Education and debate

Climate change and health: implications for research, monitoring, and policy

Andrew Haines, Anthony J McMichael

Joint Department of Primary Care and Population Sciences, Royal Free Hospital School of Medicine and University College London Medical School, London NW3 2PF

Andrew Haines, *professor of primary health care*

Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Anthony J McMichael, *professor of epidemiology*

Correspondence to: Professor Haines a.haines@ucl.ac.uk

BMJ 1997;315:870–4

The potential effects of climate change on human health are summarised in the previous article.[1] These are wide ranging and pose considerable challenges to both researchers and policy makers.[2] The complexity of these environmental processes and their impact on health necessitates a multidisciplinary approach.

In environmental health there is a close relation between epidemiological research and those monitoring activities which seek evidence of changes in the environmental or health status of populations. A distinction is usually made between monitoring and surveillance, the latter being the continuing standardised recording of the occurrence of disease. In the context of climate change and health, however, both monitoring and surveillance are needed to *(a)* identify important changes in disease incidence, health risk indicators, or health status; *(b)* determine whether these changes are likely to be the result of local, regional, or global environmental changes; *(c)* to help develop countermeasures and assess their effectiveness; and *(d)* to develop hypotheses about the potential health effects of climate change. Monitoring should also help in the detection of unexpected events.[3]

Summary points

Health risk assessment based on scenarios encompasses extrapolation, integrated mathematical modelling, and generalised assessments of the consequences of demographic, social, and economic disruption

Uncertainty is unavoidable because of such factors as the unpredictability of future industrial activities, and differences in sensitivity of disease systems and vulnerability of populations to climate change

Monitoring of health indicators and disease surveillance activities must be integrated with global observing systems currently being developed for climate change and its impact

The potentially serious effects of climate change on health heighten the urgent need for policies to limit greenhouse gas emissions

The research challenge

Since there is uncertainty about the profile and rate of future climate change it is necessary to estimate effects on health in relation to specified probable climate scenarios. This process differs in several important ways from the more familiar empirical procedure of quantitative risk assessment.[4] The latter is usually conducted in relation to some existing index of environmental exposure for which there is prior empirical evidence of direct (usually toxicological) health risks across an exposure range which includes the index exposure.

The three main approaches to health risk assessment based on scenarios are extrapolation based on specific (historical) analogue situations for some aspects of climate change; formal integrated mathematical modelling; and generalised assessments drawing on expert judgment of the range of health consequences (physical, microbiological, and psychological) of diffuse and complex demographic, social, and economic disruption.

Historical analogues

Historical analogues probably come from recent times, although earlier documented experiences may also be informative. Most useful are those situations which seem to simulate aspects of future climate change.[5] For example, epidemiologists have begun to study the regional health consequences of the worldwide climatic fluctuations associated by "teleconnection" (remote linkages) with the El Niño southern oscillation (box).

These studies can be useful in assessing the vulnerability of populations to climate change, although the relatively short time scale makes direct extrapolation to the effects of global warming on health difficult.

Integrated mathematical modelling

Integrated mathematical modelling is increasingly being used to estimate the future impact on health of climatic (and other environmental) change. It requires that each component of the sequence of climatic, environmental, and social changes in the chain of causation should be represented mathematically.[4]

Modelling skin cancer rates

A recent illustration of integrated mathematical modelling comes from a related topic—stratospheric ozone depletion, increased ultraviolet irradiation, and the impact on the incidence of skin cancer in fair skinned populations.[12] The study modelled the excess rate of skin cancer in Europe and the United States during the coming century in response to three contrasting ozone depletion scenarios: (*a*) a "business as usual" scenario of gaseous emissions destroying ozone; (*b*) one incorporating the emission restrictions required by the protocol drawn up the Montreal meeting on substances that deplete the ozone layer in 1987; and (*c*) a scenario taking account of the Montreal protocol as amended in Copenhagen in 1992.

The study used an integrated model that combined metamodels of the components of the chain linking source and risk—that is, emissions, ozone depletion, changes in ultraviolet irradiance, and dose-effect (ultraviolet-cancer) models. The uncertainties in each component were included in the model. A multiple run Monte Carlo technique was used to allow for variable combinations of these uncertainties and estimated future trajectories of skin cancer incidence rates were produced. The central estimate for the third scenario was that the skin cancer incidence would peak at an increment of approximately 10% around the middle of next century. For the other two scenarios, the cancer increments were orders of magnitude larger.

Geography of malaria

The best known mathematical modelling in relation to the health impacts of climate change has been in relation to potential shifts in the geographical range of malaria.[10-15] These first generation models, which are highly aggregated, have incorporated climate change scenarios (equations that express the average relations of mosquito and parasite biology to temperature, rainfall, and humidity[16]) and information about pre-existing levels of malaria and acquired immunity in populations in different regions. (Figure 3 in our previous paper shows the results of this model.[1])

Dealing with uncertainty

As with all forecasting, assessing the impacts of global climate change entails unavoidable uncertainties.[4 17] These uncertainties arise from the intrinsic unknown element in future trends in human industrial, demographic, and trading behaviour; from the nature of the non-linear and interactive relations within the various complex natural systems; and from the variable (and population specific) sensitivity of the health outcome to the change in climate and environment. Uncertainty also arises from the stochastic nature of the biophysical systems being modelled.

Differences in vulnerability between populations are another source of variability.[18] These occur because of the heterogeneity and changeability of human culture, social relations, and behaviour. As Balbus and Patz state: "While a given disease system may be particularly *sensitive* to the effects of climate change based on biological or physiological characteristics, the ultimate *vulnerability* of a given population to that disease may be considerably lessened by adaptive responses."[19] Some populations and geographical regions will be particularly vulnerable. For example, populations whose food supplies are insecure are vulnerable to downturns in agricultural productivity caused by climatic factors, and people living on the edge of regions where infectious diseases borne by vector organisms are endemic are

El Niño events—a partial analogue for future climate change?

- The El Niño southern oscillation is a large, irregular, unstable atmosphere-ocean system which produces relatively short-term climate changes over the Pacific region[6]
- Events related to the El Niño southern oscillation (that is, El Niño warm events and La Niña cold events) strongly influence climate variability between years and are associated with regional land and sea surface warming, changes in precipitation and in the occurrence of tropical cyclones
- These anomalies impinge primarily on countries bordering the Pacific and Indian Oceans but also affect other continents
- El Niño events can affect human health—epidemics of malaria and dengue fever are more likely to occur in the year of an El Niño event or in the year following[7-9]; the occurrence and distribution of harmful coastal algal blooms is also associated with El Niño events[10]
- Weather disasters are twice as frequent worldwide during the year of an El Niño event[11]

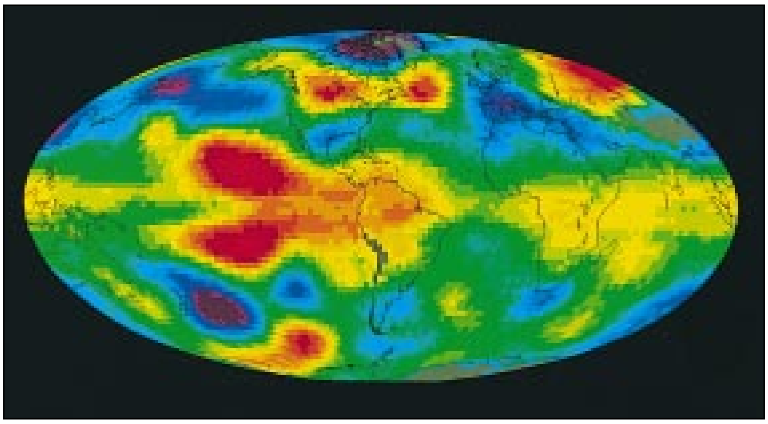
most likely to experience the early extensions in range of these diseases.

Another dimension of complexity in the assessment task results from the interplay of several environmental stresses that are coexistent. Interaction between local environmental degradation and changes on a larger scale—climate change, population growth, and loss of biodiversity—may significantly influence the effects on health. For example, local deforestation caused by increased population pressure may directly change the distribution of vector borne diseases while also causing a local increase in temperature (in addition to its contribution to a global temperature increase by depleting one of the biosphere's great carbon dioxide "sinks").[20]

Major research needs

Important research needs include the following[2]:

- Improvements in mathematical models for predicting the impact of climate change on health, including higher resolution to enable local and regional impact assessments to be made;
- Further studies to (*a*) distinguish more clearly between the effects on health of climate and of air pollution; (*b*) determine the extent to which, in different regions, a reduction in mortality related to cold might offset the impact of more frequent heatwaves; and



NASA GSFC/SCIENCE PHOTO LIBRARY

An El Niño event causes temperature anomalies in the troposphere. In this false-colour satellite image of the 1983 event, red areas correspond to regions with temperature higher than the long term average; blue areas indicate regions where temperatures are lower than the long term average

(*c*) assess the longer term health effects, if any, on populations living in locations with different climates;
- Analysis of infectious disease epidemics associated with recent regional changes in climate, using these as analogues of future climate change. For example, a systematic examination of vector borne outbreaks in regions affected by climatic events related to the El Niño southern oscillation would improve our understanding of the relations between climate and health;
- For vector borne diseases, there is a need for basic laboratory and field investigations of arthropod vector ecology and pathogen infectivity at raised temperatures and varying humidity and ecological studies on the climate sensitivity of diseases in locations at the margins of endemic areas;
- Assessment of how changes in food production—as a result of climate and weather changes, increased ultraviolet irradiation, sea level rise, changes in pest ecology, and socioeconomic shifts in land use practices—could affect human health and nutrition;
- Study of the association of extreme climatic events with global warming and the occurrence of disasters affecting large human populations;
- Modelling studies of the potential public health implications of forced migration from climatically vulnerable regions;
- Ecological studies of the range of possible public health impacts of reductions in biodiversity related to the climate;
- Assessment of the potential health impacts of strategies to mitigate greenhouse gas emissions (for example, the health risks of biomass fuels).

Monitoring for changes in health related indices

Global observation systems

The monitoring of health effects should be integrated with global observation systems that are currently under development. The Global Climate Observing System is a joint initiative of the World Meteorological Organisation and other international agencies and will encourage the development of coordinated climate observations by national and international organisations. Its coverage will exceed that of current monitoring programmes, such as Global Atmosphere Watch and World Weather Watch, which comprise a network of satellites, telecommunications and data processing facilities. The Global Ocean Observing System, operated by the Intergovernmental Oceanographic Commission of Unesco, includes monitoring of sea level rise, sea surface temperature and, eventually, biological measures such as the phytoplankton concentration. The Global Terrestrial Observing System is being established under the auspices of the United Nations Environment Programme and other international agencies. It will be used to detect and monitor response of terrestrial ecosystems to global change including new patterns of land use and climate change.[21]

Summary of methods needed to monitor the impact of climate change on health

What	Where	How
Heat stress	Urban centres in developed and developing countries	Daily mortality and morbidity data
Changes in seasonal disease patterns (eg asthma and allergies)	"Sentinel" populations at different levels	Morbidity data from primary care, hospital admissions, emergency room attendance
Natural disasters	All regions	Mortality and morbidity data
Effects on health of rise in sea level	Low lying regions	Local population surveillance
Freshwater supply	"Critical" regions especially in the interior of continents	Measures of run off, irrigation patterns, pollutant concentrations
Food supply	Critical regions	Remote sensing; measures of crop yield, food access, nutrition (from local surveys). Agricultural pest and disease surveillance
Emerging diseases	Areas of population movement or ecological change	Identification of outbreaks of "new" syndromes or diseases, population based time series, laboratory characterisation
Vector borne diseases	Margins of distribution (latitude and altitude)	Primary care data, local field surveys, communicable surveillance disease centres, remote sensing
Marine ecosystems	Oceans	Remote sensing, sampling of biotoxins, phytoplankton, essential nutrients. Epidemiology of cholera, other vibrios, shellfish and fish poisoning

Source: Haines et al[28]

Linkage between climate monitoring and health indicators

There is also a need to link the proposed monitoring of climate change and associated health indicators with the health related monitoring activities currently undertaken or sponsored by the World Health Organization and the United Nations Environment Programme. A potentially important example on the Internet is the programme for monitoring emerging diseases (proMED–http://www.healthnet.org/programs/promed.html) conducted under the auspices of the Federation of American Scientists and sponsored by the WHO. It facilitates worldwide electronic exchange of data on new, resurgent, or otherwise varying infectious diseases.

Population based data

A range of population based sources of data could be used to determine changes in potentially relevant conditions over time. For example, in Britain the weekly information return service of the Royal College of General Practitioners permits the analysis of consultation data in relation to person, episode, and doctor-patient encounter.[22] In Bangladesh, a population of around 200 000 in the Chandpur district has formed the basis of many large nutritional and epidemiological studies, showing the feasibility of long term monitoring of population health in a country that is particularly vulnerable to flooding and other extreme weather events.[23]

The Global Health Network (http://www.pitt.edu/HOME/GHNet/GHNet.html) is under development and is designed to promote connections between public health workers worldwide. It should lead to major advances in the telemonitoring of health and disease, and could incorporate methods to compensate for undercounting cases, using the capture-recapture approach developed originally by wildlife biologists to enumerate animal populations.[24]

Ecosystem monitoring

It has been increasingly recognised that ecosystems have important influences on human health—for example, through changes in key indicator species such as insects and rodents which may have both direct and indirect effects. Algal blooms in marine ecosystems can act as reservoirs for certain pathogens including

General and specific policies to reduce climate change or its impacts

General strategies (see text)

- Universal access to family planning
- Development of economic measures to reduce greenhouse gas emissions
- Public and professional education
- Improvement of population adaptation to climate change
- Encouragement of sustainable development policies

Options to reduce greenhouse gas emissions*

- Agriculture

Reduced land conversion through improved farming techniques
Improved tillage to reduce fossil fuel combustion
Improved feed use for ruminants to reduce methane emissions
Reduced biomass burning

- Forestry

Reduced deforestation with concurrent improvement in agricultural productivity (tropical forests have maximal potential for sequestering carbon)
Regeneration of degraded lands for reforestation

- Human settlements

Buildings with improved thermal integrity
Condensing furnaces and heat exchangers
Solar water heaters and insulated water storage
Financial incentives for energy conservation
Building codes and utility regulations
Planting shade trees to reduce "heat islands"
More efficient cooking stoves

- Energy supply

More efficient power generation
Natural gas turbines in place of oil or coal
Gasification of fossil fuels before combustion
Combined heat and power production and district heating
Alternative energy sources (solar, wind, geothermal energy, etc)
Coal conversion technology

- Industry

Cogeneration and steam recovery
Efficient lighting and electric motors
Alternative materials (eg, replace concrete with wood)
"Heat cascading" to use energy by-products of industrial processes
Recycling of energy intensive materials

- Transportation

Improved public transport
Facilitation of cycling and walking
Urban traffic control for shorter transit times
Car tuning programmes
Improved fuel efficient engines
Improved energy efficient design of ships and aircraft
Use of ethanol and methanol fuels

*Source: Intergovernmental Panel on Climate Change[31]

Vibrio cholerae.[25] Monitoring indicator species could help our understanding of important links between climate change and its effects on health.

Remote sensing

Remote sensing, particularly by satellites, can be used to monitor a range of variables relevant to climate change, including sea surface temperatures, algal blooms, and changes in terrestrial ecosystems. For example, vegetation indices produced by high resolution radiometry have been correlated with mortality and the population density of tsetse flies.[26] The United States's National Aeronautics and Space Administration is sponsoring research on the use of satellite information to study the distribution and control of vector borne disease.[27] Data from remote sensing may need to be validated by local data on the vector organisms and diseases of interest. The table summarises a framework for the development of monitoring systems for the health impacts of climate change.[28]

Policy implications

The implications of climate change for public policy are wide ranging. Policies to reduce changes are shown in the box. Mitigation options aim to reduce greenhouse gas emissions or to increase carbon dioxide sinks—for example, by promoting reforestation. Some options directly affect health, such as the promotion of bicycling, which would increase fitness and lower cardiovascular risk while helping to reduce carbon dioxide emissions.[29] Renewable energy sources should be assessed for their impact on health since some may have adverse consequences. Hydroelectric dams for example may cause population displacement and social disintegration.[30]

Population growth is an important driving force of climate change. It is estimated that half of the increase in carbon dioxide emission between 1992 and 2022 will be a result of population growth.[32] Although most will occur in developing countries, any growth in developed countries is an important contributor because of the much higher per capita consumption of fossil fuels. Currently only around 1% of international donor aid is spent on family planning, whereas just 2-3% would give worldwide access to contraception.[33] Policies such as these, which meet short term local needs as well as long term environmental goals, should be priorities for implementation.

Tension between the priorities of conventional economics and environmental protection has led to the development of "environmental economics."[34] This attempts to assign a market value to the otherwise uncounted costs of the adverse impact of environmental degradation. "Ecological economics" seeks to incorporate the concept of sustainability and thus to avoid compromising the health and survival of future generations.[35] There is clearly a need for greater public and professional debate about the long term consequences of climate change and the balance between the immediate economic impact of mitigation strategies and their potential to reduce the impact on health and wellbeing in the future.

Agenda 21, the principal outcome of the 1992 UN Conference on Environment and Development held in Rio de Janeiro, indicates that many countries support an integrated approach to reducing poverty and environmental degradation. Richer countries agreed—at least in principle—to increase funding to promote "sustainable development" and to transfer information and energy efficient technology and improve education and training.[36]

The WHO is now seeking to persuade governments that public health considerations are a key criterion in sustainable development.[37] There has already been international cooperation in phasing out compounds that cause stratospheric ozone depletion in the form of the updated Montreal protocol of 1987; this is an example of the use of a precautionary approach which yielded international action despite scientific uncertainty.[38] However, in the case of climate change action has not yet matched the rhetoric. The UN Framework Convention on Climate Change urges developed countries to take the lead in combatting climate change but it will be late 1997 before concrete action on greenhouse gas emissions will be agreed for the period beyond 2000, and most nations are unlikely to reach the initial target of reducing their emissions to 1990 levels by 2000.[39] The recently elected British government has announced its intention to reduce the nation's carbon dioxide emissions in 2010 to a level 20% lower than that in 1990 but clear policies need to be developed and implemented to ensure that substantial and timely reductions occur.

Meanwhile, steps could be taken to improve population adaptation to climate change—but not as alternatives to mitigation. Some, such as eliminating the breeding sites of vector organisms and improving vaccination coverage, are within the capacity of health professionals. Others, such as changes in building design to reduce heat load and improvements in flood protection mechanisms, require policy changes in other sectors.

Finally, global environmental hazards to health should feature in medical school curricula since much of the anticipated impact on health would occur within the coming decades.[40] Meanwhile, in the spirit of primary prevention, health professionals should advocate to policy makers early application of strategies to minimise climate change in order to limit the anticipated impact on health. That impact, mediated through disruption to life supporting biophysical systems, has unprecedented importance for the sustainability of human health.

We received assistance and helpful comments from Sari Kovats.

1 McMichael AJ, Haines A. Global climate change: the potential effects on health.*BMJ* 1997;315:805-9.
2 McMichael AJ, Haines A, Slooff R, Kovats S, eds. *Climate change and human health.* Geneva: World Health Organization, 1996. (WHO/EMG/96.7.)
3 Levins R. Preparing for uncertainty. *Ecosystem Health* 1995;1:47-57.
4 McMichael AJ. Integrated assessment of potential health impact of global environmental change: prospects and limitations. *Environmental Modelling and Assessment* 1997;2:1-9.
5 Patz JA, Balbus JM. Methods for assessing public health vulnerability to global climate change. *Climate Research* 1996;6:113-25.
6 Aceituno P. El Niño, the southern oscillation, and ENSO: confusing names for a complex ocean atmosphere interaction. *Bulletin of the American Meteorological Society* 1992;73:483-5.
7 Bouma MJ, van der Kaay HJ. The El Niño southern oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? *Tropical Medicine and International Health* 1993;1:86-96.
8 Bouma MJ, Sondorp HJ, van der Kaay HJ. Climate change and periodic epidemic malaria. *Lancet* 1994;343:1440.
9 Hales S, Weinstein P, Woodward A. Dengue fever epidemics in the South Pacific: driven by El Niño southern oscillation? *Lancet* 1996;348:1664-5.
10 Patz JA, Epstein PR, Burke TA, Balbus JM. Global climate change and emerging infectious diseases. *JAMA* 1996;275:217-23.
11 Dilley M, Heyman BN. ENSO and disaster: droughts, floods and El Niño/southern oscillation warm events. *Disasters* 1995;19:181-93.
12 Slaper H, Velders GJM, Daniel JS, de Gruijl FR, van der Leun JC. Estimates of ozone depletion and skin cancer incidence to examine the Vienna Convention achievements. *Nature* 1996;384:256-8.
13 Intergovernmental Panel on Climate Change. Climate change 1995. Assessment of impacts and adaptation options: human population health In: Watson RT, Zinyowera MC, Moss RH, eds. *Second assessment report of the Intergovernmental Panel on Climate Change.* New York: Cambridge University Press, 1996:561-86
14 Martens WJM, Niessen LW, Rotmans J, Jetten TH, McMichael AJ. Climate change and vector-borne disease: a global modelling perspective, *Global Environmental Change* 1995;5:195-209.
15 Martin P, Lefebvre M. Malaria and climate: sensitivity of malaria potential transmission to climate. *Ambio* 1995;24:200-7.
16 Gilles HM. Epidemiology of malaria. In: Gilles HM, Warrell DA, eds. *Bruce-Chwatt's essential malariology.* London: Edward Arnold, 1993:124-63
17 Bell W. *Foundations of futures studies: human science for a new era.* Vols I and II. New Brunswick, NJ: Transaction Press, 1997.
18 Woodward AJ, Hales S, Weinstein P. Climate change and human health in the Asia Pacific region: who will be most vulnerable. *Climate Research* (in press).
19 Balbus JM, Patz JA. Human health impacts and adaptation assessment. In: *Guidance for vulnerability and adaptation assessments.* Washington, DC: US Country Studies Program, 1996.
20 Almendares J, Sierra M, Anderson PK, Epstein PR. Health and climate change. Critical regions: a profile of Honduras. *Lancet* 1993;342:1400-2.
21 Head OW, Menaut JC, Steffen WL, eds. *Towards a global terrestrial observing system (GTOS): detecting and monitoring in terrestrial ecosystems.* Paris: Unesco and International Geosphere-Biosphere Programme; Stockholm, Royal Swedish Academy of Sciences, 1993. (Man and Biosphere Digest 14; Global Change Report 26.)
22 Fleming DM, Norbury CA, Crombie TL. *Annual and seasonal variations in the incidence of common diseases.* London: Royal College of General Practitioners, 1991. (Occasional paper No 53.)
23 International Centre for Diarrhoeal Disease Research. *Annotated bibliography of ICDDR studies in Matlab, Bangladesh.*Dhaka: International Centre for Diarrhoeal Disease 1990. (Specialised bibliography series No 14.)
24 La Porte R, Barnes E, Chang YF, Libman I. Global epidemiology and public health in the 21st century. Applications of a new technology. *Ann Epidemiol* 1996;6:162-7.
25 Epstein PR, Ford TE, Colwell RR. Marine ecosystems. *Lancet* 1993;342:1216-9.
26 Rogers DJ, Randolph SE. Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature* 1991;351:739-41.
27 Epstein PR, Rogers DJ, Slooff R. Satellite imaging and vector-borne disease. *Lancet* 1993;341:1404-6.
28 Haines A, Epstein PR, McMichael AJ on behalf of an international panel. Global health watch: monitoring the impacts of global environmental change. *Lancet* 1993;342:1464-9.
29 British Medical Association. *Cycling towards health and safety.* Oxford: Oxford University Press, 1992.
30 McCully P. *Silenced rivers.* London: Zed Books, 1997.
31 Intergovernmental Panel on Climate Change. Climate change 1995: assessment of mitigation options. In: Watson RT, Zinyowera MC, Moss RH, eds. *Second assessment report of the Intergovernmental Panel on Climate Change.* New York: Cambridge University Press, 1996:587-799.
32 Bongaarts J. Population growth and global warming. *Population and Development Review* 1992;18:229-319.
33 Potts M. The population Doomsday forecast [letter]. *J R Soc Med* 1994;87:718.
34 Pearce D. *Blueprint 2.* London: Earthscan Publications, 1991.
35 Ekins P. *New environmental and ecological economics..* London: Birkbeck College, University of London, 1995. (Queen Mary and Westfield College Public Policy Seminars.)
36 Keating M. *The earth's summit's agenda for change. A plain language version of agenda 21 and the other Rio agreements.* Geneva: Centre For Our Common Future, 1993.
37 World Health Organisation. *Health and environment in sustainable development.* Geneva: WHO, 1997.
38 United Nations Environment Programme. *Handbook for the Montreal Protocol on substances that deplete the ozone layer.* Nairobi: Ozone Secretariat, United Nations Environment Programme, 1993.
39 World Energy Council. *Climate change negotiations: COP-2 and beyond.* London: World Energy Council, 1996.
40 Peel M, Berlin A, Piachaud J, James E. Contemplating a one child world. Medical students should be taught to appraise evidence on global health. *BMJ* 1995;311:1651-2.

(Accepted 13 June 1997)

Resource allocation to health authorities: the quest for an equitable formula in Britain and Sweden

Finn Diderichsen, Eva Varde, Margaret Whitehead

In recent years countries with very different healthcare systems have been showing increasing interest in resource allocation policies based on weighted capitation. In countries whose healthcare systems have competing health insurers the main concern has been to construct capitation formulas that prevent favourable risk selection or "cherrypicking." Reforms to the American Medicare programme and Dutch healthcare proposals have stimulated renewed efforts to find a way of overcoming this problem.[1-4]

Countries with national health services, such as the United Kingdom and Sweden, have also experienced far-reaching reforms of health care, with important implications for equity in access to care.[5 6] Risk selection should be less of a problem, at least with health authority purchasing, as the population is assigned to a purchaser based on area of residence. The new role of local purchaser, however, calls for more exact methods to allocate "purchasing power," because local areas will show stronger variation in relative need than regions and counties.

We outline British experiences in attempting to devise an equitable formula then present the new model that we have developed in Sweden for Stockholm County Council. We discuss what lessons these experiences hold for other countries facing a similar challenge.

Summary points

The United Kingdom and Sweden face similar problems in how to achieve a fair allocation of resources within a purchaser-provider system

In contrast with the British formula, the new Swedish approach is based on individual level data and uses demographic and socioeconomic variables as proxy measures of healthcare need

The Swedish model incorporates actual, rather than estimated, costs of care

The resulting model allocates proportionately more resources to populations with poorer health and socioeconomic characteristics

Both the Swedish and British approaches illustrate the practical problems and the highly political nature of resource allocation

These experiences hold important lessons—not least for the growing number of other countries with a similar quest

This is the first of three articles reflecting on recent developments in healthcare policy in Sweden

Department of Public Health Sciences, Division of Social Medicine, Karolinska Institute, S-172 83 Sundbyberg, Sweden

Finn Diderichsen, *professor*

Eva Varde, *statistician*

Margaret Whitehead, *visiting fellow*

Correspondence to: Professor Diderichsen finn.diderichsen @phs.ki.se

BMJ 1997;315:875–8

British developments

In Britain serious attempts to devise more equitable mechanisms for resource allocation for the NHS date back to the 1970s, when it became clear that funding to the regions based on historical activity had perpetuated the inequalities in funding that existed before the NHS. Since then, development work has gone through three distinct phases.[7]

In the first phase the formula created by the Resource Allocation Working Party was developed for distributing resources from central government to regions. It used mortality in each area as an indicator of healthcare need.[8] The formula was in use from 1977-90 and gradually managed to redistribute resources from the metropolitan regions to the poorer regions in the north.[9]

In the second phase the argument that the measurement of need should be based on empirical data led to a new formula for weighted capitation, applied from 1991 to 1995.[10] This empirical approach was severely criticised on methodological grounds and because it seemed inequitable.[11-16]

Clearly, the Department of Health needed a more sophisticated model for allocating funds directly to local districts now that they were purchasers. It commissioned health economists at York University to develop a more sensitive, empirically based model, to be incorporated into a third allocation formula from April 1995 onwards.

The York model is based on an ecological study of small areas to identify the determinants of use of hospital services.[17 18] The need variables identified include both health and socioeconomic factors (table 1). In addition, statistical models were developed to distinguish several confounding influences on the use of services, such as the supply of hospital beds and general practitioners. The effect of applying the formula in full at the district level would be to redistribute funds towards poorer, inner city areas.[18] The Department of Health decided, however, that the full York model would apply to only 76% of funding and the new arrangements would be introduced only gradually over several years. Other adjustments for "market forces" were also added. In effect, these adjustments watered down the full potential of the York model to allocate resources equitably. As about 70% (£23bn a year) of NHS funding is distributed through these formulas, even slight adjustments can make a big difference to local allocations.

New approach in Sweden

Like Britain, Sweden has a national health service, publicly funded and provided. Of the total healthcare budget of Kr82bn (£8bn), 82% comes from regional income taxes raised by the 26 county councils responsible for administering health care.

This regional funding has until recently been distributed directly to public hospitals and primary

Distinctive features of resource allocation in Sweden and Britain

Sweden (Stockholm model)
- Need for health care is measured by demographic and socioeconomic variables rather than mortality or other health status indicators
- Analysis is based on individual level data rather than at a small area (ecological) level
- Actual, rather than estimated, relative costs of health care are used

Britain (York model[17 18]*)*
- Need is measured by mortality, self reported morbidity, and various socioeconomic variables
- Analysis based on an ecological study of small areas to identify the determinants of inpatient services
- The estimates are adjusted for the confounding influences of supply on geographic variations in use
- Estimated costs of health care are used

care centres on the basis of historical activity, adjusted for inflation. This has changed in the past four years in counties that have introduced an internal market. In particular, Stockholm has been at the forefront of the introduction of a purchaser-provider split, and associated developments in resource allocation have consequently gone further than in the other counties.

Stockholm County Council serves a population of 1.7 million with a healthcare budget of £1.6bn. Most (90%) of the county budget is distributed to nine health authorities, each covering populations of between 50 000 and 300 000.

Basis of model

The contrasting features of the Swedish and British approaches are listed in the box. Individual level analysis was chosen not only because of the practical availability of data but also because of the problems inherent in ecological analysis.[19 20]

Finding a direct indicator of health status for measuring healthcare need that could be linked to individual use of health care and cost data proved difficult. The model therefore uses various socioeconomic indicators as proxies for healthcare need, over and above that created by the demographic profile of the population. The choice was based on evidence showing that use of hospital services in Sweden was proportional to the relative

ANDREW MILNE

Is a demographic profile of the population adequate for determining health need and therefore resource allocation?

Table 1 Need variables used in the current British formula for resource allocation (York model)[18]

Need variables	General and acute model	Psychiatric model
Standardised limiting long standing illness ratio <75	✓	
Standardised mortality ratio <75	✓	✓
Proportion of economically active people who are unemployed	✓	
Proportion of people of pensionable age living alone	✓*	✓*
Proportion of dependents in single carer households	✓	
Proportion of persons in single parent households		✓
Proportion of dependants in no carer households		✓
Proportion of adult population permanently sick		✓
Proportion of population born in New Commonwealth		✓

*These variables are included in both models but with different coefficients.

need of major socioeconomic groups.[21 22] Higher use by more socially disadvantaged groups is assumed to translate into higher costs of care, for which health authorities need to be funded.

Psychiatric services, however, were used at a low level by non-Nordic immigrants, perhaps not reflecting all their needs.[22] Immigrant status was therefore excluded from the analysis. A different model was devised for primary care (not reported here).

Statistical analysis

The analysis makes use of the personal identification number, which everyone in Sweden has and which can link healthcare records with census and other socioeconomic databases. Since a new system of payment was introduced in 1994, actual costs of care billed to purchasers have also been available for each individual in the population. The analysis has four main stages.

Stage 1—We created two new databases each year, linking the records on healthcare use and related costs to data on age, sex, socioeconomic group, education, cohabitation and marital status, country of birth, and housing conditions. One database covered a 30% random sample of the country's population, containing their socioeconomic characteristics and any health care they had used. The other database included all people with inpatient care and their background variables.

Stage 2—We then tested different models (with multivariate Poisson regression of outpatient and inpatient episodes) to select the demographic and socioeconomic variables that had the greatest effect on use, controlling for other variables. The variables selected by this process for the final model were (*a*) age in 10 classes; (*b*) socioeconomic groups in four groups based on occupation and employment (education for pensioners); (*c*) cohabitation and marital status in four classes; and (*d*) housing in five classes, according to tenure and size. Sex was not included in the final model. The effect of including sex made a negligible difference to the distribution of resources as the distribution of men and women did not differ between districts.

Stage 3—A matrix was constructed in which each cell represented a unique combination of the selected variables. In each cell, weights were calculated equal to average costs per inhabitant. Separate weights were calculated for acute medical and surgical care, non-acute care, and psychiatric care. Because actual costs were not available for psychiatry, the costs for this specialty were estimated on the basis of number of bed days and outpatient visits. Table 2 shows an abridged version of the matrix.

Stage 4—A corresponding matrix with the number of inhabitants in each of the nine health authority areas was then constructed, and each individual was ascribed a weight based on their social and demographic characteristics. These weighted individuals were then summarised for each area and the budget calculated as a proportion of the total sum for the whole county council (table 3).

Table 2 Matrix (abridged version) used in Stockholm for resource allocation to hospital care showing cost (Kr per inhabitant) spent by health authorities in Stockholm County Council, 1994

	Acute and non-acute medical and surgical care		Psychiatric care	
Age of inhabitant (years)	Owner occupied home	Rented home	Owner occupied home	Rented home
0 to <1	7200*		0	0
1-24	1900	2100	400	600
25-64 cohabiting:				
Higher non-manual	3100	3600	400	800
Lower/intermediate non-manual	3700	4300	600	900
Manual	4000	4400	900	1300
Not employed	5300	6400	1400	2400
25-64 living alone:				
Higher non-manual	3600	3900	900	1600
Lower/intermediate non-manual	3600	4200	1000	2400
Manual	3900	4600	1400	3800
Not employed	5100	6400	4900	12 700
65-84 years:				
Cohabiting	13 500	16 500	500	1000
Living alone	15 400	18 200	1100	2100
≥85 years:				
Cohabiting	27 600	29 800	300	1000
Living alone	24 200	29 400	500	1000

*Split between both categories of housing.

Implementation

The model has been applied gradually in calculating health authority budgets in Stockholm County Council since 1992. Before 1996, costs were estimated from the number of admissions and bed days, whereas the 1996 budget was based on actual costs for the purchasers.

Overall, the model has allocated more resources for the care of people living in more disadvantaged socioeconomic circumstances (table 2). The resulting ranking of authorities in table 3, based on these costs, follows the known differentials in health, demographic, and socioeconomic factors in the county.[23]

The interim model used in 1995, based on estimated costs, allocated a large share of the budget to areas containing a high proportion of elderly people and people living alone. As the year unfolded, it became apparent that the interim model might have overcompensated for the costs of providing health services for elderly people. In fact, central Stockholm, with the highest proportion of elderly people, could not spend all its allocated budget, whereas the suburban areas with young families ran up budget deficits. When actual costs became available for the 1996 model, it was found that each bed day was cheaper for elderly than for younger age groups. In 1996 therefore the share of the budget was reduced for central and south Stockholm and increased for suburban areas (a shift of 1.4% of the budget) (table 3). Politically, this was seen as too great a shift to be achieved in one year. The county council therefore gave extra funds in the 1996 allocation to the authority hardest hit by the redistribution.

Insights from these developments

What are the lessons from these British and Swedish experiences? In both countries the principle has been firmly established that healthcare resources should be distributed in proportion to the relative needs of local populations. It is a step forward that serious attempts are being made to translate this principle into practice, but the quest for improvements continues.

Making best use of available data

The experiences illustrate two different ways of going about the task, largely determined by the need to make the best use of whatever data are routinely available in each country. This has led to an analysis based on area of residence (ecological analysis) in Britain and an approach based on data from individuals in Sweden. Several commentators have concluded that individual level analysis is the better option, to reduce the problems of confounding and misclassification.[24 25] The Swedish approach has made the most of the opportunity offered by newly available individual data, though this was the only practicable option for Sweden because the small numbers obtained from area based data would have made the resulting statistical models unstable. It did, however, restrict the choice of indicators of need. For example, no suitable health indicators were available that could be linked to the other individual level data.

The York model has to rely on data for small areas, not directly linked to individuals, which brings added problems of interpretation. On the other hand, with care it can include additional local data on mortality and morbidity, increasing its sensitivity to geographical variations that are not simply the sum of individual variations in the basic sociodemographic characteristics.

Proxies for need

Both the British and Swedish approaches are based on the assumption that the different needs for health care of the various sections of the population are matched by

Table 3 Per capita weighting for the nine health authorities in Stockholm County Council, according to Stockholm model, 1995-7

	Per capita weights		
Health authority	Interim model, 1995	1996	1997
Norrtälje	98.5	96.4	100.6
North east	96.7	96.6	97.9
North west	91.4	94.5	94.2
Central Stockholm	127.7	120.3	119.2
West Stockholm	99.1	100.3	98.6
South Stockholm	122.7	116.3	117.4
South west	97.6	99.5	98.1
South east	83.4	85.9	86.2
Södertälje	93.0	96.0	95.7
Whole Stockholm county	100*	100*	100*

*100=Kr9166 per inhabitant in 1995; Kr9082 per inhabitant in 1996; Kr8979 per inhabitant in 1997.

their differential use of services. But in practice the use of services is influenced not only by legitimate need but also by supply and many other socioeconomic factors, so the match is not perfect. Given the circumstances, informed judgments have to be made on the most practical solutions. The Swedish decision, for example, to leave out an indicator of "ethnic group" from the final analysis was based on the evidence that non-Nordic immigrants have higher psychiatric morbidity but a relatively low rate of use of psychiatric services. Incorporating a factor based on use by ethnic group would have led to fewer resources being allocated to health authorities with large immigrant populations.

Taking deprivation into account

Both approaches consider it essential to take social and material deprivation into account. They have both selected employment factors and living alone as important indicators of increased need for healthcare resources. Sweden has added indicators of poorer housing, and Britain has added households containing singlehanded carers (including single parents) as well as direct health indicators.

Two new relevant findings emerge from the Swedish data on differential costs of care. Firstly, the analysis of actual costs for care of different groups provides a direct demonstration of the higher costs incurred by more disadvantaged groups in the population and the need for extra resources in areas where the proportion of people from these groups is greatest. Secondly, the comparison of estimated costs in 1995 with actual costs in 1996 revealed the scale of the bias introduced when only estimated costs are used. A similar problem with estimated age-cost weights was encountered in the British formula introduced in 1991, when it was applied to populations at district level.[16]

Political reality

Both experiences illustrate the highly political nature of resource allocation. The Swedish model ran into some difficulties when quite large shifts had to be achieved in the switch from the interim model in 1995 to the full model in 1996, particularly as the overall funding per inhabitant was falling over the same period. Although full implementation was agreed for 1996, a one-off compensation, as mentioned above, was given to the authority that stood to lose the most. Agreement on full implementation for 1997 was politically easier, as the shifts in funding were not as great.

In 1995 the York model was not implemented in full in Britain because of the government's nervousness over the size and direction of the implied shifts in resources, generally from suburban towards poorer areas. Identifying two separate models (table 1) allowed room for subsequent manoeuvre. There are even suggestions now that the market forces factor, introduced into the British formula by the Department of Health, is seriously undermining the model's attempt to allocate resources according to need.[26]

This illustrates the need to ask continually whether the policy as implemented is achieving its original objectives of equitable resource allocation.

Effects of cost containment

Finally, both approaches illustrate the complications of trying to devise and implement an equitable formula in a time of cost containment, when any redistribution of resources is much more painful. Some commentators suggest that the strain imposed by the prolonged underfunding of the British NHS in the 1980s was a key factor in the decision to overhaul the original formula created in the late 1970s.[13] The drastic cuts that have had to take place in Sweden in the 1990s with the economic recession mean that the effects of resource allocation are not easy to disentangle from the effects of cutbacks.

Yet it is at just such times that efforts need to intensify. The joint effects of cutbacks and market-style reforms could be especially damaging to access to healthcare for the sections of the population in greatest need, as in a more competitive environment resources tend to flow to more prosperous areas and groups. It is important that the quest for equitable methods of resource allocation continues and is taken up by the growing number of other countries facing a similar challenge.

Funding: Stockholm County Council funded the development of the resource allocation model and its annual update.

Conflict of interest: None.

1 Newhouse J, Manning W, Keeler E, Sloss E. Adusting capitation rates using objective health measures and prior utilisation. *Health Care Financing Review* 1989;10:41-54.
2 Brown RS, Hill JW. The effects of Medicare risk HMOs on Medicare costs and service utilization. In: Luft HS, ed. *HMOs and the elderly.* Ann Arbor, MI: Health Administration Press, 1994.
3 Van Vliet RCJA, Van den Ven WPMM. Towards a capitation formula for competing health insurers. An empirical analysis. *Soc Sci Med* 1992;34:1035-48.
4 Van Vliet RCJA, Van den Ven WPMM. Capitation payments based on prior hospitalizations. *Health Economics* 1993;2:177-88.
5 Whitehead M. Who cares about equity in the NHS? *BMJ* 1994:308:1284-7.
6 Diderichsen F. Market reforms in health care and sustainability of the welfare state. *Health Policy* 1995;32:141-53.
7 Mays N. Geographical resource allocation in the English national health service 1971-94: the tension between normative and empirical approaches. *Int J Epidemiol* 1995;24:96-102.
8 Department of Health and Social Security. *Sharing resources for health in England: report of the resource allocation working party.* London: HMSO, 1976.
9 Holland W. The RAWP review: pious hopes. *Lancet* 1986;ii:1087-90.
10 NHS Management Board. *Review of the resource allocation working party formula.* London: Department of Health and Social Security, 1988.
11 Judge K, Mays N. Allocating resources for health and social care in England. *BMJ* 1994;308:1363-6.
12 Mays N. NHS resource allocation after the 1989 white paper: a critique of the research for the RAWP review. *Community Med* 1989;11:173-86.
13 Sheldon T, Carr-Hill R. Resource allocation by regression in the NHS: a critique of the RAWP review. *J R Stat Soc A* 1992;155:403-20.
14 Sheldon T, Davey Smith G, Bevan G. Weighting in the dark: resource allocation in the new NHS. *BMJ* 1993;306:835-9.
15 Royston G, Hurst J, Lister E, Stewart P. Modelling the use of health services by populations of small areas to inform the allocation of central resources to larger regions. *Socioeconomic Planning Science* 1992;26:169-80.
16 Raftery J. Capitation funding: population, age, and mortality adjustments for regional and district health authorities in England. *BMJ* 1993;307:1121-4.
17 Carr-Hill R, Sheldon T, Smith P, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: development of methods for small area analysis and use of inpatient services. *BMJ* 1994;309:1046-9.
18 Smith P, Sheldon T, Carr-Hill R, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: results and policy implications of small area analysis of use of inpatient services. *BMJ* 1994;309:1050-4.
19 Diderichsen F, Spetz C-L. *Need based planning in health care.* Stockholm: National Board of Health Care and Social Welfare, 1987. (In Swedish.)
20 Morgenstern H. Uses of ecological analysis in epidemiological research. *Am J Public Health* 1982;72:1336-44.
21 Health policy targets and need based planning. *Health care in the 1990s.* Stockholm: Allmänna Förlaget, 1984:40-1. (In Swedish.)
22 Haglund B. *Equity in care.* Stockholm: National Board of Health and Welfare (Socialstyrelsen), 1994:3. (In Swedish.)
23 Diderichsen F, ed. *Public health report 1994.* Stockholm: Stockholm County Council, 1995. (In Swedish.)
24 Benzeval M, Judge K. The determinants of hospital utilisation: implications for resource allocation in England. *Health Economics* 1994;3:105-16.
25 Carr-Hill R, Rice N, Roland M. Socioeconomic determinants of rates of consultation in general practice based on fourth national morbidity survey of general practices. *BMJ* 1996;312:1008-13.
26 Hacking J. Capitation funding: is it worth the weight? *Health Service Journal* 1996;18 Jan:26-7.

(Accepted 6 March 1997)