	35-49	50-64	65+
Smokers 1-12 cigarette/day	M	.090	.068	.052	.044	.047
	F	.124	.086	.064	.062	.056
13-22	M	.159	.154	.103	.098	.053
	F	.137	.139	.109	.089	.050
+23	M	.094	.158	.174	.133	.051
	F	.078	.111	.119	.083	.028
Ex-Smokers 1-12	M	.026	.063	.105	.147	.182
	F	.040	.072	.079	.082	.074
13-22	M	.026	.063	.105	.147	.182
	F	.040	.072	.079	.082	.074
23+	M	.026	.063	.105	.147	.182
	F	.040	.072	.079	.082	.074

levels; mild (DBP of 90-94 mm Hg and/or SBP of 140-159 mm Hg) and severe(DBP $\geq$ 95 mm Hg and/or SBP $\geq$ 160 mm Hg) The prevalence of controlled hypertensives from each of these age, sex and exposure categories is also estimated. The reference category is non-hypertensives.

Data Sources The pooled results of hypertension prevalence from the Heart Health Surveys of nine Canadian provinces[46] are used as an estimate of Ontario prevalence 1990.

Problems There is no population-based estimate of the prevalence of hypertension in Ontario in 1990. Fortunately, between 1986 and 1990 the provincial governments, in collaboration with Health and Welfare Canada, conducted population based heart health surveys in nine Canadian provinces. A similar heart health survey was conducted in Ontario but the data from that study had yet to be published when this

Ontario database for PREVENT was finalized. The pooled data from these surveys give a reasonable estimate of prevalence in Ontario.

Unfortunately, while the published data give estimates at the exact exposure cut points, the age categories do not match exactly. Re-examination of the risk data

Table 7 Estimated Hypertension Prevalence in Ontario,1990

Exposure	Sex	Age Categories			
		30-44	45-54	55-64	65+
Uncontrolled Hypertensives					
Mild	M	.11	.14	.30	.34
	F	.04	.12	.22	.35
Severe	M	.05	.10	.13	.13
	F	.00	.04	.09	.14
Controlled Hypertensives*					
Mild	M	.010	.008	.011	.025
	F	.031	.064	.012	.035
Severe	M	.004	.006	.005	.010
	F	.000	.021	.005	.014

* Ex-hypertensives, treated and controlled.

indicated that elimination of one age category and slight adjustment of ages in other categories (to match the published prevalence data) did not alter the age cut points for relative risk. These new age categories were adopted. Prevalence of uncontrolled hypertension was listed separately for systolic and diastolic blood pressure. Information on joint exposure allowed calculation of prevalence to match the combined systolic and diastolic cut point in PREVENT. Similarly, information on the number of subjects "treated and controlled" provided data for the previously exposed categories. There was no information on how many "treated and controlled" subjects had mild or severe hypertension. The proportion of mild to severe controlled

hypertensives was adjusted to match the proportion of current mild to severe uncontrolled hypertensives.

Resolution The prevalence data used are presented in Table 7.

4.8.3 Prevalence of Hypercholesterolemia

Data Required In the Dutch database prevalence of hypercholesterolemia is available for males and females in the age categories, 35-39 years, 40-44 years, 45-49 years, 50-54 years, 55-59 years and 60+ years. Prevalence data are also needed for two exposure levels; mild (total cholesterol of 5.2-6.1 mmol/L) and severe(total cholesterol $\geq$ 6.1 mmol/L). The prevalence of previous exposure to high cholesterol from each of these age, sex and exposure categories is also estimated. The reference category is normal cholesterol ($\leq$ 5.2 mmol/L total cholesterol).

Data Sources The pooled results of hypercholesterolemia prevalence from the Heart Health Surveys of nine Canadian provinces were used[47].

Problems There is no population-based estimate of the prevalence of hypercholesterolemia Ontario in 1990. As in the case of hypertension, the pooled data from the heart health surveys in other provinces give a reasonable estimate of prevalence in Ontario.

As discussed in previous sections (section 4.7.1), the relative risks chosen for the Ontario database differ from those used in the Dutch database. This was done so that the cut points and age groups available from the Canadian Heart Health Surveys could be used. This resulted in the elimination of one age group and the adjustment of others. Unfortunately, there were no data available on the proportion of people

Table 8 Hypercholesterolemia Prevalence in Ontario, 1990

Exposure	Sex	Age Categories			
		35-44	45-54	55-64	65+
Current high cholesterol					
Mild	M	.36	.32	.36	.40
	F	.23	.43	.43	.36
Severe	M	.22	.33	.25	.25
	F	.08	.17	.38	.44
Previously high cholesterol					
Mild	M	.005	.004	.006	.013
	F	.016	.032	.006	.018
Severe	M	.002	.003	.003	.005
	F	.000	.011	.003	.007

who are "treated and controlled". The use of medications to control hypercholesterolemia has only recently become widespread and the effectiveness of diet and exercise is limited. It is reasonable to assume that the proportion of the whole population "treated and controlled" is very much smaller for cholesterol than for high blood pressure. Without further information, I have made the estimate that the previously exposed proportions for high cholesterol will be half those found for hypertension.

Resolution The prevalence data used are presented in Table 8.

4.8.4 Prevalence of Alcohol Use

Data Required Prevalence of alcohol use is needed for males and females in the age categories, 20-39 and 40+ years. Prevalence data are also needed for two exposure levels: abstainers (<4 drinks per week) and excessive drinkers (21+ drinks

Table 9 Alcohol Use in Ontario, 1990

Exposure	Sex	Age Categories	
		20-39	40+
Current Drinkers			
abstainers	M	.332	.464
<4 drinks/week	F	.574	.716
excessive	M	.136	.087
>21 drinks/week	F	.020	.017
Ex-Drinkers	M		
ex-moderate	F	.000	.000
		.000	.000
ex-excessive	M	.024	.055
	F	.013	.016

per week). The prevalence of ex-drinkers from each of these age, sex and exposure categories is also needed. The reference category is moderate drinkers. This reference has been chosen to accommodate the risk pattern for alcohol and IHD. For the other diseases, moderate drinkers remain the nominal reference category, but abstainers are assigned no excess risk (Relative Risk 1.0) and thereby join the reference group in the calculations.

Data Sources The prevalence data were abstracted directly from the Ontario Health Survey data.

Resolution The prevalence data used are presented in Table 9.

4.8.5 Prevalence of Obesity

Data Required Information on the prevalence of obesity (defined as BMI$\geq$30) is needed for males and females in the age categories, 40-49 and 50+ years. Data are needed for only a single exposure category, BMI of 30 or more and an estimate of

previous exposure is also needed. The reference category is BMI < 30. Cut points currently used to define obesity in Canadian medical practice, are commonly lower (usually in the range of 25 to 27). However, for use in PREVENT, cut points for risk data and prevalence must align.

Data Sources The prevalence data are abstracted directly from the Ontario Health Survey data. Unfortunately, there is no question in the survey that would assess previous exposure.

Table 10 Prevalence of Obesity in Ontario, 1990

Exposure	Sex	Age Categories	
		40-49	50+
Current Status BMI ≥ 30	M F	.216 .219	.206 .244
Ex-Obese BMI ≥ 30	M F	.000 .000	.000 .000

Resolution The prevalence data used are presented in Table 10 using the cut point of BMI < 30. The assumption of no former obesity has been made.

4.9 Future Trends in Risk Factor Prevalence

Data Required The PREVENT database requires information on past population prevalence of each risk factor as well as estimates of future population trends of the risk factor over the maximum simulation period of fifty years. Prevalence histories and projected trends are needed in each year for both sexes, each age group and each exposure category.

The prevalence history of each risk factor goes back a number of years before the base year so that the risk to people who have recently moved from "exposed" to "no longer exposed" can be determined. The number of years of prevalence history needed is determined for each risk factor and disease by adding the lag time to the latency time.

The future population trends of each risk factor must also be estimated for fifty years forward from the base year. Two types of trends must be estimated, "Inflows" and "Outflows". Inflows trace the percent of those currently exposed changing from non-exposed to exposed. Outflows trace the percent of those currently exposed changing from exposed to ex-exposed. In theory, it is possible to trace inflows and outflows separately, however use of sequential cross-sectional surveys for input data results in only the net changes from survey to survey being available. When net change data is used, as it is in the Ontario database, increases in risk factor prevalence are modelled as inflows and decreases as outflows.

Discussion In order to estimate future trends it is first necessary to establish the past trends of any risk factor and then speculate on future trends. There is a body of research that studies patterns in the uptake of innovation[48]. An innovation can generally be defined as any idea, object, or practice that is perceived as new by members of the social system. Although there are many variations in the mathematical models that have been developed to describe patterns in uptake of innovation, they as all based on a S-shaped diffusion curve. The uptake curve starts at zero, the proportion of adoptions is quite slow at first, the adoption rate then increases through the central part of the curve and finally slows as uptake approaches

maximum. For the simple symmetric, S-shaped uptake pattern, a complementary (mirror image) decay pattern is proposed for regression of the innovation. The shape of the curve can be modelled by a logistic function. This theory is useful in projecting future trends in behaviour-based risk factors.

4.9.1 Cohort and Age-Group Options and Trends

PREVENT offers a calculation option of "cohort" or "age-group". The age-group option works best for risk factors that are age-related. Hypertension is an example of a risk factor that is largely age-dependent (it is rare in young adults and prevalence increases with age). Smoking is a cohort-dependent risk factor, as are most behaviours, and prevalence varies according to the popularity of smoking within

Table 11 Example of Age Group and Cohort Inflow and Outflow Calculations

	Risk Factor Prevalence	
Year	Age 40-49	Age 50-59
1990	50% (birth cohort 1941-1950)	53% (birth cohort 1931-1940)
2000	40% (birth cohort 1951-1960)	55% (birth cohort 1941-1950)

Age group change between 1990 and 2000 is
40 - 50 = -10% over 10 years.
This would be modelled as an outflow of 1% for each of the 10 years.

Cohort change between 1990 and 200 is
55 - 50 = 5% over 10 years.
This would be modelled as an inflow of 0.5% for each of 10 years.

specific birth cohorts among adolescents and young adults. The choice of cohort or age-group option can result in differences in output. For example, choice of the age-group option for smoking may underestimate future mortality for smoking women. This occurs because the age-group option assumes that as young women (who have relatively high smoking prevalence) in 1990 age, they will have the same smoking patterns as older women (who have low smoking prevalence) in 1990. The cohort option is preferable because it assumes that young women will continue to have the same smoking prevalence patterns established as young adults. The only files in the database that differ for the "age-group" and "cohort" options are the inflow and outflow trend files. These files are used to model future changes in risk factor prevalence (e.g. estimated changes in smoking behaviour). They are also used to adjust for differences in "age group" versus "cohort" effects. A simplified diagram of the manner in which cohort and age group inflows and outflows are calculated is presented in Table 11.

The Ontario database offers the "age-group" option for all risk factors and offers the "cohort" option only for smoking and hypertension. Ideally, both options should be offered for all risk factors. In order to offer both, high quality past prevalence information is required. Because of limitations in available data, I have followed the Dutch data set by including only two risk factors with both options.

4.9.2 Trends in Smoking

Data Required The maximum combined latency plus lag time associated with smoking is 20 years for chronic obstructive lung disease. For the inflow and outflow

files, changes in smoking prevalence (by sex, age group and exposure level) are needed for the years 1970 to 1990 and projected changes are needed for the years 1991 to 2040.

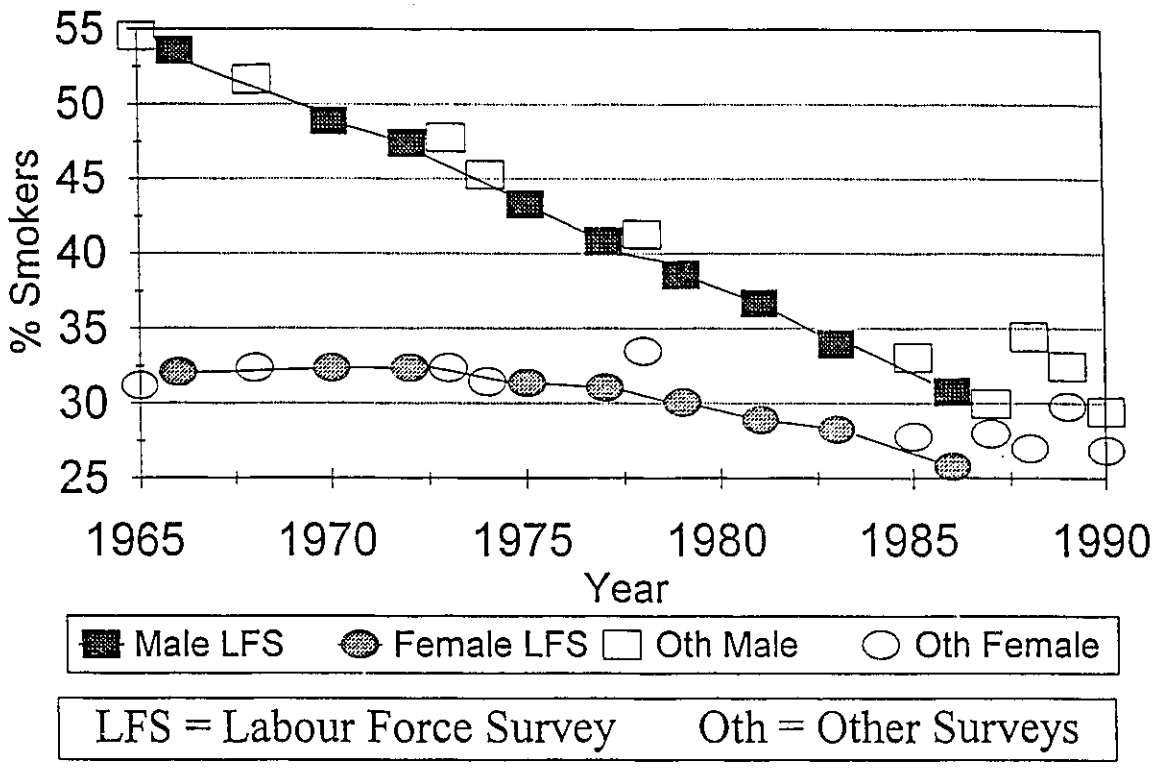
Table 12 Sources of Smoking Past Prevalence Data in Canada

Source of Data	Year(s)
Labour Force Survey of Smoking Habits	1966 1970 1972 1975 1977 1979 1981 1983 1986
Labour Force Survey Supplement	1965
Canada Health Survey	1978
Canada Fitness Survey	1981
Canadian Health Promotion Survey	1985
General Social Survey	1985
Canadian Heart Health Surveys	1988
National Alcohol and Drug Survey	1989
Canada's Health Promotion Survey	1990

Data Sources Information on general smoking prevalence (regular smoking among those 15 years and older) of Canadians is available from 1965 from a variety of sources[49, 50, 51, 52, 53, 54, 55, 56]. With the exception of 1965, data are available by sex and age group (15-19, 20-24, 25-44, 45-64 and 65+ years) but not by exposure category. The sources and years are listed in Table 12. Unfortunately, past prevalence of smoking in Ontario is not available in published reports stratified by sex and age group.

Problems Although there initially appears to be a variety of sources available from which to determine smoking trends, some problems do arise. Each survey has

Figure 9 Past Prevalence of Smoking



unique properties with respect to sampling frame, questions asked, cut points used and reporting formats. Some of the irregularity between estimates can be seen in Figure 9 which shows smoking prevalence for all regular smokers between 1965 and 1990 from the sources in Table 12.

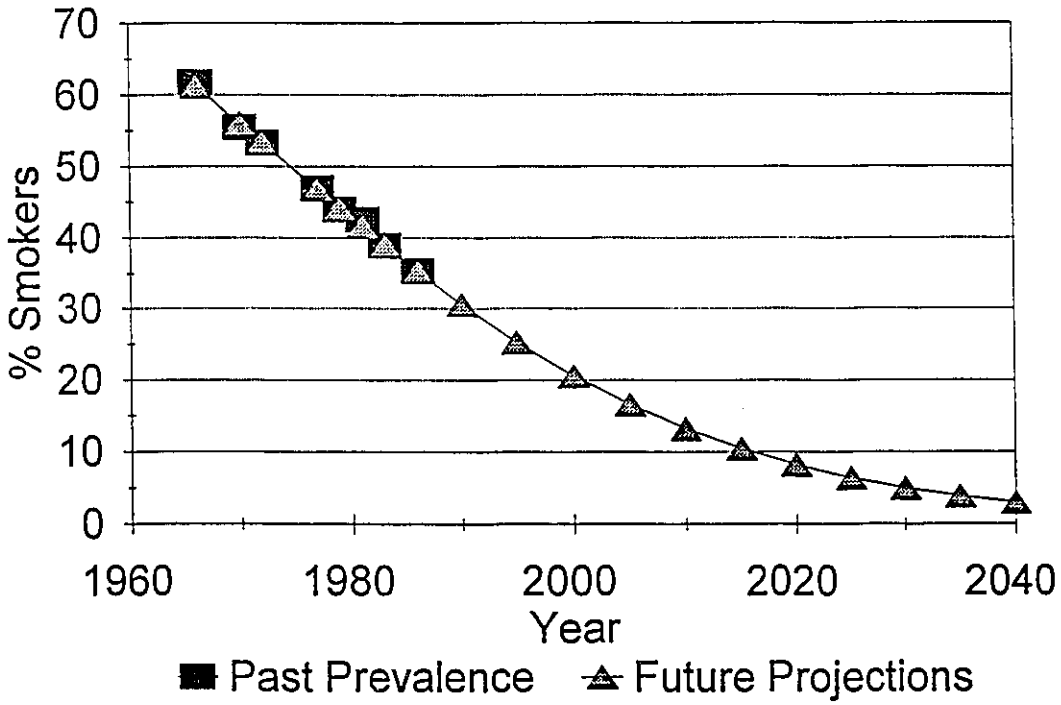
Ontario data are available in published reports of some of the studies above, but usually in a comparison of provinces under the general category of all regular smokers (no stratification by age group or exposure level).

Discussion When determining trends, the absolute prevalence value at any

particular point in time is less important than the relative change in prevalence over time. For this reason the best estimates of trends would come from a series of related surveys such as the national Labour Force Surveys.

Smoking can certainly fit within the uptake of innovation model. Smoking is already a well established behaviour and there is evidence that the proportion of smokers in the population is declining. If the innovation model is used to estimate future trends, then smoking should be modelled on the regression rather than the

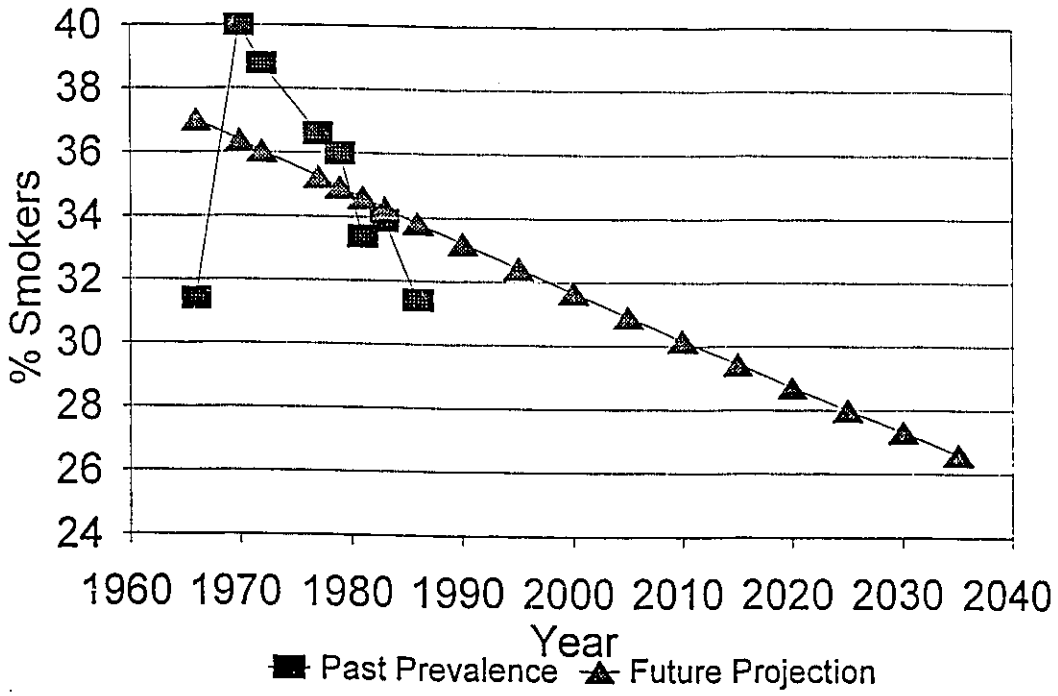
Figure 10 Past Smoking Prevalence and Projections
Canadian Males 25-44 Years



uptake part of the diffusion curve and a logistic model might be used.

Resolution Since stratified data for Ontario is not available, the national data from the Labour Force Surveys will be used to estimate smoking trends. The Labour Force Surveys report prevalence by sex, and age group (15-19, 20-24, 25-44, 45-64 and 65+). There is one extra age group used in PREVENT (20-24, 25-34, 35-49, 50-64 and 65+ years). Estimates for the age group 35-49 were averaged from the age groups 25-44 and 45-64 in the Labour Force Surveys.

Figure 11 Past Smoking Prevalence and Projections
Canadian Females 25-44 Years



For each sex and age group, available data points were plotted between 1966

Table 13 Smoking Prevalence and Age Group Trends

Date	Age Group Prevalence and Trends (% change in prevalence across years)				
	20-24 years	25-34 years	35-49 years	50-65 years	65+ years
Males					
1970	55	56	54	52	31
	(-2.5)	(-2.3)	(-2.4)	(-2.3)	(-2.3)
1980	41	43	41	40	24
	(-3.2)	(-2.8)	(-2.7)	(-2.8)	(-2.9)
1990	28	31	30	29	17
	(-3.6)	(-3.2)	(-3.3)	(-3.1)	(-2.4)
2000	18	21	20	20	13
	(-3.9)	(-3.8)	(-3.5)	(-3.5)	(-3.1)
2010	11	13	13	13	9
	(-3.6)	(-3.8)	(-3.8)	(-3.8)	(-3.3)
2020	7	8	8	8	6
	(-4.3)	(-3.8)	(-3.8)	(-3.8)	(-3.3)
2030	4	5	5	5	4
	(-5.0)	(-4.0)	(-4.0)	(-4.0)	(-2.5)
2040	2	3	3	3	3
	(-5.0)	(-4.0)	(-4.0)	(-4.0)	(-2.5)
Females					
1970	42	36	34	32	10
	(-1.0)	(-0.3)	(-0.6)	(-0.9)	(+0.1)
1980	38	35	32	29	11
	(-1.1)	(-0.6)	(-0.9)	(-1.4)	(+1.8)
1990	34	33	29	25	13
	(-1.2)	(-0.3)	(-0.7)	(-0.8)	(+0.8)
2000	30	32	27	23	14
	(-1.3)	(-0.6)	(-0.7)	(-1.3)	(+1.4)
2010	26	30	25	20	16
	(-1.2)	(-0.3)	(-0.8)	(-1.0)	(+1.3)
2020	23	29	23	18	18
	(-1.3)	(-0.7)	(-0.9)	(-1.1)	(+1.1)
2030	20	27	21	16	20
	(-1.5)	(-0.4)	(-0.4)	(-1.3)	(+1.0)
2040	17	26	20	14	22
	(-1.5)	(-0.4)	(-0.4)	(-1.3)	(+1.0)

* The percentages (in brackets) in this table represent changes in smoking prevalence per year as a percentage of those exposed. The "+" entries are increases in prevalence and are included as inflows, the "-" are decreases and included as outflows.

and 1986. These points were then used to fit a logistic function. This function was used to estimate prevalence from 1990 to 2040. Based on these projections, percent changes from year to year were determined. An example of the plotted points, the logistic curve fitted to them and the future projections are presented in Figures 10 and 11 for smoking prevalence of males and females aged 25-44. A table of prevalence between 1970 and 2040 for the all age groups is presented in Table 13. This table is

the basis for trend estimates of the "age group" option. The "age group" inflow and outflow trends are estimated by determining the percentage change from year to year, by age. For example the change in prevalence for females aged 20-24 from 2010 to 2020 (from 26% to 23%) represents a 10 year decrease of 3%. This translates to a yearly change during that period of 0.3%. The annual percent decrease in smoking prevalence is then calculated (0.3/26 x 100) as 1.2%.

The "cohort" inflow and outflow trends are estimated by determining the change from year to year, by cohort (e.g. the change in prevalence for females in the 1990 birth cohort from 2010 to 2020 expresses the ten year trend). Table 14 presents smoking prevalence data arranged by birth cohort.

The smoking inflow trend file for the "age group" option is made up of zero entries except for small increases for older women. These positive inflows reflect smoking patterns expected as the young women today, with their high smoking patterns, move into older age groups. The outflow trend file for "age group" reflect the small but steady decline in smoking prevalence among all adults over 25 years of age.

The smoking inflow file for the "cohort" option is made up of all zero entries. The "cohort" option outflows model small but consistent declines as each cohort ages. The values of these declines are estimated from the data in Table 14.

The inflow and outflow changes are assigned to all three exposure categories.

4.9.3 Trends in Hypertension

Data Required The combined latency plus lag time for hypertension associated

Table 14 Smoking Prevalence and Cohort Trends

Year	Birth Cohort				
	20-24 in 1990	25-34 in 1990	35-49 in 1990	50-64 in 1990	65+ in 1990
Male 1970			55 (-2.2)	55 (-2.5)	52 (-3.8)
1980		41 (-2.4)	43 (-3.0)	41 (-2.9)	32 (-4.7)
1990	28 (-2.5)	31 (-3.5)	30 (-3.3)	29 (-5.5)	17 (-2.4)
2000	21 (-3.8)	20 (-3.5)	20 (-3.5)	13 (-3.1)	13 (-3.1)
2010	13 (-3.8)	13 (-3.8)	13 (-5.4)	9 (-3.3)	9 (-3.3)
2020	8 (-3.8)	8 (-5.0)	6 (-3.3)	6 (-3.3)	6 (0.0)
2030	5 (-4.0)	4 (-2.5)	4 (-2.5)	4 (0.0)	
2040	3 (-4.0)	3 (-2.5)	3 (-2.5)		
Female 1970			39 (1.3)	35 (-1.1)	32 (-4.1)
1990		38 (-1.3)	34 (-1.5)	31 (-1.9)	19 (-2.6)
2000	32 (-2.2)	27 (-1.5)	25 (-2.4)	18 (-1.1)	13 (0.0)
2010	25 (-2.8)	23 (-2.2)	19 (-0.5)	16 (0.0)	16 (0.0)
2020	18 (-1.1)	18 (0.0)	18 (0.0)	18 (0.0)	18 (0.0)
2030	16 (0.0)	20 (0.0)	20 (0.0)	20 (0.0)	
2040	22 (0.0)	22 (0.0)	22 (0.0)		

The percentages (in brackets) in this table represent decreases in smoking prevalence per year as a percentage of those exposed. These decreases are expressed as outflows.

with both IHD and CVA is 5 years. For the inflow and outflow files, changes in uncontrolled hypertension prevalence (by sex, four age groups and two exposure

levels) are needed for the years 1985 to 1990 and projected changes are needed for the years 1991 to 2040.

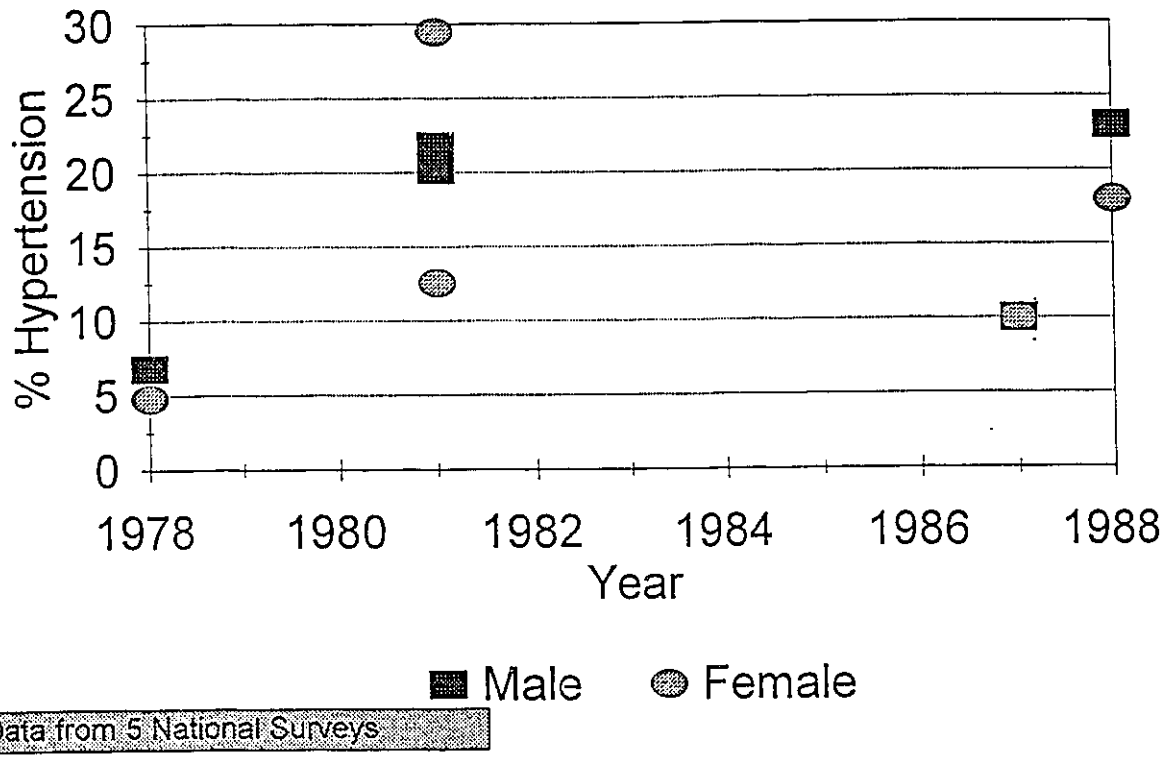
Data Sources Information on general uncontrolled hypertension prevalence (all age groups) of Canadians is available from 1978 from a few sources. The Canada Health Survey of 1978[50] reports diastolic blood pressure (DBP) of 95 mm Hg or greater and systolic blood pressure (SBP) of 145 mm Hg or greater by sex and age groups (20-24, 25-44, 45-64 and 65+ years). For all ages combined, uncontrolled hypertension (DBP$\geq$95) is reported for 6.8% of males and 4.8% of females.

The Canada Fitness Survey of 1981[51] also reported percentage of subjects by sex and age group (20-29, 30-39, 40-49. 50-59 and 60-69 years). This study defined hypertension as DBP$\geq$90 mm Hg or taking antihypertensive medication. Overall hypertension was reported as 21.8% of males and 12.6% of females.

A ten year follow up to the Nutrition Canada survey[57] reports hypertension in 20.2% of male and 29.5% of female subjects. In this study hypertension was defined as DBP$\geq$100 mm Hg or being treated by a physician. The Canadian Blood Pressure Survey of 1987[58] reported prevalence of uncontrolled high blood pressure by sex and three age groups (18-44, 45-64 and 65+ years). Uncontrolled hypertension was defined as DBP$\geq$90 mm Hg. Overall 10% (male and female) of the survey sample had uncontrolled high blood pressure and another 8% were under treatment for hypertension.

For all ages combined the Canadian Heart Health surveys (1986-1990)[54] reported uncontrolled hypertension (DBP$\geq$90 mm Hg) for 23% of men and 18% of women.

Figure 12 Past Prevalence of Uncontrolled Hypertension



Data from 5 National Surveys

Problems From the overall prevalence presented above and presented in Figure 12 it is apparent that due to differences in survey methods, age group cut points and definitions of hypertension, there is little basis for determination of trends. A second problem is the lack of data specific to Ontario.

Discussion Unfortunately, we do not have a series of related surveys from which to establish trends. A similar problem was encountered in the development of the Dutch database. Their solution was to consider the improved diagnosis and treatment of hypertension in the last decade or more, estimate it at a 1% improvement per year

and project that improvement rate over the full simulation period. This approach is the one I have adopted for outflow trends in the Ontario database. The Canadian Blood Pressure survey [58] reports that, compared to the Canada Health Survey data a decade earlier, prevalence of uncontrolled hypertension (DBP$\geq$ 90 mm Hg) had decreased from 18 to 10%.

Since hypertension is largely an age-dependent risk factor, it is reasonable to assume that, in the future, similar proportions of the population will develop hypertension at each age as those observed in 1990. This assumption would lead to no inflows in the "age group" option. In order to offer a "cohort" option, inflow trends would need to be adjusted to mimic the increasing prevalence of uncontrolled hypertension as cohorts age.

Resolution The estimate of an 8% decline over 10 years, 0.8% per year, is used as a general indicator of trend in uncontrolled hypertension. It is applied to all age groups and exposure categories for the both the "age group" and "cohort" option outflow files.

For the "age group" option, the inflow trend file is all zero values. For the cohort option, age-related inflows were estimated based on the age of the cohort during the simulation period. The estimated prevalence of uncontrolled hypertension and the percent change per year in those exposed are presented in Table 15.

The inflow and outflow trends are applied to both exposure categories.

4.9.4 Trends in Hypercholesterolemia

Data Required The latency plus lag time for high cholesterol associated with

IHD is 6 years. Trend data are needed for both sexes, in four age groups and for two exposure categories over the period from 1984 to 2040.

Data Sources In the report of the Canada Health Survey of 1978[50], prevalence of high cholesterol was reported by sex, six age groups and blood level. In general, 36.8% of males had mildly elevated cholesterol (5.2-6.5 mmol/L) and 11.8% had very elevated cholesterol levels ($\geq$ 6.6 mmol/L). For females the corresponding figures are 34.0% with mild elevation and 11.9% with very high cholesterol.

The Nutrition Canada Survey of 1971[59], also reports prevalence of a range of cholesterol levels by sex and age group. This survey also reports the same information for Ontario alone but with much smaller sample sizes. For the age group 20-39, 25.8% of males and 25.7% of females had mildly elevated cholesterol levels (5.2-6.1 mmol/L) and 12.2% of males and 6% of females had very high levels ($\geq$ 6.2 mmol/L). For the age group 45-54, 39.5% of males and 35.9% of females had mildly elevated cholesterol levels and 17.4% of males and 25.2% of females had very high levels. For the age group 65+, 38.6% of males and 40.4% of females had mild elevation and 15.2% of males and 43.5 % of females had very high cholesterol levels.

Problems Unfortunately, the Canada Health Survey, the Nutrition Canada Survey and the Canadian Heart Health Survey (source of 1990 prevalence data) all report their data for different age groups, and slightly different cut points. Another problem is the lack of data specific to Ontario.

Discussion The two past surveys do not allow estimation of trends in high cholesterol. The available Dutch data were similarly limited. Gunning-Shepers chose to assume that hypertension and high cholesterol were similar and that trends in

Table 15 Uncontrolled Hypertension Prevalence and Cohort Trends

Year	Birth Cohort			
	Born 1946-55 35-44 in 1990	Born 1936-45 45-54 in 1990	Born 1626-35 55-64 in 1990	Born 1916-25 65+ in 1990
Male 1985		23 (+2.2)	39 (+2.8)	41 (+0.2)
1990	18 (+5.5)	28 (+7.9)	50 (+0.4)	52 (0.0)
2000	28 (+7.9)	50 (+0.4)	52 (0.0)	52 (0.0)
2010	50 (+0.4)	52 (0.0)	52 (0.0)	52 (0.0)
2020	52 (0.0)	52 (0.0)	52 (0.0)	52 (0.0)
2030	52 (0.0)	52 (0.0)	52 (0.0)	
2040	52 (0.0)	52 (0.0)		
Female 1985		15 (+4.7)	33 (+3.0)	51 (+1.4)
1990	7 (21)	22 (9.5)	43 (3.5)	58 (0.0)
2000	22 (+9.5)	43 (+3.5)	58 (0.0)	58 (0.0)
2010	43 (+3.5)	58 (0.0)	58 (0.0)	58 (0.0)
2020	58 (0.0)	58 (0.0)	58 (0.0)	58 (0.0)
2030	58 (0.0)	58 (0.0)	58 (0.0)	
2040	58 (0.0)	58 (0.0)		

The percentages (in brackets) in this table represent increases in hypertension prevalence per year as a percentage of those exposed. These increases are expressed as inflows.

treatment and control would be similar. For the Dutch database the inflow and outflow files for the "age group" option in hypertension were adopted for high cholesterol. There is some evidence that total cholesterol levels have declined somewhat in the last 10 to 20 years in North America[47]. The magnitude of the

decrease is not specified. Although there is evidence that lifestyle changes and drug treatment can effectively control hypercholesterolemia, I do not believe that there is enough evidence to claim that the treatment and control is as effective as control of hypertension. Like hypertension, high cholesterol is largely an age-dependent risk factor, so it is reasonable to assume that similar proportions of the population will develop high cholesterol at similar ages to those observed in 1990. This assumption would lead to no inflows.

Resolution As a compromise between the alternative of assuming no trends towards effective treatment and control and assuming trends of the same magnitude as that of hypertension, I have arbitrarily decided to apply outflow trends of 0.4% per year (half of the values for hypertension). The inflow trend file is all zero values. The inflow and outflow trends are applied to both exposure categories.

4.9.5 Trends in Alcohol Use

Data Required The latency plus and lag for alcohol associated with cirrhosis is 5 years (maximum latency plus lag). Trend data are needed for both sexes, in two age groups and for two exposure categories over the period from 1985 to 2040.

Data Sources There have been several large surveys that investigated levels of alcohol consumption in Canada, and one has also reported levels for Ontario[50, 52, 53, 55, 60]. These past alcohol prevalences, including age groups and cut points used, as well as the source of the data are presented in Table 16. Ideally, data reported would match the age groups (20-39 years and 40+ years) and exposure category cut points for abstainers (<4 drinks per week) and excessive drinkers (22 or more drinks per

Table 16 Past Prevalence of Alcohol Use

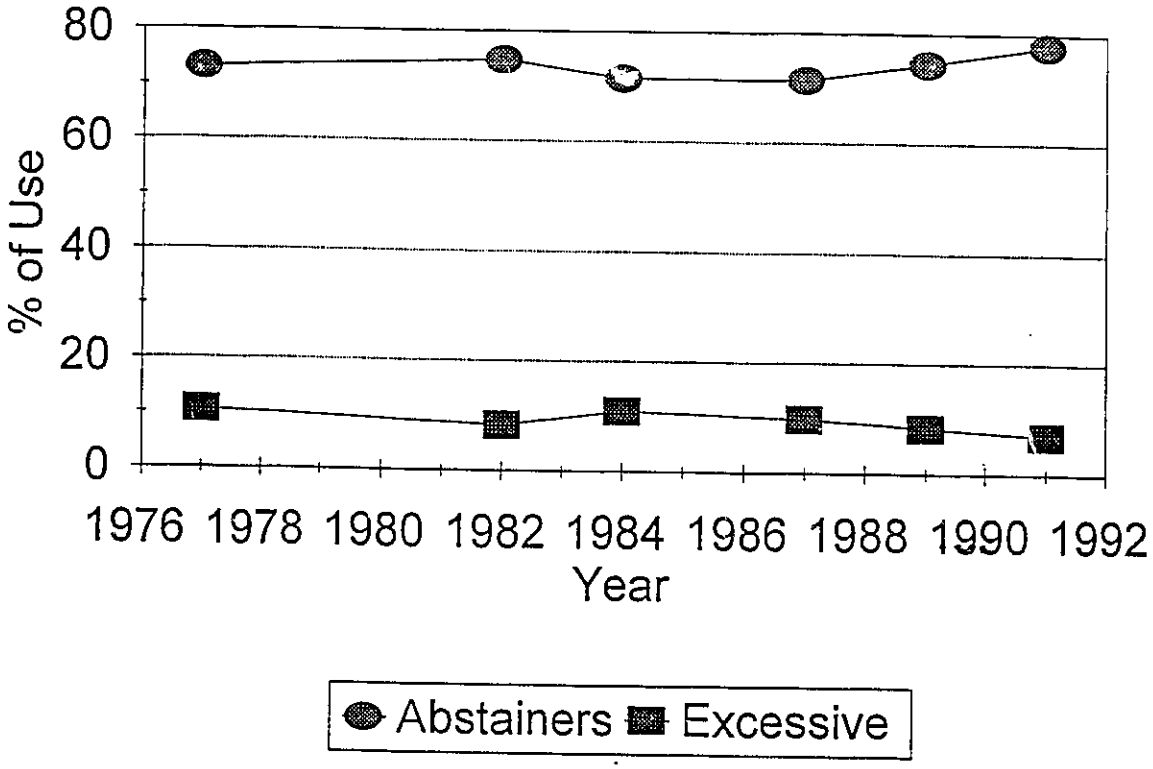
Source & Year	Age & Sex	Exposure Category	Prevalence (percent)	Actual Cut Points
Canada Health Survy 1978	20-44 Male	Abstainer	18.4	Never+occasional+<1 drink per week
	Female		39.5	
	45+ Male		24.1	
	Female		50.5	
	20-45 Male	Excessive	24.0	14+ drinks per week
	Female		5.7	
	45+ Male		15.5	
	Female		3.1	
General Social survey 1985	20-44 Male	Abstainer	26.4	Never-occasional+0 drinks per week
	Female		52.1	
	45+ Male		35.6	
	Female		59.3	
	20-44 Male	Excessive	15.5	14+ drinks per week
	Female		2.5	
	45+ Male		11.9	
	Female		1.9	
Health Promotion Survey 1985	20-44 Male	Abstainer	27.4	Never+0 drinks per week
	Female		43.4	
	45+		31.5	
	Male		50.9	
	Female		9.3	22+ drinks per week
	20-44 Male	Excessive	0.5	
	Female		4.3	
	45+ Male		0.0	
	Female			
Alcohol & Other Drugs Survey 1989	20-44 Male	Abstainers	36.8	Never +0 drinks per week
	Female		60.4	
	45+ Male		43.2	
	Female		74.0	
	20-44 Male	Excessive	11.5	14+ drinks per week
	Female		1.7	
	45+ Males		13.7	
	Females		3.6	
Addiction Research Foundation of Ontario	All			
	1977	Abstainers	73.2	Never+not more frequent than once a week Daily
		Excessive	10.7	
	1982	Abstainers	74.9	
		Excessive	8.3	
	1984	Abstainers	71.6	
		Excessive	10.9	
	1987	Abstainers	71.5	
		Excessive	9.8	
	1989	Abstainers	74.5	
		Excessive	8.3	
	1991	Abstainers	77.9	
		Excessive	6.9	

week) identified in other PREVENT data. In the construction of Table 16, age

groups and categories were aggregated to match these categories as closely as possible.

Problems Although there is a series of related surveys for Ontario, the results are not reported in appropriate age groups or by the cut points needed in PREVENT.

Figure 13 Past Prevalence of Alcohol Use



Discussion From Table 16 it is difficult to identify any trend over time. This is due in large part to differences in survey design, age groups and exposure cut points. But even if the related Addiction Research Foundation surveys are considered, Figure 13, there is little apparent trend. Between 1987 and 1991 there may be some

movement towards lower consumption, but this covers a small period of time.

Resolution On the basis of the available data, I am unwilling to predict any future change in prevalence of abstainers or excessive drinkers (as defined for this database). The inflow and outflow files will contain no changes (all zero entries).

4.9.6 Trends in Obesity

Data Required The combined latency plus lag time for obesity associated with breast cancer is 2 years. Trend data are needed for both sexes, two age groups and one exposure category over the period from 1988 to 2040.

Data Source There have been several large surveys that have reported on body weight and obesity[51, 52, 56, 59, 61]. These past prevalences, including age groups and cut points used are presented in Table 17. In the construction of this table age groups and categories were matched as closely as possible to those used in PREVENT (ages 40-49 years and 50+; BMI$\geq$30).

Problems There are no trend data available specific to Ontario, and the Canadian data available do not match on age groups and cut points.

Discussion It is difficult to identify any trend in obesity prevalence over time. This is due in part to differences in surveys. In a recent survey[61] the authors review the available data and conclude that recent levels of obesity have not changed substantially from two previous national surveys.

Resolution On the basis of data available, I am unwilling to predict any change in patterns of obesity in Ontario. The inflow and outflow files will not predict change (all entries zero).

Table 17 Past Prevalence of Obesity

Source & Year	Age & Sex	Exposure Category	Prevalence (percent)	Actual Cut Points
Nutrition Canada Survey 1971	40-64	Obese		Ponderal Index
	Male		4.2	
	Female		20.6	
	65+			
	Male		8.6	
	Female		44.1	
Canada Fitness Survey 1981	40-49	Obese		BMI $\geq$ 30
	Male		11.9	
	Female		15.9	
	50-59			
	Male		12.1	
	Female		19.6	
	60-69			
	Male		14.1	
	Female		22.0	
Health Promotion Survey 1985	45-54	Obese		BMI > 27
	Male		29.2	
	Female		17.1	
	55-64			
	Male		23.9	
	Female		22.5	
	65+			
	Male		19.9	
	Female		19.6	
Canadian Heart Health Survey 1988	45-54	Obese		BMI $\geq$ 30
	Male		18	
	Female		17	
	55-64			
	Male		22	
	Female		27	
	65-74			
	Male		17	
	Female		23	
Health Promotion Survey 1990	45-54	Obese		BMI > 27
	Male		35	
	Female		23	
	55-64			
	Male		33	
	Female		27	
	65+			
	Male		-	
	Female		22	

5.0 EXPANSION OF THE DATABASE

5.1 AIDS

With the recent onset of the AIDs pandemic, including devastating effects on morbidity and mortality in Africa and Asia[62,63] and predictions of accelerating incidence in Europe and North America[62], it seemed worthwhile to investigate the possibility of including AIDS and its risk factors in the PREVENT database for Ontario. Initially this seemed feasible. The mortality data for Aids in Ontario are known, the behavioural risk factors are open to intervention, some relative risks and odds ratios have been reported in the literature and with some assumptions the prevalence of AIDS (or HIV infection) in the general population might be estimated.

In the course of my literature review several large problems became apparent. Most researchers conclude that HIV infection leads almost inevitably to development of AIDS, although there is considerable variation in estimates of the incubation period between HIV infection and presentation of AIDS symptoms and the period between presentation with AIDS and death[64]. This is not an insurmountable problem especially if HIV infection is taken as the incidence point in the development of AIDS. PREVENT models mortality from chronic diseases, with modifiable risk factors, in large general populations using etiologic fractions calculated from relative risks that are assumed to be fixed. AIDS is an infectious disease. Almost all of the research on AIDS has been conducted in high risk, high prevalence, subgroups (primarily homosexual males and intravenous drug users)[65,66]. Very specific behavioural risk factors have been identified and the relative risks of these behaviours calculated within these high risk subgroups. However, even within subgroups such as IV drug

users, the relative risks vary with HIV seroprevalence. On a biological level, an individual is only at risk if the risk behaviour includes direct exposure to body fluids that are HIV infected. On a population level, relative risk for a given risk behaviour is then proportional to the prevalence of HIV infection in that population. An IV drug user is at much greater risk of HIV infection if he shares needles in a high prevalence area such as large city compared to a low prevalence area such as a smaller rural town. The PREVENT requirement of fixed relative risks cannot be met for risk of HIV infection or AIDS.

Also, in order to obtain accurate relative risks for a general population, studies must measure risk behaviours and determine the HIV infection status in a representative sample of that population. To my knowledge this has not been done. One large seroprevalence study has been published[67]. It involved applicants for military service in the U.S. and cannot be considered truly representative of the general population in that females, older adults, homosexuals and IV drug users are probably under represented. It was not linked to a survey of risk behaviours. The results of a logistic regression analysis reported HIV risk in terms of gender, age groups, racial origin, and urban densities. These risk factors, in reality, reflect differential HIV prevalence by gender, age group, etc. They are incompatible with the requirement of modifiable risk behaviours such as number of sexual partners, condom use and needle sharing activity which could be used in the PREVENT model.

The National Center for Health Statistics conducted pilot studies in two counties in Pennsylvania and Texas[68,69], but based on the results and problems encountered, decided against conducting a large general population study of HIV

infection and its risk factors. The problems included low prevalence of HIV infection and high rates of refusal on the risk behaviour survey. The investigators believe that refusal rates were highest among those at high risk, causing considerable (and unmeasurable) bias.

In addition to the above problems, the HIV prevalence in Ontario is not easily estimated. There have been no general population seroprevalence surveys in this province or elsewhere in Canada.

Because of these problems, especially the lack of general population studies, I have decided not to include AIDS in the Ontario Database for PREVENT.

5.2 Seat Belt Use

The inclusion of seat belt use as a risk factor for motor vehicle crash seems quite natural for PREVENT. It is a protective behaviour, used widely in the general population and its safety value is well established and well accepted. In order to include seat belt use in the Ontario database for PREVENT, an estimate of the relative risk across age and sex categories is required. Also needed are 1990 estimates of seat belt use in Ontario over a full range of age groups in the population. This prevalence information is available from the Ontario Health Survey. The age and sex specific mortality rates for motor vehicle crash are already a part of the Ontario database, for use with another risk factor.

5.2.1 Relative Risk for Seat Belt use in Motor Vehicle Crash

Data Sources The literature available on seat belts was reviewed for estimates

of risk reduction associated with seat belt use. Several reports gave risk estimates for subgroups of the population, such as the elderly. Others estimated risk under conditions too specific for general population estimates (e.g. specific to seating placement or vehicle type). Health Risk Appraisal includes risk estimates for seat belts. The Evalu-Life Report[70] uses a relative risk associated with seat belt use of 0.5 (no confidence limits are given). Evans, at the General Motors Research Laboratory, has investigated seat belt use extensively. Using the Fatal Accident Reporting System (FARS) data from the United States as the basis for a large case control study[71], he developed an measure of "Effectiveness" (actually efficacy) which appears to be equivalent to a prevented fraction (exposed). Using the following relationship, a relative risk can be estimated from this "Effectiveness" value:

$$PF(E) = 1 - RR$$

$$RR = 1\text{-}PF(E)$$

The overall seat belt effectiveness of 41% translates to a relative risk of 0.59. Research by McGee and Rhodes[72], suggests that seat belt effectiveness has been improving. Using FARS data in a stratified case control study, they estimate the odds ratio for seat belt use between 1981 and 1986 of 0.425. A study focusing on alcohol and traffic safety[73], includes seat belt use as a variable in a logistic regression analysis and reports an odds ratio of 0.23 (95% CI=0.14-0.39).

Problems Confidence intervals are generally not available for these risk estimates, although in the McGee and Rhodes study the odds ratio (0.425) was statistically significant. These estimates are good general population estimates but they do not indicate any variation for age or sex.

Resolution Overall the estimates of risk range from 0.2 to 0.6. I have chosen a relative risk of 0.55 for use in the Ontario database for PREVENT. This value is between estimates from the two large studies using FARS data. This relative risk will be assigned to both sexes and across all age groups.

I have chosen a latency of 0 because the potective effect of seat belt is immediate. The lowest lag value in PREVENT is 1 year and this will be used. All direct fatalities are likely to occur in less than one year.

5.2.2 Prevalence of Seat Belt Use

Data Required Prevalence of seat belt use information is needed for males and females in the age categories which reflect differences in use. Data are needed for only a single exposure category, seat belt use, and an estimate of "seat belt lapse"

Table 18 Prevalence of Seat Belt Use in Ontario, 1990

Exposure	Sex	Age Categories		
		15-29	30-64	65+
Current Status Seat belt use	M F	.759 .876	.856 .942	.942 .958
Ex Seat belt use	M F	.000 .000	.000 .000	.000 .000

(previous exposure) is also needed. The reference category is seat belt not used.

Data Sources The prevalence data were abstracted directly from the Ontario Health Survey data. Unfortunately, there is no question in the OHS that would assess lapse from previous use of seat belts.

Resolution Three age categories were chosen to reflect variation in seat belt use at different ages: 15-29 years, 30-64 years and 65+ years. The prevalence data used are presented in Table 18 . The assumption has been made of no lapse from previous seat belt use.

5.2.3 Trends in Seat Belt Use

Data required Estimates of the inflow and outflow trends for use of seat belts are needed for the years 1989 to 2040 by sex and three age groups (15-29, 30-64 and 65+).

Table 19 Past Prevalence of Seat Belt Use in Canada

Source & Year	Age & Sex	Exposure Category	Prevalence (percent)
Canada Health Survey 1978	15-25 25-64 65+	Seat belt use "always" or "most of time" drivers and passengers	44.1 56.6 52.5
Health Promotion Survey 1985	15-24 Male Female 25-64 Male Female 65+ Male Female	Seat belt use "always" or "most of time" drivers and passengers	 68.4 75.3 59.4 72.7 73.8 80.4
Health Promotion Survey 1990	15-24 Male Female 25-64 Male Female 65+ Male Female	Seat belt use "always" or "most of time" drivers and passengers	 64 76 72 88 81 90

Data Sources Seat belt use has been surveyed in three national surveys, the Canada Health Survey of 1978[50], the Health Promotion Survey of 1985[52] and the

Health Promotion Survey of 1990[56]. The estimates of seat belt use in age groups as close as possible to those needed are presented in Table 19. Discussion in the Canada Health Survey notes that in the four provinces where seat belt use was mandated by law, prevalence was 60% overall, compared to 16% in provinces where seat belt use was not mandatory.

Problems The past prevalence data available are for Canada rather than the province of Ontario. None of the surveys available estimate the prevalence of "seat belt lapse", so there is no indication of the prevalence of ex-exposed.

Discussion As can be seen in Table 19, seat belt use has increased substantially in all categories. In fact the 1990 prevalence estimates for Ontario are in excess of 90% in some strata. Since 1978, legislation has spread across the country with associated increase in seat belt use. Since this legislation itself and the level of enforcement of the laws affects seat belt use it is not possible to predict future trends from past prevalence of seat belt use. The uptake of innovation theory[48] indicates that behaviours will follow an uptake curve and then slowly approach a saturation level. Furthermore, with prevalence in some categories in excess of 90% in some strata (and clearly approaching the ultimate saturation level of 100%), there is little room left for increased use. The saturation level need not be the same for all strata, so there is no indication of the saturation level for younger age groups without further intervention. While the past prevalence data show a pronounced trend towards increased use of seat belts, I am unwilling to estimate continued improvements in the future because I feel there is sufficient evidence that the saturation levels are being approached and any substantial changes in the future will be dictated by levels of enforcement of seat belt

laws.

Resolution The inflow and outflow trends for seat belt use will be assigned zero values (modelling no change). There is assumed to be no "lapse" in seat belt use.

6.0 PREVENT OUTPUT

As the main PREVENT program is executed, it produces a series of user-friendly menus offering the user a range of options. These menus include choice of

Table 20 Output Options in PREVENT

PREVENT Output Options	Description
Total Mortality	Absolute number of deaths (male and female) from all causes - for trend alone and trend plus intervention.
Total Mortality Reduction	Absolute number of deaths avoided by the intervention (male and female).
Potential Years of Life Gained	Potential years of life gained by deaths avoided in each year of the simulation (male and female).
Actual Years of Life Gained	Actual years of life gained by the intervention with each year in the simulation (male and female).
Survival Curve	Survival curve for the base population (1990) for any year in the simulation with and without intervention.
Life Expectancy at Birth	Compared to 1990 life expectancy for males and females born in any year of the simulation with and without intervention.
Etiologic Fraction	Specific to risk factor and disease, by 5 year age group for males and females, over the simulation period with and without intervention.
Potential and Trend Impact Fraction	Specific to disease, by 5 year age group for males and females over the simulation period with and without intervention.
Disease-Specific Mortality Rate	By 5 year age group for males and females over simulation period with and without intervention.
Disease-Specific Mortality	Absolute number of deaths, males and female, for each year in the simulation period with and without intervention.
Disease-Specific Mortality Reduction	Absolute number of deaths avoided, male and female, by the intervention for each year in the simulation.

database (if more than one is installed), risk factor(s) to be intervened upon, and number of years to be simulated (maximum 50). The menus then offer choices regarding interventions to be modelled, incluiding general versus youngest cohort,

number of years over which intervention is to be achieved, year in which intervention is to start and the actual risk factor prevalence (by age, sex and exposure category) to be achieved by the intervention. After the user has specified these conditions, the program runs the calculations and PREVENT offers a range of output options as presented in Table 20.

Output options in PREVENT are available in tabular and graphic format. Specific examples of output will be presented in the Results section below.

7.0 RESULTS

Five specific simulations have been chosen for presentation in the results section. These simulations have been chosen to demonstrate the range of options available in PREVENT and to illustrate the flexibility and scope of the PREVENT model. Selected output options will be presented for each simulation.

7.1 Increased Smoking Simulation

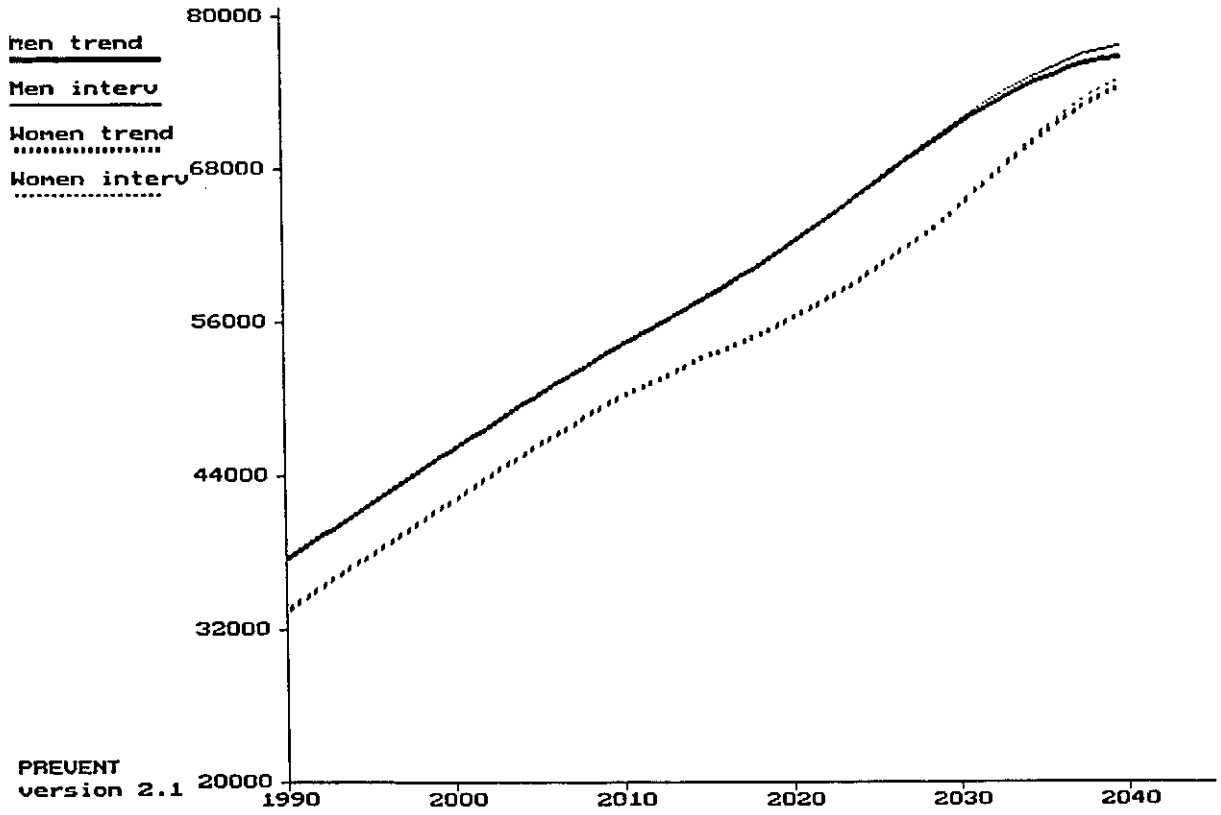
This simulation models a 50% increase in smoking among the youngest cohort; this increase is maintained as each new young cohort enters the model for the full 50 years of the simulation.

This model was chosen to demonstrate the cohort option, the effect of an increased risk factor prevalence, and the effect of one risk factor on several diseases. A resurgence in the popularity of smoking among young adults may occur. With recent changes in tobacco taxation, the cost of cigarettes in Ontario has dropped to approximately half its former level. Price of tobacco has been linked to consumption[74]. The sudden drop in tobacco price is anticipated to have two major effects. First, smokers of all ages will tend to increase their consumption. Second, the number of new smokers will increase. Since smoking is largely a behaviour established in youth, most of these new smokers will be young adults.

The choice of a large increase in prevalence among the youngest cohort (following a cohort option) will isolate and estimate the impact of increased smoking by young people.

Total Mortality Total mortality for this simulation is presented in Figure 14.

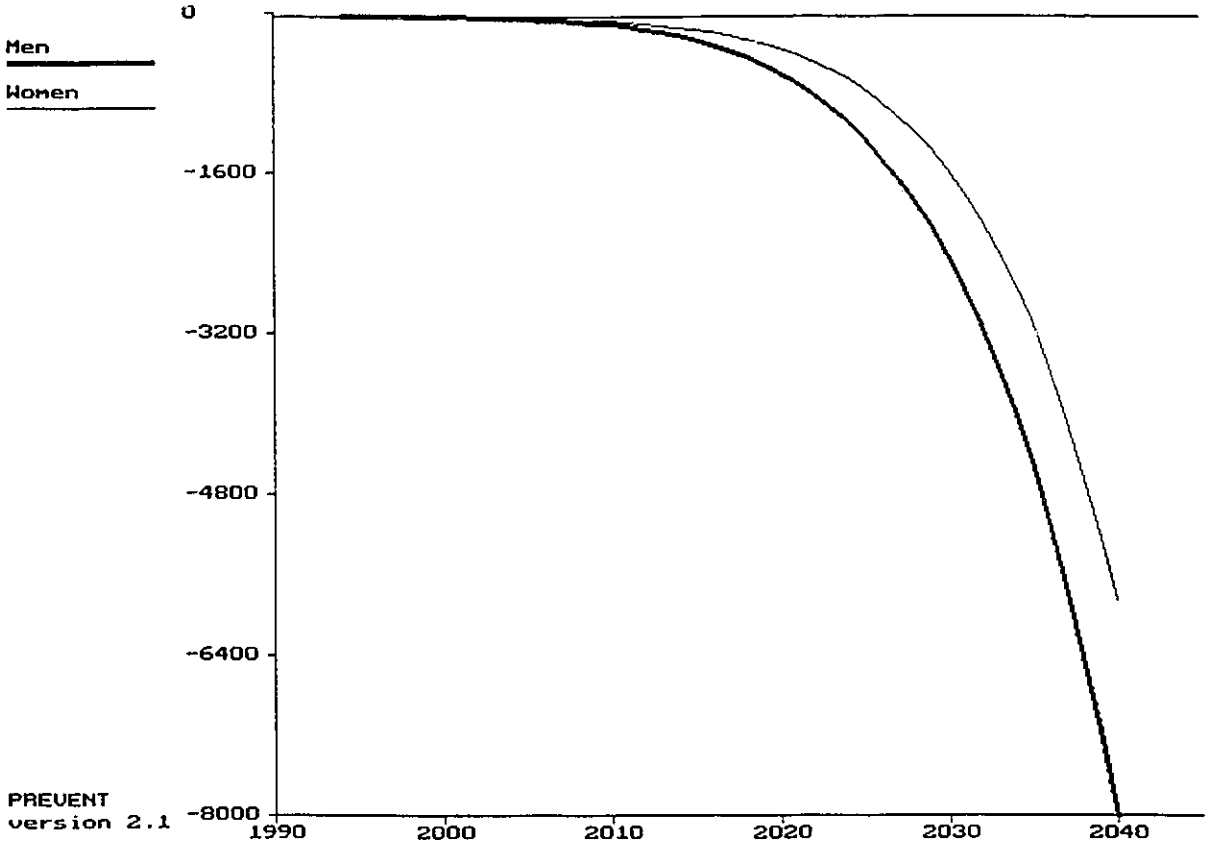
Figure 14 Total Mortality for 50% Smoking Increase Simulation



The most striking feature of this graph is the upward trend in total mortality over the simulation period. This trend exists with and without the intervention modelled and is largely due to the aging of the Ontario population projected for this period (higher mortality rates at older ages). The second feature to note is the very small overall impact of the intervention in all-cause mortality. The largest difference in deaths between trend and intervention is in 2040, with an estimated 1574 deaths additional deaths with increased prevalence of smoking.

Actual Years of Life Gained Because the intervention modelled has an negative

Figure 15 Years of Life Gained in 50% Smoking Increase Simulation

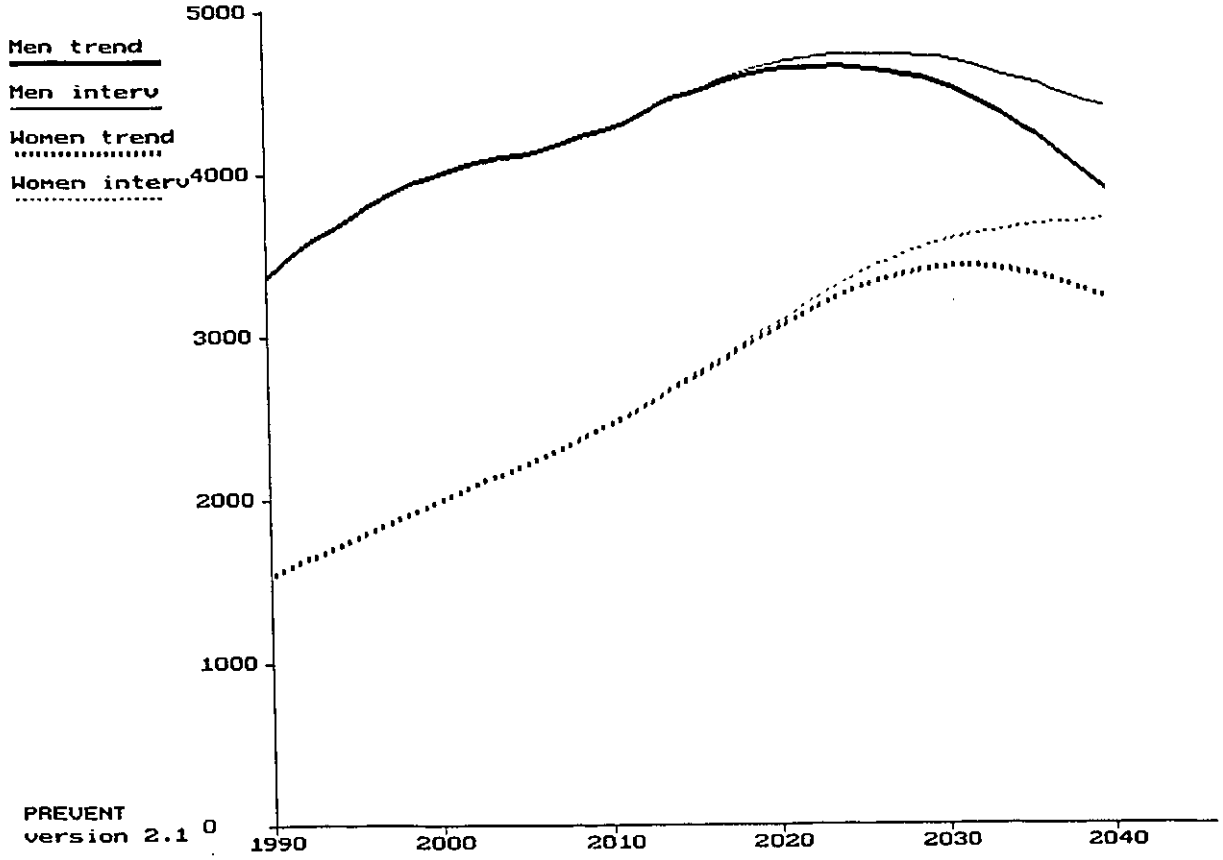


impact, this measure becomes actual years of life lost and is presented in Figure 15. By the end of the simulation period the cumulated years of life lost are 8019 for males and 5851 for females.

Life Expectancy By the end of the simulation there has been a small but measurable decrease in the life expectancy. For trend alone life expectancy at birth in 2040 is estimated at 75.67 years for males and 81.49 for females compared to 75.38 years for males and 81.22 for females under the increased smoking modelled.

Disease-Specific Mortality The disease-specific mortalities for lung cancer and

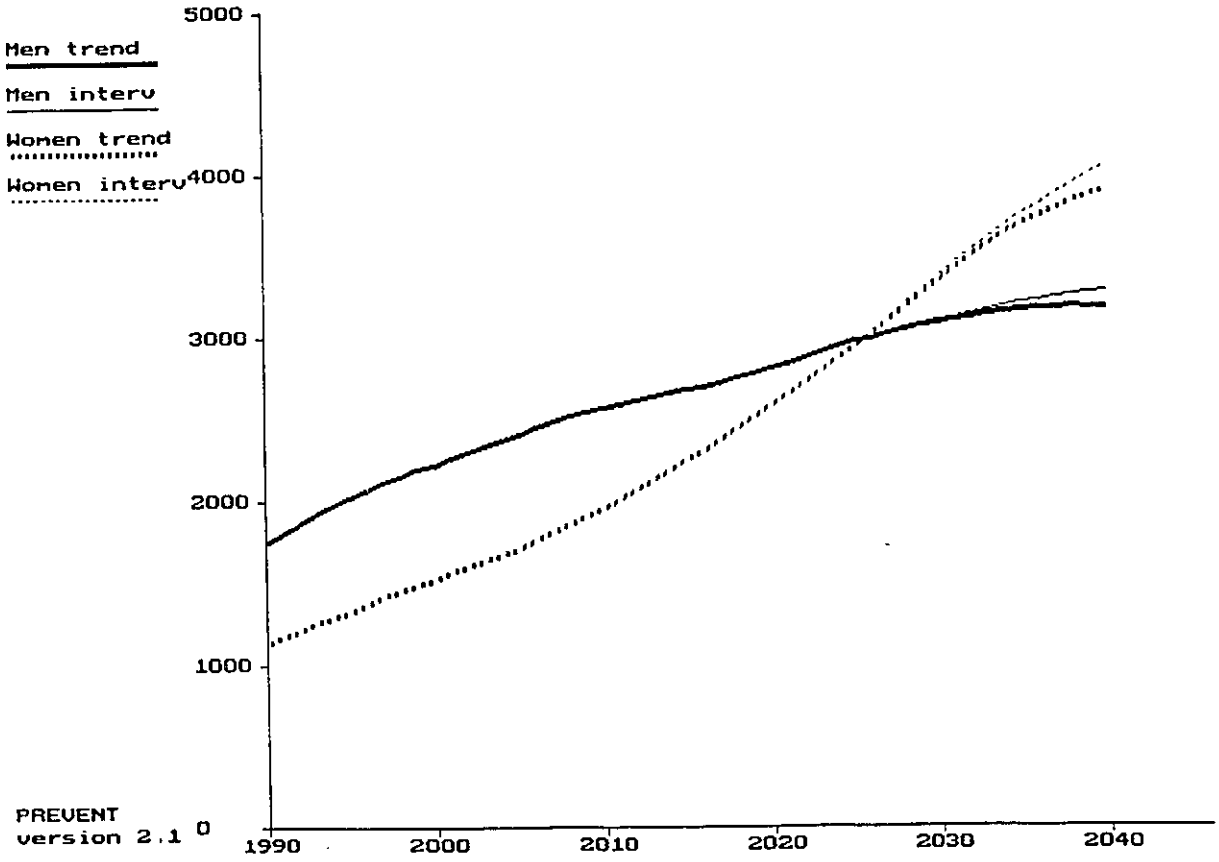
Figure 16 Lung Cancer Mortality for 50% Smoking Increase Simulation



COLD are presented in Figure 16 and 17. In all three diseases the trend is for increasing mortality rates over time. This effect is largely due to projected aging of the population. In the case of lung cancer there is some reduction after 2025. This is due to two effects; first, lung cancer rates decline in the oldest age groups (so the aging population has less effect on this outcome) and second, the overall trends in smoking are modelling a decrease as older adults stop smoking. In all three diseases the impact of increased smoking results in increased number of deaths by the end of the simulation period. In lung cancer and COLD the additional number of deaths in

the year 2040 is similar for men and women (510 and 490 deaths respectively due to lung cancer and 106 and 148 respectively due to COLD). .

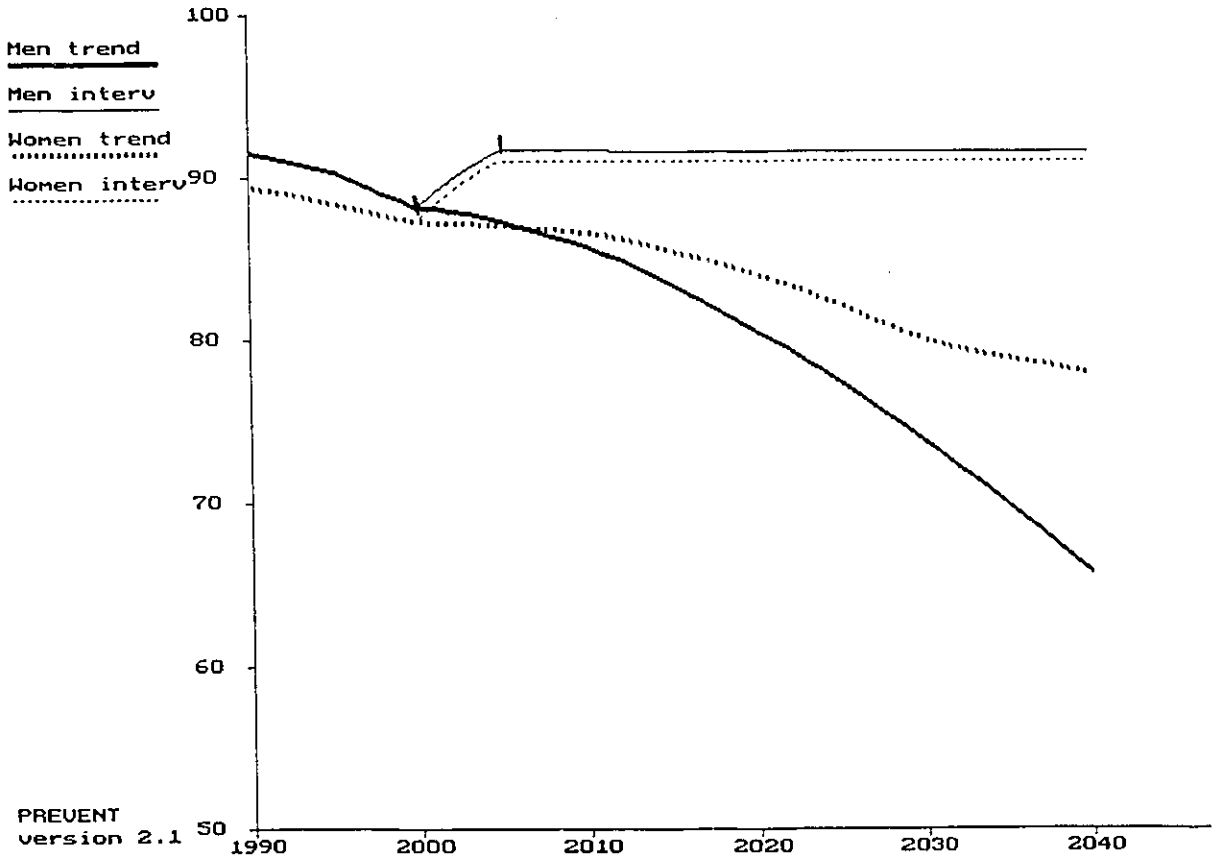
Figure 17 COLD Mortality for 50% Smoking Increase Simulation



One very interesting feature of the COLD simulation is the fact that females are modelled to overtake males in total deaths after 2025. This is due to the smoking trends modelled which have more males than females stopping smoking. At some point in the simulation the number of females smokers in the model exceeds the number of males smokers and the mortality from COLD is reflecting this. There is

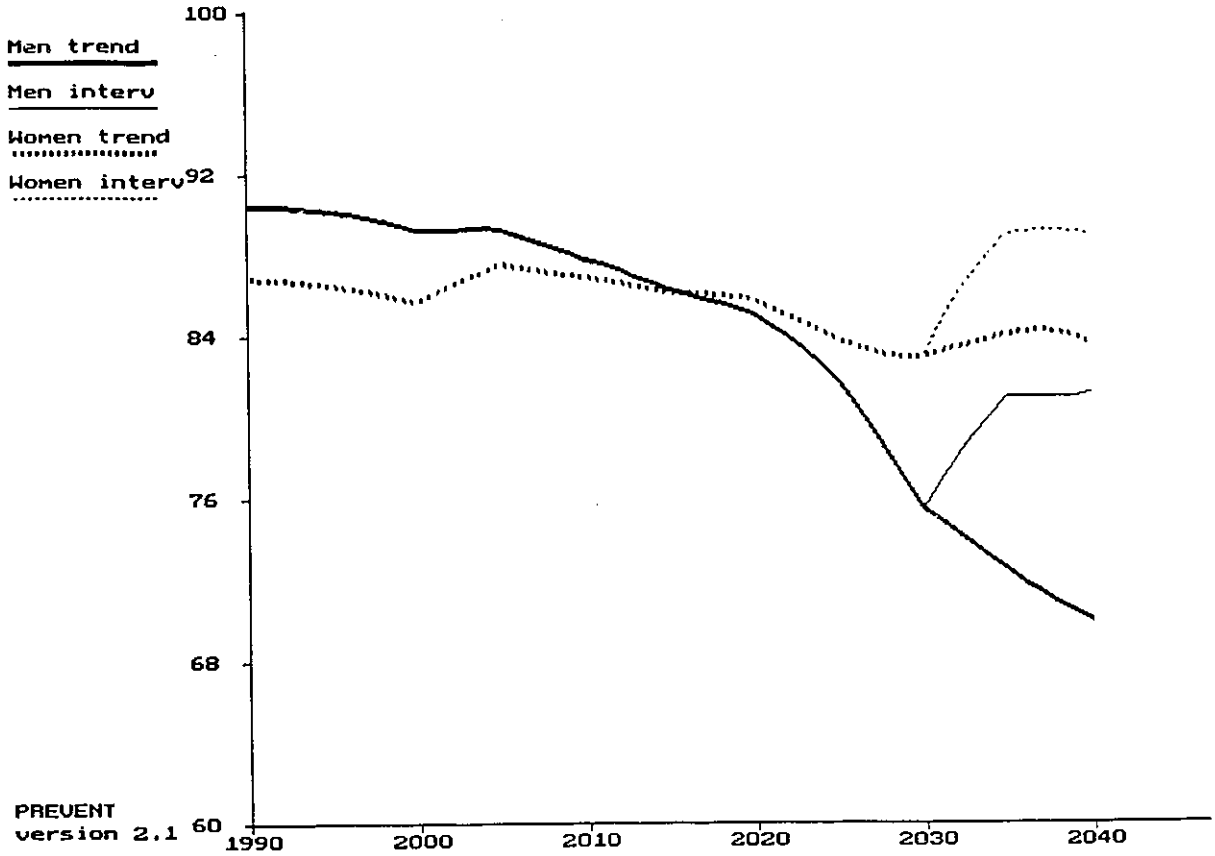
an indication that if the lung cancer simulation had extended a few years further a similar pattern would have been seen.

Figure 18 Etiologic Fraction for COLD Ages 30-34
50% Smoking Increase Simulation



Etiologic Fraction The etiologic fractions for smoking as it impacts on COLD is presented for age groups 30-34 in Figure 18 and 60-64 in Figure 19. Since the 20 year old cohort of 1990 is being intervened upon, the effect on etiologic fraction of 30-34 year olds is not seen until the 20 year old cohort of 1990 turn 30 in 2000. The

Figure 19 Etiologic Fraction for COLD Ages 60-64
50% Smoking Increase Simulation



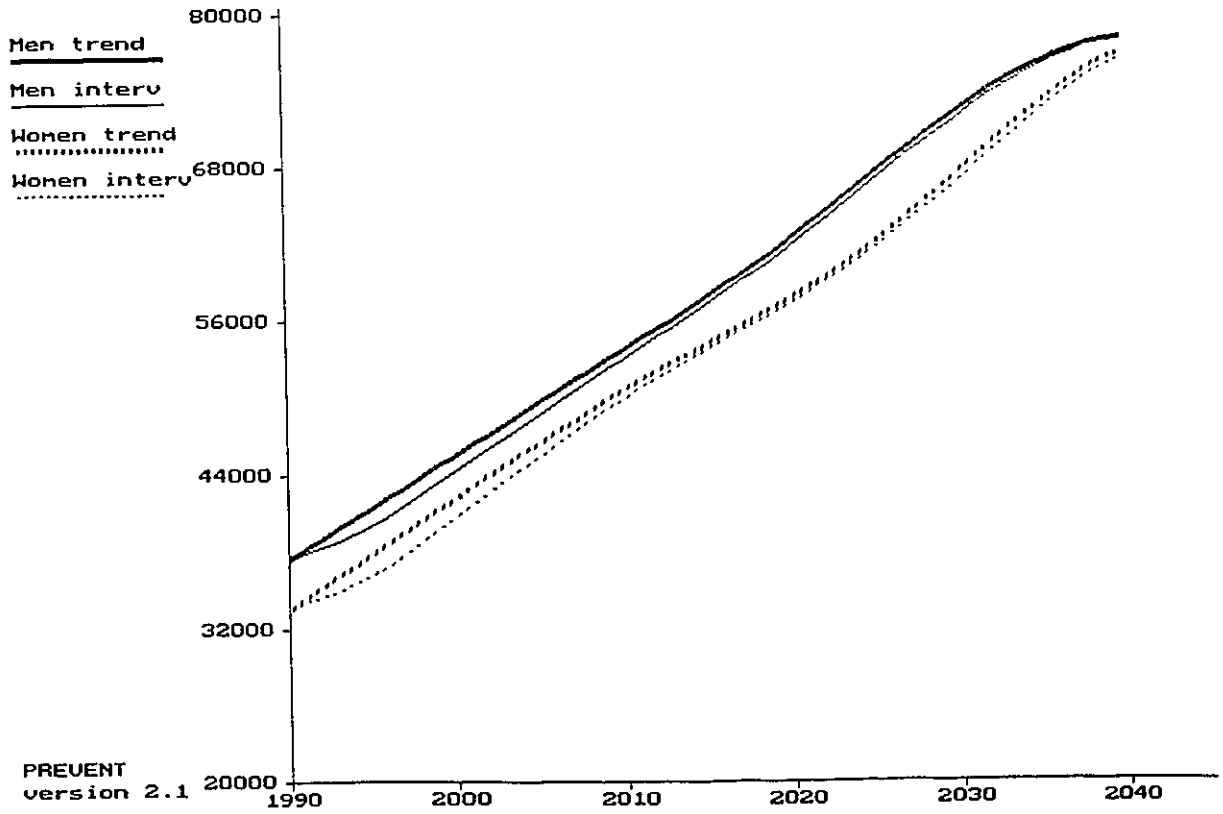
effect takes place gradually over five years, as five years of young adults intervened upon turn 30. Similarly, the effect on etiologic fraction for 60-64 year olds is not seen until the 20 year olds intervened upon turn 60 in 2030.

7.2 Decrease in Uncontrolled Hypertension and Hypercholesterolemia

This simulation models a 50% decrease in uncontrolled hypertension and a 20% decrease in hypercholesterolemia over 5 years. The intervention follows the age-group option. This simulation was chosen to demonstrate the effects of more than

one risk factor on a single disease (IHD), the effect of reductions in risk factors, and the age-group option for intervention. A larger intervention was chosen for hypertension than for hypercholesterolemia because the treatments for this risk are better established and have demonstrated efficacy for more patients.

Figure 20 Total Mortality for Hypertension and Hypercholesterolemia Decrease Simulation

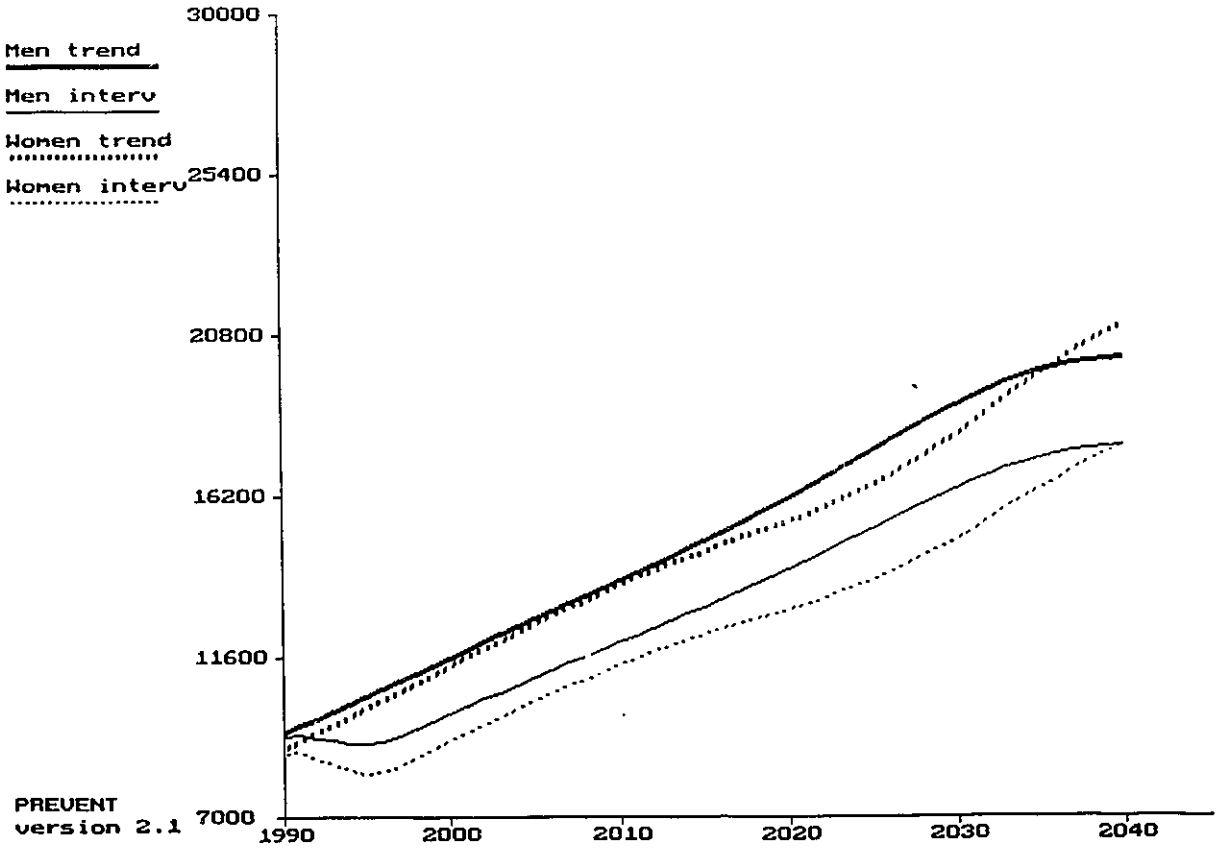


Total Mortality Total mortality is presented in Figure 20. The curve for trends alone is similar to that for the smoking intervention (section 7.1) and again the effect of the intervention modelled is relatively small. The largest difference occurs for both men (1338 deaths prevented) and women (1680 deaths prevented) in 1996. This

peak occurs in 1996 because of the interaction of two factors, the period of five years over which the intervention is spread, and the lag times of one and two years (average 1.5) operating in the model. All the latencies in this simulation are zero.

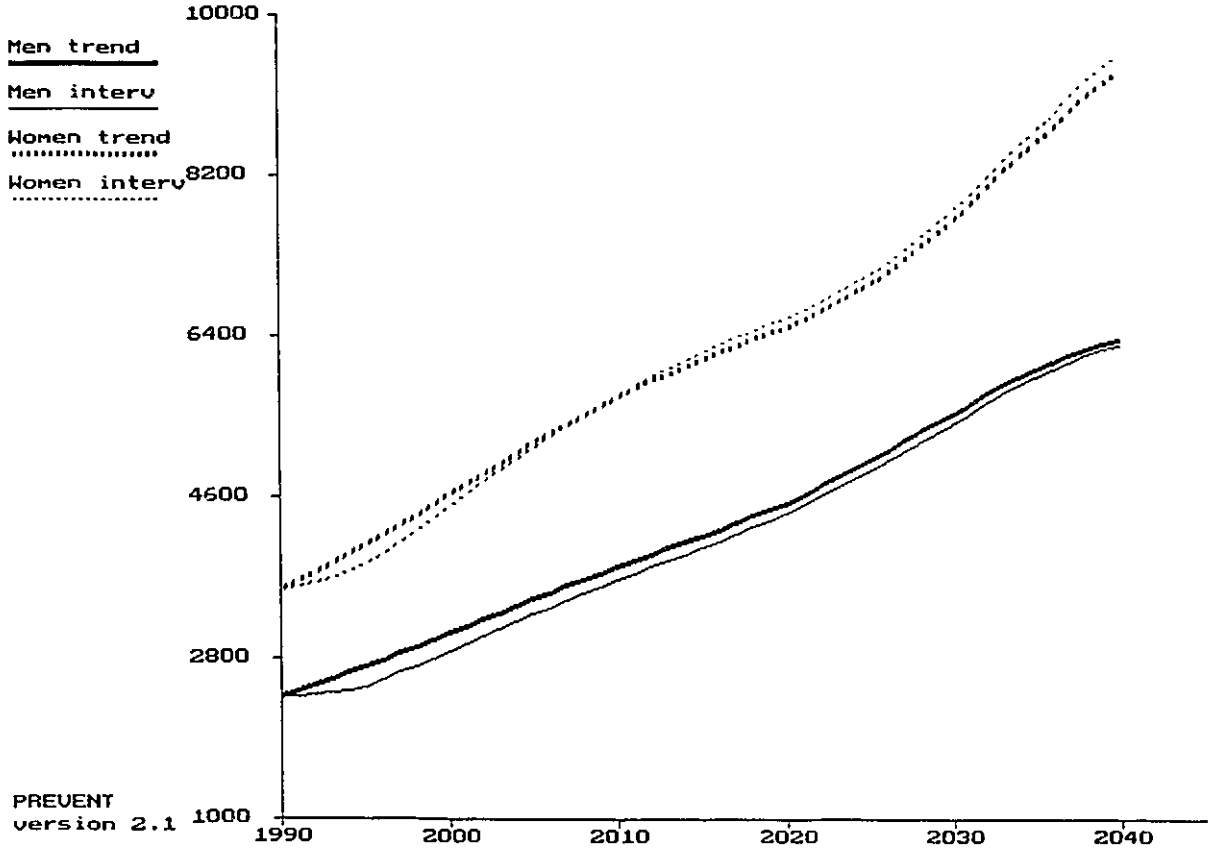
The impact of the intervention is greater for women. This is due to the fact that there is a higher prevalence of hypercholesterolemia (especially very high cholesterol levels) among older women than older men. Consequently more women than men are benefiting from the intervention modelled.

Figure 21 IHD Mortality for Hypertension and Hypercholesterolemia Decrease Simulation



Disease-Specific Mortality Disease-specific mortality for IHD and CVA are

Figure 22 CVA Mortality for Hypertension and Hypercholesterolemia Simulation



presented in Figures 21 and 22. The reduction in mortality for IHD in 1996 is 3279 in total (men and women), whereas the reduction for CVA is 443. This large difference is primarily due to the very large mortality rates for IHD. Some of the difference is also due to the fact that the intervention on both risk factors is affecting mortality due to IHD, but only hypertension is impacting on CVA. The relative risks for hypertension are slightly higher for CVA than for IHD.

In the trends alone the number of deaths for IHD in women exceeds that of men in the last few years of the simulation. This is due to the smoking trends which

model higher rates of smoking cessation for men than women and the fact that there will be progressively more older women in the population as "baby-boomers" age.

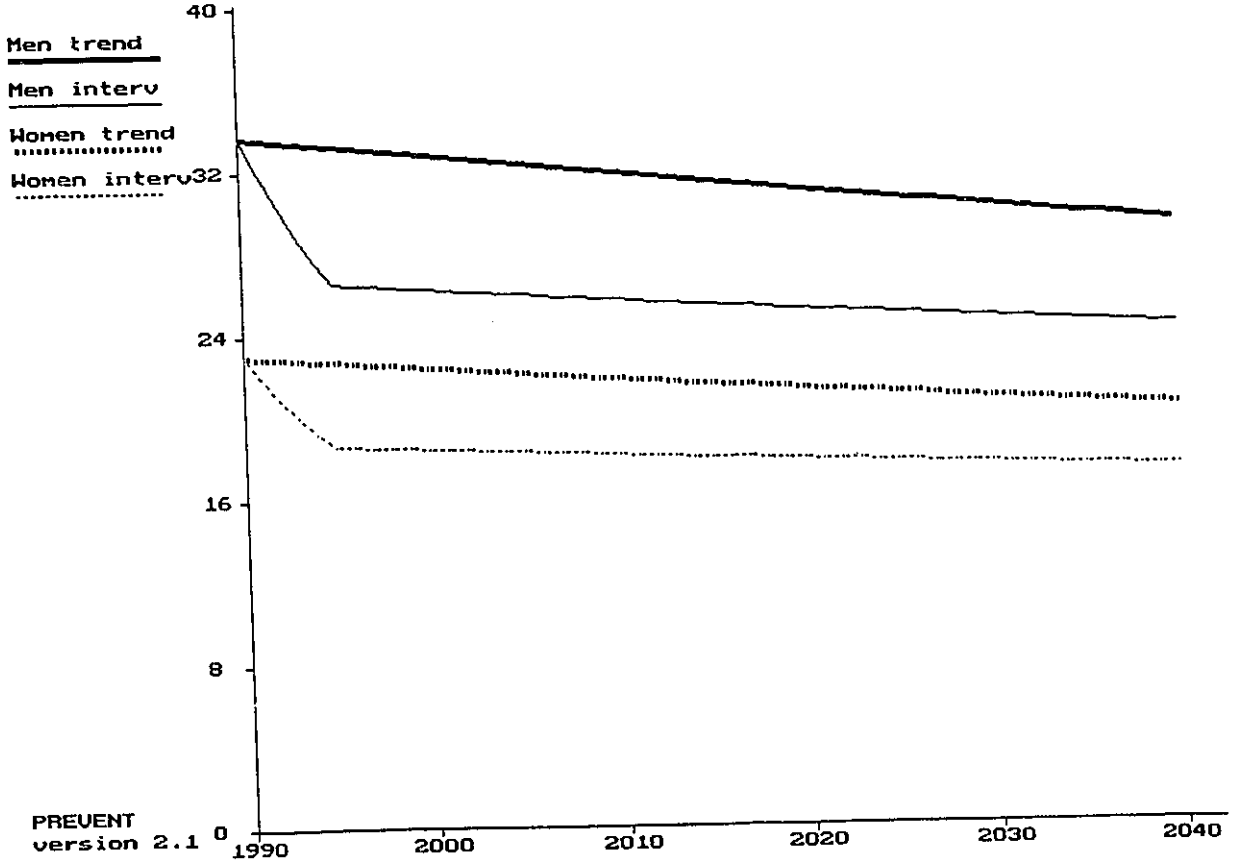
Etiologic Fraction The etiologic fraction of CVA due to hypertension for ages 60 to 64 is presented in Figure 23. The majority of the decrease due to intervention is manifested during the first six years of the simulation. The lines are showing a very slight decrease in etiologic fraction over most of the simulation for both trend and intervention. This is due to the small improvement modelled for all strata in the future prevalence of uncontrolled hypertension.

7.3 Seat Belt Use

This simulation models an increase in seat belt use to at least 95% in all strata in one year (those strata already over 95% maintain their level of use). It follows the age-group option and was chosen to demonstrate the relatively large impact on years of life gained when an intervention acts to prevent even a small number of deaths in young adults. Age-specific mortality rates for motor vehicle crash peak for 15-24 year olds of both sexes and is considerably higher for males. These ages coincide with ages of low prevalence in the use of seat belts, particularly for males. Consequently the impact of the intervention is greatest for young men.

Total Mortality Reduction The total mortality reduction due to the intervention is presented in Figure 24. The full intervention takes place within the first year of the simulation (there is no latency period), as the lag time is one year. Consequently the impact is immediate with a total of 57 deaths prevented in the first year, 50 of them for males. This is the largest number of deaths prevented for any year in the

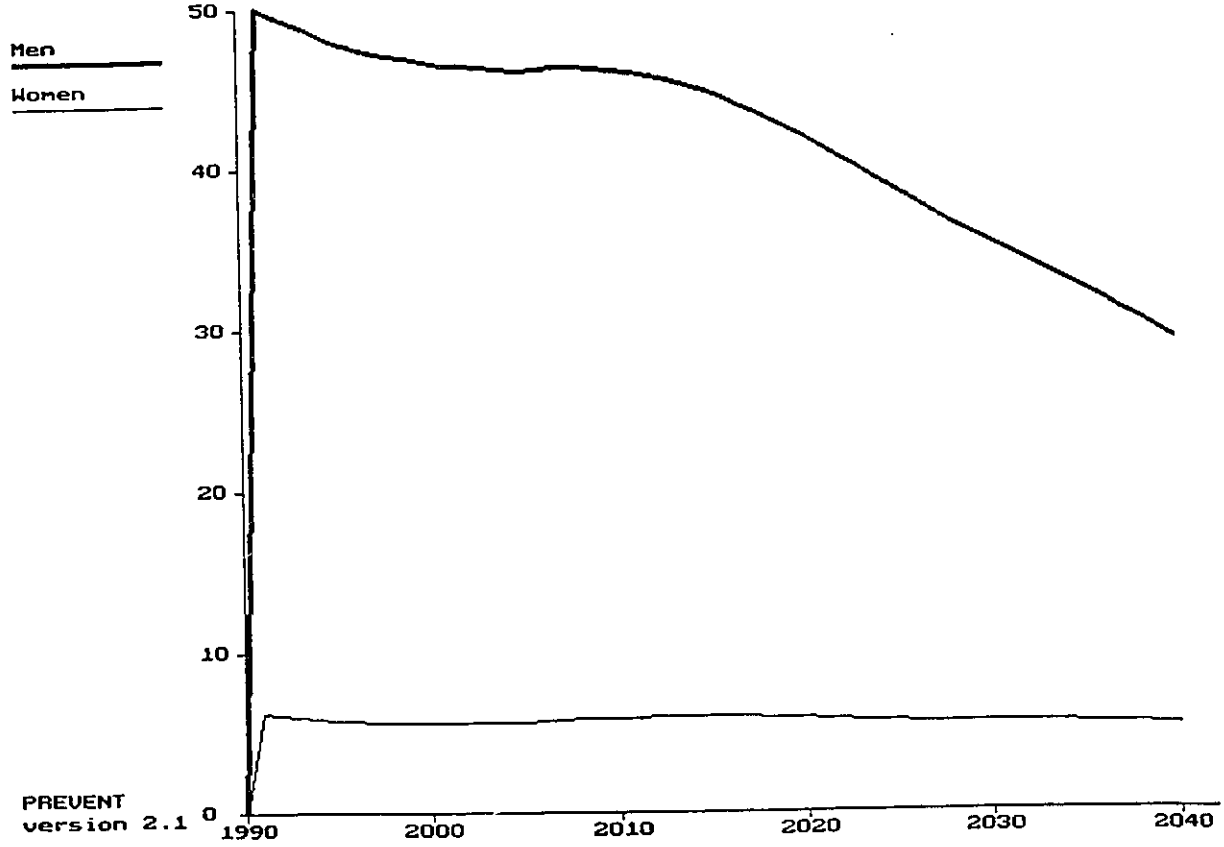
Figure 23 Etiologic Fraction for CVA Ages 60-64 for Hypertension and Hypercholesterolemia



simulation. The number of deaths prevented after the first year gradually drifts downward because PREVENT maintains the intervention only for those modelled at increased risk in the base year. The youngest age modelled at increased risk is 15 years for seat belt simulations. As the proportion of the population younger than 15 years in 1990 age, they do not acquire the benefit of the intervention and the maximum benefit in deaths prevented is not maintained.

Compared to peak mortality prevention for IHD (over 3000 deaths prevented) presented in Section 7.2, the mortality reduction in this intervention is quite small.

Figure 24 Total Mortality Reduction for 95% Seat Belt Use Simulation

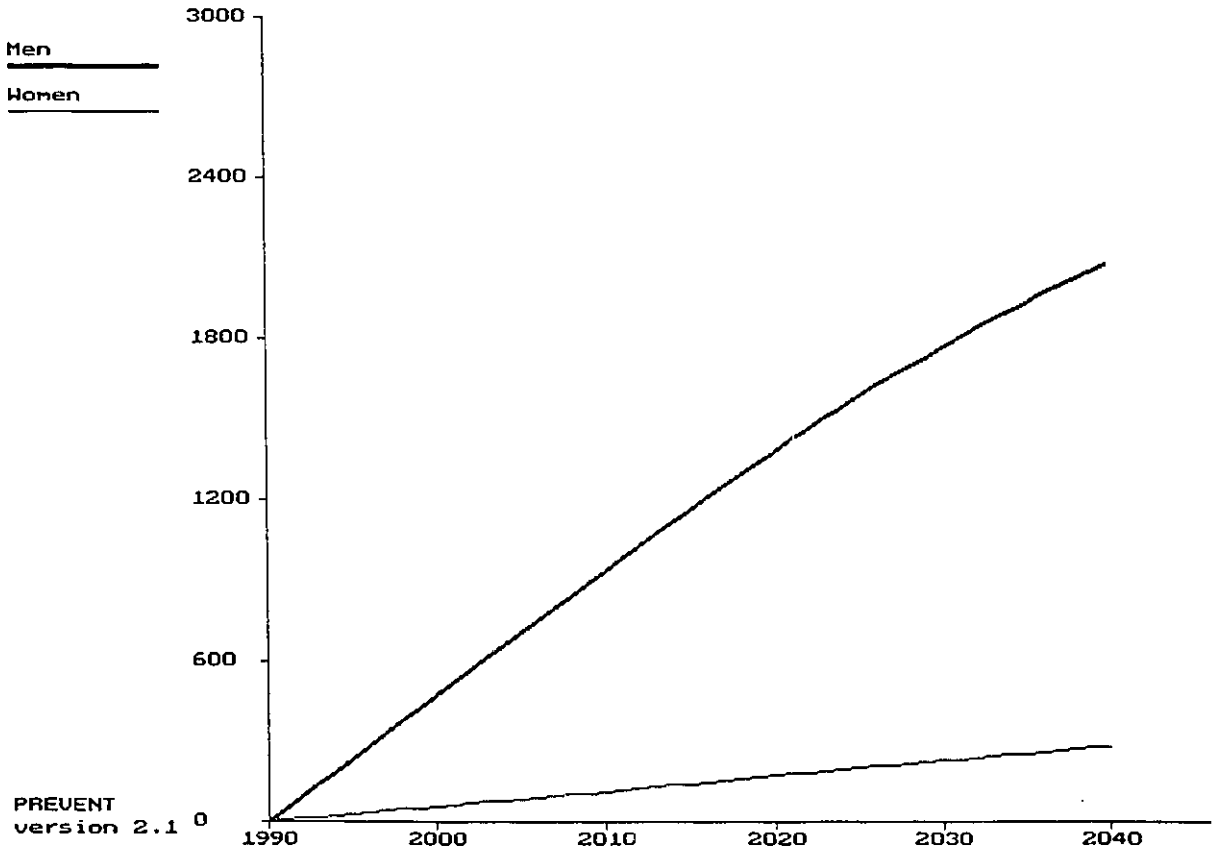


Actual Years of Life Gained In terms of actual years of life gained, as presented in Figure 25, the years of life gained exceeds 500 within ten years and is over 2000 by the end of the simulation period.

Disease-Specific Mortality Rates The disease-specific mortality rate for motor vehicle crash, as presented in Figure 26 decreases from 27.3 to 23.7/100,000 after intervention. This is a decline of 13%.

Disease-Specific Mortality Reduction The mortality reduction for motor vehicle crash is presented in Figure 27. This graph is similar to the all-cause mortality

Figure 25 Years of Life Gained for 95% Seat Belt Use Simulation.

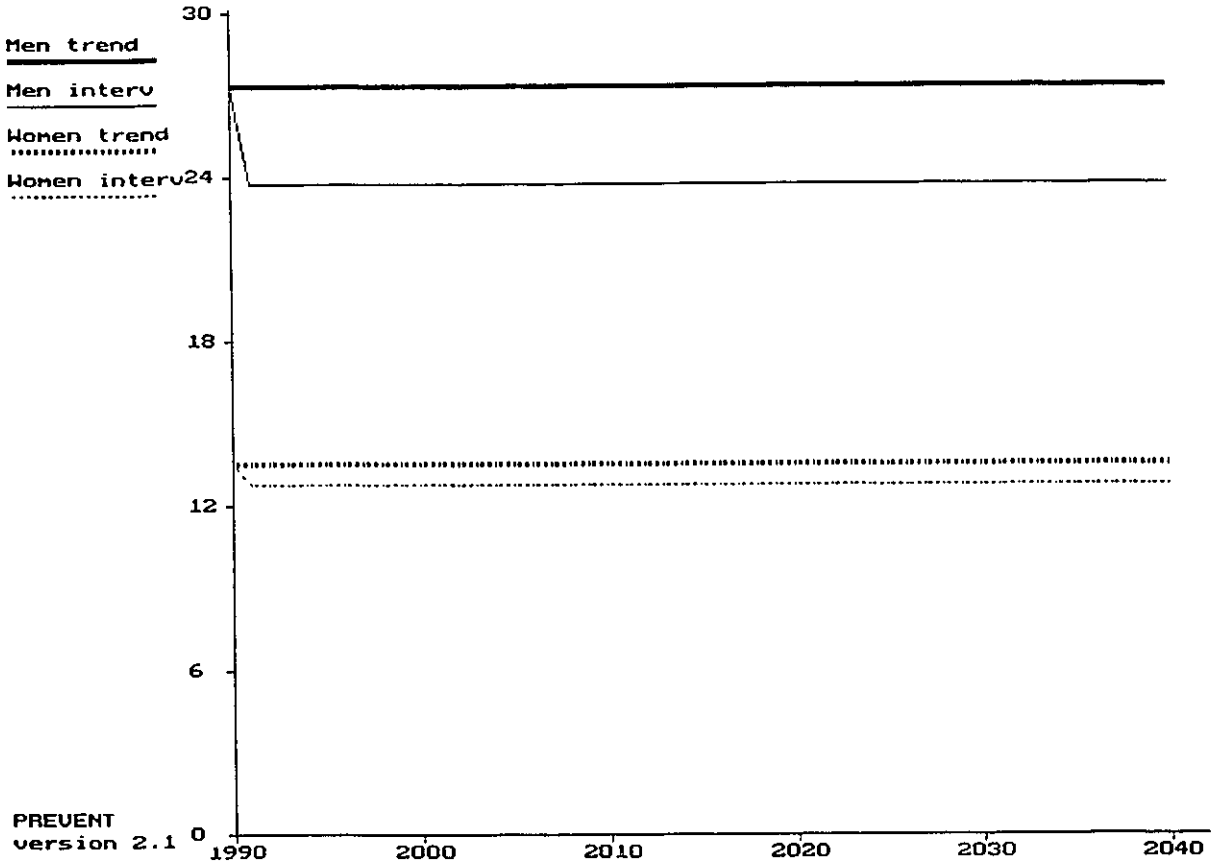


reduction. The peak effect occurs in the first year and the impact is much greater for males, but unlike all causes the effect on motor vehicles alone is maintained for the entire simulation period (not subject to mortality due to other causes).

7.4 Decreased Alcohol Use

This simulation models a 50% decrease in alcohol consumption over a 10 year period. For alcohol interventions only the age-group option is available. This simulation models the effect of one risk factor on many causes of death (IHD,

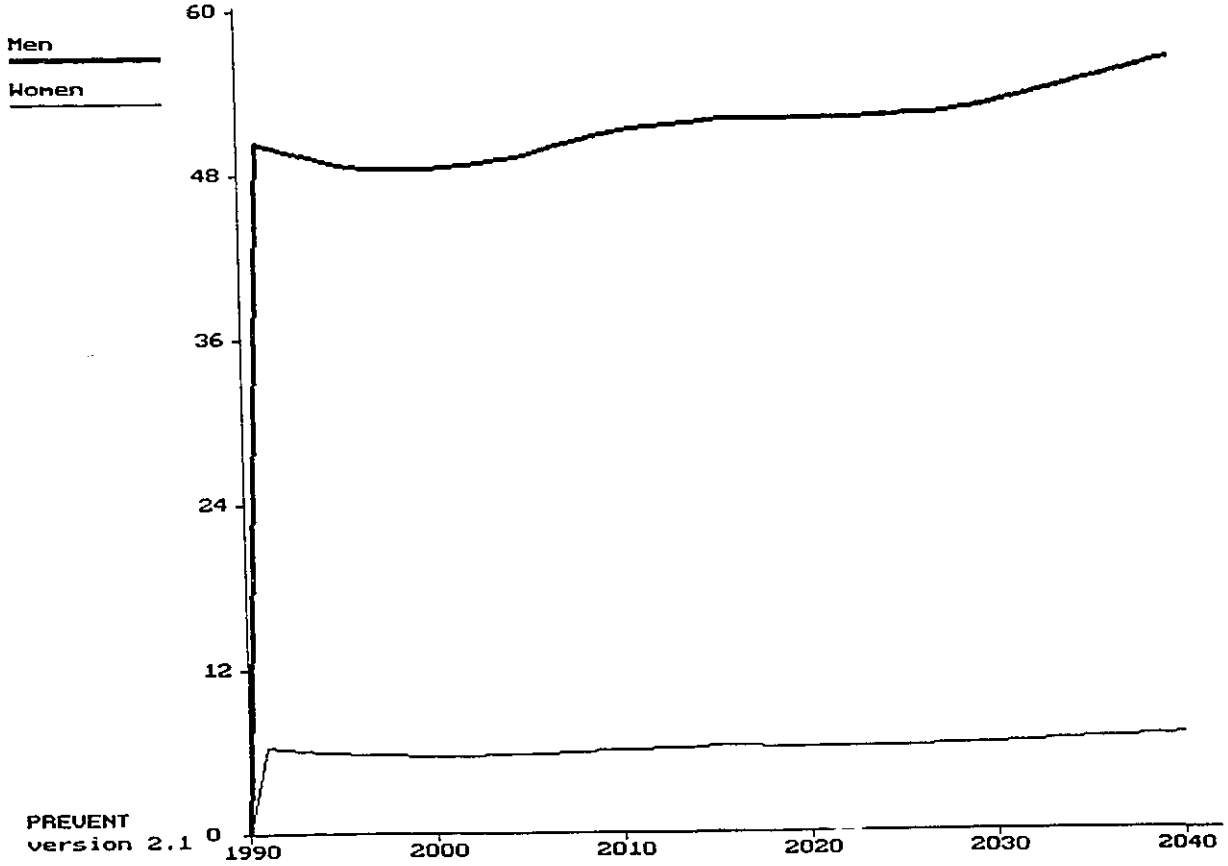
Figure 26 MVC Mortality Rates Ages 20-24 for 95% Seat Belt Use Simulation



cirrhosis, accidental fall and motor vehicle crash).

Cause-Specific Mortality Reduction The mortality reduction for the four causes of death in the model are presented in Figures 28, 29, 30 and 31. For IHD, accidental fall and motor vehicle crash, the largest initial impact is seen in the year 2000, at the end of the ten year implementation period for the intervention. For cirrhosis there is a four year period between 1990 and 1994 where no effect is seen. This corresponds to the modelled latency time of four years. At the end of the latency period the intervention is then implemented over a ten year period. For

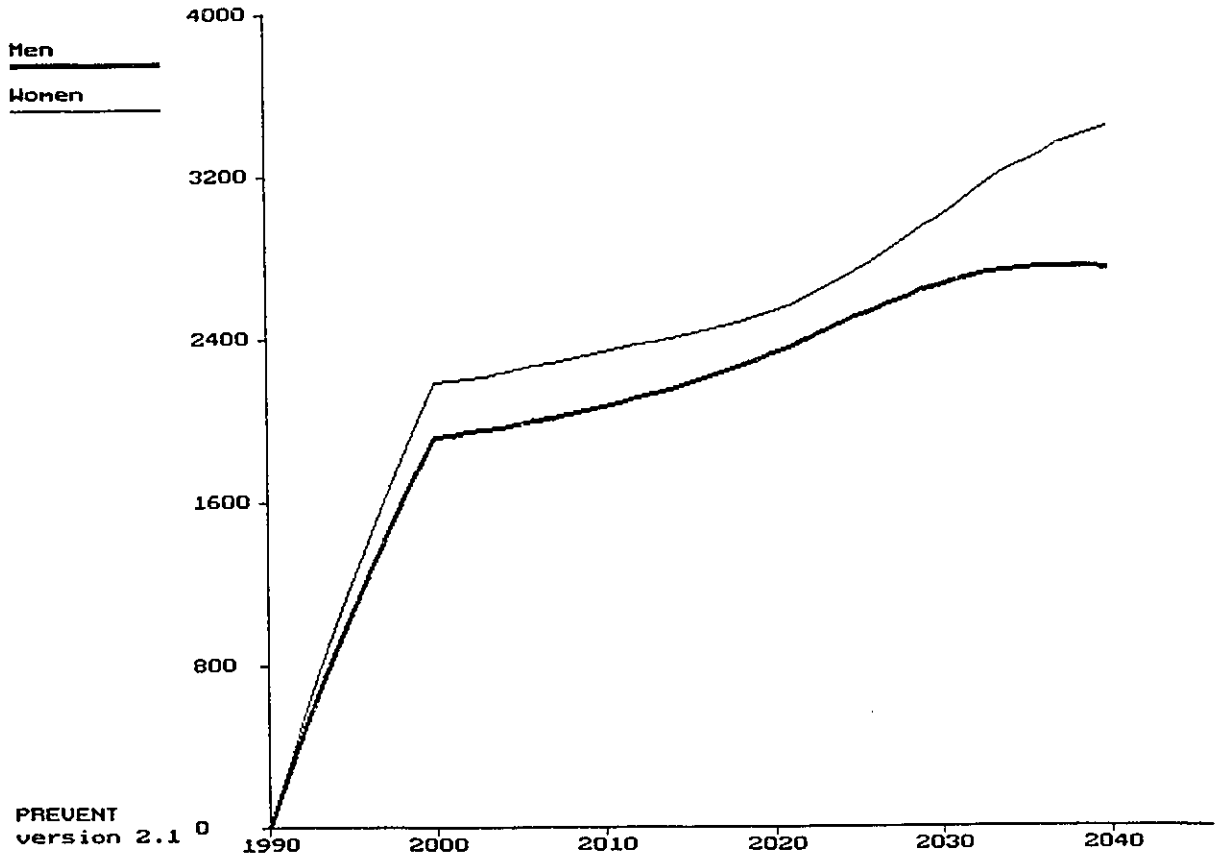
Figure 27 MVC Mortality Reduction for 95% Seat Belt Simulation



cirrhosis and motor vehicle crash the effect of the intervention is much more pronounced for men than for women. This reflects the much higher prevalence of excessive alcohol use among men compared to women.

In contrast, the impact of the intervention for IHD is greater for women than for men. In the Ontario database for PREVENT, IHD risk in relation to alcohol consumption follows a "J" shaped curve with non-drinkers at a higher risk than moderate drinkers for development of this disease. In order to accommodate alcohol as a risk for IHD in the model, moderate drinkers are chosen as the reference

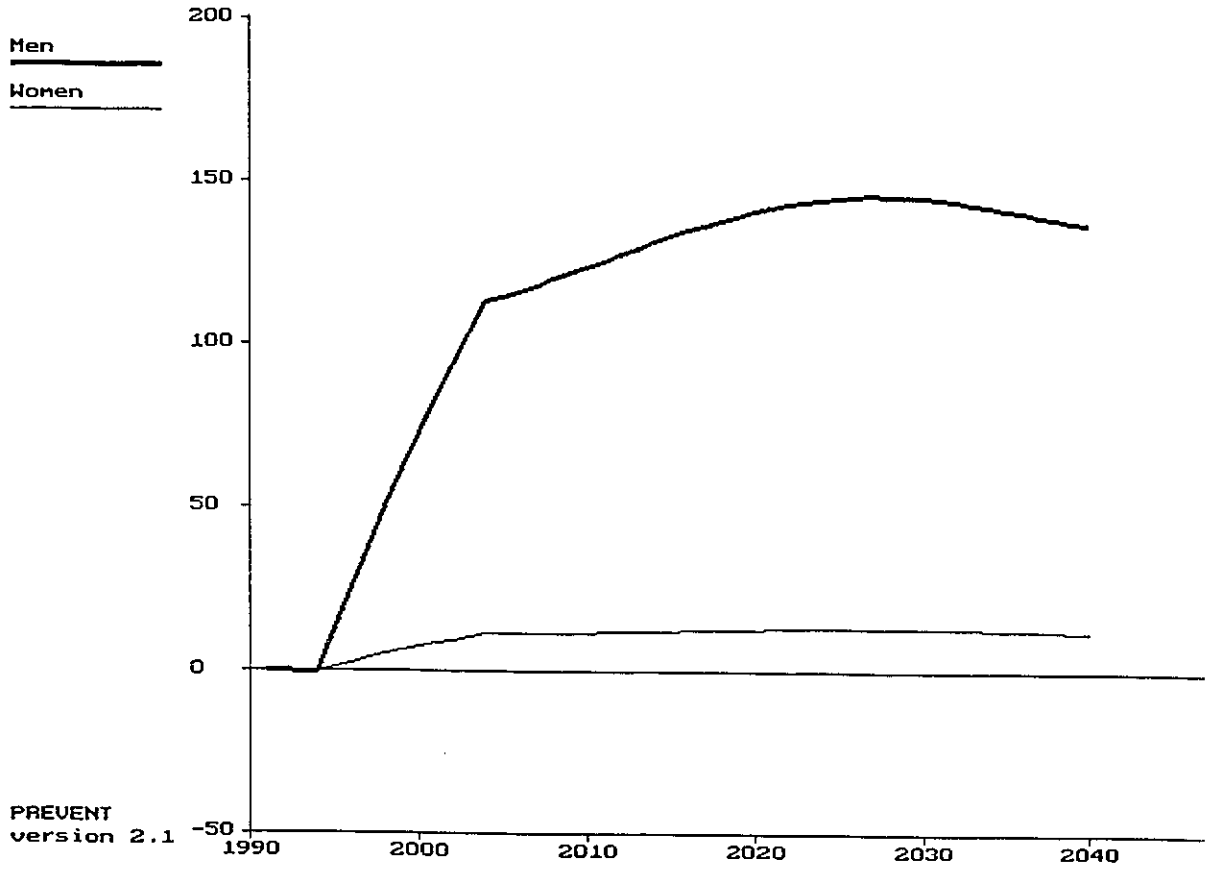
Figure 28 IHD Mortality Reduction for Reduced Alcohol Use Simulation



category. In this case when an intervention (decreased alcohol use) is modelled the program is actually modelling a change from both higher risk groups (non-drinkers and heavy drinkers) to the reference category (moderate drinkers). While the relative risk for alcohol and IHD is the same for men and women in the database, a much larger percentage of women are non-drinkers or very light drinkers. It is likely that their contribution increases the impact for women on reduced IHD mortality.

This unusual effect, non-drinking women being modelled as changing to moderate drinkers, does not influence the three other causes of death in the alcohol

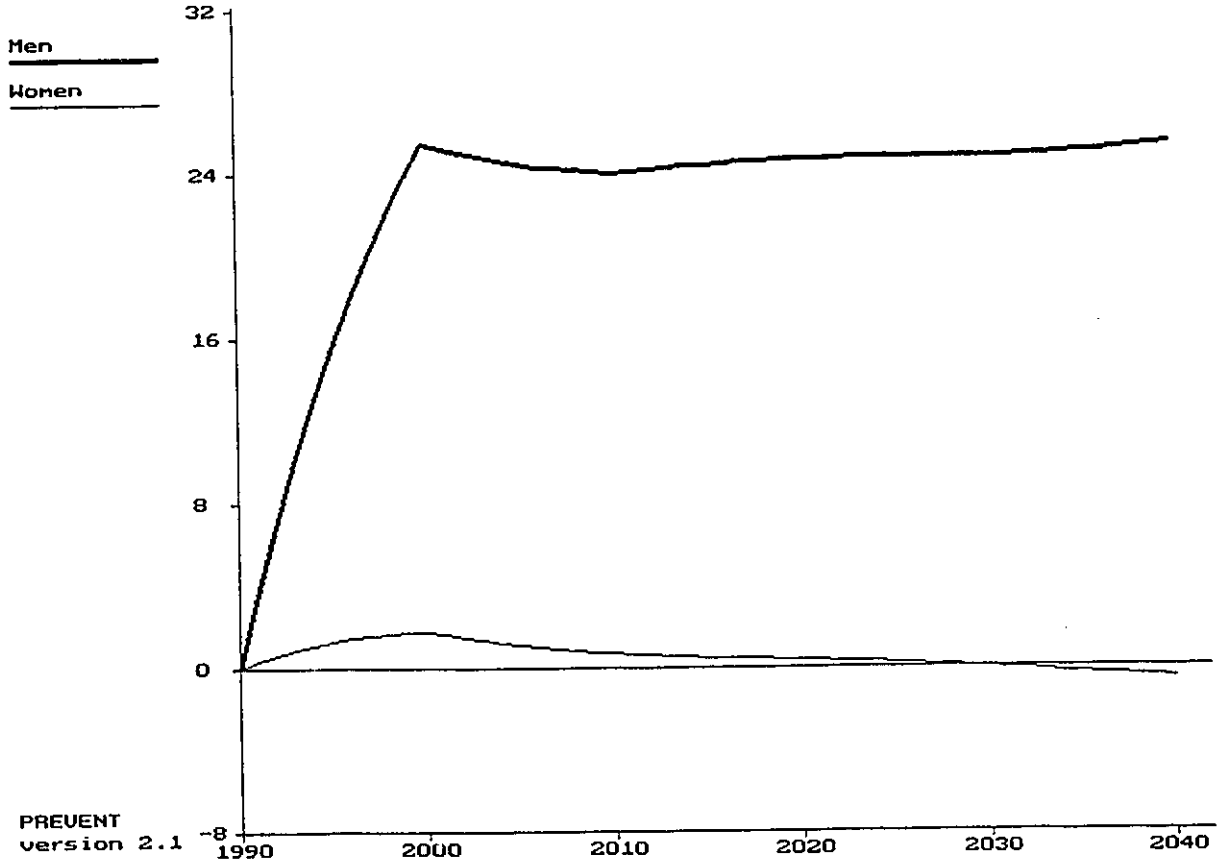
Figure 29 Cirrhosis Mortality Reduction for Reduced Alcohol Use Simulation



simulation because only heavy drinkers, not non-drinkers, are modelled as being at risk (assigned a relative risk greater than 1.0).

Although there is an initial reduction in the number of deaths due to accidental fall, this gain is eliminated and the number of deaths actually increases later in the simulation. Accidental fall has very high mortality rates in elderly age groups. The increase in mortality modelled is caused by later deaths due to falls in those who were saved from earlier death due to cirrhosis, motor vehicle crash and most importantly, IHD.

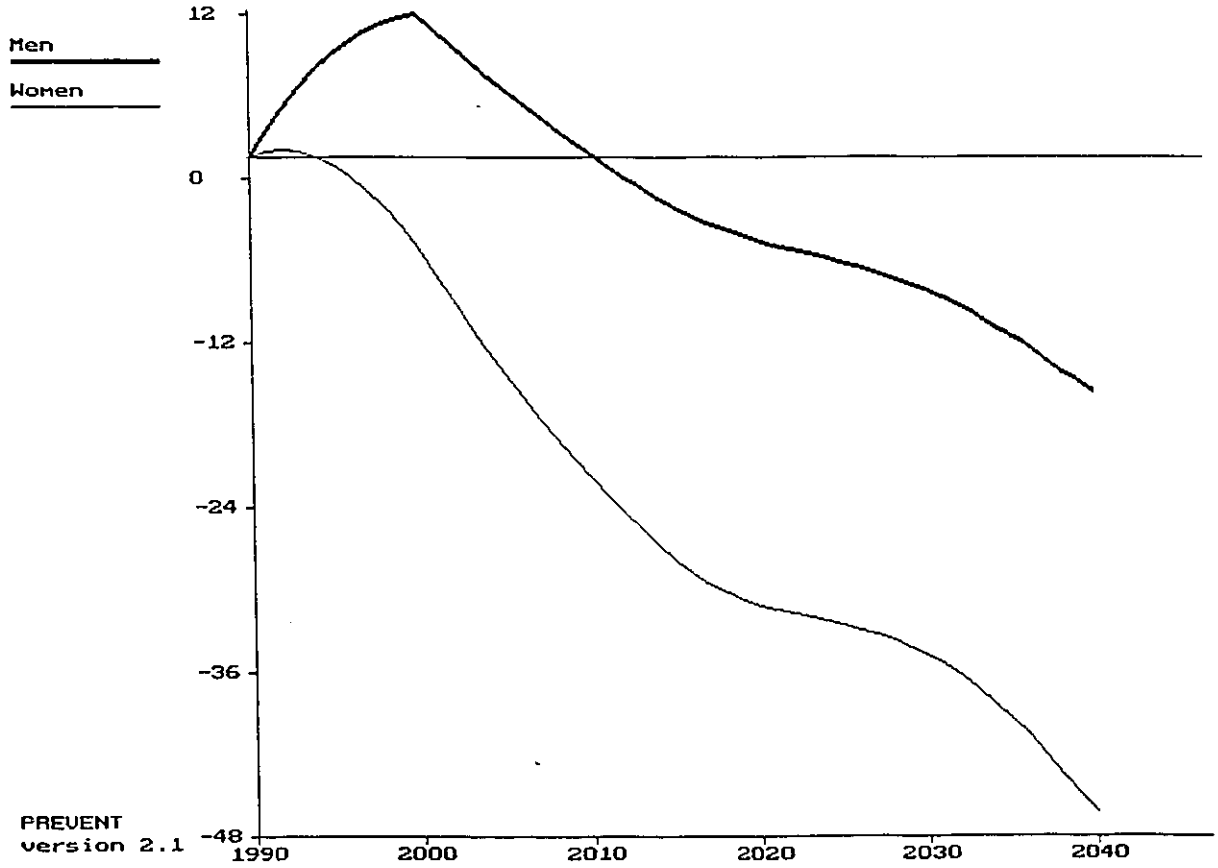
Figure 30 MVC Mortality Reduction for Reduced Alcohol Use Simulation



Survival Curve The survival curve for the years 2040 is presented in Figure 32 and shows increased survival after intervention on those over 50 years of age.

Life Expectancy at Birth This increased survival is accompanied by a slight increase in life expectancy as presented in Figure 33. The life expectancy at birth in 2040 for males is 75.46 for trend alone and 75.97 after intervention. For females life expectancy is 80.72 for trend alone and 81.16 with intervention.

Figure 31 Accidental Fall Mortality Reduction for Reduced Alcohol Use Simulation

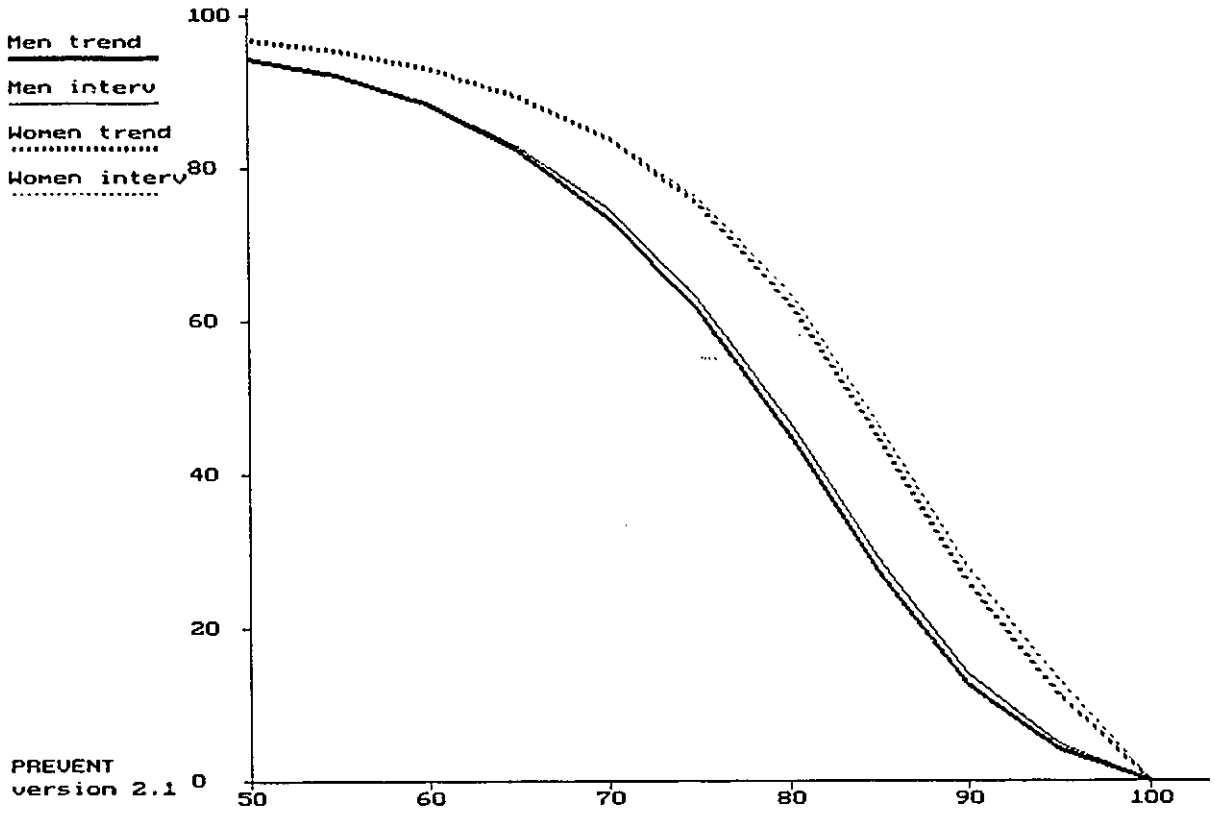


7.5 Decreased Obesity

This simulation models a 50% decrease over ten years in the proportion of the population with body mass index >30.. Only the age-group option is available for this risk factor. This simulation was chosen to demonstrate the relatively small impact due to the small relative risks modelled for obesity and breast cancer. In the Ontario database for PREVENT, women with BMI>30 and who are fifty years or older have a relative risk of 1.3 of developing breast cancer.

Disease-Specific Mortality The mortality for breast cancer is presented in Figure 34.

Figure 32 Survival Curve for Reduced Alcohol Simulation



The largest difference between trend and intervention is 96 deaths in 2040. Breast cancer is one of the leading causes of death in women, so the intervention effect of less than 100 deaths in 3200 is relatively small.

Etiologic Fraction The etiologic fraction (for trend alone) of breast cancer deaths due to obesity is five percent, as presented in Figure 35. The effect of the intervention starts in 1992 (after the one year latency period modelled) and continues over the ten year implementation period for the intervention. A reduction of 50% in obesity results in a 50% reduction in etiologic fraction.

Figure 33 Life Expectancy With and Without Reduced Alcohol Simulation

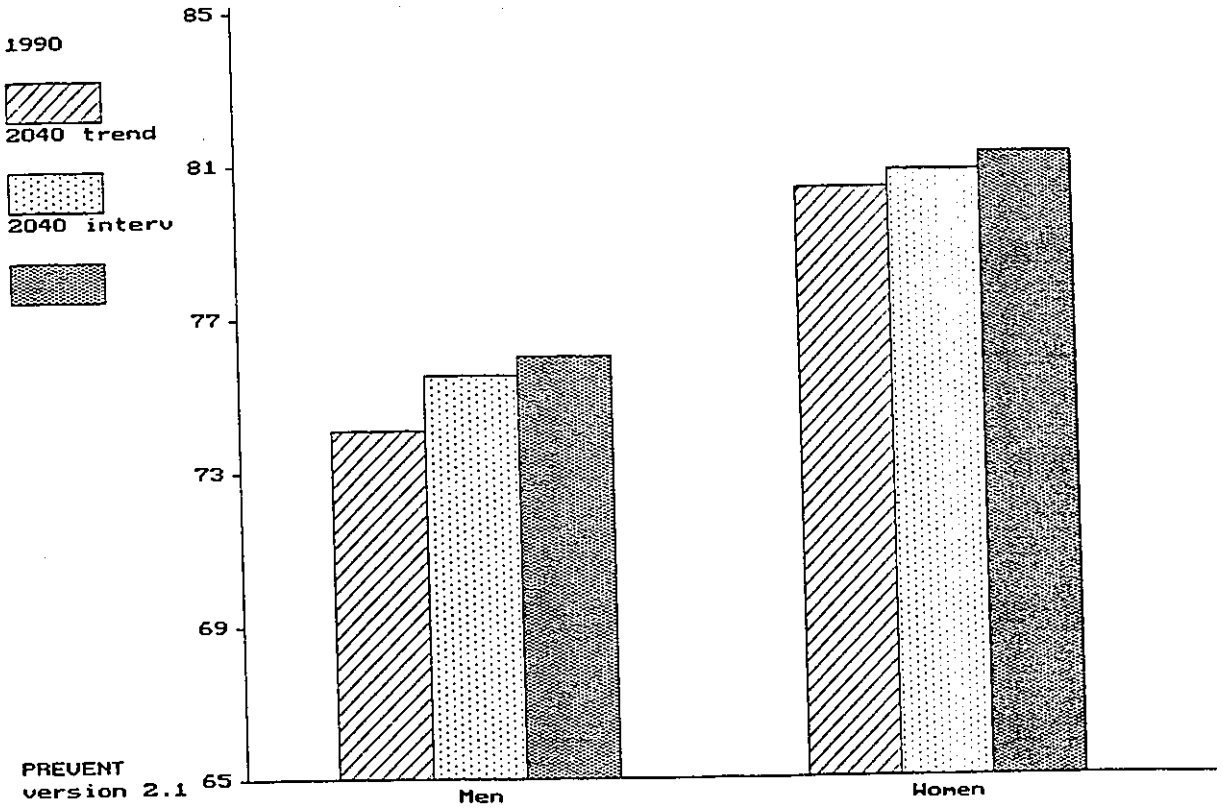


Figure 34 Breast Cancer Mortality for Reduced Obesity Simulation

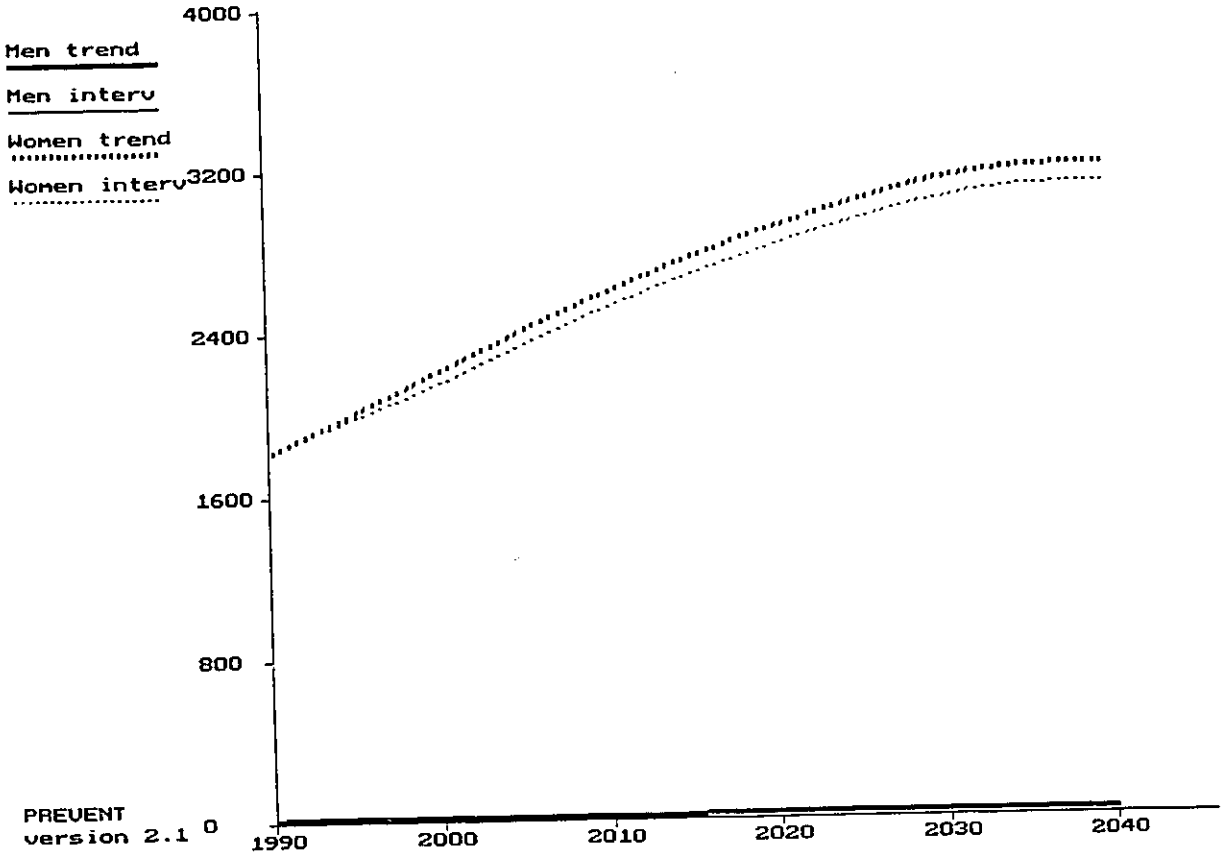
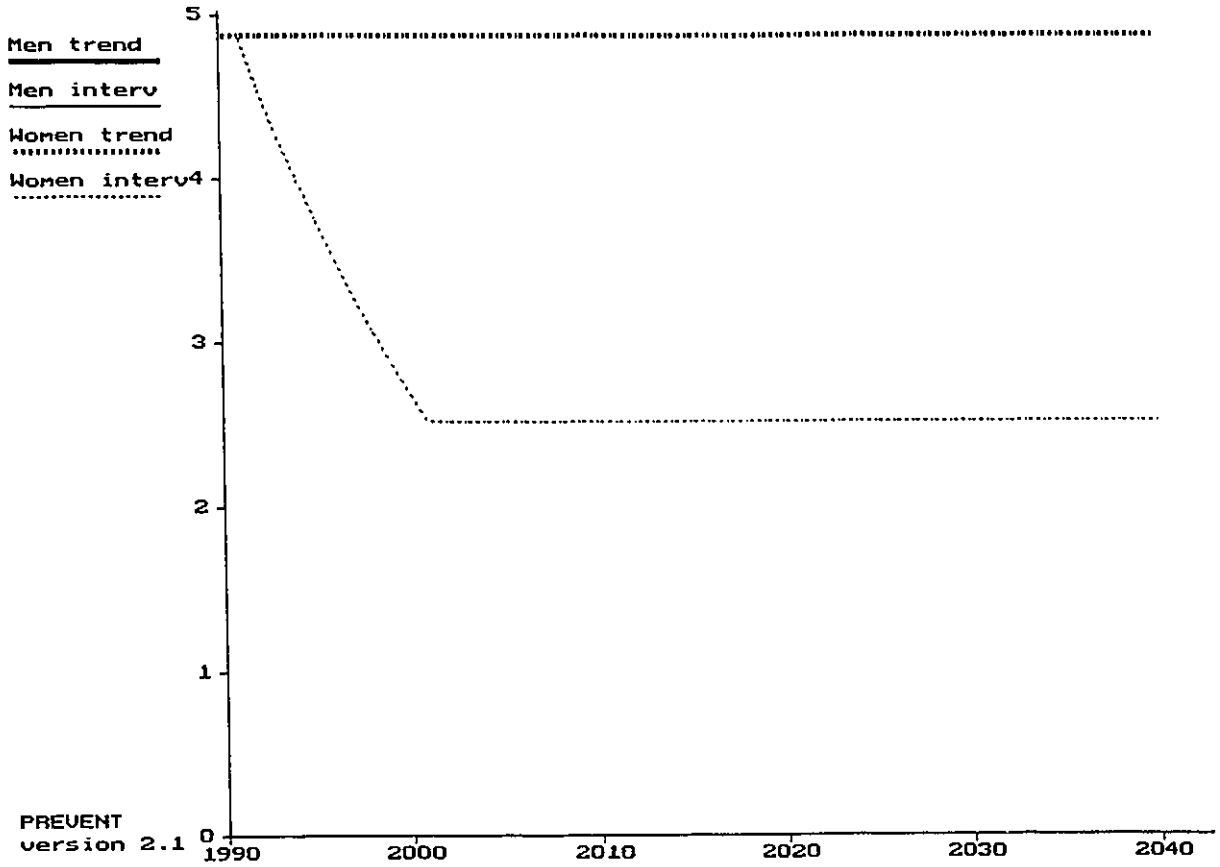


Figure 35 Etiologic Fraction for Breast Cancer Ages 60-64 for Reduced Obesity Simulation



8.0 TESTING THE DATABASE

8.1 Sensitivity Testing

In the preparation of this database many choices and judgements were made in the selection and estimation of required data. It is important to determine how much and in what ways output is affected by different choices in input data. The variables in the model can be divided into five categories.

The first category is population structure and mortality data. These are not estimates, but recorded data and, as such, have much smaller margins of error than other variables in the database. They are not amenable to sensitivity testing. The second type of information is data on relative risks for risk factors and associated diseases. The third type of variables are the time dimensions of latency and lag time. These values were selected from study data when available. The fourth category of data is prevalence information. These data were taken from Ontario-based surveys if they were available. In some cases similar national survey data were also available. The final data category is trend for future prevalence. These data were based on projection of past trends and as such probably have a wider range of error than any other type of data in the model.

For the sensitivity testing of the Ontario database, I have selected one intervention which is run under different conditions in order to assess the effect of different input data. The test intervention is an immediate 50% reduction in smoking, using the age-group option, simulated over 30 years. In the comparison of effect the all-cause mortality and mortality due to lung cancer, COLD and IHD are used.

8.1.1 Results of Test Simulation with Modified Relative Risks

Data Sources When relative risks for smoking were chosen for the Dutch data set and reviewed for the Ontario database, the choice was made from a summary of research studies. The relative risks form these studies were presented in Appendix B.

Data Used From these studies the highest and lowest relative risks were selected and used in PREVENT runs. The lowest relative risks were taken from the Japanese study by Hirayama for lung cancer, IHD and COLD. The highest relative risks were chosen from the British Physicians Study. The British Physician's Study had the largest relative risks for younger ages, but not for older age groups. For the purposes of this sensitivity testing, the high relative risks (for the young age group) in the British study were applied to all ages. The high, low and standard relative risks are presented in Table 21. The output that these relative risks produce is presented in Figures 36 and 37, which show the difference attributable to the intervention as expressed by reduction in all-cause mortality and reduction in IHD mortality.

Discussion The level of risk is seen to influence trend alone as well as intervention; however, the differences in total mortality for trend are small. All three simulations have the same mortality in the base year (as defined by the input population and mortality rates for 1990). At the end of the simulation period the difference among the three simulations is very small (varies approximately 1%): 119582, 119689 and 117904 for standard, low and high relative risks respectively.

The total mortality reduction (Figure 35) curves are very similar in shape but the magnitude of the benefit varies. Mortality reduction is higher for high relative risks and lower for the low relative risks. The mortality for standard relative risks is

Table 21 Relative Risks for Smoking used in Sensitivity Testing

Disease	Exposure Category cigarette /dy	Relative Risk				
		20-24 low std high	25-34 low sad high	35-49 low sad high	50-64 low sad high	65+ low sad high
Lung Cancer	1-12	3.5 7.0 7.8	3.5 7.0 7.8	3.5 7.0 7.8	3.5 7.0 7.8	3.5 7.0 7.8
	13-22	5.7 12 12.7	5.7 12 12.7	5.7 12 12.7	5.7 12 12.7	5.7 12 12.7
	23+	6.5 20 25.1	6.5 20 25.1	6.5 20 25.1	6.5 20 25.1	6.5 20 25.1
	residual	1.4 2.0 2.1	1.4 2.0 2.1	1.4 2.0 2.1	1.4 2.0 2.1	1.4 2.0 2.1
COLD	1-12	4.5 12 17	4.5 12 17	4.5 12 17	4.5 12 17	4.5 12 17
	13-22	11.2 25 26	11.2 25 26	11.2 25 26	11.2 25 26	11.2 25 26
	23+	17.5 30 38	17.5 30 38	17.5 30 38	17.5 30 38	17.5 30 38
	residual	1.6 2.6 2.6	1.6 2.6 2.6	1.6 2.6 2.6	1.6 2.6 2.6	1.6 2.6 2.6
IHD	1-12	1.1 3.0 3.7	1.1 3.0 3.7	1.1 3.0 3.7	1.1 1.7 3.7	1.1 1.5 3.7
	13-22	1.1 3.0 4.5	1.1 3.0 4.5	1.1 3.0 4.5	1.1 2.0 4.5	1.1 1.5 4.5
	23+	1.1 4.0 4.5	1.1 4.0 4.5	1.1 4.0 4.5	1.1 2.4 4.5	1.1 1.7 4.5
	residual	1.0 1.2 2.4	1.0 1.2 2.4	1.0 1.2 2.4	1.0 1.2 2.4	1.0 1.2 2.4

* Males and Females have the same relative risks.

between them. The peak mortality reduction for high relative risks is approximately double that for low relative risks. The shape of the curves is not greatly different but mortality in general is quite sensitive to changes in relative risk.

In the case of lung cancer mortality (Figure 36) the curves are again as expected, with higher relative risks accounting for more deaths prevented when 50% of smokers are modelled to quit. Lower relative risks have less impact on lung cancer mortality.

Figure 36 Total Mortality Reduction with Variations in Relative Risk

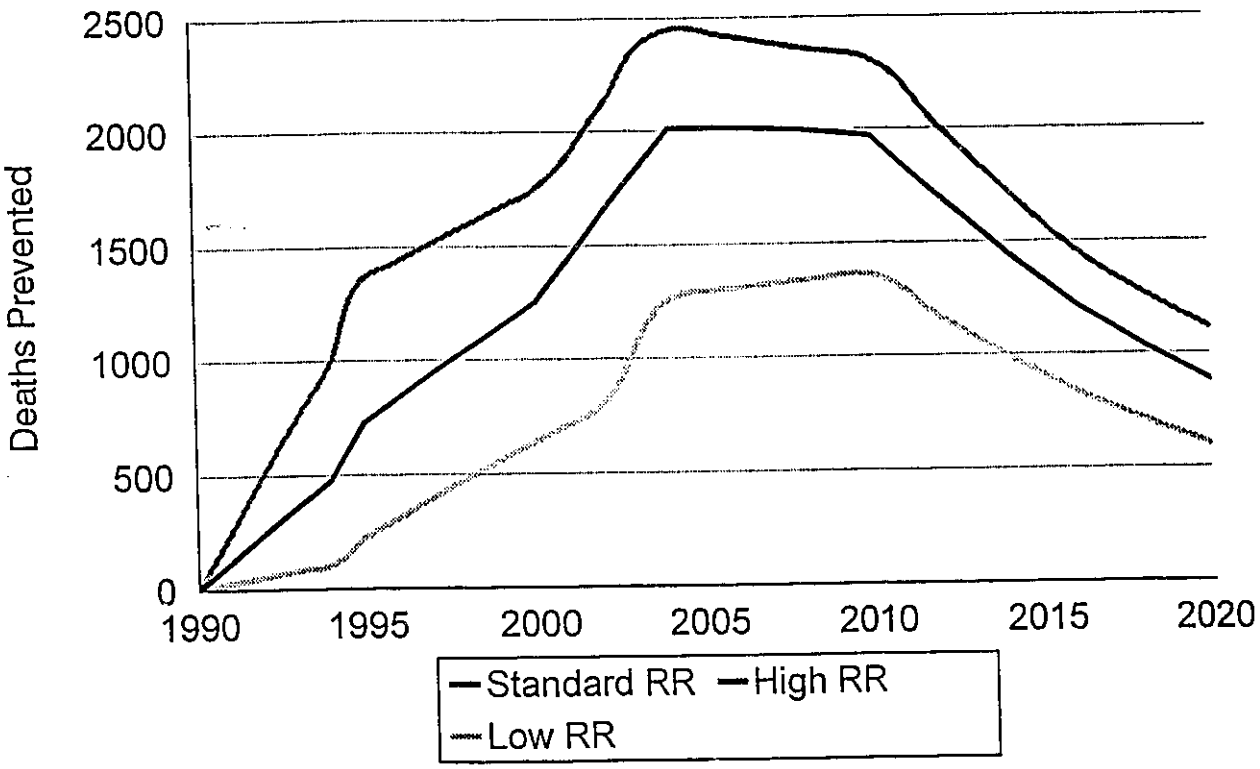
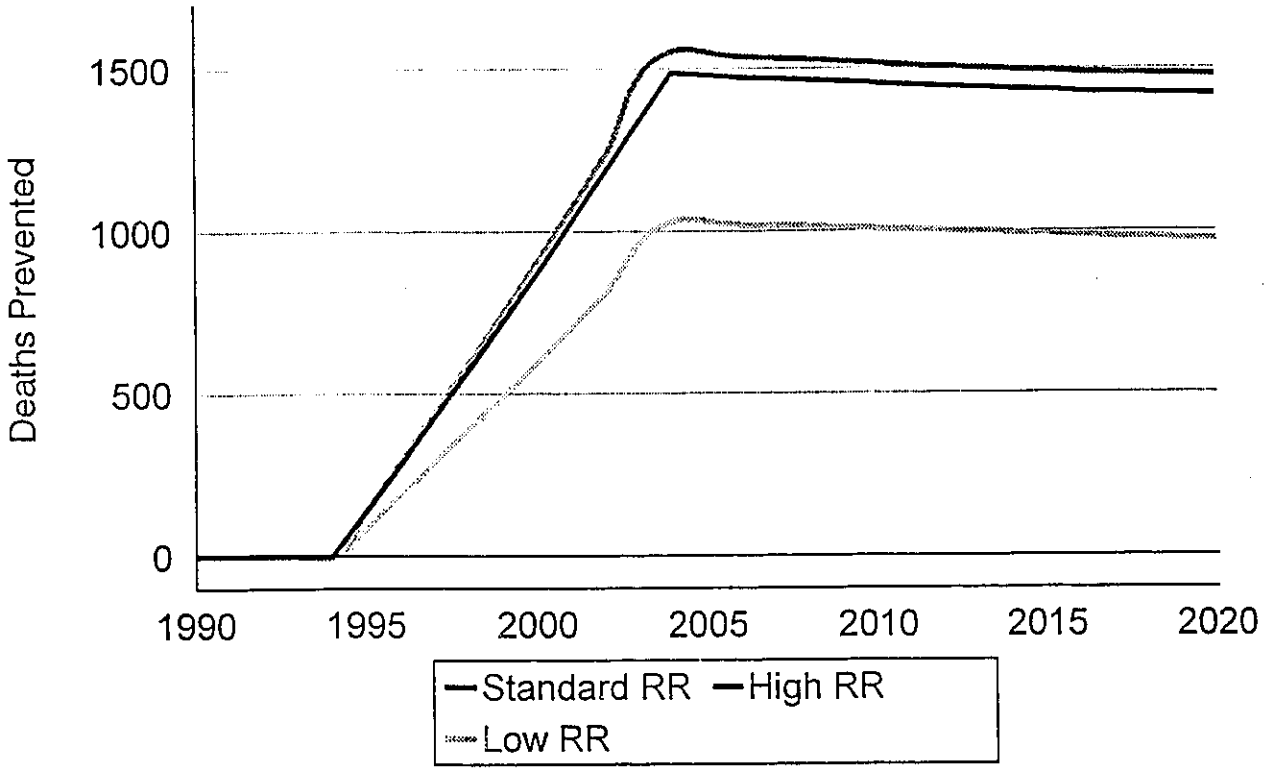


Figure 37 Lung Cancer Mortality Reduction with Variations in Relative Risk



8.1.2 Results of Test Simulation with Modified Prevalence

Data Sources Since the Ontario Health Survey was specific to the population under study, smoking prevalence from that survey was used in the Ontario database. Smoking prevalence has also been estimated nationally in the pooled data from the

Table 22 Alternate Smoking Prevalence Data for Sensitivity Testing

Exposure Category	Age group				
	20-24	25-34	35-49	50-64	65+
Smokers					
1-12 male	6.9(9.0)*	5.6(6.8)	5.8(5.2)	3.5(4.4)	3.2(4.7)
female	9.3(12.4	7.8(8.6)	5.9(6.4)	4.5(6.2)	4.6(5.6)
13-22 male	25.1(15.9)	25.2(15.4)	19.5(10.3)	11.9(9.8)	11.2(5.3)
female	21.4(13.7)	23.8(13.9)	23.4(10.9)	12.4(8.9)	6.4(5.0)
23+ male	1.3(9.4)	4.2(15.8)	6.4(13.3)	6.6(13.3)	1.6(5.1)
female	1.3(7.8)	2.4(11.1)	4.0(8.3)	0.9(8.3)	1.12(2.8)
former Smokers					
1-12 male	4.4(4.5)	4.3(7.6)	6.7(8.0)	9.6(10.0)	1.3(16.1)
female	7.0(8.1)	6.9(10.8)	5.4(10.9)	8.8(11.5)	13.7(12.2)
13-22 male	16.0(3.4)	19.4(6.4)	22.6(12.9)	32.4(17.4)	45.5(21.3)
female	16.1(3.7)	21.0(6.5)	21.3(6.4)	24.2(7.6)	19.1(5.9)
23+ male	0.8(0.0)	3.2(4.9)	7.4(10.7)	18.0(16.7)	6.5(17.1)
female	1.0(0.1)	2.1(4.2)	3.6(6.4)	1.8(5.6)	3.2(4.0)

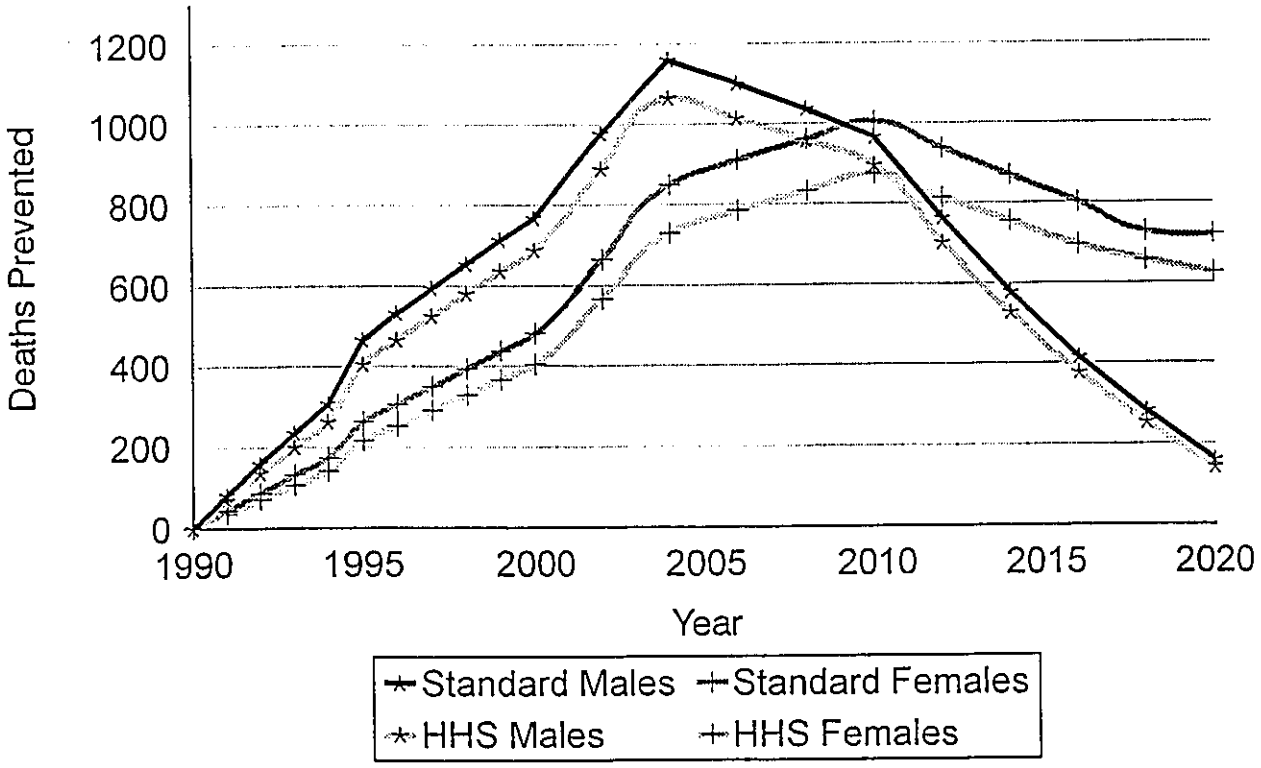
Figures in brackets are prevalence values for standard set from Ontario database.

provincial heart health surveys[75]. These estimates were used as an alternative to standard prevalence in order to show the effect of changing prevalence data.

Data Used The prevalence data adopted from the Canadian Heart Health Surveys are presented in Table 22 with the Ontario Health Survey data (standard prevalence) along side for easy comparison.

Results The outputs for the two sets of prevalence data are presented in terms

Figure 38 Total Mortality Reduction with Variation in Smoking Prevalence



of total mortality reduction for trend alone, Figure 38. In terms of all-cause mortality in trends alone the two test simulations are so similar that the curves overlap on the graph. At the end of the simulation period total mortality with the standard set is 61923 for males and 57658 for females. With the Heart Health Survey set the total mortality is 62130 for males and 57665 for females. Total mortality reduction is presented in Figure 38, and IHD mortality reduction in Figure 39.

Discussion It is difficult to assess the overall difference in the two prevalence sets. In terms of overall percentage of smokers the two surveys are quite similar. The

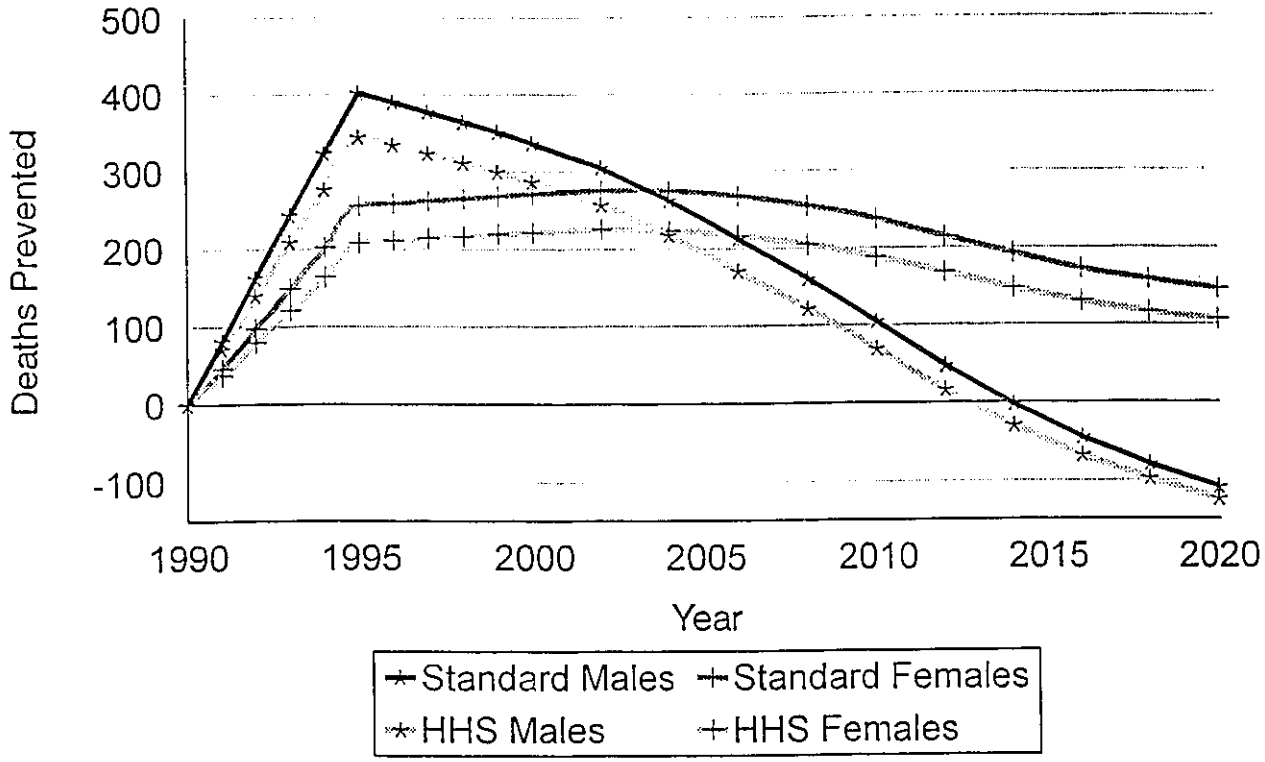
standard prevalence set has lower overall prevalence for the exposure level 13-22 cigarettes per day and higher prevalence levels for the other two exposure categories. The total mortality simulated by the test intervention is similar for both prevalence sets, both in the number of deaths and in the shape of the mortality curves. Similarly, the results for IHD mortality reduction are close in magnitude and almost identical in shape. The mortality estimated for the Heart Health Survey set is consistently a little lower. These results imply that overall the Heart Health survey prevalences are slightly lower than the standard set.

In Figure 38, total mortality reduction, the peak effect for women occurs more than five years later than for men. The shape of the last half of these curves is influenced heavily by future trends in smoking. Many more men than women are modelled to quit smoking in future years. By the year 2005 there is a smaller "pool" of smoking males than females still available for the intervention to act upon. Consequently, the mortality reduction for men starts to decline. In the case of IHD mortality reduction, this effect is even more pronounced. Prevented male mortality peaks immediately after the five year lag period. For women, the full effect is maintained because few women are modelled to quit and older women are modelled for increased smoking trends.

8.1.3 Results of Test Simulation with Modified Latency and Lag

Time dimensions are an important part of PREVENT. The disease most affected by them in the Ontario database is COLD which uses the maximum values allowed (ten years each).

Figure 39 IHD Mortality Reduction with Variation in Smoking Prevalence

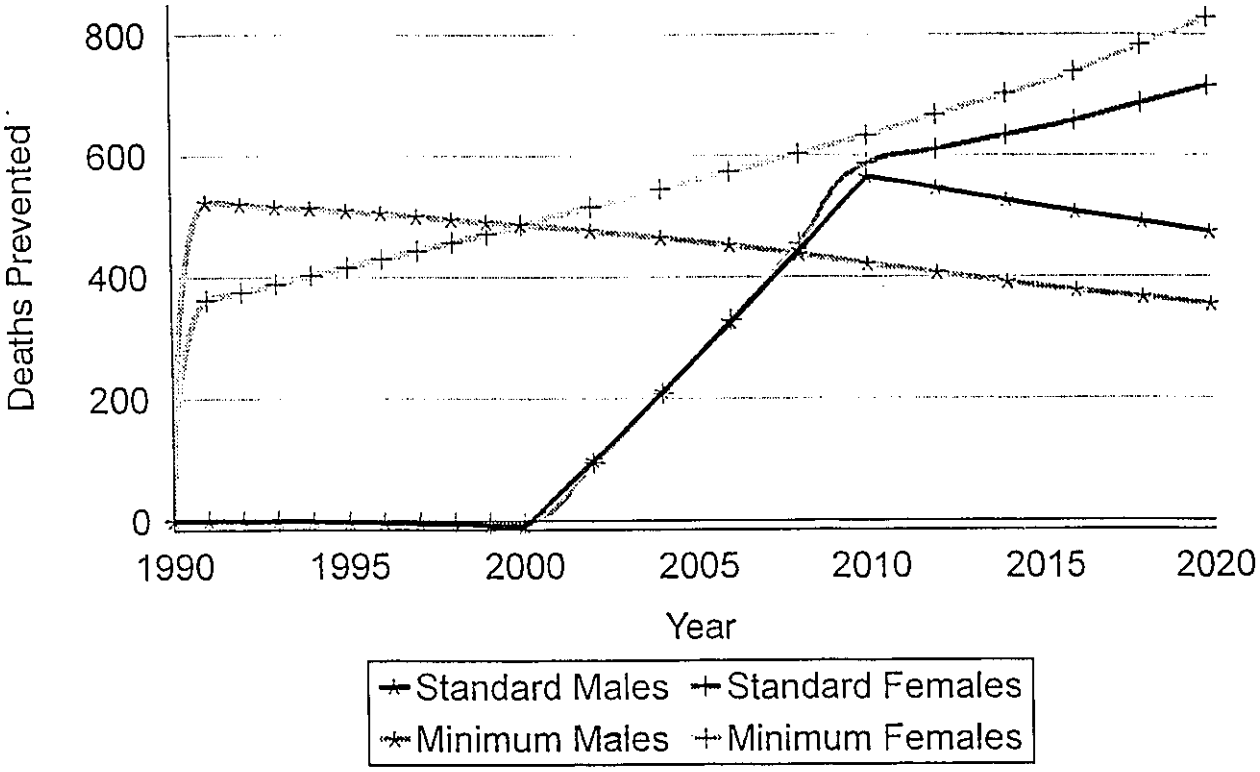


Data Used In order to show the effect of these two time dimensions the immediate 50% reduction in smoking intervention was run using the minimum values. The lag time was reduced to one year and the latency at zero (years).

Results The results are presented in Figure 40 which shows mortality reduction in COLD using both sets of time dimensions.

Discussion There is a striking difference in the shape of the curves under the two time dimension sets. In the standard set the COLD mortality is virtually unchanged for the first 10 years after the intervention (10 year latency period), then

Figure 40 COLD Mortality Reduction with Variations in Time Dimensions



over the next 10 years mortality is gradually reduced in a linear manner (10 year lag time). After that point the curve is dictated by population aging patterns and the future trends modelled for the risk factor. During the first 10 years the mortality from COLD increases from zero in the base year to eleven deaths a year in the year 2000. This slight increase is due to COLD deaths among those "saved" from IHD and lung cancer in the first few years of the simulation.

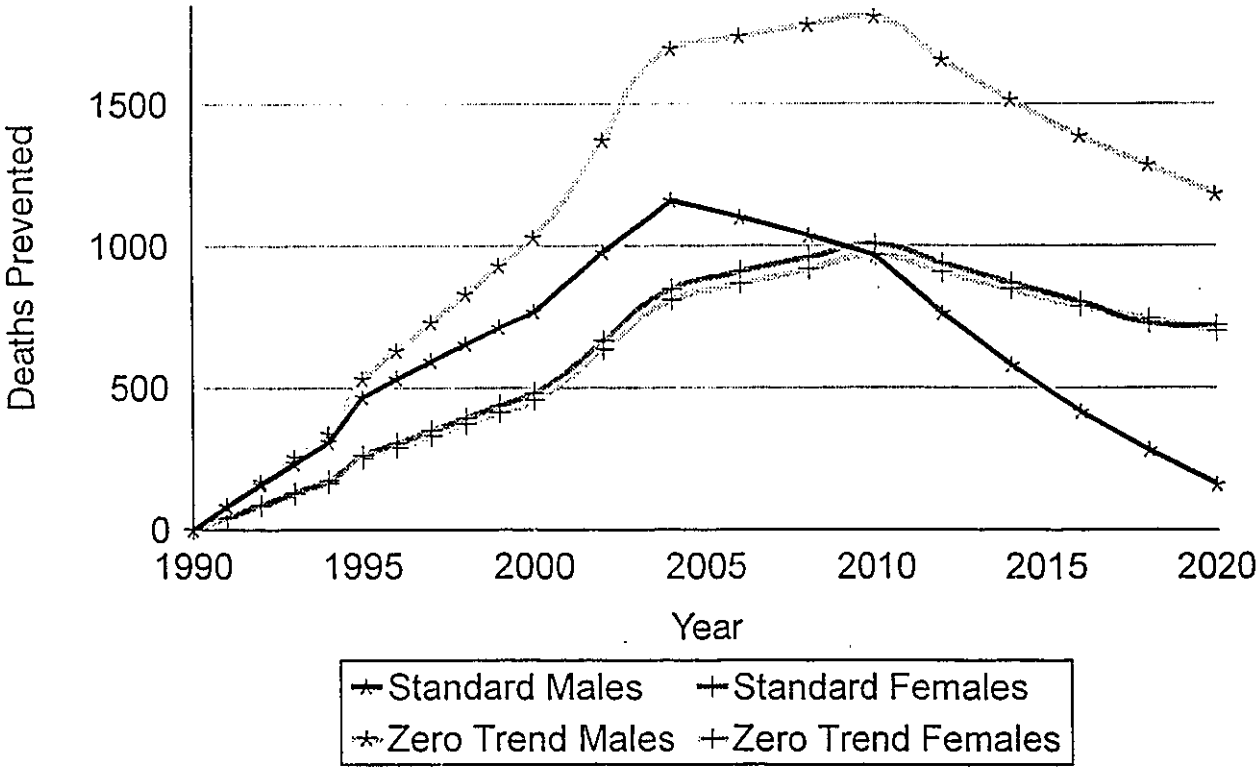
When the time dimensions are reduced to minimum the latency period disappears (modelled as zero), the mortality reduction starts immediately and reaches

a peak in the first year as modelled by the one year lag time.

8.1.4 Results of Test Simulation with Modified Trends

The category of data that is subject to the most uncertainty is the future trends. Yet based on changes in smoking behaviour in Ontario over the past twenty years it is incorrect to ignore future changes, especially in the age-group option in PREVENT. The proportion of young women who smoke is much larger than the proportion of

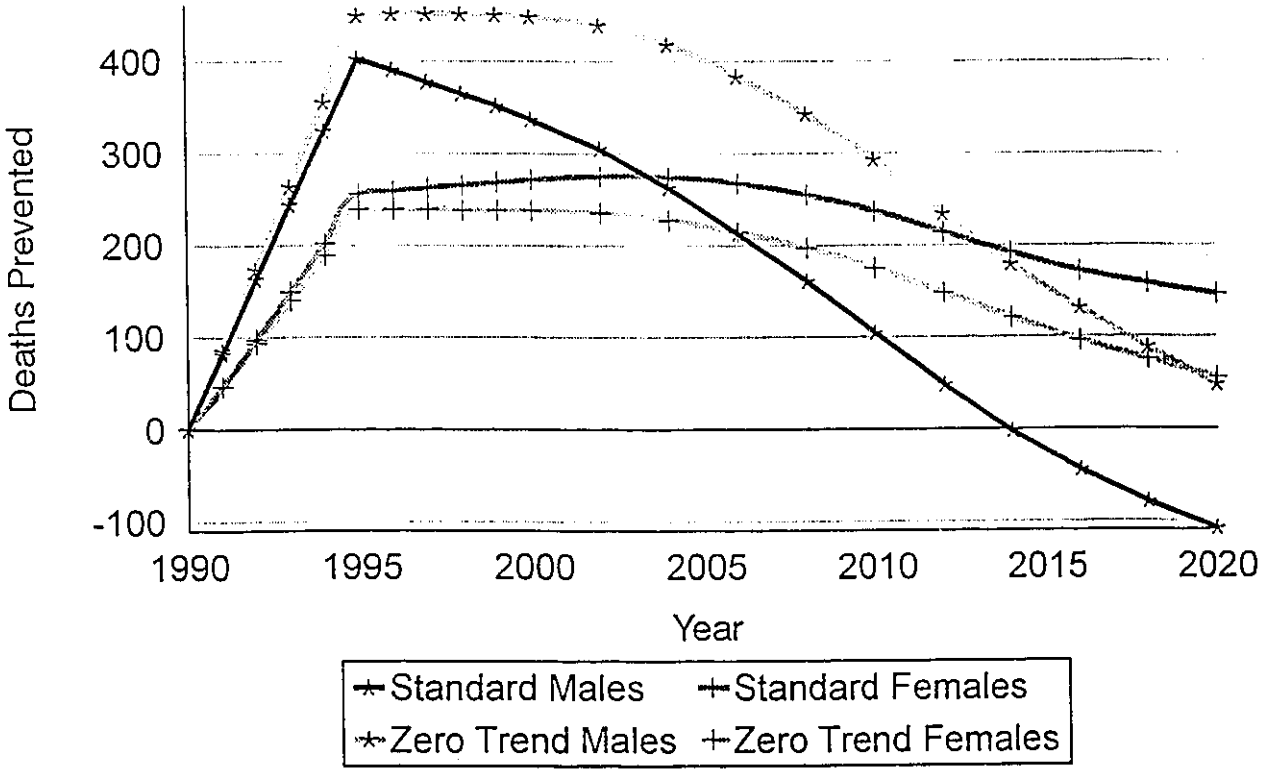
Figure 41 Total Mortality Reduction with Variation in Future Smoking Trends



older women who smoke. Thus in the age-group option, as young women in 1990 "age" the inflow must increase the proportion of smokers to model correctly.

Data Used In order to demonstrate the effect of trend data in smoking the results of the test intervention were compared using standard trends (Ontario database) and using a set of inflow and outflows with all values at zero (no change from 1990).

Figure 42 IHD Mortality Reduction with Variation in Future Smoking Trends



Results The total mortality for trend alone does not change greatly, and is not presented graphically because the curves are very nearly superimposed. At the end of the thirty year simulation period the total male mortality is 61923 in the standard set

and 64147 in the zero trend simulation. For females mortality with the standard set is 57658 and with zero trends 57621.

The results of total mortality reduction caused by the test intervention are presented in Figure 41. The IHD mortality reduction is presented in Figure 42.

Discussion The most important feature of both mortality reduction graphs (total and IHD) is the great difference in magnitude of male deaths compared to females. The shape of the curves is similar for each set of trends and the magnitudes of deaths prevented in females are surprisingly similar. The difference in male deaths is caused by the fact that the standard set of trends predicts that a large proportion of men at all ages will quit smoking. The "quit" rates projected for women are much smaller. If the zero trend simulation is compared to the standard, in each successive year more men are still smoking (providing a larger pool of males smokers to intervene upon).

8.1.5 Conclusions

The results of the sensitivity testing allow some overall conclusions. Relative risks and prevalence data do not affect the shape of the mortality reduction curves (impact of intervention), but they do affect the height of these curves. The impact of prevention (deaths avoided) varies directly with the relative risk and the prevalence of the risk factor in the population. Substitution of smoking prevalence from a second source resulted in very little change in outcome of the intervention modelled. Agreement between the two sources lends support to the veracity of the data obtained from the Ontario Health Survey.

Variations in the latency, lag time and period of implementation of an

intervention affect the shape of the mortality reduction curve (due to intervention), especially in the first years of the simulation. In contrast, the future trends modelled for risk factors affect impact of an intervention in the latter part of the simulation. Efforts at improving the database should focus on trends in risk factor prevalence. Even so, sensitivity analyses will continue to be needed to assess the impact of alternative scenarios for these trends.

8.2 Historical Testing

One method of testing the performance of models is to apply them to historical data. For the Ontario database, I have developed a second database, containing only the smoking risk factor as it affects lung cancer, COLD and IHD. Smoking was the only possible risk factor around which to construct an historical database. It is the only risk factor for which any comprehensive prevalence data are available before 1970, either on a provincial or national basis.

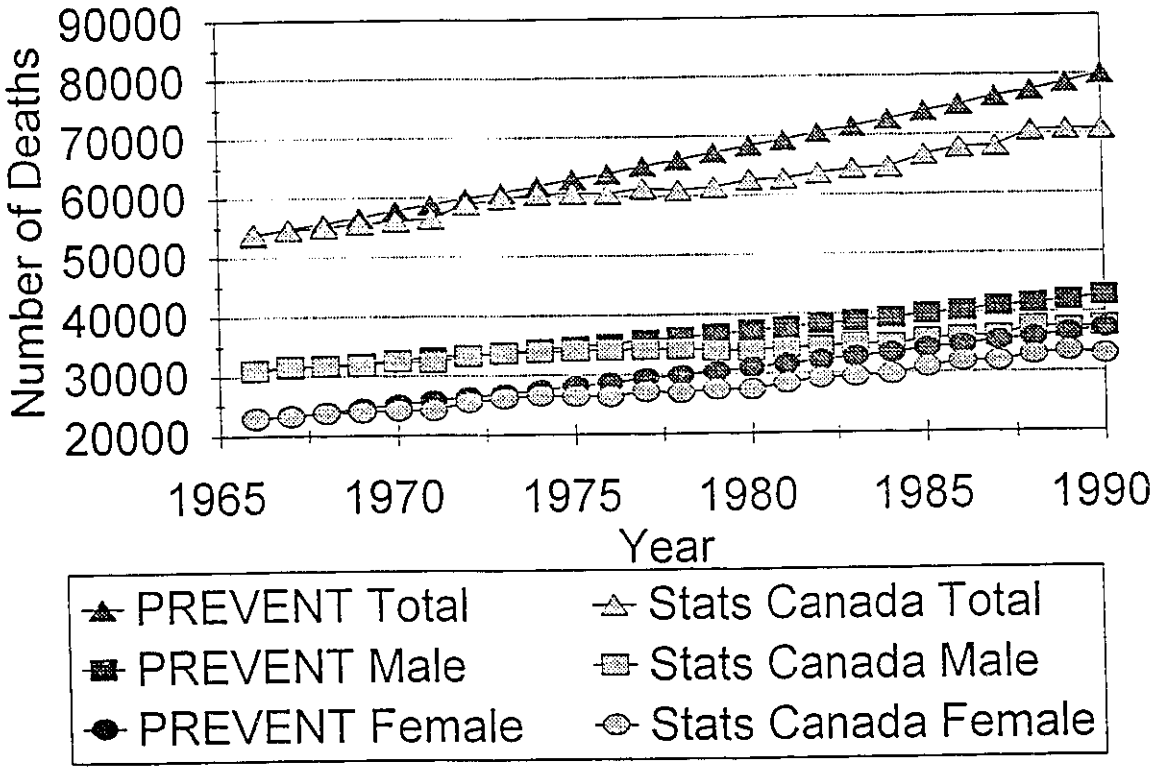
8.2.1 The Historical Database

Data Used The data elements and requirements are identical to those in the 1990 database, with the exception that they relate to the base year 1966. The base year chosen is 1966, because this is the first year for which stratified smoking prevalence data are available (part of the series of smoking surveys conducted as part of the Labour Force Surveys).

The population and mortality data for Ontario in the year 1966 were obtained from published tables produced by Statistics Canada[76]. The birth projections for 1966

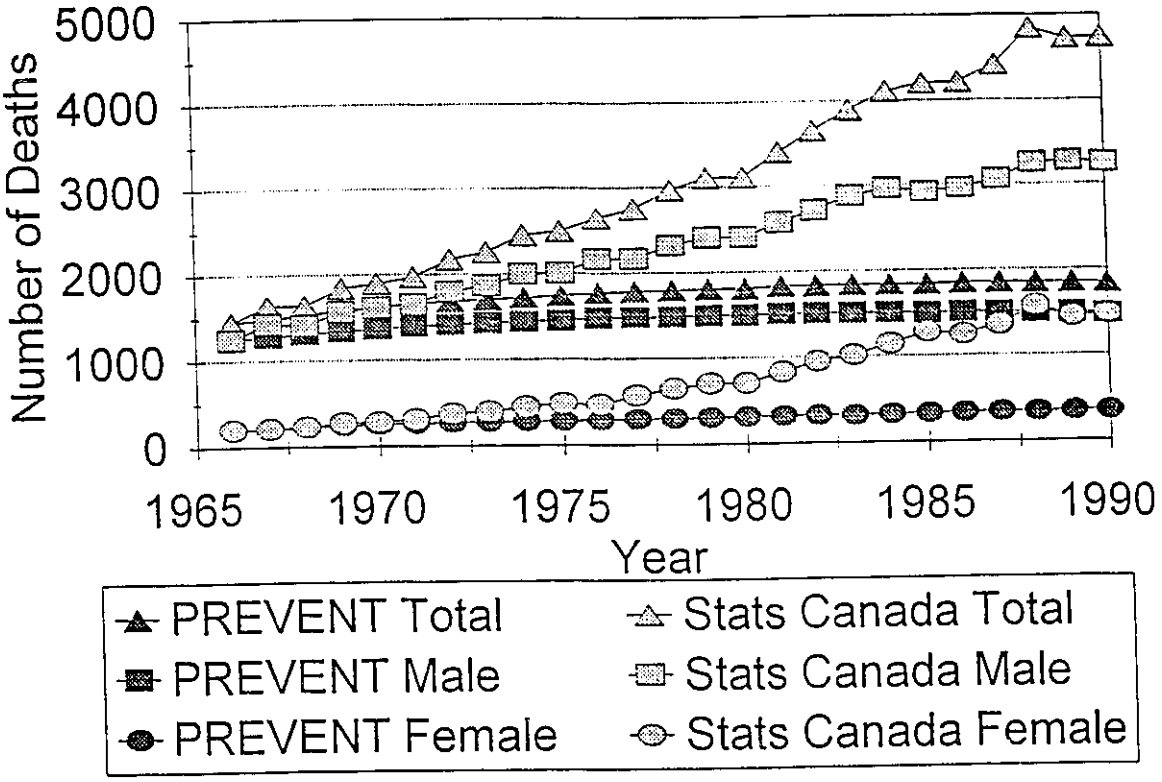
to 1990 were the actual births in Ontario for those years as reported in the same series of vital statistics tables published yearly by Statistic Canada. The birth projections for the years 1991 to 2016 were taken from the birth projection data set used in the Ontario 1990 database. Mortality rates by sex and five year age group for

Figure 43 Total Mortality Estimated by PREVENT 1966-1990



the three diseases in the model were obtained from the Laboratory Centre for Disease Control[41]. relative risks, residual risks and time dimensions chosen for the 1990 Ontario database were retained in the 1966 database.

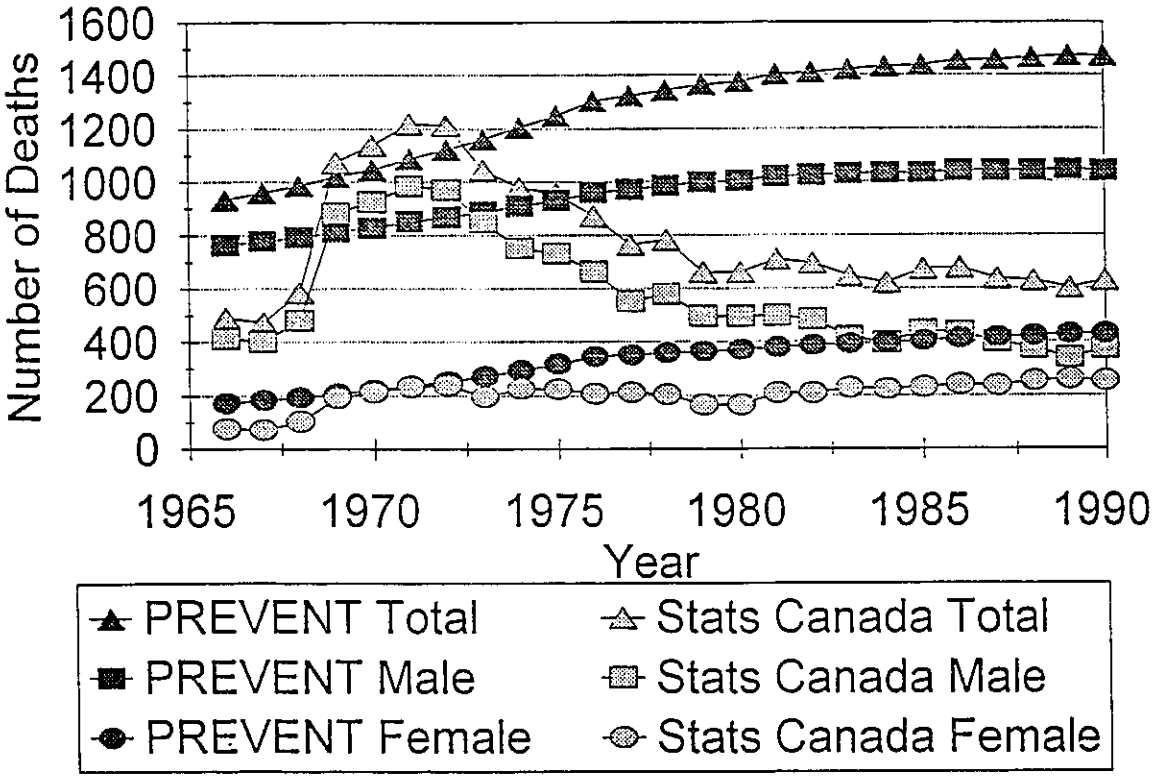
Figure 44 Lung Cancer Mortality Estimated by PREVENT 1966-1990



The prevalence data were taken from the 1966 Labour Force Survey as reported by Millar[56]. These data were national rather than specific to the province of Ontario. The trend data were estimated from changes observed in the Labour Force Survey results to 1990 and the future prevalence projections used in the 1990 database were used for the years 1991 to 2016. Smoking prevalence in Ontario for 1946 to 1966 was not available from provincial or national surveys. Estimates were taken from a paper by Ferrence who reported on cigarette smoking in Canada from 1900 to 1978[77]. There are no data available on the proportion of ex-smokers in the

population in 1966. At that time, the health risks of smoking were just beginning to be accepted. I have assumed that very few people were ex-smokers and estimated the percentage of ex-smokers as zero.

Figure 45 COLD Mortality Estimated by PREVENT 1966-1990

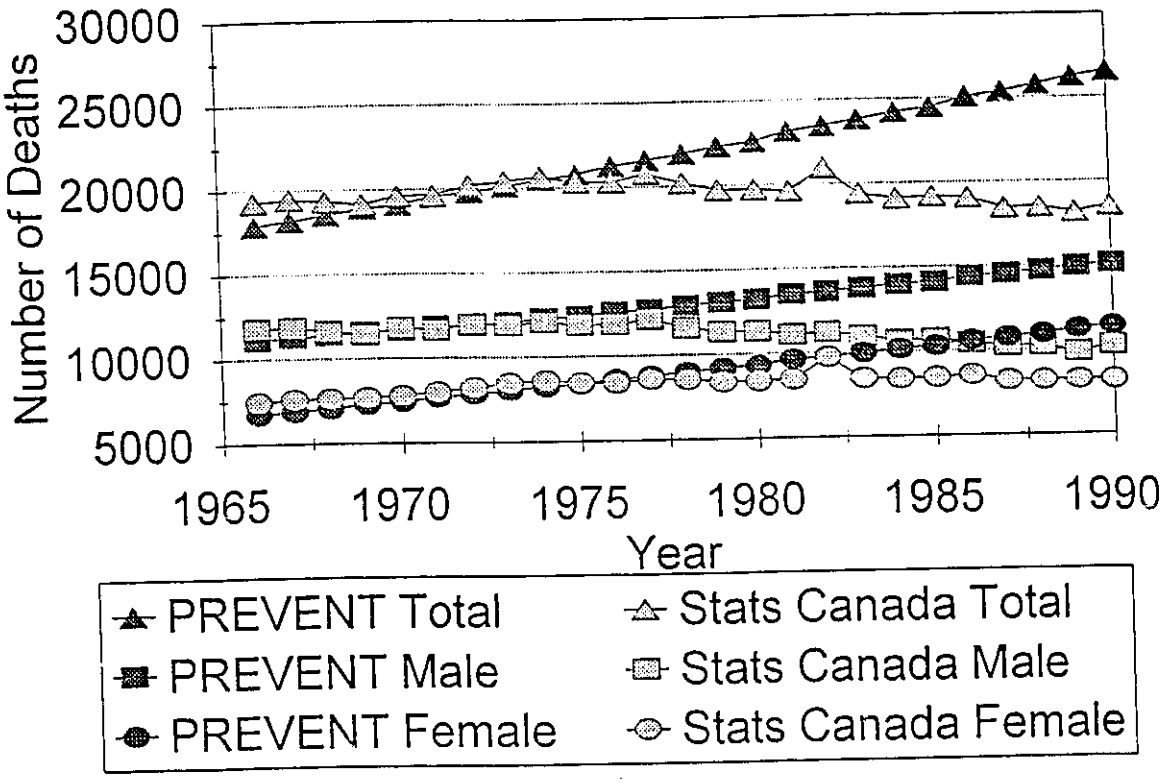


For this 1966 database only the age-group option for trends was prepared.

Results No intervention was modelled in the runs of the 1966 Ontario database. Rather, the mortality outcomes estimated by PREVENT were compared with observed mortality as published yearly by Statistics Canada in vital statistics tables. The results

for total mortality are presented in Figure 43. The results in terms of specific cause mortality for lung cancer, COLD and IHD are presented in Figures 44, 45,and 46

Figure 46 IHD Mortality Estimated by PREVENT 1966-1990



respectively.

Discussion The agreement in total mortality between the PREVENT run and observed data is quite high. As expected the agreement early in the simulation is very good, since the base year is modelled to the observed mortality data used in the input. Over the 24 year period between 1966 and 1990 the divergence gradually

increases and at the end of the period PREVENT models total mortality at 113% of that observed (113% for males, females and total). In this PREVENT simulation change in the prevalence of smoking was the only factor affecting all-cause mortality rates. Between 1966 and 1990 it is reasonable to suggest that other factors also contributed to the observed decline in mortality rates. Changes in other risk factors, some of them included in the 1990 database for PREVENT, but not modelled in this simulation could have led to reduced mortality. Improvements in the treatment of many diseases may also have resulted in delayed mortality. This argument is particularly valid for the observed reduction in age-specific mortality rates for IHD: during this period the development and use of coronary artery bypass grafting, angioplasty, beta blockers, ACE inhibitors and thrombolytic agents have led to much improved prognosis for IHD patients. The factors in the database itself that could be contributing to the over estimation are: relative risks that are too high, over estimation of smoking prevalence, or over estimation of smoking prevalence during the period (as modelled by the smoking trend data).

PREVENT is not as accurate in estimating lung cancer mortality. In this case PREVENT systematically underestimates mortality, predicting 38.8% of observed all mortality, 45.5 % of observed male mortality and only 24.1% of female mortality. While error can enter the PREVENT simulation from any one of the elements in the data set, large differential error between males and females is unusual. In this case it is likely due to changes in the trends of lung cancer mortality over the simulation period. Between 1966 and 1990 the age-specific mortality rates for lung cancer have increased substantially. As an example the observed lung cancer mortality rate for

males aged 60-64 years increased from 170.1 in 1966 to 251.7 in 1990 (a 48% increase). For females aged 60-64 years the observed mortality rate of 16.9 in 1966 rose to 99.6 in 1990, an increase of 489%. The PREVENT program works on the assumption that mortality rates are stable and it applies those of the base year to all successive years in the simulation. In this PREVENT simulation, mortality rates change only in response to changes in smoking prevalence. For the age group 60-64 PREVENT output estimates disease specific mortality rates declining from 170.5 in 1966 to 129.1 in 1990 for males. For females the rates increase from 16.9 in 1966 to 21.0 in 1990. These changes correspond almost exactly to the changes in smoking trends being modelled in the inflow and outflow files of the data set, but with a delay for the latency and lag times. The difficulty in this PREVENT simulation is likely due to its inability to model a lead time. If a long lead time was modelled for gradual acquisition of smoking risk (especially for the two lung diseases), then the delay between changes in prevalence and mortality rates would be longer. This would better reflect the increases in smoking prevalence that occurred in the 1940's and 1950's and this increase is likely to still be influencing mortality rates between 1966 and 1990..

The PREVENT program generally over estimates COLD mortality, especially for men. The shape of the observed mortality curve is irregular particularly during the first few years of observations. This irregularity is due to changes in the spectrum of diseases reported under COLD by Statistics Canada. There was a large change in 1969 and data before that date are not comparable. These reporting changes make interpretation difficult.

PREVENT also over estimates IHD mortality. At the end of the 24 year simulation period PREVENT over estimates all IHD mortality by 43%, male mortality by 47%, and female mortality by 37%. In this case the observed IHD mortality rates have declined for both men (41% of 1966 rates) and women (38% of 1966 rates) over the simulation period. Again this decrease in IHD mortality rates accounts for at least some of the difference between estimated and observed mortality. However, other factors could also be contributing. In the 1966 database the risk factors hypertension and hypercholesterolemia are not included, although identification and treatment of both has improved. Also, the identification and treatment of IHD itself has also improved.

Conclusion It is possible to make only a few conclusions and these should be accepted with caution for the following reasons. First, the mortality rates modelled for the years 1967 to 1990 are very dependent on the future trends as modelled by inflows and outflows in the database. These trends are based on estimates of smoking for the years 1945 to 1965 and not on population survey data (as in the 1990 database). Second, the 1966 database includes a single risk factor, smoking.

The 1966 Ontario database for PREVENT projected a reasonable estimate of all-cause mortality, as observed in the Ontario population between 1966 and 1990. The 1966 database overestimated IHD mortality and underestimated lung cancer mortality. There are some possible reasons for the discrepancy in these estimates. Improvements in the treatment of IHD over this period may have significantly improved prognosis for IHD patients while during the same period treatments for lung cancer have failed to improve prognosis. It is also possible that the lack of a lead

time in PREVENT has resulted in lung cancer atributable to very high smoking prevalence among men in the 1950s being modelled too soon in the simulation.

9.0 DISCUSSION

The variations in outcome modelled by PREVENT in the five simulations presented in section 7.0 (Results) illustrate the complexity and interdependency of factors affecting population health. Although not developed for this purpose, PREVENT is useful in illustrating this complexity. For example, PREVENT clearly demonstrates that substantial reductions in mortality can be achieved by interventions, even when such effects are masked by the large general increases in mortality expected in the Ontario population that are due to overall aging (of the peak population cohort born between 1945 and 1965). In general, PREVENT is an excellent working model of applied epidemiology.

9.1 Assumptions and Limitations in PREVENT

In its development several large theoretical assumptions were necessary in order to bring the underlying population attributable risk measure into a multifactorial, stratified mortality model. The first major assumption made is that combinations of risk factors interact with each other in a multiplicative manner and that exposures to risk factors are independently distributed in the population under study. Walter discussed the problems of attributable risk estimation in multifactorial situations[78]. He suggests that it is reasonable to assume independence of exposure to risk factors in certain situations where specific data are lacking[79]. Kupper and Hogan[80] review the calculation of risk interaction under the assumption of multiplicative combination. Spasoff and McDowell[20] explored options in calculation of the effects of joint exposure to several risk factors (including many used in the

Ontario database for PREVENT) and concluded that the multiplicative model is suitable for estimation of cancer risks.

The actual structure and calculations in the PREVENT programming impose other assumptions on the model. The relative risks and residual risks used in the input data are assumed to be fixed within each stratum. The program further assumes that as proportions of the population change strata (due to aging or change in risk factor exposure status), they acquire the risk associated with the new strata. In a similar manner the mortality rates, both all-cause and disease-specific, are specified in the input for the base year only. Throughout the 50 year simulation the base year mortality rates are applied to the population under study. This assumption can lead to substantial error regarding the performance of the historical database, as discussed in section 8.2.1.

The situation for population structure is somewhat better, in that birth projections allow for flexibility in the future population structures in the model.

While the latency and lag times incorporated in the model allow for the realistic modelling of mortality for changes in risk factors that have a delayed effect, the PREVENT model does not similarly model a lead time. Such a lead time would model a corresponding delay and/or gradual onset of full risk for proportions of the population that become exposed. Although PREVENT was primarily designed to model interventions expected to increase overall health benefits, it can also be used to model increases in risk factor prevalence when other factors indicate that this may occur. The simulation of increased smoking in response to declining cigarette prices, as presented in the results section, is an example of this. Also, the Ontario data

model inflow of older women into smoking strata over the simulation period. In both these situations the inclusion of adjustments for lead time would improve the program.

9.2 Limitations in Data Available for Input

Not only PREVENT but most of the other comprehensive population health models reviewed in section 1.3 are limited by available input data. Kotva [81] discusses general aspects of the data-model relationship, reviews the problems of using unverified epidemiologic simulation models and strongly emphasizes the need for national health information systems. The limitations of available data are quickly recognized when the international literature is searched for risk estimations jointly stratified by sex, age group and exposure level. Any limitations in this area can be quickly compounded by any disagreement in stratification and cut points used in available risk factor exposure data. In the development of the Ontario database for PREVENT, I was extremely fortunate to be able to request prevalence data from the Ontario Health Survey in the categories exactly matching the cut points used in the relative risks. Access to raw data from population surveys and large prospective epidemiologic studies is very limited. Even if such raw data were available, their manipulation and interpretation require advanced computing skills not available to most public health researchers.

9.3 Other Considerations

One of the theoretical considerations for users of population health models is

the utility of the model in dealing with competing risks. Cornfield[82] identified two aspects of competing risk that have been referred to extensively, the formal effect and the empirical effect. The formal effect considers the fact that if a person avoids death from one cause (as in the case of an intervention) he is then susceptible to death form other diseases. The empirical effect considers the possibility that a person who avoids death from one cause will have altered susceptibility to other causes. The phenomenon of competing risk has been extensively studied and computational models developed, as by Wong[83]. While PREVENT does not specifically adjust calculations to incorporate competing risk, its overall structure handles the problem well, especially if only the formal effect is considered. In the PREVENT model, the numbers "saved" by an intervention re-enter the general population pool and are susceptible to all other causes based the risks and mortality associated with that stratum of the population they occupy after the intervention. An example of this is found in section 7.4 of the results, where an intervention in alcohol use causes slight increases in the number of deaths due to accidental fall late in the simulation. This is due to deaths from accidental fall (high mortality rates in very old age) acting on that portion of the population which was earlier "saved" from death due to the other three causes of death affected by alcohol use.

The empirical effect is also partially dealt with in the PREVENT model. The proportion of the population that is "saved" from death due to any one cause is subsequently acted upon by all other diseases according to specific risks and mortality associated with their strata. Thus persons "saved" are often differentially susceptible to other causes of death due to residual risks associated with cessation of the risk

behaviour intervened upon. An example of this is presented in section 8.1.3 where differences in the latency periods for the three diseases associated with smoking result in slight increases in deaths due to COLD. This occurs because risk of IHD death decreases immediately (no latency period) and deaths from this cause are prevented, whereas risk reduction from COLD is delayed for a latency period of 10 years. That proportion of the population immediately "saved" by lower risk of IHD remain at increased risk of COLD for a full ten years.

10.0 POTENTIAL USE FOR ONTARIO PREVENT DATABASE

In health policy and planning there has been a call for a shift away from segmented and isolated programming in prevention and health services towards more comprehensive and integrated programming to achieve maximum health in the population. One of the prime goals for this shift is to control health care costs. There is increasing demand that policies and programs demonstrate their worth in terms of overall health outcomes.

Disease prevention has long been recognized as an important means of maximizing health, but unfortunately preventive programs are very hard to evaluate, much less estimate expected outcomes in advance[84]. In order to compete for limited resources, preventive programs must be able to use available information on risk factors and associated diseases to realistically predict the expected health benefits. In other words, they must accurately model their outcomes.

PREVENT was developed to model the benefits of intervention programs. It is useful in predicting both population trends (with no intervention) and in modelling interventions on modifiable risk factors. PREVENT was designed as a user-friendly, interactive, computer program so that users could produce informative outputs without having highly specialized computer skills. PREVENT has several features which help insure realistic outcomes. They include use of time dimensions and the multifactorial capacity. The latter allows simultaneous modelling of cause-specific and cumulative mortality outcomes. Because intervention outcomes are always compared with non-intervention results, the potential benefits of an intervention are evident even when future aging in the population indicates increased mortality in spite of intervention.

PREVENT has been successfully used in Europe in the following ways[85]. In the Netherlands a number of community health services use it to determine priorities in program planning and policy making. It has been used as a teaching tool to demonstrate the interrelationship between mortality rates, risk factor prevalence and demography. PREVENT has also been used to extrapolate health benefits of regional intervention programs to the national level (for the Netherlands).

The European experience with PREVENT has encountered two limitations. First users who try to construct new databases have experienced difficulty in finding and preparing input data. Secondly, adequate interpretation of results of PREVENT runs requires a basic understanding of the isolated epidemiologic elements used in the program and their interrelationship within PREVENT.

The Ontario database for PREVENT offers the opportunity for similar uses here. Regional health units can easily modify the population file in the database and create one reflecting their own communities if they by assume provincial prevalence and mortality rates apply. Furthermore, Ontario data could be used to model national intervention outcomes, with a similar change in the population file.

11.0 CONCLUSIONS

The Ontario database for PREVENT offers the opportunity for users to model the effects on the province of interventions on six risk factors as they affect eight causes of death. The results of runs using this database demonstrate its worth and flexibility in modelling both increases and decreases in risk factor prevalence, cohort and age-group interventions, and the cumulative impact of risk factors such as alcohol use on several causes of death. The use of latency and lag times realistically models the delayed impact of interventions on risk factors such as smoking.

Historical testing, using the 1966 Ontario database, demonstrates good agreement in total all-cause mortality between PREVENT and observed deaths in Ontario for the period 1966 to 1990.

It would be useful if future research could establish the extent to which the major assumptions in the model hold: independent distribution of risk factors in the population and multiplicative combination of risks. The database could also be refined and improved if more specific and accurate input data were to become available.

12.0 REFERENCES

1. Lalonde M. A new perspective on the health of Canadians: a working document. Ottawa: Minister of Supply and Services Canada, 1978.

2. Report of the Ontario Health Review Panel. Towards a shared direction for health in Ontario. Toronto, 1987.

3. Report on the Panel of Health Goals for Ontario. Health for all Ontario. Toronto, 1987.

4. Rose G. The strategy of preventive medicine. Oxford: Oxford University Press, 1992: 12.

5. Multiple risk factor intervention trial research group. Multiple risk factor intervention trial, risk factor changes and mortality results. Journal of the American Medical Association 1982; 248: 1465-77.

6. World Health Organization European collaborative group. European collaborative trial of multifactorial prevention of coronary heart disease: final report on the 6-year results. Lancet 1986; 1: 869-72.

7. Dawber, TR. The Framingham study. The epidemiology of atherosclerotic disease. Harvard University Press, Cambridge, Mass., London, 1980.

8. Doll R, Hill AB. Smoking and carcinoma of the lung. British Medical Journal; 2: 739.

9. Puska P, Salonen J, Tuomilehto J. The North Karelia project. Preventive Medicine 1983; 12: 191-5.

10. Wiley JA, Camacho TC. Life-style and future health: evidence from Alameda county study. Preventive Medicine 1980; 9: 1-21.

11. Chigan EN. Introduction. In: Morgenstern W, Chigan E, Prokhorskas R, Rusnak M, Schettker G, eds. Models of noncommunicable diseases: health status and health service requirements. Berlin: Springer-Verlag, 1992:1-3.

12. Gunning-Schepers, L. The health benefits of prevention: a simulation approach. Rotterdam: Erasmus University, 1988.

13. Peron Y, Strohmenger C. Demographic and health indicators: presentation and interpretation. Ottawa: Statistics Canada, 1985: 140-141.

14. McGinnis JM, Harrell JA, Artz LM, Files AA. Objectives-based strategies for disease-prevention. In:Holland WW, Detels R, Knox G. Oxford textbook of public health, second edition, volume 3: Applications in public health. Oxford: Oxford University Press,1991:127-9.

15. Walter SD. Calculation of attributable risks from epidemiologic data. International Journal of Epidemiology 1978; 7(2): 175-82.

16. Last JM, editor. A dictionary of epidemiology. Second edition. New York: Oxford University Press, 1988: 101.

17. Sturmans F, Mulder GH, Valkenburg HA. Estimation of the possible effect of interventive measures in the area of ischemic heart disease by the attributable risk percentage. American Journal of Epidemiology 1977; 105 (3):281-9.

18. Ouellet BL, Romeder J-M, Lance J-M. Premature mortality attributable to smoking and hazardous drinking in Canada. American Journal of Epidemiology 1979; 109(4): 451-63.

19. Morgenstern H, Bursic ES. A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population. Journal of Community Health 1982; 7: 292-309.

20. Spasoff RA, McDowell IW. Estimating the combined effect of several disease precursors in health risk appraisal. American Journal of Preventive Medicine 1987; 3: 182-9.

21. Eddy DM. CAN*TROL: a computer model for designing national cancer control strategies. Bull. Cancer 1987;74:323-32.

22. Shultz JM, Novotny TE, Rice DP. Quantifying the disease impact of cigarette smoking with SAMMEC II software. Public Health Reports; 106(3):326-33.

23. Shultz JM, Rice DP, Parker DL, Goodman RA, Stroh G, Chalmers N. Quantifying the disease impact of alcohol with ARDI software. Public Health Reports;106(4):443-50.

24. Weinstein MC, Coxson PG, Williams LW, Pass TM, Stason WB. Forecasting coronary heart disease incidence, mortality and cost: the coronary heart disease policy model. American Journal of Public Health 1987; 77 (11): 1417-26.

25. US Department of Health, Education and Welfare. The Framingham study: an epidemiologic investigation of cardiovascular disease. Some characteristics related to the incidence of cardiovascular disease and death, 18-year follow-up. Section 30. Bethesda, MD: Public health Service, 1973.

26. Grover SA, Abrahamowicz M, Joseph L, Brewer C, Coupal L, Suissa S. The benefits of treating hyperlipidemia to prevent coronary heart disease: estimating changes in life expectancy and morbidity. Journal of the American Medical Association 1992; 267(6): 816-22.

27. Milsum JH, Jones WN. Population attributable risk for health promotion policy making: a management tool. Canadian Journal of Public Health 1988; 79: 181-8.

28. Spasoff RA, McDowell IW. Further methodologic considerations, final report of HHA Implementation Project. Ottawa: Department of Epidemiology and Community Medicine, University of Ottawa, 1984.

29. Zhuo Z, Gatewood LC and Ackerman E. A Monte Carlo simulations program for coronary heart disease. In Fourteenth Symposium on Computer Applications in Medical Care (SCAMC) Conference. Los Alamos, California: IEEE Computer Society, 1990: 303-7.

30. Zhuo Z, Ackerman E, Gatewood L, Kottke T, Wu S-C and Park H-A. Polychotomous multivariate models for coronary heart disease simulation: I. tests of a logistic model. International Journal of Bio-Medical Computing. 1991; 27(2):133-48.

31. Zhuo Z, Ackerman E and Gatewood L. An expert system for simulation of coronary heart disease risk factor intervention. In: Proceedings - the Annual Symposium on Computer Applications in Health Care. 1991: 674-8.

32. Manton KG and Stallard E. Cross-sectional estimates of active life expectancy for the U.S. elderly and oldest-old populations. Journal of Gerontolgy 1991; 46: S170-82.

33. Manton KG and Stallard E. Chronic disease modelling: measurement and evaluation of the risks of chronic disease processes. London: Charles Griffin Ltd., 1988.

34. Wolfson, MC. POHEM - A framework for understanding and modelling the health of human populations. In: Proceedings of the Bureau of Census 1992 Annual Research Conference. Arlington, Virginia: U.S. Department of Commerce, 1992:261-82.

35. Wolfson MC, Manton KG. A review of models of population health expectancy: a micro-simulation perspective. In: Research Paper Series, Analytical Studies Branch. Ottawa: Statistics Canada, 1992. (Paper No.45)

36. Post censal annual estimates of population by marital status, age, sex and components of growth for Canada, provinces and territories, June 1, 1990. Ottawa: Statistics Canada 1991; Catalogue 91-210 Volume 7 (Table 2):30-3.

37. Life tables, Canada and provinces, 1985-87. Ottawa: Statistics Canada, 1990; Health Reports Supplement No. 13 1990; Volume 2 (No.4): 36-9.

38. Population projections for canada, provinces and territories, 1989-2011. Ottawa: Statistics Canada, 1991: Catalogue 91-520.

39. Unpublished data on components of population growth. Statistics Canada, 1991.

40. Live births by sex, Canada and provinces, 1931-1989. Ottawa: Statistics Canada,1991; Health Reports, Supplement No. 14.

41. Deaths 1990. Ottawa: Statistics Canada, 1992; Health Reports, Supplement No. 15, Volume 4, No.1.

42. Specific cause mortality rates, by 5 year age group and sex, for Ontario in 1990. Unpublished data. Laboratory Centre for Disease Control, Ottawa, 1991.

43. Mortality: summary list of causes 1990. Ottawa: Statistics Canada, 1992; Supplement No.12, Volume 4 No.1, Catalogue 82-003S.

44. Cutson TM. Falls in the elderly. American Family Physician 1994; 49(1): 149-56.

45. Jacobs DJ, Blackburn H, Higgins M, et al. Report of the conference on low blood cholesterol: mortality associations. Circulation 1992; 86: 1046-60.

46. Joffres MR, Hamet P, Rabkin SW, et al. Prevalence, control and awareness of high blood pressure among Canadian adults. Canadian Medical Association Journal 1992; 146(11): 1997-2005.

47. Conolly PW, MacLean DR, Horlick L, et al. Plasma lipids and lipoproteins and the prevalence of risk for coronary heart disease in Canadian adults. Canadian Medical Association Journal 1992; 146(11): 1977-87.

48. Mahajan V, Peterson RA. Models for innovation diffusion. Beverly Hills: Sage Publications, 1985.

49. Millar WJ. The smoking behaviour of Canadians -1986. Ottawa: Minister of Supply and Services Canada , 1988. (Health and Welfare Canada Catalogue No. 39-66/1988).

50. Health and Welfare Canada and Statistics Canada. The health of Canadians: report of the Canada Health Survey. Ottawa: Minister of Supply and Services, 1981.

51. Fitness Canada. Fitness and lifestyle in Canada (Canada Fitness Survey). Ottawa: Health and Welfare Canada, 1983.

52. Rootman I, Warren R, Stephens T, et al. (eds.). Health and Welfare Canada's Health Promotion Survey:technical report. Ottawa: minister of Supply and Services, 1985.

53. Health and social support 1985. Ottawa: Statistics Canada, 1987. (General Social Survey analysis series Cat No. 11-612).

54. MacLean DR, Petrasovits A, Nargundkar M, et al. Canadian heart health surveys: a profile of cardiovascular risk. Survey methods and data analysis. Canadian Medical Association Journal 1992; 146(11): 1969-74.

55. Health and Welfare Canada. Alcohol and other drug use by Canadians: a national alcohol and other drug survey (1989) technical report. Ottawa, Minister of Supply and Services, 1992.

56. Health and Welfare Canada. Canada's health promotion survey 1990 technical report. Ottawa, Supply and Services Canada, 1993.

57. Johansen H, Semenciw R, Morrison H. et. al. Important risk factors for death in adults: a 10-year follow-up of the nutrition Canada survey cohort. Canadian Medical Association Journal 1987; 136: 824-9.

58. Health and Welfare Canada. Main findings of the Canadian Blood Pressure Survey. Ottawa, Minister of Supply and Services, 1989.

59. Department of National health and Welfare. Nutrition Canada, nutrition: a national priority. Ottawa: Information Canada, 1973.

60. Williams B, Chang K, Truong M. Annual Sourcebook,Ontario profile 1992: alcohol and other drugs. Toronto: Addiction Research Fondation of Ontario, 1992.

61. Reeder BA, Angel A, Ledoux M, et al. Obesity and its relation to cardiovascular disease risk factors in Canadian Adults. Canadian Medical Association Journal 1992; 146(11): 2009-19.

62. Chin J, Mann JM. HIV infections and AIDS in the 1990s. Annual Review of Public Health 1990; 11: 127-42.

63. Kaslow RA, Francis DP. Epidemiology: general considerations. In AIDS: Occurrence in populations. pp. 87-97.

64. Taylor JMG, Chon Y. Smoothing grouped bivariate data to obtain the incubation period distribution of AIDS. Statistics in Medicine 1994; 13: 969-81.

65. Winkelstein W, Padian NS, Rutherford G and Jaffe HW. Homosexual men. In: Kaslow RA, Francis DP, eds. The epidemiology of AIDS: expression, occurrence and control of human immunodeficiency virus type I infection. New York: Oxford University Press, 1989: pp. 117-35.

66. Friedland G. Parenteral Drug Users. In: Kaslow RA, Francis DP, eds. The epidemiology of AIDS: expression, occurrence and control of human immunodeficiency virus type I infection. New York: Oxford University Press, 1989: pp. 153-78.

67. Burke DS, Brundage JF, Heibald JR, et al. Human immunodefficiency virus infections among civilian applicants for United States military service, October 1985 to March 1986: demographic factors associated with seropositivity. New England Journal of Medicine 1987; 317(3): 131-6.

68. Centers for disease Control. Pilot study of a household survey to determine HIV seroprevalence. Morbidity and Mortality Weekly Report. 1990; 40: 1-5.

69. McQuillan GM, Ezzati-Rice TM, Siller AB, Visscher W and Hurley P. Risk behavior and correlates of risk for HIV infection in the Dallas county household HIV survey. American Journal of Public Health. 1994; 84: 747-53.

70. Spasoff RA, McDowell IW, Wright PA, Dunkley GC. Evalu-Life. Proposals for a revised final report of a contract submitted to the Department of Health and Welfare, Canada. Ottawa: University of Ottawa, 1981.

71. Evans L. The effectiveness of safety belts in preventing fatalities. Accident Analysis and Prevention 1986; 18: 229-41.

72. McGee DL and Rhodes P. Estimating trend in the effectiveness of seat belts in saving lives, 1975-1985. Statistics in Medicine 1989; 8: 379-85.

73. Connolly MA and Kinball AW. Alcohol and traffic safety: a sensitivity analysis of data from composite sources. Accident Analysis and Prevention 1989; 21: 1-31.

74. Kaiserman MJ and Rogers B. Forty year trends in Canadian tobacco sales. Canadian Journal of Public Health 1992; 83(6): 404-6.

75. Stachenko SJ, Reeder BA, Lindsay E, et al. Smoking prevalence and associated risk factors in Canadian adults. Canadian Medical Association Journal 1992; 146(11):1989-96.

76. Vital statistics 1966. Ottawa, Canada: Dominion Bureau of Statistics, 1968. (Cat 84-202).

77. Ferrence, RG. Sex differences in cigarette smoking in Canada 1900-1978: a reconstructed cohort study. Canadian Journal of Public Health 1988; 79:160-165.

78. Walter SD, Holford TR. Additive, multiplicative, and other models for disease risks. American Journal of Epidemiology 1978; 108(5): 341-6.

79. Walter SD. Effects of interaction, confounding and observational error on attributable risk estimation. American Journal of Epidemiology 1983; 117: 598-604.

80. Kupper LL and Hogan MD. Interaction in epidemiologic studies. American Journal of Epidemiology 1978; 108 (6): 447-53.

81. Kotva M. Some conceptual problems in the simulation of epidemiologic processes. In: Morgenstern W, Chigan E, Prokhorskas R, et al. Models of noncommunicable diseases: health status and health service requirements. Berlin: Springer-Verlag, 1992: 5-10.

82. Cornfield J. The estimation of the probability of developing a disease in the presence of competing risks. Journal of the American Public Health Association, 1957; 47: 601.

83. Wong O. A competing risk model based on the life table procedure in epidemiologic studies. International Journal of Epidemiology 1977; 6(2): 53-9.

84. Russell LB. The cost effectiveness of preventive services. In: Goldbloom RB, Lawrence RS, eds. Preventing disease beyond the rhetoric. New York: Springer-Verlag, 1990: pp 435-7.

85. Gunning-Schepers L, Barendregt JAM and van der Maar, P. PREVENT , a model to estimate the health benefits of prevention. In: Morgenstern W, Chigan E, Prokhorskas R, et al. Models of noncommunicable disease: health status and health service requirements. Berlin: Springer-Verlag, 1992: 55-69.

APPENDIX A

DEFINITIONS AND CALCULATIONS OF PREVENT[1]

In the text of the thesis the following abbreviations and equations are used. The definitions and the equations are presented here with all the necessary subscripts.

IDR is the incidence density ratio, the ratio of the incidence density among exposed over the incidence density among never exposed. In the text of this thesis it is used interchangeably with relative risk (RR). The IDR is age, sex, risk factor, exposure category and disease specific. The IDR of the never exposed is usually 1 by definition.

Remnant IDR is the lowest Incidence Density Ratio that can be achieved after cessation of exposure. It is risk factor and disease specific but is equal for all exposure categories.

LAG is the time it takes (in years) after cessation of exposure, for the Incidence Density Ratio associated with a certain exposure category, to reach the remnant IDR level, through linear reduction. LAG is risk factor and disease specific but is assumed to be equal for all age, sex and exposure categories.

Lead time is the time it takes (in years) after first exposure to a risk factor to reach the full relative risk associated with that exposure category. It is not used in the Prevent model.

LAT is the latency period, the time (in years) between incidence and mortality. It is disease and risk factor specific but assumed to be equal for all age, and sex categories, and to remain equal over time for each disease/risk factor combination.

Exposure Category is a stratification by severity of exposure.

Ex-exposure Level is a stratification by the remaining effect of exposure after cessation of exposure.

P is the proportion of the population exposed to a certain risk factor. P is always age, sex and risk factor specific. It can be stratified further by exposure category and ex-exposure level. Within one age, sex specific subpopulation $\Sigma P = 1$.

PIDR is a intermediate variable, used to calculate EF's, TIF's and PIF's. It is risk factor, disease, sex, and age specific, it is time dependent, and there is a set of PIDR's for both the reference and the intervention population.

$$PIRD_{t}^{rjzsA} = \sum_{n=1}^{cn} \sum_{i=0}^{ID} P_{t-LAT^{n}}^{rjsAni} IDR^{rzsAni}$$

Where:

- P: proportion of the population
- IDR: Incidence Density Ratio
- cn: total number of exposure categories;
- n: index for exposure category;
- r: index for risk factor;
- ID: total number of ex-exposure levels;
- i: index for ex-exposure level:
- j=O,1: index for reference (0) or intervention population (1).
- A: index for age;
- s: index for sex;
- z: index for disease;
- t: index for time.

EF the etiologic fraction is the proportion of incident cases attributable to the prevalence of a risk factor in a population at a certain moment in time. It is always age, sex, risk factor and disease specific, and time dependent.

$$EF_t^{rjzsA} = \frac{PIDR_t^{rjzsA} - 1}{PIDR_t^{rjzsA}}$$

Where:

- r: index for risk factor;
- j=O,1: index for reference (0) or intervention population (1).
- A: index for age;
- s: index for sex;
- z: index for disease.
- t: index for time.

TIF is the trend impact fraction, the incident cases prevented at a certain moment in time. By an autonomous change in risk factor prevalence, as a proportion of the incident cases that

would have occurred at that time in the absence of change. TIF is initially always age, sex. risk factor and disease specific and time dependent. In a second stage TIF's are aggregated over risk factors.

$$TIF_t^{rzsA} = \frac{PIDR_t^{r0zsA} - PIDR_t^{r1zsA}}{PIDR_t^{r0zsA}}$$

Where:

- r: index for risk factor;
- j=O: index for reference population;
- A: index for age:
- s: index for sex:
- z: index for disease.
- t: index for time.

PIF is the potential impact fraction, the incident cases prevented at a certain moment in time, by an intervention to reduce risk factor prevalence, as a proportion of the incident cases that would have occurred at that time in the absence of the intervention. PIF is initially always age, sex, risk factor and disease specific and time dependent. In a second stage PIF's are aggregated over risk factors.

$$PIF_t^{rzsA} = \frac{PIDR_t^{r0zsA} - PIDR_t^{r1zsA}}{PIDR_0^{r0zsA}}$$

Where:
- r: index for risk factor;
- j=O,1: index for reference (O) or intervention population (1).
- A: index for age;
- s: index for sex;

- z: index for disease.
- t: index for time.

Disease specific mortality is the absolute number of deaths due to a specific disease at each point in time. This can be differentiated by sex.

Total mortality is the absolute number of deaths at each point in time. This can be differentiated by sex.

(Disease specific) Mortality difference is the difference in the total number of deaths between the intervention and the reference population, at each point in time. This can be differentiated by sex.

PYLG are the potential years of life gained, the absolute number of deaths avoided by an intervention, at a point in time, multiplied by the current life expectancy associated with the age at death.

AYLG are the actual years of life gained, the difference between the total population in the intervention and in the reference scenario, at a point in time.

References

1. Adapted from Gunning-Schepers L. The health benefits of prevention. A simulation approach. Rotterdam: Erasmus University, 1988.

APPENDIX B

The following is a review of the relative risks from international studies that were considered by Dr. Louise Gunning-Schepers in choosing relative risks for the Dutch database for PREVENT. These relative risks have been adopted in the Ontario database.

Smoking as a risk of Lung Cancer

A 1983 report on the health consequences of smoking[1], published by the U.S. Department of Health, reviewed available research on smoking and lung cancer . They presented data on smoking, as a risk for lung cancer, from eight prospective studies as presented in Table 1.

There are data available from four of these studies on the risks of lung cancer after smoking cessation, as presented in Table 2. Selection of relative risks for PREVENT was influenced heavily by the larger and longer studies: the American Cancer Society 25-State study, the British Physicians study, the Swedish Study, and the U.S. Veterans Study. In these studies, mortality ratios were not stratified by age group, but there is a definite increase in risk associated with increasing exposure level.

For the first four or five years after smoking cessation, studies report mortality ratios higher than those for current smokers. This effect is thought to be due to inclusion of people who quit smoking after lung cancer has already developed. Minimum

Table 1: Lung Cancer Mortality Ratios, by Sex and Amount smoked, from Eight Prospective Studies

Study	Population	Follow up (years)	Age Range (years)	Cigarettes pre day	Male Mortality Ratio	Female Mortality Ratio
ACS 25-State Study Hammond et al.	1,000,000 subjects/ 562,671 females in 25 U.S. states	12	35-84	none 1-9 10-19 20-39 40+	1.00 4.62 8.62 14.69 18.71	1.00 1.30 2.40 4.90 7.50
British Physicians Study Doll,Hill,Peto,Pike	40,000 subjects/- 6,000 females British Physicians	20-22	20-85+	none 1-14 15-24 25+	1.00 7.80 12.70 25.10	1.00 1.28 6.41 29.71
Swedish Study Cederlof,Friberg,- Hrubec,Lorich	55,000 subjects/- 27,700 females Probability sample Swedish population	10	18-69	none 1-7 8-15 16+	1.00 2.30 8.80 13.70	1.00 1.80 11.30
Japanese Study Hirayama	265,000 subjects/142,857 females Total population 9 Japanese Health Districts	13	40+	none 1-19 20-39 40+	1.00 3.49 5.69 6.45	1.00 1.90 4.20
U.S. Veterans Study Dorn,Kahn,Rogot	290,000 subjects/ <1% female U.S. Veterans	16	35-84	none 1-9 10-20 21-39 40+	1.00 3.89 9.63 16.70 23.70	
ACS 9-State Study Hammond,Horn	187,000 males in 9 U.S. states	4	50-69	none 1-9 10-19 20+	1.00 8.00 10.50 23.40	
Canadian Veterans Best,Josie,Walker	92,000 subjects/14,000 females	6	30-90	none 1-9 10-19 20+	1.00 9.50 15.80 17.30	
California Study Weir, Dunn, Linden, Breslow	68,000 males in various occupations	5-8	33-64	none ½ pk 1 pk 1½ pk	1.00 3.72 9.05 9.56	

rates are not achieved before 10 years of smoking cessation.

The relative risks of lung cancer associated with smoking that were actually chosen for PREVENT are presented in Table 3. Estimates of lag times, latency times

Table 2: Reduction in Mortality Ratio for Lung Cancer after Smoking Cessation, by Number of Years Since Cessation

Study	Cessation Period (years)	Number of Cigarettes (per day)	Mortality Ratio
British Physicians Study	1-4		16.0
	5-9		5.9
	10-14		5.3
	15+		2.0
	current smokers		14.0
U.S. Veterans Study	1-4		18.83
	5-9		7.73
	10-14		4.71
	15-19		4.81
	20+		2.10
	current smokers		11.28
Japanese Study	1-4		4.65
	5-9		2.50
	10+		1.35
	current smokers		3.76
ACS 25-State Study	<1	1-19	7.20
	1-4	1-19	4.60
	5-9	1-19	1.00
	10+	1-19	.40
	<1	20+	29.13
	1-4	20+	12.00
	5-9	20+	7.20
	10+	20+	1.06
	current smokers		13.67

and residual risks are also presented. The relative risks chosen were the same for both genders and across all age groups. Prevent does not assign an increase in lung cancer risk immediately following smoking cessation.

Smoking as a Risk for Ischemic Heart Disease

A 1979 report on smoking and health[2], published by the U.S. Department of Health, reviewed available research on smoking and ischemic heart disease (IHD). This report reviewed data on smoking as a risk for IHD from fourteen prospective

Table 3: Relative Risks of Smoking Used In PREVENT

Disease	Lung Cancer	COLD	IHD
Exposure Level	1-12 13-22 23+	1-12 13-22 23+	1-12 13-22 23+
Relative Risks Age <35 35-49 50-64 65+			3.0 3.0 4.0 3.0 3.0 4.0 1.7 2.0 2.4 1.5 1.5 1.7
Relative Risks All ages	7 12 20	12 25 30	
Latency	4	10	0
Lag Time	10	10	5
Residual Risk	2.0	2.6	1.2

studies, and nine of which are presented in Table 4.

A few of these studies also reported risk ratios for IHD after smoking cessation. The residual risk results of four studies are presented in Table 5. The residual risks of IHD, as reported in four studies, show a reduction in risk following smoking cessation, even in the first year after smoking stopped. Risk ratios declined to levels near those of nonsmokers after five years. The relative risks chosen for PREVENT are identical for men and women in the same age and exposure categories.

The relative risks, lag times, latency tines and residual risks of IHD associated with smoking that were chosen for PREVENT are presented in Table 3.

Smoking as a Risk for Chronic Obstructive Lung Disease

A 1984 report on smoking and chronic obstructive lung disease (COLD)[3], published by the U.S. Department of Health, reviewed available research on smoking and

Table 4: Mortality Ratios for IHD related to Smoking from Selected Prospective Studies

Study	Population	F.Up years	Cig. per day	Mortality Ratio for IHD Age Range				
ACS 9-State Study Hammond and Horn 1958	187,783 white males aged 50-69 years in 9 U.S. states	3.5		All	50-4	55-9	60-5	65-9
			smokers	1.70				
			1-9	1.29	1.93	1.85	1.66	1.41
			10-19	1.89	1.38	1.38	1.17	1.27
			20-39	2.20	2.00	2.04	1.91	1.58
			40+	2.41	2.51	2.47	1.92	1.56
Doyle et al. 1964	2,282 males,30-62yr (Framingham) 1,913 males,30-35yr (Albany)	10	smokers	2.40				
			<20	2.00				
			20	1.70				
			>20	3.50				
British Physicians Study Doll and Hill 1964	41,000 British Physicians	10		All	35-44	45-64	65-84	
			smokers	1.35				
			1-14	1.29	3.73	1.40	1.71	
			15-24	1.27	4.45	1.73	1.27	
			25+	1.43	1.36	1.92	1.58	
Swiss Physicians Study 1965 Strobel and Gsell	3,749 Swiss Physicians	9	none	1.00				
			≤20	1.48				
			>20	1.76				
Canadian Veterans Study Best 1966	78,000 male Canadian Veterans	6		All	30-49	50-69	70+	
			smokers	1.60				
			<10	1.55	0.97	1.56	1.71	
			10-20	1.58	1.45	1.67	1.29	
			>20	1.78	1.85	1.76	1.73	
U.S. Veterans Study Kahn 1966	U.S. male Veterans 2,265,674 person years	8.5	smokers	1.74				
			1-9	1.39				
			10-20	1.78				
			21-39	1.84				
			40+	2.00				
Japanese Study Hirayama 1967	265,118 Japanese adults over age 40	1	1-25	1.13				
			>25	1.00				
Kannel et al. 1968	5,127 males and females age 30-59	12	none	1.00				
			>20	2.20				
Hammond and Garfinkel 1969	358,534 males 445,875 females age 40-70 at entry	6	Males	All	40-49	50-59	60-69	70+
			1-9	1.27	1.60	1.59	1.48	1.14
			10-19	1.60	2.59	2.13	1.82	1.41
			20-30	1.73	3.76	2.40	1.91	1.49
			>40	1.77	5.51	2.79	1.79	1.47
			Females					
			1-9		1.31	1.15	1.04	0.76
			10-19		2.08	2.37	1.79	0.98
			20-30		3.62	2.68	2.08	1.27
			>40		3.31	3.73	2.02	

COLD. This report reviewed available data on smoking as a risk for COLD and the

Table 5: Reduction in Mortality Ratio for Ischemic Heart Disease after Smoking Cessation, by Number of Years Since Cessation

Study	Cessation Period (years)	Number of Cigarettes (per day)	Mortality Ratio
Jenkins et al. 1968 U.S.A.	Never smoked Current smokers Former smokers		1.00 2.36 2.15
Hammond and Garfinkel 1969 U.S.A.	Never smoked	1-19	1.00
	Current smoker		1.90
	stopped <1		1.62
	1-4		1.22
	5-9		1.26
	10-19		0.96
	20+		1.08
	All former smokers		1.16
		>20	
	Current smoker		2.55
	stopped <1		1.61
	1-4		1.51
	5-9		1.16
	10-19		1.25
	20+		1.05
	All former smokers		1.28
Shapiro et al 1969 U.S.A.	Never smoked Current smoker Stopped ≤5 yrs		1.00 1.87 0.76 (for MI not IHD)
Pooling Project American Heart Association 1970 U.S.A	Never smoked Former smoker		1.00 0.80

result of six prospective studies are presented in Table 6. Only one of these studies reported risk ratios for IHD after smoking cessation as presented in Table 7.

Smoking cessation has very little short term effect on risk ratios. There is very little risk reduction in the first ten years after cessation; but this may be partially due to inclusion of people already experiencing COLD symptoms. The risk decreases slowly over 20 to 30 years after smoking cessation and always remains much greater than that for nonsmokers.

Table 6: Chronic Obstructive Lung Disease Mortality Ratios, for Men and Women, by Number of Cigarettes Smoked

Study	Cigarettes per day	Mortality Ratio Male	Female	COLD Disease Classification
British Physicians	none	1.00	1.00	chronic bronchitis & emphysema
	1-14	17.00	10.50	
	15-24	26.00	28.50	
	25+	38.00	32.00	
U.S. Veterans	none	1.00		chronic bronchitis
	1-9	3.36		
	10-20	4.51		
	21-39	4.57		
	40+	8.31		
	none	1.00		emphysema
	1-9	5.33		
	10-19	14.04		
	21-39	17.04		
	40+	25.34		
	none	1.00		both
	1-9	4.84		
	10-19	11.23		
	21-39	17.45		
	40+	21.98		
Canadian Veterans	none	1.00		chronic bronchitis
	1-9	7.02		
	10-20	13.65		
	21+	14.63		
	none	1.00		emphysema
	1-9	4.81		
	10-20	6.12		
	21+	6.93		
Japanese	none	1.00	1.00	emphysema
	<100,000[1]	0.51	2.28	
	<200,000	2.57	3.14	
	>300,000	1.93	10.93	
California men in various occupations	none[2]	1.00		emphysema
	½ pk/dy	8.18		
	1 pk/dy	11.80		
	1½ pk/dy	20.86		
ACS 9-State Study	none	1.00		All pulmonary disease other than cancer
	1-9	1.67		
	10-19	3.00		
	20+	3.64		

[1] Lifetime estimate of number of cigarettes [2] Nonsmoker includes pipe and cigar smokers

The relative risks, lag times, latency times and residual risks of COLD associated with smoking that were chosen for PREVENT are presented in Table 3.

Table 7: Reduction in Mortality Ratio for COLD after Smoking Cessation, by Number of Years Cessation and Number of Cigarettes Smoked

Smoking Status		Cigarettes per day				
		0	<10	10-20	21-39	>30
Nonsmoker		1.00				
Ex-smoker			1.64	5.35	7.68	9.91
Current Smoker			4.84	11.23	17.45	21.98
		Years of Cessation				
Nonsmoker	Current Smoker	<5	5-9	10-14	15-20	>20
1.00	12.07	11.66	14.35	10.19	5.66	2.64

Hypertension as a Risk for Ischemic Heart Disease

Table 8 presents mortality ratios for IHD deaths, associated with hypertension,

Table 8: Mortality Ratios for Ischemic Heart Disease associated with Hypertension - Selected Prospective Studies

Study	Size	Follow up (years)	Severity	Reported IHD Mortality Risk Ratios			
				35-44	45-54	55-64	65+
Evans County	3102 males	7	SBP mild	2.52	1.54	1.37	1.78
			SBP severe	1.43	1.61	2.02	2.34
			DBP mild		0.36	1.88	1.46
			DBP severe	2.26	1.73	1.63	1.95
				30-39	40-49	50-59	
Framingham	5209	26	Mild male	1.88	1.32	1.65	
			Mild female	1.63	1.78	1.60	
			Severe male	2.34	1.76	1.88	
			Sev Female	2.95	2.69	2.71	
National Cooperative Pooling Project	7342	10	DBP	All ages			
			85-94	1.7			
			95-104	1.9			
			105+	3.6			

in three studies[4,6]. Unfortunately, none of these studies reported residual risks after control of hypertension.

Following World Health Organization (WHO) guidelines, PREVENT has adopted the following categories: mild hypertension is defined as a diastolic blood pressure (DBP) of 90-94 mm Hg and/or a systolic blood pressure of 140-159 mm Hg; severe hypertension is defined as DBP $\geq$ 95 mm Hg and/or SBP $\geq$ 160 mm Hg. Unless otherwise defined, classification of hypertension as mild or severe generally follows these definitions in PREVENT.

Of the three longitudinal studies presented in Table 8, the Framingham Study has the longest follow up period. It also reports risk ratios for males and females separately. For these reasons, relative risks for PREVENT were chosen from this study. Paucity of information does not allow objective choice of residual risk, latency or lag times. Reduction of risk is therefore assumed to occur immediately, upon control of hypertension. The relative risks, lag times, latency times and residual risks for IHD, associated with hypertension, that were chosen for PREVENT are presented in Table 9. There is assumed to be no residual risk after control of mild hypertension. There is assumed to be no latency in risk factor reduction.

Hypertension as a Risk for Cerebrovascular Accident

Table 10 presents mortality ratios for cerebrovascular accident (CVA) deaths, associated with hypertension, from the Framingham study[6]. The Framingham Study was chosen as the source of PREVENT data because the reported mortality ratios are

Table 9: Relative Risks for Diseases Associated with Hypertension

Disease	Ischemic Heart Disease		Cerebrovascular Disease	
Severity	mild	severe	mild	severe
< 45 years				
Male	1.9	2.3	3.5	5.0
Female	1.7	2.9	2.0	3.5
≥ 45 years				
Male	1.6	1.8	1.8	3.0
Female	1.6	2.7	1.5	3.0
Latency	0		0	
Lag time	2		1	
Residual Risk (all ages)	1.0	1.6	1.5	1.5

Table 10: Mortality Ratios for Cerebrovascular Disease Associated with Hypertension

Severity of Hypertension	30-39 years		40-49 years		50-59 years	
	Males	Females	Males	Females	Males	Females
Mild	3.86	2.06	1.78	1.96	2.02	1.49
Severe	10.20	3.38	3.27	3.46	2.98	2.70

available stratified by sex, age group and exposure level. One study, involving pooled data, has made general observations on risk reduction after control of hypertension. They report risk reductions as high as 50% in the studies they reviewed. As with hypertension and IHD, PREVENT assumes that the reduction of risk of CVA is immediate upon control of hypertension.

The relative risks for CVA, associated with hypertension, that were chosen for PREVENT are presented in Table 9. There is no residual risk assigned after control of mild hypertension. Latency and lag times chosen were zero and one year

respectively, the lowest values accepted by the program.

Obesity as a Risk of Breast Cancer

The mortality risk of breast cancer associated with obesity has been reported by De Waard[5], and is supported by other sources[6]. DeWaard has reported a relative risk of 1.3 for women with a Body Mass Index (BMI, or Quetelet Index) of 30 or more. Mortality data indicate that the risk of breast cancer gradually increases with age, particularly after menopause. There is no information available to help set latency or lag times. There is no discussion in the literature concerning any residual risk for obese individuals after weight reduction. Women aged 50 or older and with BMI of 30 or greater, are assigned a relative risk of 1.3. Women less than 50 years of age and men are not at risk and are assigned a relative risk of 1.0. The latency period for obesity and breast cancer is set at 5 years with a lag period of one year. The residual risk is set at 1.0.

Alcohol as a Risk for Ischemic Heart Disease

One of the largest general population longitudinal, databases which records alcohol use patterns is the Kaiser-Permanente (Health Plan members) in California. A ten year follow up of this database[7] reported mortality from various causes by alcohol consumption.

There were no reports to help establish latency and lag times. There was no discussion of any residual effects related to either abstinence or to excessive alcohol

consumption after individuals changed from one of these categories to take up a moderate alcohol consumption habit.

For the PREVENT database "abstainers" are defined as individuals who consume less than four drinks per week on average, moderate drinkers consume 4 to 21 drinks per week, and excessive drinkers consume 22 or more drinks per week. In accordance with the reported literature, the moderate drinkers are presumed to be at lowest risk of IHD and are assigned a relative risk of 1.0.

Because of the lack of information across age groups, the relative risks for IHD associated with alcohol use will be assumed to hold for all adult age groups. Women will be assigned the same risk magnitudes as men. The relative risks,

Table 11 Relative Risks for Diseases Associated with Alcohol Consumption

Disease Exposure Category	IHD		Cirrhosis		Accidental Fall		M Vehicle Crash	
	abstain	excess	abstain	excess	abstain	excess	abstain	excess
Relative Risk	2.0	2.0	1.0	9.0	1.0	2.0	1.0	2.0
Residual Risk	1.0	1.0	1.0	3.0	1.0	1.0	1.0	1.0
Latency (years)	0		4		0		0	
Lag time (years)	1		1		1		1	

The reference category is "moderate" for all diseases. The relative risks are the same for all age groups and both sexes.

residual risks, latency and lag times are presented in Table 11.

Alcohol as a Risk for Cirrhosis

The ten year follow up on the Kaiser-Permanente database was also used to

estimate the risk of cirrhosis associated with excess alcohol consumption. The lower limit of excessive drinkers used for risk of IHD, 22 or more drinks a week, tends to be low compared to alcohol consumption levels defined as "alcoholic" in literature reports. The cut point used in the Kaiser-Permanente study for excessive drinker was more than 5 drinks daily, and is closer to the cut point chosen for PREVENT. It is well established that cirrhosis only develops after years of excessive alcohol consumption and that liver damage is cumulative and not readily reversible. For this reason the residual risk after moderation in consumption habit remains high.

The PREVENT program requires that all diseases related to a single risk factor use the same reference (lowest risk) category. Normally this reference category is for those not exposed or with the lowest possible exposure. Because of the U shaped relationship between alcohol and risk of IHD, it is necessary to define the reference category for all diseases related to alcohol as "moderate" to accommodate IHD in the model. Although, the risk of cirrhosis is usually defined in comparison to abstainers, the risk among moderate drinkers is also very low and moderate drinkers can be used as the reference category as long as the abstainers are also assigned to minimum risk (RR 1.0). The relative risks, residual risks, latency and lag times are presented in Table 11.

Alcohol as a Risk for Accidental Fall

When the Dutch database was developed there was little available information linking average daily or weekly levels of alcohol consumption (as available from

population surveys) to risk of accidental fall. The Kaiser-Permanente ten year follow-up study gives a mortality estimate for accidents in general, but groups motor vehicle crash and other accidents together.

Although moderate drinkers are defined as the reference category, abstainers are assigned the same minimum relative risk of 1.0. The relative risks assigned are for all age groups and both sexes. The relative risks, residual risks, latency and lag times are presented in Table 11.

Alcohol as a Risk for Motor Vehicle Crash

As in the case of accidental fall there is little available information linking average daily or weekly levels of alcohol consumption (as available from population surveys) to risk of motor vehicle crash. The Dutch database again used the Kaiser-Permanente ten year follow up study as a source of mortality estimates for accidents in general.

The relative risks assigned are for all age groups and both sexes. The lag time is set at one year and the latency at 0 (the minimum values accepted by the PREVENT program). The residual risk set at 1.0. The relative risks can be compared with other alcohol related diseases in Table 11.

References

1. The health consequences of smoking. Cancer. A report of the Surgeon General. Washington: USDHHS,1982.

2. Smoking and health, a report of the Surgeon General. Washington: USDHHS, 1979.

3. The health consequences of smoking. Chronic obstructive lung disease. A report of the Surgeon General. Washington: USDHHS, 1984.

4. Cassel JC. Summary of major findings of the Evans county cardiovascular studies. Arch Int Med 1971; 128:887-889.

5. DeWaard F. Body size, body mass and cancer of the breast. In: Alan R, editor. Dietary fat and cancer. Liss Inc., 1986.

6. Doll R, Peto R. The causes of cancer. Quantitative estimates of avoidable risk of cancer in the US today. New York, Oxford University Press, 1981.

7. Klatsky AL, Friedman GD, Siegelaub AB. Alcohol and mortality. Annals of Internal Medicine 1981; 95:139-45.