



FORECASTING MORBIDITY, MORTALITY & DISABILITY IN NONCOMMUNICABLE DISEASES

Summary of Discussions at an Informal Workshop Cambridge, Mass. USA 22-23 October, 1997

The Informal Workshop was held jointly by the Ageing and Health Unit, World Health Organization (WHO) and the Burden of Disease Unit, Center for Population & Development Studies, Harvard University School of Public Health. The National Institute on Aging (NIA) at the National Institutes of Health (NIH) in the United States was a major funding source for the meeting. The following summary follows the meeting agenda which is provided as Attachment I. This report should be considered an attempt to provide a concise synopsis of the various presentations and discussions which occurred during the meeting. Though many of the officials in attendance at the meeting were affiliated with governmental and non-governmental organizations and institutions, none of the views expressed here should be considered official positions or statements of policy. This was an informal and lively assemblage of producers and users of projection models for noncommunicable diseases. This environment was intended to encourage a substantive as well as candid exchange of ideas amongst the participants.

A list of participants who were registered to attend the meeting is enclosed as Attachment II. Not all of the persons listed on this attachment were physically present at the meeting.

October 22, 1997

Opening of the meeting:

The meeting was opened by Ed Dowd who made the following points:

WHO has interests in projection models which attempt to estimate future health events since WHO is adopting a focus aimed at optimizing the cost-effectiveness of interventions. In this context systematic forecasting models:

- 1) Help to pinpoint which health issues are most important nationally and regionally and hence require interventions.
- 2) Facilitate production of cost estimates for interventions and health outcomes.
- 3) Provide estimates and projections of the consequences of interventions.

This document is not issued to the general public, and all rights are reserved by the World Health Organization (WHO). The document may not be reviewed, abstracted, quoted, reproduced or translated, in part or in whole, without the prior written permission of WHO. No part of this document may be stored in a retrieval system or transmitted in any form or by any means – electronic, mechanical or other – without the prior written permission of WHO.

The views expressed in documents by named authors are solely the responsibility of those authors.

There have been ongoing working groups investigating projection models at WHO. However, communicable diseases and noncommunicable diseases of young and middle-aged persons have been the primary focus of these groups. With the publication of the *Global Burden of Disease (GBD)* there has been an increased interest in the impact of ageing on the anticipated burden of non-communicable diseases. This workshop has the aim of facilitating the development of further work in the projection methodology of models designed for forecasting non-communicable disease trends, and the impact of interventions.

Participant self-introductions:

Each participant who was physically present at the opening of the meeting then identified themselves. Addresses and affiliations are provided as Attachment II. Persons present at the opening were:

| | | | |
|---------------------|---------------|-----------------|--------------|
| A. Acharya | J. Barendregt | J.M. Bertholet | T. Buettner |
| D. Cutler | J. Dowd | L. Gunning | W. Harlan |
| A.E.M. de Hollander | H. King | P.D. Johnson | M. McKenna |
| C.J.L. Murray | S. Preston | J.A. Schuttinga | K. Shibuye |
| E.J. Sondik | R. Suzman | M. Wolfson | M. Weinstein |
| D. Wise | | | |

Opening session chaired by Sam Preston:

Comments by Sam Preston:

Dr. Preston introduced himself as a demographer and economist who has had a long interest in mortality trends at the population level. He noted that most of the funding for this meeting was furnished by the National Institute on Aging. Therefore, he asked Dr. Richard Suzman to provide a few comments prior to convening the formal presentations listed on the agenda.

Comments by Richard Suzman:

Dr. Suzman noted that the GBD had a considerable impact at the National Institutes of Health. One of the most interesting epiphanies precipitated by the findings in this report was the realization within segments of the biological community of the enormous importance of social and economic factors in the production of health. The rapid changes in the global rank order of chronic versus acute disease forecast by the GBD was on a time scale incommensurate with any hypothesis of "genetic drift."

However, two major concerns arose as a result of the report:

- 1) There was concern that the impact of diabetes mellitus was underestimated, especially on the incidence and mortality associated with cardiovascular disease. This is because comorbid states were not incorporated into the models and only underlying causes were used to attribute mortality.
- 2) There did not appear to be any way within the framework of GBD to look at the "what if" impact on chronic disease prevalence and outcomes of biomedical and behavioral oriented interventions.

Concurrent with the publication of the GBD two other events took place:

- 1) Ken Manton published a series of papers based on the National Long Term Care Survey (NLTCs) which documented a profound decreasing trend in the incidence and prevalence of disability in older Americans. Many researchers were profoundly surprised by these findings, since they had been expecting the opposite -- that increasing survival at older ages would be accompanied by an expansion of the prevalence of disability.
- 2) A summit meeting in Denver was held where population ageing made it on the agenda. This summit called for more research on ageing.

The result of these two events is that there is an increasing appreciation of the value of demography and population research. One of the anticipated outcomes of the current meeting is a better delineation of the data that should be collected to facilitate surveillance, modeling, and projection research. Such research helps to elucidate the future impact of population ageing on society and the health care system. Also, there is a recognition that several existing datasets, could facilitate such research. This meeting provides an opportunity to bring together investigators who are interested in modeling and projection research to better define these data requirements.

Presentation by Chris Murray:

Dr. Murray presented the basic structure and theory behind the projection model which was used to produce the estimates for the *Global Burden of Disease* (GBD). There were 3 major objectives of the GBD projection methodology.

- A) Develop plausible projection scenarios for the burden of disease for the entire globe, specific regions, and selected populations.

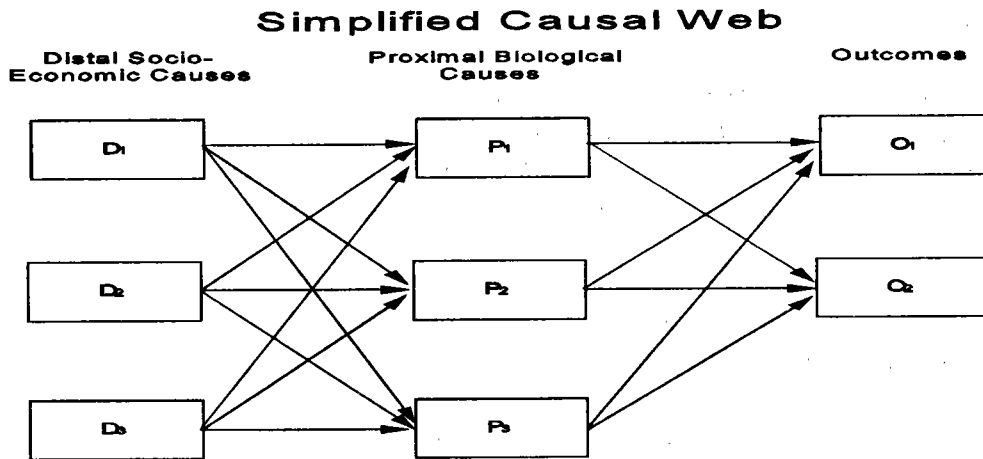


Figure 1

- B) Analyze the impact of different socio-economic and health care policies
- C) Project low probability but relevant scenarios (e.g., drug resistance, new diseases)

An underlying dynamic of any projection model is that the projections can be self-defeating: If the projections influence decision makers, the basis on which the projections are developed may be undermined.

A simplified causal web provides a basic framework for understanding most projection model approaches (see Figure 1).

Medical technology is viewed as a proximal cause under this schema. However, there are actually 3 types of causes operating at the distal and proximal levels-1) Known and observed, 2) known but unobserved, 3) Unknown causes (see Figure 2).

Frequently the “unknown” cause is “left-over” as residual variance and is assigned to the relevant modeler’s favorite cause.

With such a framework for understanding projection models, most modelers take one of two basic approaches:

- 1) A Proximal Causes Approach where the focus is on known and observed proximal causes. The functional relationships and parameter values for proximal causes are estimated from small, intensively studied populations such as Framingham. Individual level relationships and projections are developed using micro simulation, macro simulation or other analytic approaches. Unknown or known but unobserved proximal causes may be included by using distal causes as proxies.

The advantages to this approach are:

- a) describes complex non-linear multiple risk factor relationships and non-

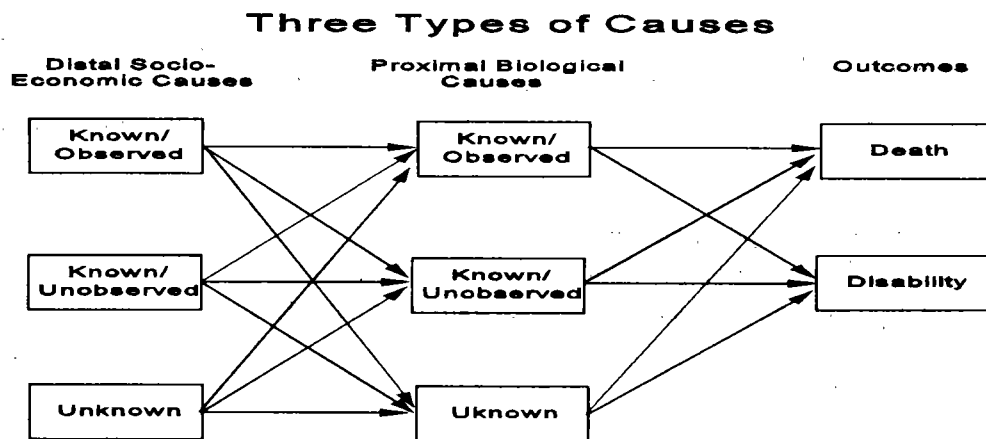


Figure 2

- b) independent risk factor distributions well.
- b) It is possible to evaluate exogenous changes in proximal causes through health interventions.

The disadvantages are:

- a) Low generalizability since the results are based on small populations.
- b) The distal causes are ignored and if they are entered they are included only as proxies for missing proximals.
- c) Models are rarely validated by fitting them to trends in the past.

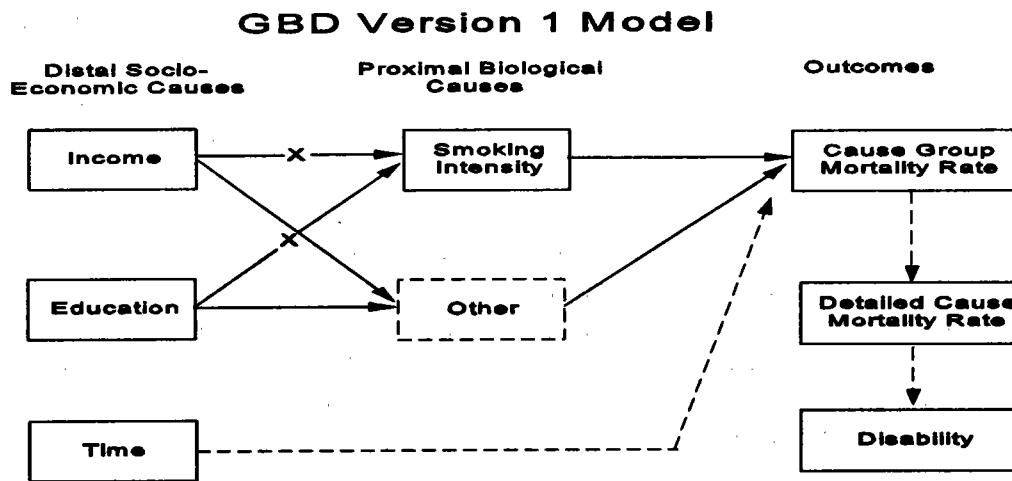


Figure 3

- 2) The Distal Causes Approach where the focus on the relationship is between distal causes and outcomes without specifying how distal causes mediate proximal causes pathways. The parameter estimates are based on fitting observed aggregate trends using various methods.

The advantages to this approach are:

- a) It is simple and possesses minimal data requirements.
- b) The modeler can fit to panels of aggregate mortality data.

The disadvantages:

- a) Difficult to capture distributional effects within populations.
- b) Cannot model the impact of health interventions that act directly on proximal causes.
- c) The model provides no insight into the pathways through which changes are mediated.

The GBD attempted to explore a variety of distal and proximal causes and outcomes. In this way both socio-economic changes and health interventions could be evaluated. However, because of data limitations only 3 distal causes were included, and one specific proximal cause (smoking intensity- see Figure 3).

The basic form of the model used in the GBD was:

$$\ln M_{a,k,i} = \beta_1 \ln Y + \beta_2 (\ln Y)^2 + \beta_3 \ln E + \beta_4 \ln SI + \beta_5 T + C_{a,k,i}$$

Where M is the mortality rate in age group a , sex k , and cause I . Y , E and T denote GDP per capita, education and time, respectively. The causes were grouped into 9 cause clusters for 14 age and sex groups using mortality data from 47 countries for the years 1950-1990. For selected causes where the regression models explained only a small fraction of the variance alternative assumptions were imposed. The model predictions for 1990 were compared to the actual burden and scalars were calculated. The scalars were assumed to be constant for projections from 1990 to 2020.

Baseline, pessimistic and optimistic projections of income, average years of education, and smoking were developed. Detailed cause-specific mortality rates were estimated as a function of the trends in group cause mortality using ICD-9 data from 67 countries. Some causes were assumed to be a constant fraction of the cause group mortality.

Age, sex and cause specific mortality rates were projected for each region for 1995, 2000, 20005, 2010, 2015 and 2020. An independent projection model was used for HIV since this disease did not exist prior to the early 1980s. Tuberculosis (TB) was also broken out for India and Africa to allow for the interaction between HIV and TB in these regions.

In order to provide estimates for disability as well as mortality the ratio of the age, sex and region specific Years Lost to Disability (YLD) versus Years of Life Lost (YLL) was kept constant for causes where $YLD/YLL < 10$. For selected causes (e.g. STDs) it was assumed that YLDs declined as a scalar (<1) of the decline in YLLs. For many neuro-psychiatric causes and musculoskeletal diseases which cause large amounts of disability, but little mortality, constant YLD rates were assumed. The trends for these conditions were strictly determined by changes in population demographics.

Using World Bank fertility projections and GBD mortality projections optimistic, pessimistic and baseline estimates of population trends were developed. Region, cause and time specific YLLs, YLDs, deaths and DALYs were produced using the projected rates and populations.

In the future the GBD process needs to incorporate more information on proximal factors in order to facilitate evaluation of different health interventions. This requires specification of the relationships between distal and proximal risk factors and explicit assignment of parameters for unknown and unobserved distal and proximal causes. The process would involve using data from intensively studied populations to inform the choice of functional forms for proximal causes and parameter values. These would be validated by fitting them to real aggregate data trends. The GBD process also needs to expand the number of cause groups for more detail, re-estimate the relationship between detailed causes and cause groups, and develop more flexible models of the relationship between YLDs and YLLs (i.e. the case-fatality relationship).

The current plan for such an analysis has the basic form depicted in Figure 4.

The current information available to facilitate the estimation process for such a model includes:

- 1) The prior distributions and form and parameter values for A_1 and A_2 are well described.
- 2) There is relatively complete information for smoking intensity, but little for physical activity from 1950s or 1960s.
- 3) Little prior information on $B_1(Y,t)$, $B_2(Y,t)$, $B_4(E,t)$, or $B_5(E,t)$.
- 4) Information regarding income (Y) and education is relatively complete.

Theoretically, O_D can be modeled if $O_{D,t} = f(O_{D,t-1})$; $O_{D,t} = f^*(O_{D,0})$.

The modeling would involve choosing a plausible functional form for each relationship and a set of N parameters for the model and values of parameters when uncertain. Then develop a likelihood function for outcomes and perhaps the proximal values. Then fit a "set" of models and make a "set" of projections. With this "set" of models a constrained maximum likelihood

Extending the GBD Model

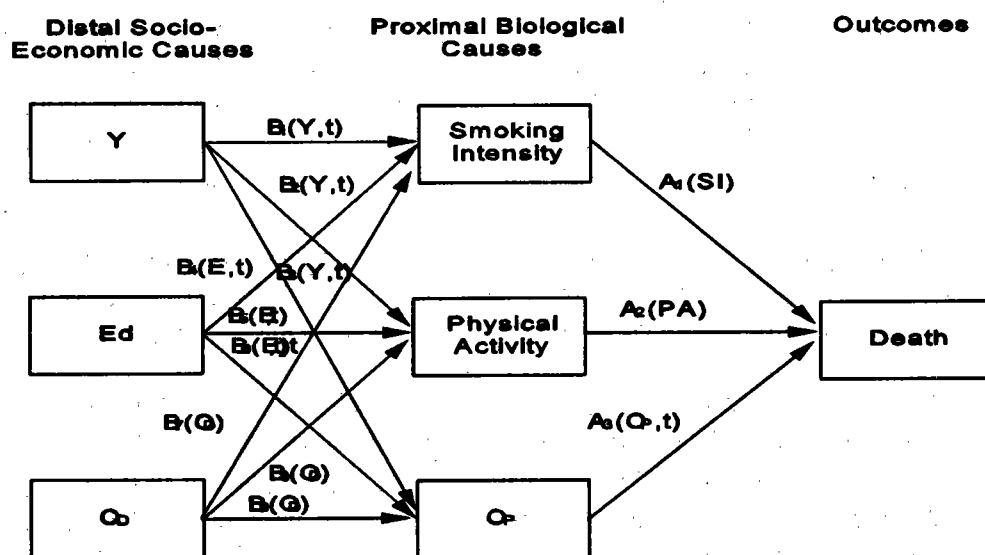


Figure 4

parameter vector and 10% likelihood interval for parameter vectors can be estimated. These set of vectors would be used to generate a range of projections for each set of socio-economic and proximal cause scenarios.

“Carve out” models for specific diseases would be required--especially HIV and TB. The advantage is that knowledge not captured in trends or adequately reflected in the data could be incorporated into these models (e.g. war). However, there is a danger of re-creating advocacy positions around these diseases by detaching them from the constraints of the internally consistent, comprehensive models.

One of the acknowledged weaknesses of the original GBD projection model is the fixed nature of the relationship imposed on the YLD-YLL relationship. There clearly needs to be some modeling of the relationship between YLDs and YLLs for causes where the YLD/YLL ratio is < 10 . Cause specific analyses are needed to estimate trends where $YLD/YLL > 10$. Finally, some incorporation of the variability in access to efficacious medical care needs to be included, and better consideration of how such treatment can change the disability weight associated with certain conditions.

Discussion following Dr. Murray’s presentation:

Sam Preston, the first designated discussant raised the following issues:

Dynamics of social inequality and economic disparity (including: material deprivation, social hierarchy, social disruption) appear to be ignored in the models. However, social inequality may not be a powerful determinant in trends. For example, the decrease in mortality in the U.S. during this century has been very constant despite great variability in the socio-economic conditions.

Many of the models looking at socio-economic issues use income, whereas wealth may be a more important indicator.

Fertility as a determinant of health trends is not clearly considered.

Michael Wolfson, the second designated discussant made the following comments:

Functional form of the tobacco and death relationship is not well understood:

(a) we know that smoking is generally estimated as a contemporaneous flow rate risk factor for CHD, while a lagged cumulative exposure is the best specification for lung cancer, and;

(b) other RFs like education are missing from major studies like Framingham. Thus, to the extent that the relationship is understood, it has different mathematical forms for different diseases.

In turn, this means that any aggregate relationship, such as that in the GBD analysis, must suffer from specification error.

Essentially, and for a host of good reasons, the GBD analysis is aggregate level. However, there is more than enough epidemiological and other analysis to indicate that well-understood micro-level relationships, like the LC and CHD risk functions just noted, are often not well

represented by simple aggregate functional forms. It was therefore suggested that an important methodological endeavor would be to examine "when is aggregation bias safely ignorable". For example, one could take a disease cluster like diabetes and CHD where there was considerable discussion, write down something closer to a "true" model at the micro or individual level using, say, one country's data, generate a representative sample of complete individual lifepaths using microsimulation methods, then aggregate the data into the form found in most countries, estimate a set of aggregate relationships using the GBD style of functional forms, and then test to see whether there is a serious resulting bias or not. This exercise could also be used to test other kinds of "specification bias" such as the effects of omitted variables (e.g. education, or another more conventional risk factor -- where there is good evidence that the RF is significant, but there are no aggregate level data available for most countries), and the effects of choice of functional form (e.g. ignoring the influence of co-morbidity and other competing risks).

Other comments made in the open discussion included:

The models should be simple and cause-specific models tend to explain trends more poorly than aggregate models. The cause specific models are usually more pessimistic.

The models need to include factors which the responsible authorities (i.e. health ministers) can influence to assess the potential "pay-off" from various and often competing interventions.

Murray's response to some of these issues were:

Inequality is a major focus of much of the current GBD work. Especially in the U.S. It is clear that the distribution, not just the absolute level of income, does matter for overall mortality. However, when viewed from a cause-specific perspective it is much harder to detect effect of income inequality. For homicide in the U.S. it is the major determinant. However, the major determinants for most other causes appear to be age, tobacco, HIV and childhood infectious disease mortality.

Models which are not specifically oriented towards policy issues still get a great deal of attention from policy makers. Recognition of overall trends in diseases leads to other models focused on interventions to change the projected course of events.

Time as an indicator of technology change and health care is clearly inadequate and needs more work.

The tobacco variable used in the GBD projections was a cumulative measure based on lung cancer mortality, and fits the data much better than cross sectional surveys which do not account for two major dynamics:

- 1) Changes in the amount of tobacco in cigarettes over time (generally decreasing).

- 2) In areas with large decreases in tobacco use 'past users' generally begin to report themselves as 'never users', hence contaminating the "unexposed" group.

Dr. Manton was scheduled to present the next session. This was to be an overview of various modeling methods. In his absence Mr. Dowd, Dr Gunning-Schepers , Dr. Wolfson and Dr Barendregt presented the conceptual basis and discussed computational and/or policy issues of specific models :

Presentation by Ed Dowd:

There are a number of issues which must be considered when judging the appropriateness of models for projecting health related conditions:

1. The purpose of the projection:

- a) Prediction of stable systems (general population mortality or population size over short time periods)
- b) Prediction of unstable systems (cause-specific morbidity and/or mortality, or highly aggregated groupings over longer time frames)
- c) Modeling of specific short or long term health outcomes which are expected to result from effective interventions (e.g. smoking cessation programs on C.D. mortality in Brazil)

2. What type of data are available:

- a) Complex projection models can be derived which are able to address most of the requirements outlined in 1 above if highly desegregated data are available. These data must include information on morbidity, disability and mortality from an efficient vital statistics system. Models of adequate complexity also require supplementary information from nationally representative surveys on epidemiologic, economic and demographic variables.
- b) Some but not all of the data described in *a*).
- c) None of the data described in *a*).

The big question is how much of the data available from countries with all the data sources described in 2 a can be utilized in situations where only 2b and 2c obtain.

There are 3 general types of projection models:

- 1) Empirical models where matrix multiplications of age, period and/or cohort rates are used to forecast health rates. These may be complex in detail but do not contain direct information on physiological or behavioral processes.
- 2) Empirical models with risk factors which include risk factor values as predictive variables. However, the risk factors are treated in a static sense with no direct indication of how changes in the risk factor over time will influence disease outcomes.
- 3) State-space models are characterized by the mathematical integration of the stochastic dynamics of risk factors jointly with other variables. This implies the modeling of longitudinal data sets containing a significant number of the state space variables in order to examine dependencies.

The Stochastic Process Model (SPM):

SPM consists of a system of two inter-related sets of equations. The first equation tracks risk factor changes over time as a first order autoregressive equation. The equation includes the value of an individual's risk factor at time t as a function of his/her risk factor values and age at time $(t-1)$, and a component which represents the attenuation of the regression relationship over time and with age. The second equation represents the risk(hazard)of mortality at time t as a quadratic function of the individual's risk factors at time t (providing the individual has survived to time t).

The quadratic function is modified by a Gompertz coefficient which represents the portion of the hazard not accounted for by the observed risk factors. The hazard function determines the survivorship of an individual of age A in the tracking equations. The predominant source of the data for modeling individual risk factor distributions and their trajectories and for fitting the quadratic risk equations is based on the Framingham cohort study. Adjustments to other national situations are then incorporated into the two equations for adaptation to individual country risk factor and cause-specific morbidity and mortality levels. The risk factor relationships are maintained based on mean levels and distributions of the risk factor in the country under study. The quadratic risk equations are fitted to a Gompertz function estimated from the specific country cause-specific mortality data.

When this modeling procedure was applied to China the results for population projections were essentially identical to GBD and United Nations estimates.

Discussion following Ed Dowd's presentation:

The presented model concentrated on population projections as the major outcome. This does not seem to be the best indicator for evaluating health.

Dowd agreed but stated that when the SPM and GBD projections were applied to all-cause or cause-specific mortality, age- and sex- specific projections were often quite different even when projections for both sexes and all ages were similar.

Presentation by Louise Gunning-Schepers:

Projections have been a large part of the health policy development in Holland. The Dutch approach has been to use different models depending on the questions being asked (e.g. area vs. disease specific). It is important in this context that the model let people think they understand what the model is doing (i.e. high face validity). Models also clearly present the impact of demographic changes which can be difficult to explain in words.

One barrier to the development of comprehensive and coherent modeling approaches has been that such a process requires linking a variety of research traditions and disciplines. Overcoming differing terminology and existing mistrust between these disciplines can be difficult. Also, linking the costs of health care to demographic changes in ways that policy makers can understand is difficult. For example, the growth in aggregate costs with the increase in the overall population which has been followed by an increase in costs secondary to aging despite a stable population size.

The issues which have been addressed by the Dutch include:

- 1) Relationships in population health
- 2) Two major types of research:
 - a) Etiologic research
 - stroke mortality reduction in relationship to ischemic heart disease case fatality
 - the relationship of social inequality and smoking with health.
 - b) Evaluation research
 - cost-effectiveness of technologies with complex cause structures.
- 3) Assess the dynamic of time lags
 - effects of risk factor exposure by period and birth cohort

Policy issues which have been addressed include:

- 1) Future health status and projected Burden of Disease
- 2) Future health care needs for services, training and budget growth
- 3) Optimizing policy interventions in terms of health benefits
 - * Causal Webs
 - * Interaction between diseases
 - * Time dimensions

Policy relevant variables of population health are:

- * Demography - absolute and time related effects
- * Risk factor exposure - there are options to explore cross-sectional versus cohort approaches
- * Potential interventions - preventive versus curative
- * Disease specific outcomes - this restores epidemiology as the backbone of public health
- * Choices of weights
 - Health status
 - Time preferences (discounting)
 - Age preferences

Presentation by Michael Wolfson :

Dr. Wolfson's presentation involved a description of the modeling approach used in the Health Analysis and Modeling Group(HAMG) at Statistics Canada, called POHEM -- short for POpulation HEalth Model.

POHEM is a micro- (people level) simulation model which creates representative but synthetic human populations at birth and, through the use of computer simulation techniques, ages them while exposing them to risk factors and diseases, as well as other major socio-economic life events.

By way of introduction, Dr. Wolfson noted that POHEM has been developed by an analytical group in Statistics Canada with longstanding experience in constructing and disseminating for general use microsimulation models, e.g. for income taxes, public pensions, and post-secondary student loans.

The major objectives of POHEM include:

- 1) providing a coherent framework for an overall system of health statistics, where things "added up", where there was a more substantial focus on health outcomes in comparison to the current predominance of data on the inputs and throughputs to health care, and where there were explicit micro-foundations so that distributional ; and
- 2) providing a capacity to answer "what if" questions relating to the burdens of disease and risk factors, both proximal and distal -- and including burdens both in terms of costs of treatment and care, and in terms of health status.

The content of POHEM includes detailed, often quite multivariate,

statistical descriptions of transition dynamics (e.g. hazard functions) for:

A) Socio-demographic characteristics over the life course such as nuptiality (e.g. marriage and divorce), fertility, labour force participation, and income;

B) Risk factors, including cholesterol, cigarettes, body mass index, blood pressure, radon, bone mineral density, HRT; and

C) The "natural histories" of major disease processes, including CHD (built in part on Weinstein's model), lung and breast cancer (drawing, e.g., on Canada's extensive cancer registry data and selective chart review studies), arthritis, and osteoporosis (including hip fracture sequelae).

POHEM has been written in a new software environment called ModGen (which is available to other researchers), a C++ pre-compiler designed to facilitate writing agent-based continuous time, discrete-event monte carlo microsimulation models. It is accompanied by a program called Bio-browse for exploring individual life course trajectories generated by POHEM (as well as Statistics Canada's other longitudinal dynamic models).

An example given of a POHEM simulation result was the effects of smoking on the ages at onset and rates of mortality from lung cancer and CHD.

Concerns noted by Dr. Wolfson with POHEM included the complexity and computer time required for model simulations (many hours on a PC), and the challenges posed to analysts when there are complex sets of causal pathways. POHEM offers a unique opportunity to address these challenges because, for example, competing risks can be modelled explicitly (particularly in continuous rather than discrete time). For example, Dr. Wolfson's unit is working with data from the Alameda county cohort study to disaggregate the interactions and subgroup effects of risk factors, and develop models that represent more fully the complex "web of causality".

Discussion concerning Dr. Wolfson's presentation:

Question: Has POHEM been validated to data from the past?

Answer: There has been limited validation since POHEM has been run as a period model. Incorporating data to develop cohort effects involves

much more data from the past, which are generally unavailable. Work is substantially advanced to transform POHEM into an overlapping cohort model. This will allow explicit validation.

Question: Should very proximal factors (e.g. atherosclerotic plaques) be included in the model.

Answer: Such modeling exercises are limited by: a) the availability of accurate data, b) determining a "stopping rule" to avoid ridiculous reductionism.

One topic of discussion after Dr. Wolfson's presentation revolved around a slide he showed which demonstrated a relatively small increase in life-expectancy with the elimination of smoking (1 or 2 years, considering only effects via lung cancer and CHD). Dr. Murray felt that this may have resulted from a bias mentioned earlier in the conference where over time past smokers have become more likely to identify themselves as non-smoking - thus biasing the effect of current and past smokers towards the null since actual smokers are being included in the baseline, never smoker, group. However, Dr. Wolfson pointed out that the model was mutually consistent with actual distributions of smoking from recent population surveys, mortality rates by cause, and risk functions from Framingham and the lung cancer literature.

Presentation by J. Barendregt:

Initial work done with Louise Gunning-Schepers resulted in the dynamic, multiple risk factor, multiple causes of death model called *Prevent*. Subsequently, another project was initiated to develop a model which was dubbed *NIMPH*, for Netherlands Integrated Model of Public Health. Like the *Prevent* model *NIMPH* is dynamic and handles multiple risk factors. The *NIMPH* has additional features, the most important of which is morbidity. *NIMPH* describes disease processes explicitly from incidence through several stages of varying severity and on to cure, death, or death from other causes than the disease of interest. In addition to multiple, multi-stage disease, *NIMPH* models:

- A) Heterogeneity of disease risk and survival
- B) Mortality selection
- C) Comorbidity

A number of diseases are modeled, including:

- 1) Heart disease - borrowing from the Weinstein model, but adding heart failure
- 2) Stroke
- 3) Specific cancers including: Lung, breast, prostate and colorectal.

These disease models can run in single disease mode, with just mortality from all other causes, or can run simultaneously, together describing a considerable part of morbidity and mortality. In the latter mode in particular the *NIMPH* model is very complex, with input to match. In fact, we increasingly realized, the model is too complex. It defies description in the context of an article, and therefore tends to remain a black box. At the same time it is often too simplistic to answer interesting research questions. Like any model *NIMPH* makes simplifying assumptions, but since it was developed as a general purpose model the assumptions may conflict with a particular research question.

The above considerations resulted in a changed in emphasis for the modelers. They no longer pursued the development of a general model of public health. They now put the research question first. They then develop a model that is fit to answer it, and these are often simple models, like multi-state life tables. The methodology developed for *NIMPH* is used to enhance the original *Prevent* model to create *Prevent Plus*. *Prevent Plus* adds morbidity to *Prevent*, like *NIMPH*, but uses one stage disease models and ignores heterogeneity and mortality selection. This makes it much simpler than *NIMPH*, and better fit to illustrate issues on the general level of public health.

Discussion concerning J. Barendregt's presentation:

In response to a question about whether Barendregt and his colleagues routinely fit models to the past he suggested that the effort of adding variables to be able to meticulously "fit" the past was not worth the effort in the context of the overall projection results. Also, most of the models Barendregt works with deal with specific diseases one at a time. When considering a "menu" of potential health interventions which may affect a variety of specific diseases, as well as overall mortality, it would be very difficult to include all of these interventions in an overall mortality model. Finally, the Murray model leaves out the disease states which lead to deaths. This multi-stage approach is maintained in Barendregt's approach.

Three principles of modeling were then stated by Barendregt:

- 1) Use more information only when needed.
- 2) Be even handed.
- 3) Do not take forever.

On the issue of including comorbidities in his models Barendregt made the following points:

Knowledge about the dynamics of comorbidity (i.e. how diseases overlap and interact) is very limited and the complexity of the models would make them very difficult to estimate. Considering comorbidities raises a number of issues:

- 1) When should comorbid conditions be treated as risk factors (e.g. Diabetes)?
- 2) Many diseases share common pathologic pathways.
- 3) Some diseases affect the cost of treating other diseases in the same person, but knowledge about this is piecemeal.
- 4) Some persons have higher susceptibility to independent diseases because of social status, occupation, etc.

A major argument against the inclusion of comorbidities is that experience suggests when the calculations are performed at the population level the overall impact is very small on major results, but consideration of comorbidity may have an influence on individual diseases.

A criticism which was levied against Murray's approach was that choices surrounding the factors to include in models depend on the policies being addressed. Murray's model does not include incidence, but anchors on the relationship between "risk factors" and mortality. Without incidence in the model it is impossible to differentiate the discrete effects of prevention and treatment.

Finally, Dr. Gunning noted that models assume smoothness. However, some events never fit smooth systems and are completely anomalous (e.g. HIV). "Carve-outs" for these events are essential and distinguish a "model" from the "modeling program."

Presentation by T. Buettner:

Dr. Buettner presented projections of the world, and world regional populations. The anchor for these United Nations (U.N.) projections is the relationship between life expectancy and fertility. As life expectancy increases fertility decreases. When this relationship is plotted on a scattergram with countries as the units of measure, one outlier is Saudi Arabia which has a relatively high life expectancy with high fertility as well.

The models used by the U.N. for population projections are very general and do not examine detailed events. However, recent revisions of the world population estimates (these are revised every two years) project a universal shift in the population distribution to older age groups. The oldest age groups are actually growing the fastest. One analysis indicated that currently in Asia

persons ages 85-89 actually outnumber persons 80-84. This has prompted the U.N. to move the upper bound on groups for which population estimates are performed beyond 75 years.

The basic process for forecasting involves:

- 1) Establishing estimates for age specific mortality, fertility and migration at baseline. These are based on life tables.
 - Recent life tables have been projected to include 5 year age groups to 100 years based on systems developed by Sam Preston at the University of Pennsylvania.
 - Patterns of mortality are adjusted using the Lee and Carter approach.
- 2) Recent deviations from past tendencies in life-expectancy (actual decreases for men) has required "breaking-out" Eastern Europe from the standard projection models. This is because pathways or trajectories of life expectancy are usually based on past trends. The recent events in Eastern Europe clearly reflect major deviations from the past.

Discussion concerning T. Buettner's presentation:

Most of the discussion revolved around technical issues such as the upper age bound on the life tables and definitions of infant mortality in Russia. The fundamental issue underlying most of this discussion was that there is immense variability in the quality and consistency of demographic data from around the world. These limits on the data put practical constraints on what can be accomplished when producing world-wide population projections.

Presentation by Milton Weinstein:

Dr. Weinstein briefly described the Coronary Heart Disease Model. This is a deterministic model with 3 sectors:

- 1) A population with no heart disease.
 - Based on risk factor profiles and intervention scenarios this group can transition to 2).
- 2) A population which had coronary heart disease in the last month.
 - This group can transition to death or survival and move to 3).
- 3) A population which had coronary heart disease for greater than one month.

Risk factors included in the model are smoking, LDL cholesterol, HDL cholesterol, body mass index and hypertension. The relationships between risk factors and outcomes are derived from the Framingham data.

The uses of the model have included:

- Predict the effect of changing risk factors.
- Perform cost-effectiveness analyses
- Perform a back validation from 1980 to 1990 to attempt to explain the observed decline in cardiovascular mortality during that period. This analysis suggested that much of the decline in mortality was attributable to treatment rather than prevention. Registry and hospital discharge data suggested that incident disease increased while mortality declined \Rightarrow suggesting a decline in the case fatality rate. This was attributed to improvements in treatment.

Discussion concerning M. Weinstein's presentation:

Several issues were raised about the model's heavy dependence on the Framingham data. The parameter effects associated with this cohort have changed over time. Weinstein acknowledged these limitations, but noted that Framingham has many different variables than most other cohort studies, and only the β s from Framingham are used, not the α s.

Questions were also raised about the inclusion of other risk factors in the model. Weinstein noted that the model has mostly been used to assess the potential impact of these other risk factors (e.g. aspirin, hormone replacement therapy).

A final point was made that in many ways the specification of model parameters appears more important than the mathematical fitting of parameters to data. For example, the specification of the model form in the Coronary Heart Disease Model requires emphasizing the impact of treatment when data sources indicate rising incidence and declining mortality.

Presentation by Hilary King:

Dr. King's presentation focused on recent epidemiologic estimates of the prevalence of diabetes mellitus (DM) in the world. The estimates are based on 25 field surveys from around the world. The age-specific estimates from these surveys were then applied to world population projections. Variables considered in this estimation process were age and urban versus rural.

Prevalence of DM for persons age 20+ years.

| <u>World Region</u> | <u>1995</u> | <u>2025</u> |
|----------------------|-------------|-------------|
| Developed Countries | 6.0% | 7.6% |
| Developing Countries | 3.3% | 4.9% |

Current estimates suggest a higher prevalence in Former Socialist Economies of Europe (8+%) than in the Established Market Economy countries (7%).

The number of people with diabetes will increase in the developed world from 51 to 72 million, and in the developing world from 84 to 228 million. The largest increases will be in Middle Eastern Crescent countries and India. Currently, the three countries with the most people with DM are India, China and the USA.

The male to female ratio of DM in developing countries is ~1. However, it is only ~0.8 in developed countries. The urban to rural ratio is very high in Latin America. When stratified by age the greatest number is in the 45-64 year old age group for the developing countries, whereas the 65+ years group is the largest in developed countries.

Three groups have performed projections: 1) GBD, 2) Australia, and 3) WHO. At the global level, the results for all three groups are similar.

The major problem with these estimates is that they are based on extrapolation from relatively limited survey data. This leads to imprecise estimates for individual countries. The next step will be to link these estimates prevalence estimates of individual complications. In the developed world the major complication is coronary artery heart disease. In Africa the major complications are infection, coma and foot problems. Research on complications in developing countries is definitely needed.

Discussion concerning H. King's presentation:

The major topic raised during the discussion was the effect of intrauterine malnutrition as a risk factor for adult onset DM and coronary artery heart disease in populations where individuals become well nourished as adults (a.k.a. the *Barker hypothesis* since D. J. Barker has been the major proponent). If this hypothesis has merit, then it was proposed that one could suspect that projections of DM based on current prevalence may underestimate the future burden. Dr. King pointed out that if the *Barker hypothesis* is true, then on a population basis the greatest effect of fetal malnourishment on adult rates of DM has probably already been observed. Previous generations of pregnant women, and their fetuses, were generally less nourished than current generations. Therefore, this factor may already be having an effect on well-fed, contemporary adults who were undernourished as fetuses - whereas future generations are protected due to currently superior prenatal supplies of nutrients.

October 23, 1997

Opening discussion chaired by Louise Gunning-Schepers:

The first item for discussion on the opening of the second day of the meeting related to the projection of risk factor distributions and effects. Models based on this approach assume that there are longitudinal studies which have examined the relationship between these influences and health outcomes. A major question is: Can the results from these studies be applied to situations where no data exist? - And - What adjustments need to be made?

The issue which dominated the discussion of this topic was the modeling of the trajectory of risk-factors in populations. Most of the models assume age-specific trajectories based on longitudinal studies such as Framingham, and then apply these rates of change to baseline information available from specific countries. In Europe four countries have performed such analyses and shown that there are different trajectories of risk factors within cohorts. These findings require adjusting the data based on past analyses. This heterogeneity suggests that models which rely on the associations between "distal" determinants and outcomes may be more reliable since there appears to be limited generalizability of information from intensely studied cohorts.

However, there is a very large problem with highly generalizable models which relate "distal" indicators of socio-economic development to health such as was developed for the GBD. These models rely on an assumption of stable case-fatality ratios in order to "back track" from mortality to incidence and the burden of disability. This weakness is particularly acute for prevalent, low mortality conditions (e.g. neuropsychiatric conditions and musculoskeletal).

Final session of meeting:

The final session consisted of each participant providing a series of recommendations for Dr. Murray regarding his anticipated attempts to refine the GBD models:

- 1) In the near future there will be an almost universal epidemiologic transition from a world where health outcomes are dominated by infectious diseases and maternal-infant dynamics to the sequelae associated with an ageing population. In this context there appears to be a growing disparity between the current focus of public health and health ministries and the anticipated problems. A projection model which highlights this disparity by incorporating the influence of health care utilization will be essential.
- 2) In a developed country like the U.S. the progress in medical technology is perceived as a major "driver" of costs. A model which appropriately assesses the *direct* impact of this technology on health outcomes is most relevant in this setting. Continued support for medical technology research is increasingly contingent on the elucidation of the relationship between technology, costs and health outcomes.
- 3) Validation of models should not just involve fitting it to a large "panel" of information. The models can be augmented by using detailed information from a variety of different

countries. For example, the results of using the *POHEM* in Canada can be compared to EME and USA results using the GBD approach.

- 4) Risk factors which may be more relevant to issues surrounding life-style habits which affect health include: anthropometry, blood sugar, family history etc. Much of this data exists in the surveys on DM mentioned in Dr. King's talk and are available for incorporation into any future modeling projects.
- 5) The current GBD model does not adequately distinguish between prediction and "what if" questions. For "what if" modeling a major consideration is face validity. The model has to make sense to people who are experts in the disease being modeled. The biological mechanisms which underly the relationships between social determinants and disease are poorly understood and are frequently disparaged as merely "statistical." This is a criticism of the face validity associated with the model.
- 6) The modeling of the GBD has previously used DALYs as an end-point measure. This measure involves many ethical and social value choices beyond technical issues of measuring incidence, disability and case-fatality which were discussed at this Workshop. Issues surrounding age-weighting, discounting and health-state preference weighting are also important to the ultimate results of the model and need to be addressed in a future workshop composed of a different range of subject-matter experts.
- 7) Developing models for countries where data are sparse will probably have minimal value to policy makers. Therefore, it would be best to concentrate on countries where data are available.
- 8) There is an exponential growth in the number of candidate risk factors for projection models. Wealth versus income, genetics, and growth in nursing home populations in developed countries are all novel issues worthy of modeling. In this context there needs to be greater emphasis on the importance, and potential heterogeneity, of the case-fatality ratios. With the epidemiologic transition the case-fatality ratio undoubtedly improves, with a resultant increase in the prevalence of disability. An assessment of the impact of medical technology on this dynamic is essential as this technology, and its costs, diffuses globally.
- 9) The fundamental issue for any modeling project is the availability of quality data. Proprietary issues frequently limit the availability of important datasets such as MONICA, MRFIT and the Honolulu Heart Study. The creation of a "library" of data for modeling efforts must be supported. This is particularly important as more data are collected on the impact of ageing, and models are developed to assess the relationship between ageing populations and trends in health.

- 10) The gathering of epidemiologic data (e.g. prevalence, duration, case-fatality etc) is the highest priority activity. Weighting for disability has proven much more consistent than the epidemiologic estimates.
- 11) In the context of a general model which includes distal and proximal determinants of disease incidence and mortality, special "carve outs" for specific diseases are inevitable. Such "carve outs" require the use of dynamic population models. Simply multiplying rates of disease by the population projections are inadequate (*Note: A dynamic model was used in the GBD estimation process*).

Adjournment:

The meeting was adjourned by Ed Dowd at 1:00 PM.

Attachment I

PROJECTIONS WORKSHOP FORECASTING MORBIDITY, MORTALITY & DISABILITY IN NONCOMMUNICABLE DISEASES

**Center for Population & Development Studies
Harvard University**

October 22-23, 1997

Agenda

Wednesday, October 22

- 9:30 - 9:45 Short introduction as to aims and goals of the informal workshop. C. J. L. Murray, GBD, Harvard, E. Dowd, WHO, R. Suzman, NIA/NIH
- 9:45 - 10:00 Short self-introduction of each participant.
- 10:00 - 10:45 Presentation by C.J.L.Murray of GBD program of current concerns in forecasting disease burdens, particularly those due to NCD's in a wide variety of country conditions, and in their expectations as to specific outcomes of the workshop, relating to these concerns.
- 10:45 - 11:15 Discussants: S. Preston, M. Wolfson, Open discussion
- 11:15 - 12:00 Presentations by E.Dowd,L.Gunning-Schepers,M.Wolfson, J.Barendregt:
An Overview of Chronic Disease Models in Adult and Elderly Populations
- 12:00 - 12:30 Discussants:C.J.L.Murray, Open discussion.
- 12:30- 13:30 LUNCH

Specific risk factor and specific disease projections issues:

- 13:40 - 14:20 a) simultaneous use of general macro-level and specific micro-level risk factors in projecting single and multiple disease outcomes (co-morbidities and competing mortality risks);
- 14:20 -15:00 b) the appropriate model structure to project disease burden in the absence of micro-level risk factors, e.g., for musculoskeletal diseases, injuries and violence, certain mental disorders:
- 15:00 -15:40 c) the use of micro-simulation projection methodologies in situations where appropriate data or mathematical elaboration is missing to allow for parametric representation of the full structure of the model;
- 15:40-16:10 Coffee

Attachment I

- 16:10 -17: 10 d) the introduction of preventive and clinical components of interventions, their economic benefits and costs and their appropriate health status indicators in forecasting model structure, in order to project the cost-effectiveness of the intervention strategy;
- 17:10 -17:30 e) problems related to diseases at specific sites which may require special modeling considerations, e.g., certain cancer sites, diabetes (as a risk factor for other diseases e.g., CVD, ESRD, blindness)
- 19:00 DINNER

Thursday, October 23

Continuation of open discussion sessions

- 8:30 - 9:10 f) the transfer of risk factor trajectories from populations in which longitudinal studies of disease have been carried out to those where no studies, fragmentary studies or only single or serial cross-sectional studies exist;
- 9:10 - 9:50 g) the imputation of data from other sources where national data is unavailable or insufficient to allow disease specific projections;
- 9:50 - 10:30 h) problems of outcome measures for predominantly non- fatal diseases, i.e., co-morbidity, disability weights and selective mortality from other causes.
- 10:30 - 11:00 Coffee
- 11:00-12.30 Integration of individual item discussions into a set of recommendations for forecasting the outcomes of interventions on risk factors for noncommunicable diseases.
- 13.00 Closure of the Workshop

Attachment II

LIST OF PARTICIPANTS

Dr J. Barendregt
Unit of Public Health
Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam
Netherlands

Tel: +31-10 408-77-14
E-mail: barendregt@mgz.fgz.eur.nl
Fax: +31-10 436-68-31

Dr J.-M. Berthelot
Health Analysis and Modelling Group
Social and Economic Studies
Statistics Canada, RCH-11A
Ottawa, Canada

E-mail: berthel@statcan.ca
Tel: 001-613 951-37-60
Fax: 001-613 951-56-43

Mr T. Buettner
Population Estimates and Projections Section
Population Division/DESPA
United Nations
New York, N.Y., 10017
USA

E-mail: buettner@un.org
Tel: 001-212 963-32-09
Fax: 001-212 963-21-47

Dr D. Cutler
Harvard University
Department of Economics
Littauer Center
Cambridge, MA 02138
USA

Tel: 001-617 868-39-00 ext.366
E-mail: dcutler@nber.org
Fax: 001-617 868-27-42

Mr J. Dowd
Scientist, Ageing and Health Programme
World Health Organization
Avenue Appia, 20
1211 Geneva 27
Switzerland

Tel: +41-22 791-34-85
E-mail: dowdj@who.ch
Fax: +41-22 791-48-39

Dr L. Gunning-Schepers
University of Amsterdam
Institute of Social Medicine
Academic Medical Centre
Meibergdreef 15
1105 AZ Amsterdam
Netherlands

Tel: +31-20 566-48-92
E-mail: l.j.gunning@amc.uva.nl
Fax: +31-20 697-23-16

Dr W. Harlan
Associate Director for Disease Prevention
National Institutes of Health
Building 1, Room 260
1 Center Drive
MSC 01 74
Bethesda, MD 20892-0174
USA

E-mail: wh27v@nih.gov
Fax: 001-301 402-25-17

Dr A.E.M. de Hollander
Department of Chronic Diseases &
Environmental Epidemiology
National Institute of Public Health & Environment
P.O. Box 1
3720 BA Bilthoven
Netherlands

Tel: +31-30 274-32-22 or 38-04
E-mail: AEM.de.Hollander@rivm.nl
Fax: +31-30 274-44-07

Mr P.D. Johnson
Chief, Information Resources Branch
International Programs Center
Population Division
US Bureau of the Census
Washington D.C. 20233-8860
USA

Tel: 001-301 457-14-03
E-mail: peterj@census.gov
Fax: 001-301 457-15-39

Dr H. King
Diabetes Mellitus Programme
World Health Organization
Avenue Appia, 20
1211 Geneva 27
Switzerland

Tel: +41-22 791-34-72
E-mail: kingh@who.ch
Fax: +41-22 791-07-46

Dr M. McKenna
Medical Epidemiologist
Center for Chronic Diseases, Prevention
and Health Promotion
4770 Budford Highway MSK 45
Atlanta, GA
USA

Email: mtm1@cdc.gov
Fax: 001-770 488-59-64

Dr C. Murray
Burden of Disease Unit
Harvard School of Public Health
9 Bow Street
Cambridge, MA 02138
USA

Tel: 001-617 495-84-98
E-mail: cmurray@hsph.harvard.edu
Fax: 001-617 496-32-27

Dr S. Preston
Population Studies Center
University of Pennsylvania
Philadelphia
USA

E-mail: spreston@pop.upenn.edu
Fax: 001-215 898-21-24

Dr J.A. Schuttinga
Economist
OESP/OD, National Institutes of Health
6000 Executive Boulevard, Suite 608
Rockville, MD 20892-7010
USA

Tel: 001-301 496-22-29
E-mail: js4lz@nih.gov
Fax: 001-301 480-92-86

Dr E.J. Sondik, Ph.D.
Director
National Centre for Health Statistics
Presidential Building, Room 1140
6525 Belcrest Road
Hyattsville, MD 20782
USA

Tel: 001-301 436-70-16
Fax: 001-301 436-52-02

Dr R. Suzman
Director, Office of Demograph of Aging
Behavioral & Social Research Program
NIA/NIH
7201 Wisconsin Avenue
Gateway Building, Room 2C234
Bethesda, MD 20892
USA

E-mail: suzmanr@gw.nia.nih.gov
Fax: 001-301 402-00-51

Dr M. Weinstein
Department of Health Policy and Management
Harvard School of Public Health
718 Huntington Avenue
Boston, MA 02115

Tel: 001-617 432-08-05
E-mail: mcw@hsph.harvard.edu

USA

Fax: 001-617 432-01-90

Dr D. Wise
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138-5398
USA

Tel: 001-617 868-39-00
E-mail: dwise@nber.org
Fax: 001-617 868-27-42

Dr M. Wolfson
Canadian Institute for Advanced Research
Statistics Canada
Social and Economic Studies, RHC-24A
K1A 0T6 Ottawa
Canada

Tel: 001-613 951-82-16
E-mail: wolfson@statcan.ca
Fax: 001-613 951-56-43