

**THE DISTRIBUTION AND REDISTRIBUTION OF HEALTH
RESOURCES IN SOUTH AFRICA**

Alexander Marius Van Den Heever

August 1990

**Submitted to the University of Cape Town in Fulfilment of the Requirements for the Degree
of Master of Arts in Economics**

The University of Cape Town has been given
the right to reproduce this thesis in whole
or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.


I, Alexander Marius van den Heever, submit this thesis for the degree of Master of Arts in Economics. I claim that this is my original work and that it has not been submitted in this or in a similar form for a degree at any University.

ABSTRACT

This thesis is intended as a broad examination of the distribution of health resources in South Africa. Issues both macro and micro in nature have been covered to provide a perspective that would be lacking in a narrower study. Although the title refers to a redistribution of resources, the intention of this thesis is to stress the importance of providing appropriate health measures rather than merely apportioning existing facilities evenly. This realization is insufficient, however, if it is not accompanied by the introduction and utilization of analytical approaches for identifying resource selection priorities.

The influences on health status are many. In defining appropriate measures to improve health status it is important to be aware of the limitations of medical-care. Chapter three involves a cross-sectional regression analysis of various countries in order to examine the influences certain variables have on health status. This study suggests the need for an integrated approach to improving the health of a population. Merely focusing on medical care will only have a limited affect. However, this does not mean that medical-care is not important. It must just be provided in an appropriate manner. The rest of the thesis evaluates health-care resource distribution in South Africa.

The existing distribution of health-care resources in South Africa is ill-suited to the existing health status of the population. There is a bias toward urban based curative facilities. Furthermore, the location of facilities has been based on racial criteria, whereby some areas have sufficient resources for their needs while others do not. Two methods of identifying how these issues should be dealt with are produced in this thesis.



The first deals with a method for adjusting the broad distribution of funds toward those areas where need is greatest. The suggestion put forward by this thesis is that a formula be developed that would be able both to define need on a geographical basis, and to allocate resources based on that need. The formula would be used to allocate government health expenditure. This section is based on a formula that was developed in the United Kingdom.

The second deals with a method for defining appropriate medical interventions on the micro level. It is called cost-effectiveness analysis (CEA). CEA is used for micro-economic decision-making where a choice has to be made between at least two alternatives for attaining a particular objective. Furthermore, CEA evaluates projects or programmes that are on-going in nature. It should be noted that CEA can also evaluate non-medical interventions to solve a particular health problem.

In order to indicate the type of information that a CEA can provide, an investigation into cervical cancer procedures used on black females was produced. The entire black female population of South Africa was examined. A computer simulation of incidence and mortality rates of the disease was used to evaluate various scenarios. The results indicate that significant gains can be made by introducing cervical cancer screening on a large scale in South Africa.

A major priority of this thesis was to stress the importance of using economic criteria to assist in making decisions concerning health-care resource allocations. Very little work of this nature is produced in South Africa. Hopefully this will not always be the case.

ACKNOWLEDGEMENTS

Thanks are due to my supervisor Prof. Francis Wilson and to Diane McIntyre who assisted by commenting on an earlier draft of this thesis. I would further like to thank Dr. Estiaan Calitz for his interest and assistance in the early stages of this thesis and for making the facilities of the Central Economic Advisory Service available to me. The Central Economic Advisory Service also supported me financially while working on the thesis and for that I am extremely grateful. For the access provided to the Department of National Health and Population Development I would like to thank Dr. Kotze, chief director of Strategic Planning. Dr. Gladwin Leiman from the South African Institute of Medical Research commented on an earlier draft of the Cervical Cancer study and her assistance was much appreciated.

TABLE OF CONTENTS

CHAPTER	PAGE
1 Introduction	1
1.1 Motivation for Thesis	1
1.2 Objectives of Thesis	2
1.3 Structure of Thesis	3
1.4 Disclaimer	6
2 The Role of Government in the Provision of Health Care	8
2.1 Introduction	8
2.2 Market Failure, the Case for Intervention	9
2.2.1 Externalities and the Inverse Care Law	10
2.2.2 Cost and Expenditure Increases in Health-Care Spending, the Problem of Perverse Incentives	13
2.3 Centralised Control	19
2.4 Conclusion	22
3 Health Priorities	24
3.1 Introduction	24
3.2 Economic Theory of Health Investment and Planning	26
3.3 Bivariate Analysis	29
3.3.1 Results	30
3.4 Discussion of Results	32
3.4.1 Medical Variables	32
3.4.2 Social Variables	33
3.4.3 Economic Variables	40
3.5 Multivariate Analysis	41
3.6 Conclusion	45

4	Health in South Africa: a Brief Overview	57
4.1	Introduction	51
4.2	Mortality Profiles	51
4.2.1	Blacks	53
4.2.2	Whites	62
4.2.3	Coloureds	65
4.2.4	Asians	69
4.3	Conclusion	71
5	The South African Public Health Sector	73
5.1	Introduction	73
5.2	Structure of the Public Health Sector	73
5.3	Maldistribution of Health Resources	78
5.3.1	Geographical Distribution of Health Care Spending	78
5.3.2	Racial Distribution of Health Care Resources	79
5.4	Conclusion	84
6	The Geographical Distribution of South African Health Resources	85
6.1	Introduction	85
6.2	The Need to Redistribute Health Resources	86
6.3	Need Versus Demand	87
6.4	Weighting Population for Age and Sex	95
6.5	Adjusting Population for Morbidity	99
6.5.1	Case-Load	99
6.5.2	Potential Years of Life Lost	99
6.5.3	Health Status Indices (HSI)	101
6.5.4	Mortality Rates	101
6.6	Using the Formula to Indicate Relative Need for In-Patient Services on a Geographical Basis	107
6.7	Further Aspects of the Formula Approach	111
6.7.1	Service Categories	111
6.7.2	Targeting Resource Needs	113

6.7.3	Differential Costs	114
6.8	Ideas on the Practical Application of the Formula Approach	114
6.9	Regions	116
6.10	Cross-Border Flows	119
6.11	Conclusion	121
7	Cost-Effectiveness Analysis in the Health Sector	124
7.1	Introduction	124
7.2	Cost-Benefit and Cost-Effectiveness Analysis	125
7.2.1	Cost-Benefit Analysis	125
7.2.2	Cost-Effectiveness Analysis	125
7.3	"Output" or "Outcome" Measurement in CEA	126
7.3.1	Quality of Life	127
7.3.1.a	Health Status Indices	127
7.3.1.b	Valuation of Death as an Outcome	133
7.3.2	Duration of Health States	134
7.3.3	Relative Valuation of a Life Year	135
7.3.4	Discounting Life Years	136
7.4	Costs and Cost-Effectiveness Analysis	137
7.4.1	Indirect Costs	137
7.4.1.a	The Value of Life	138
7.4.2	Discounting Costs	141
7.5	Conclusion	142
8	A Cost-Effectiveness Analysis of a Potential Comprehensive National Screening Programme for Cervical Cancer in South Africa Using a Computer Simulation	
8.1	Introduction	145
8.1.1	Objective of Chapter	145
8.1.2	Motivation for Study	145
8.1.3	Description of Study	146

8.1.4	Hypotheses	146
8.1.5	Structure of Chapter	147
8.2	Cervical Cancer	147
8.2.1	Cervical Intraepithelial Neoplasia	147
8.2.2	Invasive Cervical Carcinoma	149
8.3	Causes of Cervical Cancer	150
8.4	Treatment and Prevention	153
8.4.1	Screening	154
8.4.2	Treatment of Invasive Cervical Cancer	155
8.5	Cost-Effectiveness of Screening (Various Findings)	156
8.6	Cervical Cancer in South African Black Females	158
8.7	The Model	159
8.7.1	Assumptions	160
8.7.1.a	Progression Rates	160
8.7.1.b	Costs	161
8.7.1.c	Screening Programme	163
8.7.1.d	Population Estimates	164
8.7.1.e	Discount Rate	164
8.7.1.f	Life Expectancy	165
8.7.2	Methodology	165
8.7.2.a	Scenarios	165
8.7.2.b	Outcome Measures	165
8.7.2.c	Direct Costs	166
8.7.2.d	Indirect Costs	167
8.7.3	Discussion of Results	167
8.7.3.a	Mortality	167
8.7.3.b	Incidence of Invasive Cervical Cancer	168
8.7.3.c	Potential Years of Life Lost	170
8.7.3.d	Potential Years of Life Saved	172
8.7.3.e	Direct Costs	174
8.7.3.f	Direct and Indirect Costs	176
8.7.3.g	Cost-Effectiveness	178

LIST OF ILLUSTRATIONS

FIGURE	TITLE	PAGE
Figure 3.1	Determining the Stock of Health	28
Figure 4.1	Black Mortality (Males)a	53
Figure 4.2	Black Mortality (Males)b	53
Figure 4.3	Black Mortality (Females)a	54
Figure 4.4	Black Mortality (Females)b	54
Figure 4.5	White Mortality (Males)a	62
Figure 4.6	White Mortality (Males)b	62
Figure 4.7	White Mortality (Females)a	63
Figure 4.8	White Mortality (Females)b	63
Figure 4.9	Coloured Mortality (Males)a	65
Figure 4.10	Coloured Mortality (Males)b	65
Figure 4.11	Coloured Mortality (Females)a	66
Figure 4.12	Coloured Mortality (Females)b	66
Figure 4.13	Asian Mortality (Males)a	69
Figure 4.14	Asian Mortality (Males)b	69
Figure 4.15	Asian Mortality (Females)a	70
Figure 4.16	Asian Mortality (Females)b	70
Figure 5.1	Financial Structure of the South African Public Health Sector	75
Figure 6.1	Supply Induced Demand	91
Figure 6.2	Expected Inpatient Bed Days	100
Figure 6.3	Percentage Weights (PYLL, SMRs and Crude Population)	107

Figure 6.4	Regional Population Compared to Health Expenditure	110
Figure 6.5	Black Male Mortality (Urban/Rural)	117
Figure 6.6	Black Female Mortality (Urban/Rural)	118
Figure 7.1	A Diagrammatic Representation of CEA	143
Figure 8.1	Mortality From Cervical Cancer	168
Figure 8.2	Incidence of Invasive Cervical Cancer	169
Figure 8.3	Potential Years of Life Lost From Cervical Cancer	171
Figure 8.4	Potential Years of Life Saved	173
Figure 8.5	Direct Cost of Treating Cervical Cancer	175
Figure 8.6	Direct and Indirect Cost of Treating Cervical Cancer	177
Figure 8.7	Cost-Effectiveness of Treating Cervical Cancer	179

LIST OF TABLES

TABLE	TITLE	PAGE
Table 2.1	Percentage of Total Household Consumption Spent on Medical Care	14
Table 3.1	Results of Ordinary Linear Regression	30
Table 3.2	The Change in the Variable Required to Result in a Ten Year Improvement in Life- Expectancy	32
Table 3.3	Results of Multiple Linear Regression (Medical Variables)	42
Table 3.4	Results of Multiple Linear Regression (Social Variables)	42
Table 3.5	Results of Multiple Linear Regression (Economic Variables)	43
Table 3.6	Results of Multiple Linear Regression (Final Results A)	43
Table 3.7	Results of Multiple Linear Regression (Final Results B)	43
Table 4.1	Estimated Quantification of the Homeless Population in the Urban Townships of the PWV	59
Table 4.2	Population Living in Informal Structures in South Africa	60

Table 5.1	The Geographical Distribution of Public Health Expenditure	79
Table 5.2	South African Population and Bed Data	80
Table 5.3	Bed Occupancy in South Africa	81
Table 5.4	Lengths of Stay in Inpatient Facilities	82
Table 6.1	National Hospital Inpatient Utilization Rates (United Kingdom)	97
Table 6.2	Population Adjusted for Bed Utilization by Age and Sex	98
Table 6.3	Populations Weighted for Bed Utilization and SMRs Compared to Total Health Expenditure by Region	109
Table 6.4	Estimated Proportions of Revenue Expenditure on Each Service Category (United Kingdom)	112
Table 7.1	Preference Ratings for Function Levels	129
Table 7.2a	Transformed Valuations for 29 Health States	131
Table 7.2b	Classification of States of Illness	132
Table 8.1	Mortality From Cervical Cancer (for 1990-94)	168
Table 8.2	The Incidence of Invasive Cervical Cancer (for 1990-94)	170
Table 8.3	The Potential Years of Life Lost from Cervical Cancer (for 1990-94)	172
Table 8.4	The Potential Years of Life Saved (for 1990-94)	174

Table 8.5	The Direct Cost of Treating Cervical Cancer (for 1990-94)	176
Table 8.6	The Direct and Indirect Cost of Treating Cervical Cancer (for 1990-94)	178
Table 8.7	The Cost Per Potential Year of Life Saved (for 1990-94)	180

GLOSSARY

Carcinoma:	cancer.
Condition:	medical classification for cause of morbidity or mortality.
Fragmentation:	term used to describe the dysfunctional character of South Africa's decentralised health services.
Infant Mortality Rate:	combines neonatal and post-neonatal mortality per 1000 live births.
Inpatient facilities:	hospital facilities for over-night patients.
Intervention:	medical or non-medical method of dealing with a medical condition.
Morbidity:	prevalence of disease.
Mortality:	number of deaths in a given period.
Neonatal Mortality:	all mortality within 27 days of birth.
Neoplasm:	cancer.
Outpatient facilities:	hospital facilities for patients not requiring overnight treatment.
Post-Neonatal Mortality:	all mortality from 28 days of life up to but not including 12 months.
Procedure:	specific medical treatment.

	8.7.3.h Overall Comment	180
8.8	Conclusion	181
9	Conclusion	182
9.1	Results and Findings	182
	9.1.1 Need-Based Resource Allocation on the Macro Level	182
	9.1.2 Nature of Health Status	184
	9.1.3 Cost-Effectiveness Analysis	185
9.2	Areas for Further Research	185
	Appendix A	188
	Appendix B	196
	Appendix C	207
	Bibliography	218

CHAPTER ONE

INTRODUCTION

1.1 Motivation for Thesis

Despite the important contribution that applied economic research can make toward improving the allocation of resources in the health sector, such work is rarely requested or produced in South Africa. This oversight could be largely attributed to ignorance, both by economists and health professionals, of the potential gains that can be made through the development of health economics as a specialized field of research. This thesis, therefore, attempts to play a minor role in amending this situation. Although the topic specifically refers to resource allocation, one of the main objectives of this thesis is to highlight those areas of health care that can be subjected to economic evaluation. It is intended to be a demonstration as well as practical research.

As health care resources in South Africa have been misallocated, for both political and economic reasons (see Chapter Five), the application of a methodology for resource distribution will inevitably indicate the need for a redistribution of resources, and a re-evaluation of existing technology. However, the information required merely to recognize a need is insufficient for the purposes of satisfying that need. It is necessary to find methods of processing information in such a way that it can supply answers to problems of resource allocation. Furthermore, the data provided should be consistent with such an objective. As the field of health economics is not established in South Africa very few statistics are generated from the health sector that can be processed to make decisions on resource allocation. This places a heavy constraint on affective economic analysis in this area. However, as economic methods of analysis become established, and their importance noted, the motivation for routine generation of the required data can be provided.

The allocation of resources at present is strongly affected by allocations made in the past. If the past allocations were inappropriate, or are no longer relevant to the present circumstances, the link between past and present decisions must be broken. The use of incremental budgeting results in an inertia whereby decisions do not respond to the changing needs of a dynamic environment. The nature and effectiveness of various methods of health care and their appropriateness in certain environments are important factors which are not given adequate attention in South Africa. The problem is often not one of rationing limited resources, but rather one of selecting appropriate facilities. Providing more medical care will not appreciably alter the health status of any community, rich or poor (see Chapter Three). Illness in many instances is symptomatic of the socioeconomic environment (see Chapters Three and Four). However, this does not mean that medical facilities do not play a significant role in dealing with health problems. They must be provided according to the circumstances and requirements of the community. Chapter Eight provides an example of how the appropriate application of medical resources can reduce the incidence and mortality from cervical cancer, where the socioeconomic environment of the population examined greatly contributed to the extent of this disease. Were the socioeconomic environment to change, so too would the nature of the medical services provided.

Such evaluations of the health care sector in South Africa are not carried out at present. It is this omission that provided the motivation for this thesis.

1.2 Objectives of Thesis

This thesis has the following objectives as its task:

- 1 To establish where the responsibility for providing health care services lies;

- 2 To evaluate the factors contributing to health status;
- 3 To evaluate the present health situation in South Africa, as well as the existing structure of health care services; and
- 4 To develop two basic methodologies to deal with resource allocation, one on the macro and the other on the micro level.

1.3 Structure of Thesis

In order to establish a basis for the approaches used in this thesis, it was necessary to provide certain background information. The thesis can therefore be divided into two categories, the first dealing with some of the more general aspects of health status and health economics (Chapters Two to Five), and the second concentrating on specific analytical approaches for allocating resources in the health sector (Chapters Six to Eight). The various chapters are now discussed briefly.

Chapter Two establishes a basis for public sector involvement in the allocation and distribution of health-care resources. This chapter provides the theoretical basis for health resources to be allocated on a *need* rather than a *demand* basis. The emphasis is on the superiority of public as opposed to private sector provision of health care. These considerations are important if note is taken of the existing South African government's moves towards privatization and increased user-charging.

Chapter Three analyses the primary constituents of *health status*. This is done through the use of a cross-sectional econometric study of low- and upper-middle income countries. Although the study does not use South African data, the results provide information that can be extrapolated to the South African

situation. The importance of socioeconomic influences on health status¹ should be compared to the relative unimportance of medical variables. This result highlights the "safety-net" function of health care² variables. The nature of health-care services provided should be regarded as a result of the health status of a community or population, rather than a cause of it. This aspect is important to methodologies presented in Chapters Six to Eight. Never is an approach to health-care developed with the intent of altering health status, as this is not possible unless the socioeconomic environment changes. Health-care facilities and services should be provided according to need, and should alter as the nature of the need alters.

Chapter Four briefly discusses of the health situation in South Africa. Mortality profiles according to "race" group are examined. The focus is placed on the widely diverse problems facing the various populations as a result of their socioeconomic circumstances. The contrasting nature of the health care requirements of blacks and coloureds compared with whites and Asians, due to very different socioeconomic environments, is indicated in this study. This chapter places some of the aspects concerning health status, (discussed in Chapter Three) in the South African context. It also serves as a background to the study discussed in Chapter Six.

Chapter Five briefly examines the existing distribution of health-care resources in South Africa, and indicates the inappropriate nature of this distribution. Access to adequate health care is a function of both "race" and income in South Africa. This access is limited through both the "racial" and geographical distribution of resources. The existing public provision of health care is found wanting on the basis of the following aspects:

-
- 1 The term "health status" can be defined in two ways. The first merely refers to the risk of falling ill generally, ie. as a consequence of socioeconomic living conditions, the second refers to the status of an individual as a consequence of having fallen ill. The latter usage of the term can be seen in Chapter Seven with reference to the "health status index". This is the only occasion in the thesis that the term "health status" is used in this context.
 - 2 Throughout this thesis references to "health care" resources relate to medical goods and services.

- a the excessive fragmentation of the health sector;
- b the use of "racial" criteria to allocate resources;
- c the unequal distribution of resources according to "race";
- d the inappropriate nature of the health services provided (ie. inadequate out-patient and primary health care facilities coupled with an excessive³ provision of in-patient hospital facilities); and
- e the unequal geographical distribution of health care resources (this is also strongly linked to the "racial" bias).

This chapter therefore describes various aspects of the historical misallocation of resources that are to be addressed in chapter six.

Chapter Six looks at a methodology for the macro distribution of health-care resources based on an approach developed in the United Kingdom. Essentially, the methodology is need-based and is designed to adapt to changing environments and changing needs. The focus is on the financial distribution of health care resources on a geographical basis. Unfortunately, such a methodology leaves out many important considerations that are beyond the scope of this thesis. Nevertheless, the approach is discussed, more as a starting point for further research than as the last word on the topic. However, a major point of this methodology is that the country is considered as a whole, rather than as consisting of "racially" defined populations and regions. The needs of populations adjusted for age, sex and morbidity become the basis for resource allocation. As this approach is useful only at the macro level, a more specific type of analysis is required to define the most appropriate interventions and procedures to use on the micro level. This, therefore, becomes the focus of Chapters Seven and Eight.

Chapter Seven involves a discussion of cost-effectiveness analysis (CEA). Cost-effectiveness analysis, as used in the health sector, is a useful method for facilitating decision-making between alternative medical

3 Excessive where whites are concerned (see Chapter Five).

procedures for treating the same disease. This chapter describes the approach and some of the developments in outcome evaluation. Chapter Seven is intended as an introduction to the study carried out in Chapter Eight.

Chapter Eight is a cost-effectiveness analysis of alternative methods for combatting cervical cancer in black South African females. The objective of the study was to emphasize the usefulness of CEA as a technique for making choices in health. It also makes a contribution to an existing debate in South Africa on the alternative approaches to dealing with cervical cancer in the black female population. The study concerns the examination of a broadly curative with a broadly preventive scenario. In this case the preventive scenario attempts to simulate the results of facilitating the early detection of the disease through the use of the Papanicolaou smear. This study examines the possible benefits of introducing a comprehensive national screening programme for black women. The results indicate that there are gains to be made in terms of lives saved, total costs and cost-effectiveness if the envisaged screening programme were introduced.

The various studies presented in this thesis are intended to provide an insight into an important field of research that, at present, is not given much attention by the economics profession in South Africa. What should be indicated in this thesis is the ability of economic research to save lives and money.

1.4 Disclaimer

As various sections of this thesis deal with analyses that focus, or separately define, "race" groups, it is necessary to clarify the reasons for doing so.

The method of official data collection in South Africa has been according to "race", up to and including the 1985 census. It is acknowledged that racial classification has enabled the South African authorities to

circumscribe the freedom of association, and the right of access to publically provided goods and services, ie. education and health facilities, for certain "race" categories.^{4 5} The terms "black", "coloured", "white", and "Asian", are used primarily out of convenience in this thesis, as these classifications do present a usable degree of social stratification. However, this author in no way supports or condones the present or past usage of "racial" classification in South Africa.

4 M. E. West, E. A. Boonzaier, "Population Groups, Politics and Medical Science", South African Medical Journal, September, 1989, vol.76, pp.185-186.

5 D. E. Bourne, "Nomenclature in a Pigmentocracy - a Scientists Dilemma", South African Medical Journal, September, 1989, vol.76, p.185.

CHAPTER TWO

THE ROLE OF GOVERNMENT IN THE PROVISION OF HEALTH CARE

2.1 Introduction

Over the past few years South Africa has become accustomed to a popular phrase: the Health Crisis. It may be comforting to know that much of the world has for years been as familiar with this term. Crises have affected very different health care systems, and health sectors, despite phenomenal improvements in medical technology and knowledge. It is too easy to conclude, therefore, that the general incidence of crises worldwide, within various health systems is the result of insoluble general problems, and the only answer is to muddle through as best one can. This chapter rejects that option and adopts the attitude that the problems can be solved, provided the root causes can be identified.

The intention of this chapter is to evaluate the nature of a government's responsibility for providing health care. There is a lot that can be learned from the experience of other countries. The basic assumption used in this chapter is that any nation must be concerned with the health of its population and with the health system established in the interests of that objective. The question of where responsibility should lie for over-seeing health services within a particular nation is thus a given. The nature of the chosen system, however, is not. This chapter consequently surveys the issues, possibilities, and failures within the various approaches opted for internationally. Although this chapter refers to "government" in general terms, the responsibility of the South African government is inferred. The various aspects examined internationally are chosen for their applicability to the South African context. Private and public sector solutions to problems within the context of mixed economies are given priority. The view taken in this chapter is that sufficient evidence exists to justify significant interference, on a centrally planned basis.

2.2 Market Failure, the Case for Intervention

Adam Smith, in the book *The Wealth of Nations*¹, developed the first concise arguments against state interference in the freedom of individuals to pursue their own interests. Smith asserted that where the state abstained from interference the *"sovereign is completely absolved from a duty, in the attempting to perform which he must always be exposed to enumerable delusions, and for the proper performance of which no human wisdom or knowledge could ever be sufficient; the duty of superintending the industry of private people, and of directing it towards the employments most suitable to the interest of society."*

Nevertheless Smith relates three instances where the "sovereign" has a duty to make its presence felt within a system of "natural liberty". These are:

- 1 The duty of protecting society from violence and invasion of other independent societies.
- 2 The duty of protecting as far as possible every member of society from the injustice or oppression of every other member of it, or the duty of establishing an exact administration of justice.
- 3 The duty of erecting and maintaining certain public works and certain public institutions which it can never be for the interest of any individual, or small number of individuals, to erect or maintain, because the profit could never repay the expense to any individual or small number of individuals, though it may frequently do much more than repay it to a great society.

Of the three, the third criterion for state intervention, described by Smith, is of primary significance to the health sector. Health care is unfortunately one of those industries that is incompatible with the contingencies of a simple free market system.

1 A. Smith, *The Wealth of Nations*, William Berton Publisher, London, Encyclopedia Britannica, Book 4 and 5, pp.182-422.

Firstly, what the free market for health would not provide can generate significant negative externalities, ie. external economic effects that escape the price mechanism.²

Secondly, many of the methods of internalizing these externalities, within the private sector context, can generate severe deficiencies in the nature of the health service provided. An example is the medical-aid scheme. Due to the high costs of medical care, it is in the interests of individuals to share the risk of incurring these costs amongst as large a group of people as possible. Although this means paying a premium, the financial burden is sufficiently small as to minimally affect the income of the medical aid member. Unfortunately such a system creates many problems.

These two broad problem areas in allowing the market to allocate health care resources are discussed in the next two sections. The first section examines with the concept of externalities and the *inverse care law*, while the second deals with the notion of perverse incentives generated within the private sector health industry.

2.1.1 Externalities and the Inverse Care Law

Allowing the market to ration health resources affects what goods and services are supplied, which in turn affects who will buy them. This process is strongly influenced by the desire of medical professionals to achieve target incomes. The medical goods and services that are inconsistent with the physicians' desire to maintain or reach a particular target income will consequently not be provided, regardless of the possible benefits that could accrue to society. An example is the provision of preventive medical procedures to low-income populations. As the procedures are preventive, the at risk population is unlikely to present itself for treatment (as they are not yet ill), especially if there is a fee involved. In such an instance the

2 B. P. Herber, D. Richard, Modern Public Finance, 3rd ed. Irwin inc., Homewood, Illinois 60430, 1975, p.36.

private sector would not be prepared to involve itself, as the major benefits of such procedures would accrue elsewhere in society. Essentially, preventive medicine for low income populations is, understandably, outside the private health sector's repertoire. Inoculations of poor communities against polio, tetanus, measles and smallpox, can result in minimal levels of mortality, where before entire communities were being wiped out. Clearly this is an area for direct state involvement, and is not contentious.

The need for direct state intervention in the health sector is also required where the provision of care is biased due to the distribution of income. The existence of a negative correlation between the need for health care, and the quantity of resources allocated, is so consistent that it has come to be described as a law. Negative socioeconomic conditions associated with poor sanitation, over-crowded conditions, malnutrition, etc., tend to generate higher levels of illness than wealthier populations residing in better socioeconomic environments. However, as ability to pay is the criterion for gaining access to a market oriented health system, those most in need are unable to make the required demands on the system to provide them with the needed care. Wealthier populations usually require curative care, as many preventable diseases and afflictions no longer affect them. The curative facilities are very expensive and mostly biased towards diseases associated with over-consumption, and longer life spans (see Chapter Four and the mortality profiles for whites in South Africa). A very visible indication of the "inverse care law" is the uneven distribution of resources that tends to occur between rural and urban populations. In many developing countries the rural population exists in worse socioeconomic conditions than the urban population. However, most health care resources are concentrated in urban areas, even within countries where the state controls the health industry.

The World Bank³ has found that in most developing countries seventy percent or more of government

3 World Bank (a), Financing Health in Developing Countries, The World Bank Washington DC, USA, 1987, p.22.

spending goes to urban hospital-based care. In China, Columbia, Indonesia, and Malaysia, the average health sector subsidy accruing to urban households is around five times that of rural resident.⁴ Self-financing rural health programmes in Senegal and the Gambia have been added on to existing free urban facilities, resulting in rural residents having to pay for low level care from a village health worker in addition to subsidizing free and more sophisticated urban hospitals.

In Peru, national per-capita public and private health expenditures amounted to about \$59 in 1980.

"Although 27% of the total population in 1981 lived in the metropolitan area of Lima, the percentage of total public spending that was spent there increased from 37 to 47 percent from 1970 to 1981. ... Of total public expenditures for health care in 1980 and 1981, 87 percent was spent on curative care."⁵

Such problems, difficult as they are to solve, will only be exacerbated by allowing the market to allocate resources. Although the inverse care law persists even in health systems with extensive public sector involvement, the problem is one for the public sector to solve. The United States of America (USA) attempted to solve aspects of this problem within the context of a free-market approach to medical care. The results were less than favourable.

In 1965 the Medicaid and Medicare Acts were passed in the USA. Medicaid was for the poor, and Medicare for the aged. These two Acts represented the government's attempt to resolve the problem of unequal access to health care experienced by people disadvantaged financially, on account of income and age. Previously medical facilities were made available almost exclusively on an ability-to-pay basis. Medicaid and Medicare made the government responsible for the payment of all the medical requirements for those people qualifying for such assistance. However, these services had to be bought from the private

4 World Bank (a), p.23.

5 World Bank (a), p.23.

sector which solved one problem but intensified another.

*"The euphoria of equality was short-lived as the general inflation of the economy affected the health services economy as well. The inflation in the health services economy was exceeding that of the general economy exacerbated by increased use by elderly and poor. It should be noted that the increased use by the elderly and poor was intended, but when the price tag for greater equity came in the Federal and State governments were shocked."*⁶

The difficulties that arose in the USA with Medicaid and Medicare were the result of perverse incentives created within the medical care profession itself. Through the existence of third-party payers, coupled with the asymmetry of information between patient and doctor, ample opportunity and motive exist to "cheat" the system. Patient confidentiality, furthermore, prevents third-party payers from being able to scrutinize the actions of physicians. The following section discusses the problems created by perverse incentives within the private health sector.

2.2.2 Cost and Expenditure Increases in Health Care Spending, the Problem of Perverse Incentives

Prices for health services, in most countries, have a tendency to rise faster than their general inflation rates. At the same time, total expenditure on health services has a tendency to grow faster than Gross National Product (GNP). Table 2.1 shows the level of per capita income and the percentage of household expenditure for various countries, ranging from low income to industrial market economies. The prominence attached to medical-care expenditure clearly becomes greater as countries become more wealthy.

6 O. W. Anderson, "Issues in the Health Services of the United States", Success and Crisis in National Health Systems, eds. M. G. Field J. B. McKinlay, Routledge, New York, 1989, pp.49-69, p.62.

Table 2.1

PERCENTAGE OF TOTAL HOUSEHOLD CONSUMPTION SPENT ON MEDICAL CARE

	% Household Expenditure On Health	GNP Per Capita (\$ 1984)
Low Income Countries		
Ethiopia	3%	110
Malawi	2%	180
Zaire	3%	140
Mali	5%	140
Madagascar	1%	260
Tanzania	1%	210
Upper Middle Income Countries		
Brazil	6%	1720
Malaysia	5%	1980
Mexico	5%	2040
Uruguay	6%	1980
Hungary	5%	2100
Industrial Market Economies		
Spain	7%	4440
New Zealand	9%	7730
Italy	10%	6420
Sweden	11%	11860
United States	14%	15390
Switzerland	15%	16330

World Development Report 1988, Published by Oxford University Press. p.33.

In the USA the major increases in costs and expenditure coincided with the introduction of large-scale third-party payment schemes for medical services. The reasons for Switzerland's high percentage of household expenditure going to medical care is similar to that of the USA. The Swiss system is

characterized by a high degree of decentralized decision-making, coupled with reimbursement of medical costs by various schemes. Very small areas are entrusted to the state.⁷

However, as countries become wealthier, various other factors contribute to the increased expenditure on health care. They are as follows:⁸

- 1 Demographic changes resulting in an ever older population (the aged require considerably more health care than the young);
- 2 The high degree of utilization of services (largely prompted by supplier induced demand; see Chapter Six);
- 3 The increasing physician density;
- 4 The technology revolution in health care that has started and is not controlled; and
- 5 The increasing demand and expectations of a public partly stimulated by growing mass-media attention.

A country such as Korea, which grew exceedingly fast, found itself experiencing similar problems to those found in the USA and Switzerland. Due to their private-sector orientation, price and expenditure increases have been very large. As a proportion of GNP, health-care expenditures have increased from 2,5% in 1970

7 P. Lehman, F. Gutzwiller, J. F. Martin, "The Swiss Health System: the paradox of ungovernability and efficacy", Success and Crisis in National Health Systems, eds. M. G. Field, J. B. McKinlay, Routledge, New York, 1989, pp.25-48.

8 P. Lehman et al, p.45.

to 5% in 1989.⁹ If the average annual growth in Gross Domestic Product (GDP) for the period 1965 to 1980 of 9,9%, and for the period 1980 to 1986 of 8,2%¹⁰ is taken into account, there must have been a significant increase in health expenditure. Korea has an income elasticity of demand for health of about 1,6.¹¹ The World Bank provides the following explanation for this situation:

*"One reason why income elasticity of demand for health services has been so high in Korea in recent years is the growing availability of health insurance. Health expenditures show a clear break in trend around 1977 when health insurance was introduced; since then expenditures have grown at a much faster rate than in the pre-insurance period ..."*¹²

The increased expenditure in Korea has mainly been the result of higher utilization rates, rather than fee increases for doctors or for medical services. According to research into utilization patterns, insurance has led to an approximate 80% increase in utilization. Insured patients also show a tendency to use large general hospitals, even for primary care.¹³ The resulting expansion of expensive hospital services causes major increases in expenditure without costs themselves having to rise significantly.

Much of the cost and expenditure problem can be attributed to methods of financing health care. Medical aid schemes, and private uncovered patients, rely on fee-for-service methods of payment. Fee-for-service payment creates an incentive for the medical profession to increase turnover. Furthermore, if a particular private sector medical establishment has invested in expensive diagnostic equipment (expensive to buy and

9 World Bank (b), Korea Health Insurance and the Health Sector, Country Department 11, Asia Regional Office, Document of the World Bank, June 14, 1989, RP/No 7412-KO, par.13, p.iii.

10 World Bank (c), World Development Report, 1988, Oxford University Press.

11 World Bank (b), p.iii, par.15.

12 World Bank (b), p.iii, par.16.

13 World Bank (b), p.iii, par.17.

expensive to run), patients can be subjected to unnecessary procedures in order to recoup the costs of the investment, and subsequently to make a healthy profit.

The USA, with its extensive private health sector, has been at the forefront of trying out methods to make market-related systems work. Many states passed certificates of need laws to regulate the supply of hospital beds. Rate review laws were also passed to regulate the charges hospitals made for their services, ie. price control. Also by the early 1980s the Federal Government froze physicians' fees for Medicare and the individual States did the same for Medicaid. There were 467 Diagnostic Related Groups (DRGs) established in the early 1980s by the Federal Government. A DRG is a group of procedures related to a particular diagnosis which requires similar usage of resources. The setting of average length-of-stay standards and average cost per case was thus facilitated, for certain diagnoses. If hospital costs exceeded the average national costs they or the patients paid the excess; if the cost was less hospitals kept the surplus. However, it is very complex and cumbersome to maintain DRGs. Furthermore, it leads to premature discharging of patients (often elderly). In fact DRGs involve an extremely stringent form of direct control that can negatively influence the actions of physicians.¹⁴

Another method used to contain cost and expenditure increases in the USA has been the introduction of Health Maintenance Organizations (HMOs) (also termed "group practice prepayment plans"). Prepayment to an HMO (usually a fairly comprehensive medical facility) creates an incentive to provide only necessary services. Under-treatment is contained (theoretically) by means of HMOs competing for members with other types of scheme, or other HMOs. Approximately 12 percent of the USA population are members of such organizations. Although HMOs are part of a market-oriented health model, they also exercise a significant degree of control over the discretion and actions of individual physicians. Also HMOs are more exposed to the risk of market failure than normal private hospitals. In addition, the risk of exposure to

14 O. W. Anderson, p.63.

adverse selection¹⁵ often leads to patients most in need of regular medical treatment being refused membership.

Despite the many methods used to control medical costs and expenditure in the USA, costs and overall expenditure have continued to rise. Evans et al¹⁶ makes the following comment on the American system from an article in the *New England Journal of Medicine*:

"So far, however, market forces have been less successful in containing the growth of health care expenditures than were the regulatory efforts of the 1970s¹⁷. The one major success in this field, prospective payment and diagnostic related groups, is virtually a pure regulatory intervention, despite its being occasionally clothed in market rhetoric. At the same time, the proportion of the American population with no insurance coverage or grossly inadequate coverage is believed to be both large and growing, and there is increasing uneasiness about the effect of market forces in health care on the interests of both patients and providers."

There is therefore strong evidence to suggest the necessity for an approach to providing health care that is not determined by market forces. The choice lies between broad centralized control in a nationalized health system, and stringent micro-level control in a private sector system.

-
- 15 Patients who know that they are going to require a lot of medical treatment (especially chronic cases and the aged) are happy to belong to any prepayment or insurance-type plan, as this lowers their total costs of medical treatment. People likely to require minimal medical care would not necessarily wish to belong to a medical scheme, as the prepayment or premium would exceed what they would pay for medical treatment without being a member.
- 16 R. G. Evans, J. Lomas, M.L. Barer, R. J. Labelle, C. Fooks, G. L. Stoddart, G. M. Anderson, D. Feeny, A. Gafni, G. W. Torrance, W. G. Tholl, "Controlling Health Expenditures - The Canadian Reality" *New England Journal of Medicine*, special article, March 2, 1989, vol.320, no.9, pp.571-577.
- 17 Although these price controls were successful they eventually had to be removed due to objections from the medical profession. After being lifted, however, price movements soon made up for lost ground. (Author's note)

2.3 Centralized Control

Nationalized health systems such as exist in the United Kingdom and New Zealand, show a much greater ability to contain both health costs and expenditures. They are also able to define overall health objectives, and carry them out. An example is the Resource Allocation Working Party (RAWP) formula, to be described in Chapter Six, whereby the health authorities in the United Kingdom were able to considerably reduce inequality in the distribution of health resources.

Cost control is also facilitated by virtue of capitation-based funding (ie. funds are allocated according to population size for which a particular health facility or authority is responsible, rather than on a fee-for-service basis).¹⁸ There is therefore an incentive to conserve (on the part of doctors) rather than over-treat.

An interesting variation of centralized control is the health system of Canada. It has managed to address many of the inconsistencies that are apparent in the USA system, without creating significant negative externalities. Prior to 1971 the Canadian health system was similar to that of the USA. Health care expenditure consumed virtually identical shares of national income in both countries (7,4% in Canada and 7,6% in the USA). However, by 1987 the percentage share for Canada stood at 8,6%, whereas for the USA estimates were in excess of 11%.¹⁹

Evans et al²⁰ attribute this difference to the form of funding adopted by Canada from 1971 onwards. The Canadian system involves the use of national health insurance which is used to finance health services for the entire population, without any user-charging. In other words, there is universal coverage where the

18 The prepayment method used to fund HMOs is also a method of capitation-based funding.

19 R. G. Evans et al, p.571-572.

20 R. G. Evans et al, p.571.

funding of the health system is placed under central control. The resulting benefits of the system can be summarized as follows:

*"... overall health expenditures have been constrained to a stable share of national income and universality of coverage (without user charges) eliminates the problems of uncompensated care, individual burdens of catastrophic illness, and uninsured populations."*²¹

According to Evans et al,²² virtually the entire expenditure saving can be accounted for by three components:

- 1 Insurance overhead, or costs of prepayment and administration ;
- 2 Payments to hospitals; and
- 3 Payments for physicians' services.

In fact Canadians spent less per capita to administer universal comprehensive coverage than Americans spent to administer Medicare and Medicaid alone. Furthermore, once a tax system²³ is in place the additional costs of raising more funds are minimal.²⁴

"There are no marketing expenses, no costs of estimating risk status in order to set differential premiums or decide upon whom to cover, and no shareholder profits; the process

21 R. G. Evans et al, p.571.

22 R. G. Evans et al, p.572.

23 In a system of universal national health insurance, the "premium" paid is nothing more than an earmarked tax. It should also be noted that in Canada coverage is not conditional upon payment of the "premium". The "premium" is income related and not risk related, ie. the earmarked tax is progressive.

24 R. G. Evans et al, p.573.

*of claims payment, although not free of costs, is greatly simplified and much cheaper. In this area it is obvious that the public sector is more efficient and less costly than the private sector... .*²⁵

In Canada there is also centralized control of capital expenditures. New facilities, equipment, major renovations, etc. can all be funded from a variety of sources which include private loans. However, all these expenses require the approval of the Ministry of Health through the provincial agency, which usually contributes the major share of financing.²⁶ Controlling these expenditures has resulted in the prevention of excessive expansion of expensive (to run) facilities. Furthermore, annual global budgets are also negotiated between the ministries and individual hospitals. It should be noted, however, that individual physicians in the Canadian system are not subjected to any substantial intervention by hospital management or third parties. Canadian physicians are actually far more autonomous than their American counterparts. In the USA system:

*"In the absence of a centralized bargaining mechanism between physicians and payers - a single source payer - there is no way to limit the overall levels of billings. Hence the leap from constraints on capacity (as by a certificate of need) straight to the level of minute scrutiny of the behaviour of individual physicians or the treatment of specific diseases."*²⁷

The advantages of centralized control of health care provision are many. However, it should be noted that many predominantly nationalized health systems do not always function smoothly. One of the major problem areas occurs in straight-forward tax financed systems (ie. there is no earmarked tax). The availability of revenue for financing the health budget is often at the discretion of politicians with more on

25 R. G. Evans et al, p.573.

26 R. G. Evans et al, p.573.

27 R. G. Evans et al, p.576.

their minds than medical care. Arbitrary reductions in budget allocations can create severe complications in a nationalized health system, or sector, as is occurring in Britain, New Zealand and South Africa. Arbitrary cuts in allocations very rarely take into account priorities, in terms of manpower requirements and capital expenditure. The benefit of the Canadian system is that the tax is earmarked for health, allowing greater stability in health care financing. It is also feasible that tax-payers would not object to increases in the "health tax", if more funds were required, as people would be aware of where the money is to be spent. Increases in general taxation, on the other hand, usually involve some political sacrifice on the part of the government of the day.

2.4 Conclusion

Much of the evidence suggests a need for some form of centralized control in the provision of health services, regardless of the prevailing economic system, or the wealth of the country. Allowing the free market to allocate health resources fails to account for: externalities; the needs of underdeveloped or developing communities; the needs of the poor and the aged; and the perverse incentives generated by the methods of funding health care. Rutton²⁸ summarizes the choices as follows:

"... it is useful to distinguish between the "needologist" paradigm and the liberal market perspective when describing possible future roads to cost containment²⁹ in the health care sector. In both options one encounters difficult problems. In the former there is the complex problem of developing an efficient public choice mechanism. The latter confronts the government with the problem of having to devise incentive schemes on the basis of incomplete

28 F. Rutton, "Invited discussion" A. Griffiths, Z. Bankowski, Economics and Health Policy, XIIIth CIOMS Round Table Conference, Published jointly by the Council for International Organization of Medical Services and the Sandoz Institute for Health and Socio-economic Studies, Geneva, 1980, pp.266-269, pp.268- 269.

29 Although only cost-containment is mentioned here, the relevance of this quotation extends to all aspects of inefficiency emanating from national health care delivery systems.

information on the behaviour of health care providers."

This thesis adopts the "needologist" paradigm. The view is taken that it is possible to define an efficient public choice mechanism that would be an improvement on market related options. Although a comprehensive analysis of these issues is beyond the scope of this thesis, some of the major aspects have been highlighted in order to substantiate why particular methodologies have been chosen for evaluation in the rest of the thesis.

CHAPTER THREE

HEALTH PRIORITIES

3.1 Introduction

It "... remains a valid observation that planning for the health sector in developing countries retains its enclave character. The central planning authorities have failed to build into their models the acknowledged interactive relationship between the health sector and other sectors, and have failed to recognize these relationships in making resource allocations to health. Nor have the health planners been able to elucidate the production function for health, or to specify the contribution to the social welfare function in a manner that convincingly substantiates their claim for more resources. The outcome of this mutual incomprehension is that, in the great majority of cases, planning for the health sector is an independent exercise in sub-optimization, carried out by health sector professionals within resource constraints defined in an essentially arbitrary manner by central economic planners."¹

The severe attitude of this author toward central planning authorities reflects the frustration that economists, working in the field of health, feel about the conventional wisdom which is retained when approaching problems concerning health improvement within a country or community. The inability of health planners to define effectively the underlying relationship between medical, social, and economic variables and health status results in a reliance on a "fire-fighting" approach to dealing with health problems.

1 M. Wheeler, "Health Sector Planning and National Development Planning", The Economics of Developing countries, eds. K. Lee, A. Mills, Oxford University Press, 1983, pp.190-207, p.190.

This chapter is concerned with an examination of the contribution made by various variables toward improving health status. In this respect a model, developed by Grossman², will be presented. This model illustrates a more incisive view of health and its contributing elements. Although the model itself is not tested here³, its assumptions concerning what can be termed a human capital approach to health are examined further. This is in the form of a regression analysis.

The regression analysis is presented in three stages. The first is in the form of a bivariate analysis where the individual relationship between life-expectancy (the dependant variable) and selected variables are examined. The next two stages involve multivariate relationships between the variables and life-expectancy where eventually a simple model is developed which can go some way toward defining appropriate policies for improving the health status of a developing country. In the bivariate analysis a comparison is made between low-income and upper middle-income countries as defined by the World Bank. According to the World Bank South Africa is defined as an upper middle-income country.⁴ However, a significant percentage of the population live in conditions consistent with a low-income country (see Chapter Four). It is for this reason that low-income countries are focused on in this section. The upper-middle income countries are included as they indicate, to a certain extent, how different variables influence health status as countries get wealthier. The multivariate regression analysis only concerns low-income countries. The statistical package used was TSP version 4.

2 R. J. Vogel, N. T. Greenspan, "Health Planning in the Sudan", Human Capital and Development, eds. Salkever, Sirageldin, Sorkin, JAI press inc. London England, 1983, vol.3, pp.271-309, pp.271-278.

3 For a fairly complete empirical examination of Grossman's model see: A. Wagstaff, "The Demand for Health, Some New Empirical Evidence" Journal of Health Economics, 1986, Vol.5, pp.195-233.

4 World Bank, World Development Report 1988, Oxford University Press.

3.2 The Economic Theory of Health Investment and Planning

In examining demand for various health-care resources Grossman came to the conclusion that what was being demanded was not the resources but healthy time. Health is thus viewed as a durable capital stock that produces an output of healthy time. Medical care is only one of the variables entered into the production function for healthy time and, therefore, the demand for medical care is likewise a derived demand. Housing, diet, income, and environmental variables are also part of the production function for the output of healthy time.⁵

The gross investment production function for health in period (i) is as follows:

$$I_i = M_i g(t_i; E_i) \quad (1)$$

Where:

- I = gross investment
- M = medical care
- t = healthy time divided by medical care
- E = the stock of human capital

The marginal monetary rate of return on an investment in health is given by:

$$Y_i = (W_i G_i) / C_i - 1 \quad (2)$$

Where:

- Y = the monetary rate of return on an investment in health or the marginal efficiency of health capital (MEC),
- W = the wage rate

⁵ R. J. Vogel et al, p.272.

$G =$ the marginal product of health capital

$C =$ the marginal cost of gross investment (I) in health

Setting the marginal utility of healthy days or the marginal disutility of sick days equal to zero, the optimal amount of health capital in period (i) can be written as:

$$W_i G_i / C_i - 1 = Y_i = r - C_i - 1 + D_i \quad (3)$$

Where: $r =$ the rate of interest

$D_i =$ the rate of depreciation of the health stock in period i.

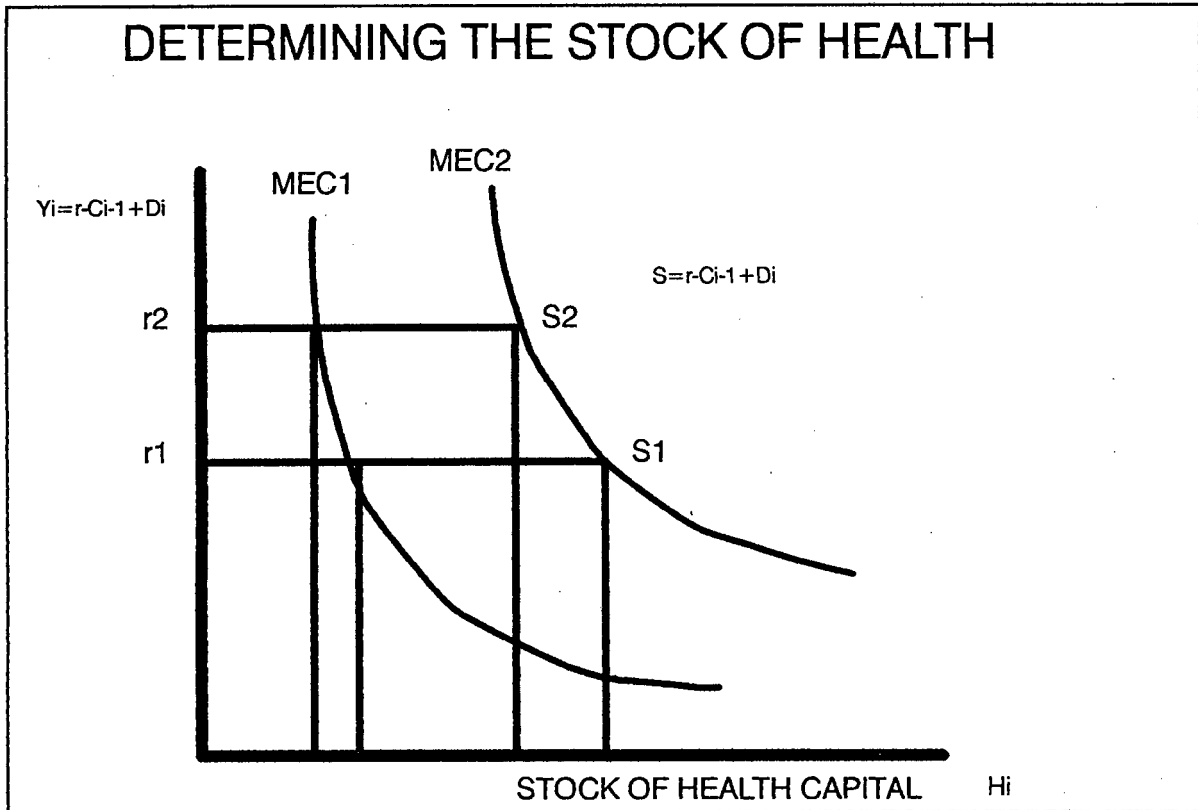
The MEC is the demand curve indicating the relationship between the stock of health and the rate of return on an investment in health. The supply curve $S=r-C_i-1+D_i$ indicates the relationship between the stock of health and the cost of health capital. The supply curve is infinitely elastic because the cost of capital is independent of the stock.

The model assumes the following:

- a individuals inherit an initial stock of health capital that is subject to depreciation over time over the life cycle, and can be increased by investment; and
- b the MEC curve slopes downwards because of the diminishing marginal productivity of health capital.

Thus, as age increases, so the rate of depreciation on the stock of health causes upward shifts in the supply schedule. Increases in the wage rate, education and socioeconomic factors in general, cause outward shifts in the MEC schedule.

Figure 3.1



This diagram is based on: M. Grossman, "On the Concept of Health, Capital and the Demand for Health," *Journal of the Political Economy*, March/April 1972, vol.80, no.2, p.232 & 242; and R.G. Vogel, N.J. Greenspan, "Health Planning in the Sudan", *Human Capital and Development*, eds. Salkever, Sirageldin, Sorkin, JAI press inc., London England, 1983, vol.3, pp.271-309, pp.274 & 275.

If the model is used to examine a household, direct inputs into the household production function for health would include such things as: own time of the consumer, medical care, diet, exercise, recreation and housing. In the case of a developing country, environmental variables such as purity of the water supply,

sanitation facilities, bathing water, adequate food supplies, etc. should be included. Negative externalities would be expected to shift the MEC curve inward. Thus a developed, first-world community, would be expected to be in the position MEC2 in figure 3.1, with a higher rate of return on investment at all levels of the stock of capital.

On the supply side, anything that would result in a shortage of capital in general, or a small initial endowment of health stock, (for example due to a mother's poor prenatal care), would cause the supply schedule for health stock to rise. This can be seen in figure 3.1 as a movement in the supply schedule from S1 to S2.

According to this model, it is obviously in the interests of a developing country to shift its MEC schedule outwards as far as possible, thus increasing the efficiency of attaining a given stock of health. The problem comes in establishing which variable or variables to manipulate in order to best achieve this goal at least cost. Should the focus be on increasing GNP or education, or on environmental factors or medical facilities? The following sections therefore examine some of these variables in more depth.

3.3 Bivariate Analysis

The following variables were selected to be regressed against life-expectancy. For elaboration on the data used see appendix A.

a Dependant variable

Life Expectancy (1986)

b Medical Indicators

Population per doctor (1981)

POPDOC

Population per nurse (1981)

POPNUK

Government expenditure on health (1986)

GOVHTH

Private expenditure on health (1986)

CONMED

Total expenditure on health (1986)

MEDEXP

c Social Indicators

Total expenditure on education (1986)	EDUEXP
Private expenditure on education (1986)	CONED
Government expenditure on education (1986)	GOVED
Government expenditure on social services (1986)	GOVSOC
Daily per capita calorie intake (1985)	CALOR
Infant Mortality Rate (1986)	IMR
Crude Birth Rate per 1000 population (1986)	BIRTH
Total Fertility Rate (1986)	FERT
Urbanization Rate (1980-85)	URB85
Urban population as a percentage of total (1985)	URBPOP
Percentage children in primary education (1985) ⁶	EDUPT
Percentage females in primary education (1985)	EDUPF
Percentage males in primary education (1985)	EDUPM
Percentage children in senior education (1985)	EDUST
Percentage females in senior education (1985)	EDUSF
Percentage males in senior education (1985)	EDUSM
Percentage population in tertiary education (1985)	EDUTT

d Economic Indicators

Per Capita Gross National Product (GNP) (1986)	GNPPC
Average annual growth rate in GNP (1965-86)	GNPGR

3.3.1 Results

Table 3.1

RESULTS OF ORDINARY LINEAR REGRESSIONS

MEDICAL VARIABLES						
Variable	Coef	Adj R2	Tstat	Coef	Adj R2	Tstat
	Upper-Middle-Income Countries			Low-Income Countries		
POPDOC	-0.001	0.221	-3.270	-0.0002	0.295	-3.450
POPNUK	-0.007	0.398	-4.340	-0.002	0.230	-2.970
CONMED	0.301	-0.067	0.560*	0.059	0.190	2.676
GOVHTH	1.062	0.039	1.470*	0.418	0.197	2.710
MEDEXP	0.342	-0.033	0.800*	0.057	0.210	2.850

6 All the percentages relating to education refer to the percentage of children of the relevant school-going age who are enrolled at school.

SOCIAL VARIABLES						
GOVSOC	0.674	0.306	3.714	0.205	0.249	3.099
CALOR	0.008	0.293	4.193	0.011	0.388	4.183
IMR	-0.207	0.886	-17.692	-0.197	0.792	-10.012
BIRTH	-0.492	0.685	-9.382	-0.577	0.573	-5.991
CONED	0.470	-0.045	0.727*	0.069	0.147	2.341
EDUEXP	-0.114	-0.088	-0.334*	0.045	0.152	2.381
FERT	-3.135	0.665	-8.960	-3.449	0.561	-5.845
EDUTT	0.102	-0.026	0.851*	0.285	0.102	1.989
URBPOP	0.184	0.356	4.807	0.240	0.249	3.099
URB85	-3.576	0.427	-3.031	-0.577	-0.022	-0.666
EDUSM	0.158	0.290	3.966	0.192	0.274	3.287
EDUSF	0.200	0.504	6.130	0.287	0.591	6.214
EDUST	0.191	0.420	5.208	0.264	0.476	4.958
EDUPM	-0.037	-0.023	-0.446*	0.166	0.415	4.407
EDUPF	0.062	-0.012	0.766*	0.151	0.481	5.005
EDUPT	-0.005	-0.028	-0.056*	0.163	0.467	4.874
GOVED	-0.602	0.028	-1.355*	0.096	0.115	2.090
ECONOMIC VARIABLES						
GNPPC	0.002	0.134	2.682	0.014	0.473	4.931
GNPGR	0.227	-0.090	0.297*	1.350	0.189	2.656
* Not significant at the 95% confidence level						

The number of observations for all low-income country variables is 27. For middle-income countries the number of observations for many of the variables are different due to the data not being complete for all the countries examined. The number of observations for each variable for middle-income countries are as follows:

POPDOC (35), POPNUR (28), CONMED (12), GOVHTH (30), MEDEXP (12), GOVSOC (30), CALOR (41), IMR (41), BIRTH (41), CONED (12), EDUEXP (12), FERT (41), EDUTT (12), URBPOP (41), URB85 (12), EDUSM (37), EDUSF (37), EDUST (37), EDUPM (37), EDUPF (37), EDUPT (37), GOVED (30), GNPPC (41), GNPGR (12).

Table 3.2

THE CHANGE IN THE VARIABLE REQUIRED TO RESULT IN A TEN YEAR IMPROVEMENT IN LIFE-EXPECTANCY

Variable	R2	Coef	Change	R2	Coef	Change
	Upper-Middle-Income Countries			Low-Income Countries		
IMR	0.90	-0.200	-50	0.79	-0.200	-51
GOVSOC	0.30	0.670	15	0.25	0.200	49
POPNUK	0.40	-0.007	-1 450	0.23	-0.002	-5 000
POPDOC	0.20	-0.001	-8 264	0.30	-0.000	-50 000
URBPOP	0.35	0.180	54	0.25	0.240	42
EDUSM	0.30	0.160	63	0.27	0.200	52
EDUSF	0.50	0.200	50	0.60	0.300	35
EDUST	0.40	0.200	52	0.48	0.260	38
FERT	0.70	-3.130	-3.2	0.56	-0.450	-2.9
BIRTH	0.70	-0.500	-20	0.60	-0.600	-17
CALOR	0.30	0.008	1 252	0.40	0.010	909
GNPPC	0.13	0.002	6 289	0.50	0.014	694

3.4 Discussion of Results

3.4.1 Medical Variables

The results indicate a fairly weak relationship between the medical variables and life-expectancy. In upper-middle-income countries the variables CONMED (private consumption expenditure on medical goods and services), GOVHTH (government expenditure on health) and MEDEXP (total expenditure on health) are all insignificant. In low-income countries the above variables all have adjusted R2s of around 0.2, and are significant. The more interesting variables, POPDOC (ratio of population to doctor) and POPNUR (ratio of population to nurses), appear more significant in both upper-middle- and low-income countries. However, the coefficients, which have the expected signs, are rather small.

In table 3.2 it is shown what the change in the X variable (in this case POPDOC and POPNUR) would

need to be in order to affect a ten year improvement in life-expectancy in the two types of country examined. The results in table 3.2 were based on the regression results indicated in table 3.1. In the case of POPDOC the regression results suggest that in order to achieve a ten year improvement in life-expectancy, a low-income country would have to decrease the doctor patient ratio by 50 000, and in upper-middle income countries this ratio would have to decrease by 8264. Although this indicates a more effective relationship between health status and medical indicators in upper-middle income than in low-income countries, decreasing the ratios by the suggested amounts will result in both types of country having doctor/patient ratios of more than one doctor per person.

Examining POPNUR in this connection reveals similar results. Therefore, it appears that these medical variables are significant in a statistical sense, but quite weak in their influence on health status. However, medically related investments seem to achieve greater gains in upper-middle-income countries than they do in low-income countries.

3.4.2 Social Variables

As several "social variables" were included in the study, only the more interesting results are discussed.

Government expenditure on socioeconomic goods and services:

Of all the various expenditure variables examined, the most significant is GOVSOC (government expenditure on socioeconomic goods and services).⁷ It has an adjusted R² of 0.3 in upper-middle income

7 This variable is more fully defined in the World Development Report of 1988 as follows:

"... public expenditure on housing, such as income - related schemes; on provision and support of housing and slum clearance activities; on community development; and on compensation to the sick and temporarily disabled for loss of income; on payments to the elderly; the permanently disabled, and the unemployed; and on family, maternity, and child allowances. They also include the cost of welfare

countries, and 0.249 in low-income countries. According to table 3.2, in order to achieve a ten year gain in life-expectancy in a low-income country it would require an increase in this type of expenditure equivalent to 49% of GNP, while in upper-middle income countries only an increase equivalent to 15% of GNP is required. Although the full implementation of what is suggested here is not really feasible, the result does indicate the premise suggested in Grossman's model, that the marginal efficiency of capital invested in health improves, the further the MEC curve moves outward. Such spending can be an important method of improving the efficiency of health inputs such as medical goods and services.

Calorie intake:

The variable CALOR (average daily per capita intake of calories), is shown to be significant in both upper-middle- and low-income countries. In low-income countries CALOR is more highly correlated with life-expectancy (0.388) than is the case with upper-middle-income countries (0.29). According to table 3.2, it will take a lower average daily per capita increase in calories (909 calories) to achieve a ten year improvement in life-expectancy, as opposed to the 1251 calories increase required in upper-middle-income countries.

CALOR can effect health status in many ways. Increases in calorie intake for young children, pregnant women, and mothers nursing children, can improve the health status of the affected children into adulthood. These results suggest that impressive gains can be made in improving the health status of a low-income population by improving calorie intake.

services, such as care of the aged, the disabled, and children; as well as the cost of general administration, regulation and research associated with security and welfare services."

Urbanization:

The percentage of population urbanized has an R2 of 0.356 for upper-middle income countries, and 0.25 for low-income countries, when regressed against life-expectancy. In both cases the coefficient is positive, and the regression is significant. According to table 3.2 it will take a 54.4% increase in overall urbanization to increase life-expectancy by ten years in upper-middle income countries, whereas in low-income countries a 42% increase will achieve the same. It is suggested that urbanization contributes to life-expectancy, and that it contributes more to life-expectancy in low-income countries than it does in upper-middle income countries.

These results therefore indicate the potential benefits to be gained from allowing populations to concentrate. The provision of basic social goods and services are made cheaper and can reach more people. The relationship between life-expectancy and urbanization should be regarded as complex, as improvements in health conditions achieved through this process are actually associated with the enhanced ability of other variables (eg. education, housing, running water, medical facilities, etc.) to improve health. This is an important policy variable with special significance for South Africa. The transition to a more developed and healthier country can be greatly enhanced if more attention is given to the urbanization process. Emphasis should be placed on ensuring that facilities of a socioeconomic nature are provided where needed as this process gains momentum.

Infant mortality:⁸

The infant mortality rate (IMR) shows a significant correlation with life-expectancy, with an R2 of 0.9 for

⁸ The results obtained here should be regarded with some caution as the World Development Report includes IMRs in the calculation of their life-expectancy values. For a discussion on the reliability of international cross-sectional mortality see C.J.L. Murray, "A Critical Review of International Mortality Data", *Soc. Sci. Med.*, Vol.25, No.7, 1987, pp.773-781.

upper-middle income countries, and around 0.8 for low-income countries. Table 3.2 indicates that a ten-year improvement in life-expectancy would be associated with a drop in the IMR of around 50 (per 10 000 live births), for both types of country. However, this does not mean the same thing for both countries. For low-income countries this would reduce the infant mortality rate to between 60 and 70 (per 10 000 live births), while in upper-middle income countries the same drop would have leave IMRs at under 10 (per 10 000 live births) on average. The relationship between IMRs and life-expectancy cannot be regarded as direct. Rather, conditions which affect the one variable, probably affect the other to a similar extent.

Fertility rates:

The above comments could be applied to fertility rates as well. Fertility rates and birth rates are both highly correlated with life-expectancy in both upper-middle and low income countries. For the sake of convenience only fertility rates are discussed here. A direct relationship between fertility and health status is most unlikely as evidence indicates that variables affecting fertility are complex. The suggestion that causality leads from a drop in fertility to an increase in income, and consequently in health status, should be regarded with caution, especially when there is evidence to the contrary.^{9 10 11} Due to the importance of the health status - fertility relationship, it is worth discussing it in more depth.

Among the variables found to be highly correlated with fertility rates are female education levels.^{12 13}This

9 R. Repetto, Economic Equality in Developing Countries, John Hopkins University Press, Baltimore and London, 1979.

10 S. Findley, A. C. Carr, Patterns of Urban-Rural Fertility Differentials in Developing Countries, General Electric Company, Tempo Centre for Advanced Studies, California Washington D. C., 1978.

11 S. Scrimshaw, "Infant Mortality and Behaviour in the Regulation of Family Size", Fertility Decline in Less Developed Countries, eds., N. Eberstadt, Praeger, USA, 1981.

12 R. Repetto, p.113.

is an interesting result considering the strong correlation in the present study between female education levels and life-expectancy. Income levels also appear to play a significant role in the reduction of fertility rates.¹⁴ Repetto elaborates as follows:

*Income "... changes affect fertility in the long run through their effects on the demand for other goods and services and through the reallocation of household resources. Once these indirect affects are taken into consideration ... income effects in fertility appear stronger. In particular, the results illustrate the point that income growth at low income levels raises household demand for education and the like. It follows, therefore, that interventions need not operate only by increasing the supply of schools and other services. Interventions that add to incomes of low-income households will raise the demand for education and other services, and thus will always be effective. The impact of income, educational, and occupational variables supports the thesis that greater socioeconomic equality results in lower population fertility rates."*¹⁵

The relationship between the basic household unit and the economy can also be regarded as being of primary importance in determining fertility levels. If the household remains the primary production unit of the economy, the incentive to have large families will be the rational choice of the parents. If, on the other hand, the production process is separated from the household, ie. into firms, where children are unable to participate in the production process and become a burden to their parents, fertility rates could be expected to decline, as it would no longer be rational to have large families.

13 S. Findley et al, p.53.

14 R. Repetto, p.116.

15 R. Repetto, p.116.

It has also been asserted that there is a relationship between IMRs and fertility. Reppetto¹⁶ found, in a time series analysis, that the fall in child mortality contributed to the decline in fertility experienced in Korea over the past 20 years. The coefficient was weak but the reliability was high. Reppetto asserts that there is a direct relationship between infant mortality and fertility rates. This suggests that a large part of the explanation for high fertility rates lies in the replacement factor, or in the anticipation of a certain number of children dying. However, this view is disputed by Scrimshaw¹⁷ who takes the view that high infant mortality may in fact be the result of parents deliberately manipulating family size to suit their ends, ie. a form of crude birth control.

Parents "... will not always go to extremes to save the life of a child, and may in fact invest more time, attention, and resources in some children than in others. Their actions directly influence ultimate family size as some children have a lower probability of surviving when they are relatively unwanted, even subconsciously."

This establishes the possibility that family size is a matter of choice, albeit sometimes subconscious. Focusing merely on infant mortality as the policy variable may achieve very little in the way of both decreased mortality, and decreased family sizes. High fertility rates are probably neither a cause, nor a result of high infant mortality. Both are symptomatic of the socioeconomic environment. The following observation illustrates this point.

"The correlation between education and fertility is weak among rural women because the first act undertaken to enhance her own or her children's chances for upward mobility is often migration to the city, leaving behind women with low aspirations. This may also explain why

16 R. Reppetto, p.115.

17 S. Scrimshaw, p.310.

*over three-quarters of those with some secondary education live in the cities: either they moved so they could be educated at secondary level, or having secondary education their aspirations could only be realized in the city. Once in the city, a woman's aspirations motivate her to find out about and use methods to limit her family size to a number she can raise in the way she wants.*¹⁸

It is therefore possible that the above discussion is part of the explanation for the failure of attempts to exogenously determine family sizes, ie. through family planning programmes experienced in many countries.¹⁹ Merely dealing with the symptoms of the problem prolongs the problem and, in the case of fertility, even leaves the symptom unaffected.

Education:

Education, especially secondary school education, has strong correlations with life-expectancy. Female secondary school enrollment appears to be the most influential variable in both upper-middle- and low-income countries. This may be a reflection of the extremely important function females have in the development of their children. The better the mother is educated the better able she is to give her child a good start in life.

Table 3.2 indicates that in order to achieve a ten year improvement in life-expectancy, upper-middle-income countries would have to increase female enrollment at secondary school level by

18 S. Findley et al, p.53.

19 Peterson in an evaluation of fertility programs finds very little evidence to suggest that family planning programs affect fertility rates in any way. W. Peterson, "American Efforts to Reduce the Fertility of Less Developed Countries", Fertility Decline in Less Developed Countries, ed., N. Eberstadt, Praeger, USA, 1981, pp.337-362.

around 50%, while in low income countries a 34% improvement is indicated. Whereas such an improvement in female enrollment would take upper-middle- income countries beyond 100%, in low income countries such an improvement would not move female enrollment over 50%. In many instances female enrollment is under 10% of the population old enough to be enrolled. The same can be said for secondary education as a whole. Therefore, it appears as that improvements in this variable can have a positive effect on health status. It should be borne in mind, however, that the way to improve educational levels is as much dependant on the demand that is created for them, as it is on the availability of educational establishments.

General:

It should be noted that variables such as URBPOP, EDUSM, FERT, BIRTH, and CALOR, are more important contributors to health status in low-income countries than they are in upper-middle-income countries. However, the reverse is true of the medical variables. Explained in terms of Grossman's model, in upper-middle-income countries the stock of health capital is already very high, and further outward movements in the MEC curve are achieved only at considerable expense. On the other hand, variables that benefit from the improved efficiency in producing "healthy time", such as POPDOC and POPNUR, become more effective in upper-middle-income countries.

3.4.3 Economic Variables

The two variables examined under this heading were GNPPC (per capita gross national product) and GNPGR (growth in the gross national product). GNPPC proved to be the only significant variable for both upper-middle and low income countries. However, the R² for upper-middle-income countries is only 0.13, while for low income countries it explains life-expectancy somewhat better with an R² of 0.473. Table 3.2

indicates that in order to achieve a ten year improvement in life-expectancy, upper-middle-income countries would have to improve their average per capita GNP by \$6289. Low-income countries would only require an average GNP per capita increase of around \$694 to achieve a similar gain in life-expectancy.

Grosse and Perry²⁰, in a study including data on 90 developing countries, over different periods, found that:

"The slopes of the lines are different, with that for 1960 being somewhat steeper. This might indicate that in 1970 increases in life-expectancy as a function of economic level required higher economic level per year added than in 1960. In 1960-65, 10 years of additional life expectancy at birth "required" a gain of \$300 per capita. For 1970-75, a gain of 10 years "required" a rise of \$500 per capita, a 60% increase"

It is clear that increases in income can play an important role in improving the efficiency of investment in health capital. It is evident that the wealthier the country becomes, the less this variable contributes to improved health status.

These and the previous results suggest that improvements in income, whether provided through the guise of social goods or through enhanced earnings, will result in major improvements in the health status of populations with low incomes.

3.5 Multivariate Analysis

The multiple regression analysis produced here was performed in order to discern, in terms of a simple linear model, what the prime determinants of health status are in low-income countries. The results are

20 R.N. Grosse, H.Perry, "Correlates of Life Expectancy in Less Developed Countries", Human Capital and Development, eds. Salkever, Sirageldin, Sorkin, JAI press inc. London England, 1983, p.225.

presented in two stages. The first stage is presented in tables 3.3, 3.4 and 3.5 where the best explanatory variables for life-expectancy within the categories "medical", "social" and "economic" respectively, are shown. The second stage (tables 3.6 and 3.7) presents the final results with medical, social and economic variables considered within the same regression. In terms of an equation the model would be as follows:

$$LE = B0 + B1(\text{medical}) + B2(\text{social}) + B3(\text{economic})$$

The intermediate results are as follows:

Table 3.3

RESULTS OF MULTIPLE LINEAR REGRESSION (MEDICAL VARIABLES)

Low-Income Countries (27 Observations)

Variable	Coefficient	Tstat
C	58.932	29.959
MEDEXP	0.035	1.873
POPDOC	-0.0001	-2.065
POPNUK	-0.0008	-1.379
R2	0.478	
ADJUSTED R2	0.410	

Table 3.4

RESULTS OF MULTIPLE LINEAR REGRESSION (SOCIAL VARIABLES)

Low-Income Countries (27 Observations)

Variable	Coefficient	Tstat
C	34.890	8.025
CALOR	0.001	3.355
EDUSF	0.228	5.364
R2	0.730	
Adjusted R2	0.710	

Table 3.5

**RESULTS OF MULTIPLE LINEAR REGRESSION (ECONOMIC VARIABLES)
Low-Income Countries (27 Observations)**

Variable	Coefficient	Tstat
C	49.347	25.257
GNPPC	0.013	3.734
GNPGR	0.394	0.811
R2	0.500	
Adjusted R2	0.465	

Final Results:

Table 3.6

**RESULTS OF MULTIPLE LINEAR REGRESSION (FINAL RESULTS A)
Low-Income Countries (27 Observations)**

Variable	Coefficient	Tstat
C	40.26	8.319
POPDOC	-3.877D-05	-0.821
CALOR	0.004	1.661
EDUSF	0.189	4.471
GNPPC	0.006	2.265
R2	0.790	
Adjusted R2	0.760	

Table 3.7

**RESULTS OF MULTIPLE LINEAR REGRESSION (FINAL RESULTS B)
Low-Income Countries (27 Observations)**

Variable	Coefficient	Tstat
C	38.303	9.160
CALOR	0.004	1.930
EDUSF	0.198	4.890
GNPPC	0.006	2.473
R2	0.790	
Adjusted R2	0.760	

The results indicate the very low explanatory power of the medical variables in the variation of life expectancy between low-income countries. MEDEXP (total expenditure on medical goods and services), POPDOC (ratio of population to doctors) and POPNUR (ratio of nurses to population) proved to be insignificant. Although POPDOC was significant in the results indicated in table 3.3, it proved to be insignificant when considered with social and economic variables (final results A, table 3.6). Table 3.7 shows the best model explaining life expectancy in low-income countries when all medical variables have been excluded. The prominence of the social variables is important to note, especially EDUSF (female enrollment in secondary school). Male enrollment in secondary school was also significant, but had to be left out due to multicollinearity. Total enrollment at secondary school level also remained significant, but female secondary education was retained since it explained more of the variation in life expectancy than did all the other education variables. GNPPC remains significant, but adds very little to the overall equation.

On the whole these results are confirmed by findings in other studies of a similar nature. Grosse and Perry²¹ found that in their multivariate analysis 90% of the variation in life-expectancy in the period 1970-75 was explained by three variables in order from most to least important: literacy (social variable); urban water reticulation (social variable); and percent of labour force in agriculture (economic variable). Furthermore, it was further found that health expenditures, facilities, and personnel variables did not add to the amount of variation explained. The limited explanatory value of per capita income is also noted.

Fuchs²² in commenting on the studies performed at the National Bureau of Economic Research on interstate differences in mortality in the United States of America, mentions that variations in the number

21 R.N. Grosse, et al, p.240.

22 V. R. Fuchs, "Some Economic Aspects of Mortality in Developed Countries" The Economics of Health and Medical Care, ed. M. Perlman, Proceedings of a Conference Held by the International Economic Association at Tokyo. Macmillan, Great Britain, 1974, p.181.

of physicians per capita and expenditures for physicians' services have only a slight effect on mortality. The inclusion of the number of physicians per capita in an age-specific mortality regression across developed countries in 1956 produced similar results. It was also found that the relationship between mortality and income for adults is tending to disappear, except for those at the lowest income levels.

In studies by Newhouse and Friedlander, and Millar and Stokes,²³ the relationship between the health of people living in different areas and the quantity of medical resources was examined. Controlling for demographic and socioeconomic characteristics both studies concluded that there was virtually no relationship between health status and health resources.

3.6 Conclusion

This chapter had as its primary motive the establishment of a broad approach to the determinants of health status. Although this investigation evaluated the correlations between certain variables and health status on an inter-country basis, the intended goal was to use these results to explain events on an intra-country level more specifically South Africa.

This study found, as others have, very strong relationships between social variables and life expectancy, with medical variables only playing a minor role. Per capita GNP appears to play a fairly significant role in low-income communities. In South Africa, many socially- and economically-deprived communities stand to gain much through policies that tend to raise their income levels, both in terms of earnings and the provision of social goods and services. South Africa has for many years followed policies that have systematically deprived various communities, on a racial basis, from the ability to earn sufficient income to advance beyond a subsistence mode of existence.

23 A. C. Enthoven, "The Behaviour of Health Care Agents: Provider Behaviour" Health Economics and Health Economics, eds. van der Gaag, Perlman, North Holland Publishing Company, Amsterdam, 1981, pp.173-188.

Furthermore, racial exclusion from access to adequate education, a crucial variable in the health status equation, has further diminished the efficiency of investment in the stock of health for many South Africans. The policies that were applied are the exact reverse of those that this study, and many others like it, would suggest as appropriate for a developing country. The process of urbanization should have been assisted rather than "forcibly" resisted at great expense to the taxpayer. Education could have been provided extensively for the fast urbanizing population. In 1953 per capita annual expenditure on white education was R128 compared to around R18 for blacks. Over the next ten years, while the amount allocated to whites increased on a per capita basis, it actually decreased for blacks to around R12 in nominal terms.²⁴

"Verwoerd justified the low level of state expenditure on black schooling on the grounds that blacks should themselves be largely responsible for financing their own education. Was it not well known that "spoon-feeding" was not appreciated? The best results would be achieved if blacks helped themselves by looking after their own."²⁵

Even during the boom years of the 1960s blacks did not succeed in improving their incomes. A harsh but accurate reflection on this contradiction comes from Hepple where he suggests that Verwoerd, as Minister of Native Affairs, had achieved "something of a miracle - he had pegged the African people to poverty in the midst of the greatest economic boom in the country's history. His tight administration, aided by ferocious police power and harsh labour and apartheid laws, had successfully prevented Africans from attaining their aim of participating in the social and cultural benefits of an industrial civilization."²⁶

24 K. Kenney, Architect of Apartheid H. F. Verwoerd - an Appraisal, National Book Printers, Cape Town, 1980, p.119.

25 K. Kenney, p.120.

26 K. Kenney, p.121.

In addition, the destabilization of established communities through the process of forced removals on a massive scale, has contributed to the poverty and destitution of vast segments of the South African population. It is estimated that by 1983 the total amount of people dealt with in this manner stood at around 3 548 900 people.²⁷ Almost no compensation was given to the people whose land was expropriated in this manner, and the areas to which they were removed had, on the whole, no infrastructure, no housing, and were economically non-viable.²⁸

The inability of blacks in South Africa to determine adequate incomes both through choice of employment and collective bargaining effectively placed many employers in a position of monopsony power over their workers. Workers could only be recruited in the homelands, and were allocated jobs. If a worker lost his job, he was forced to return to his "homeland" of "origin" where he could then apply to be allocated to another job outside of the homeland. Very little work was, or is, available in the so-called homelands. The attitude of the South African Government, until recently, is reflected in the following General Circular sent out by the Department of Bantu Administration in 1967:

"It is accepted Government policy that the Bantu are only temporarily resident in the European areas of the Republic for as long as they offer their labour there. As soon as they become, for one reason or another, no longer fit to work or superfluous in the labour market, they are expected to return to their country of origin or the territory of the national unit where they fit ethnically if they were not born and bred in the homeland.(General Circular, 1967)"²⁹

The immensity of the present housing backlog (see Chapter Four) is a direct consequence of the now

27 L. Platzky, C. Walker, The Surplus People. Forced Removals in South Africa, Raven Press, Johannesburg, 1985, p.11.

28 L. Platzky et al, pp.351-368.

29 L. Platzky et al, p.65.

discarded government policy which, in 1968, decreed that family housing for blacks only be built in the homelands. Today attempts to reduce the backlog are hampered by government policies which limit the availability of land for black townships and by the inability of most township residents to afford privately-built homes. The problem can be expected to worsen considerably unless greater priority is given to providing more land and low-cost housing.³⁰

In terms of the study described in this chapter, the following elements should be considered with reference to the overall improvement of health conditions in South Africa.

An important policy variable that should be discussed is expenditure on medical goods and services. In a developing country, it is essential that facilities of this nature should be appropriate to the needs of that population. As the bulk of medical procedures respond to, rather than alter, health status, careful consideration should be given to the most appropriate method of addressing health problems. Medical facilities are very costly, on the whole, with very high recurrent expenditure required to keep them functioning adequately. Therefore it is imperative that methods be found for evaluating alternative interventions, both medical and non-medical, so that the most effective plans are implemented.

The above regression analysis revealed that government expenditure on social goods and services appeared to indicate a more significant correlation with life expectancy than either government, private or total expenditure on medical goods and services, in both low and upper-middle income countries. Clearly, therefore, where choices have to be made, it is crucial to define criteria that facilitate this kind of decision making. The later discussions of cost-effectiveness analysis (Chapter Seven and Eight) explore the possibilities surrounding the selection of appropriate criteria.

30 H. Mashabela, Townships of the PWV, South African Institute of Race Relations, Aurora Printers, Pretoria, 1988, p.13.

Other important policy variables are the following:

- a Education, especially female education, beyond primary school level is clearly indicated as important in improving overall health levels. The relationship with health is both direct and indirect. Many benefits, external to health, are also gained through improved education.
- b Adequate nutrition, especially for young children, is essential in the improvement of health status. Nutritional deficiencies at young ages retard both the physical and the intellectual development of a child. It is essential, therefore, that methods be found to ensure adequate levels of nutrition in the developing child for the entire population. School-feeding schemes could play a useful role both in ensuring the attendance of pupils from low income groups, and in improving the health status of the child.
- c Urbanization can have a significant influence in the improvement of life expectancy in developing countries. The overall level of urbanization should be allowed to increase without hindrance. The provision of basic needs for recently urbanized populations, ie. housing, water, power, and appropriate medical care, should be provided. The economies of scale generated by the concentration of the population in this way would make it easier and cheaper to provide for these needs.
- d High fertility rates, often assumed to be the cause of continued poverty in a low income country, can be regarded more as a reflection of the economic production process rather than a cause of it. The fairly strong relationship discovered between life expectancy and fertility rates should not be regarded as a direct one. Considering the fertility rate as a policy variable could prove to be a waste of time and money. The economic function of high fertility sub-groups should rather be focused upon. If this relationship can be changed, then fertility rates will decline. If it is not

changed then no change in fertility rates will take place even if one is dealing with an urbanized population. Nevertheless it is interesting to note that the total fertility rate (TFR) for blacks in South Africa varies dramatically according to area. It is 2.8 in urban areas, 4.2 in semi-rural areas, 5.7 in rural areas (which includes the non-independent homelands) and 6.3 in the independent homelands.³¹

Finally, what should be determined from this chapter is the complexity of health status. In order to effect changes in health status through direct intervention, care should be taken to avoid superficial policies that do not necessarily achieve the intended objectives.

31 South African Institute of Race Relations, Race Relations Survey 1988/89, Johannesburg, 1989 p.150. These figures were originally obtained from Hansard (R), 11, cols.6932-6935, 21 April.

CHAPTER FOUR

HEALTH IN SOUTH AFRICA: A BRIEF OVERVIEW

4.1 Introduction

The intention of this chapter is to provide a basic picture of the health status of the South African population. Most of the people in South Africa live in adverse socioeconomic conditions that increase the probability of ill health. The strong relationship between health status and socioeconomic conditions was indicated in Chapter Three. The nature of the health care system that is superimposed on such an environment is a primary concern of this thesis. As the socioeconomic conditions change, so does the need for particular health services. This chapter therefore highlights the nature of the health problem in order that appropriate health care delivery systems can be better identified.

4.2 Mortality Profiles

Usually, in attempting to evaluate morbidity on a large scale, mortality, in the form of crude, standardized, infant, and child mortality rates are used as proxy measures (see Chapter Six for a more in-depth explanation of proxy measures for morbidity). This is necessary as morbidity cannot be measured directly. Unfortunately in South Africa, even the collection of mortality data on blacks, who constitute approximately two-thirds of the population, are too inaccurate for any but the most crude of analytical studies.

However, a method that can be used to gain some idea of the broad characteristics affecting health status between blacks and other race groups, using existing mortality data, is to examine the different reasons for mortality by age and sex. The differences in morbidity can then be surmised from the prevalence of

particular causes of death. Figures 4.1 to 4.16 show graphically the mortality profiles for blacks, whites, coloureds, and Asians. Mortality data for the years 1985 to 1987 is used for blacks in order to eliminate, as far as possible, errors in data collection in specific years. Mortality data for whites, coloureds, and Asians, is for the year 1986.¹ The causes of death will be indicated using 16 of the 17 chapter headings of the International Classification of Diseases code. Although some of the conditions indicated in figures 4.1 to 4.16 are unclear, the intention of this exercise is to highlight the most important causes of mortality.

The use of inter-racial comparisons of mortality are not entirely satisfactory, as these reveal only a part of the picture. Condition-specific mortality according to income-group would also prove useful. However, at present the Central Statistical Services only provides this kind of information according to race group.

1 All the mortality data comes from the Central Statistical Services:
 "Deaths of Blacks 1985", RP/No 07-05-08
 "Deaths of Blacks 1986", RP/No 03-10-01
 "Deaths of Blacks 1987", RP/No 03-10-01
 "Deaths of Whites, Coloureds and Asians", RP/No.03-09-01, gov. printer.

4.2.1 Blacks

Figure 4.1

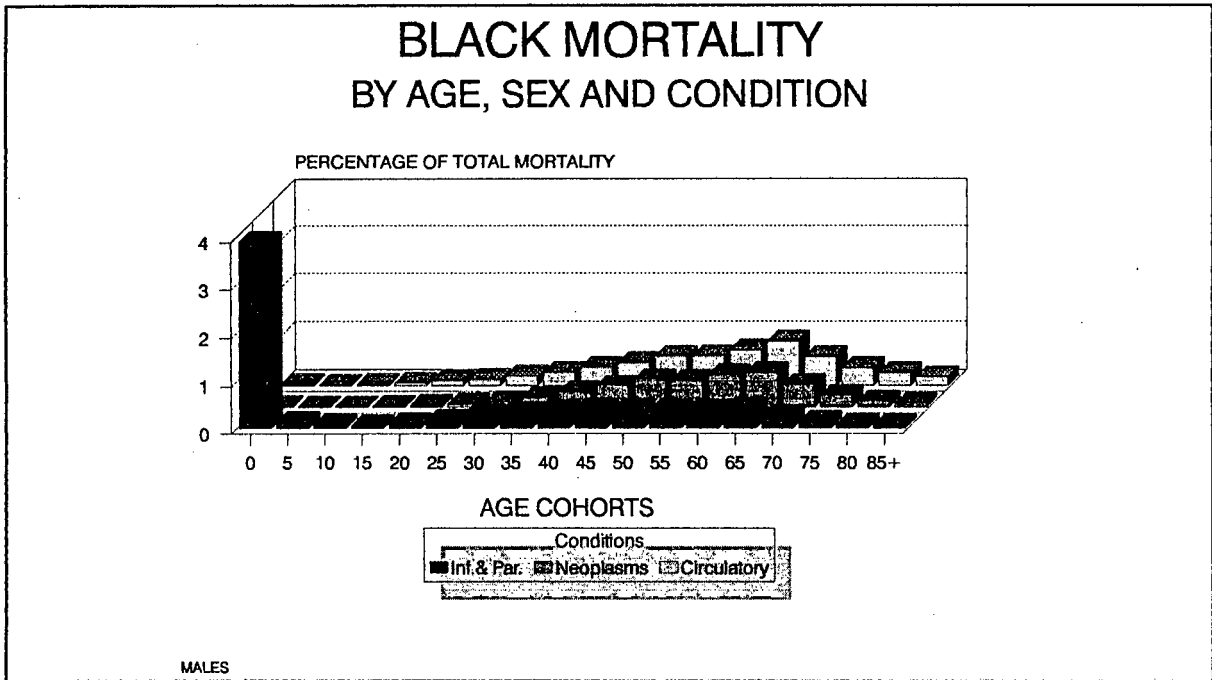


Figure 4.2

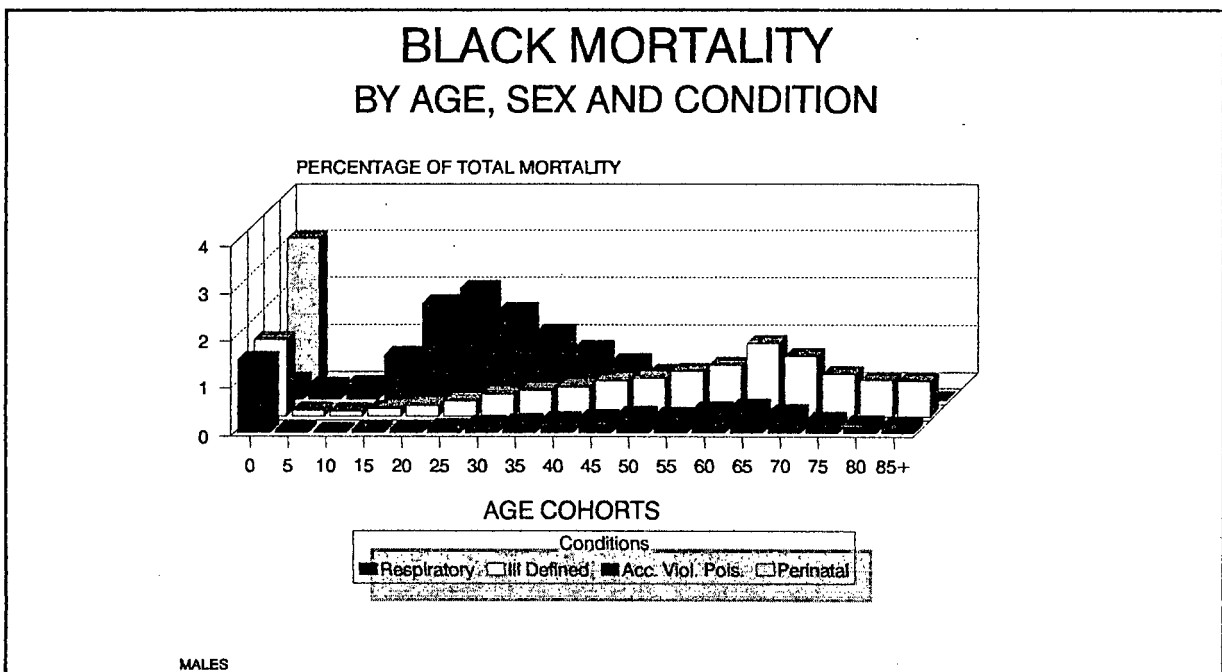


Figure 4.3

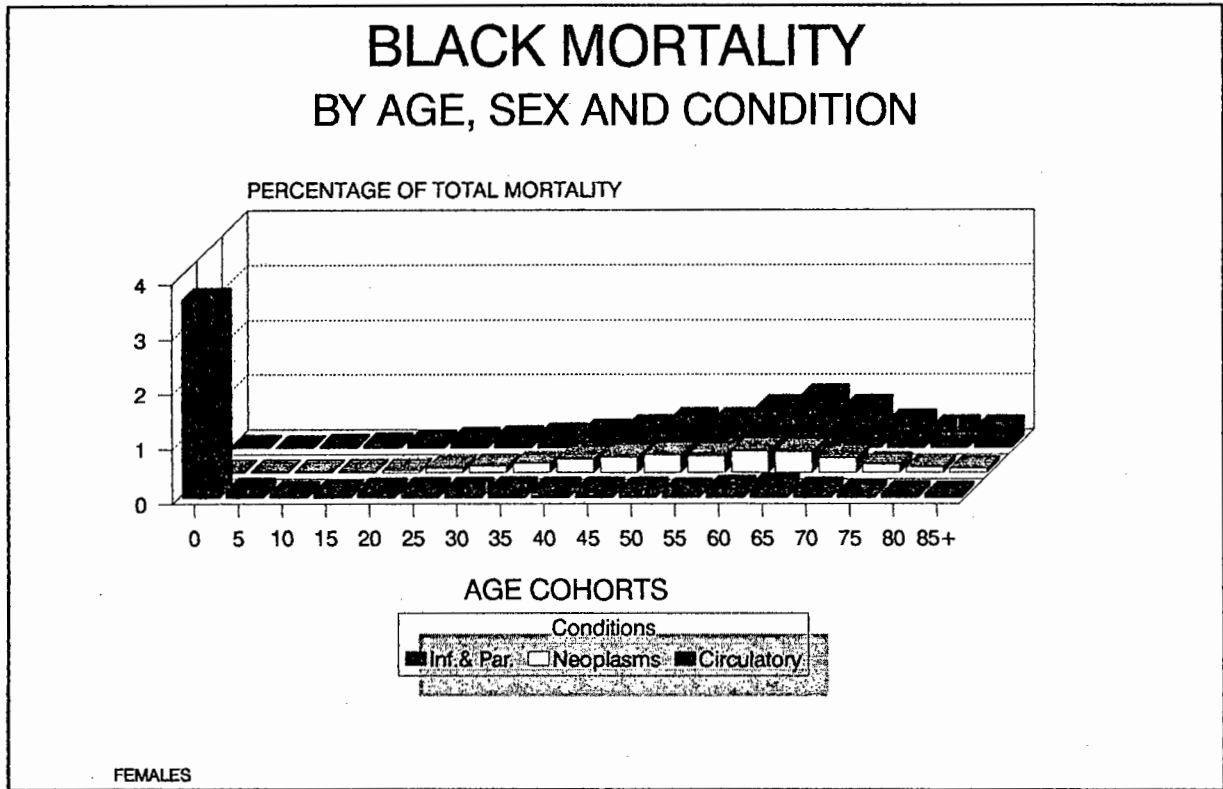
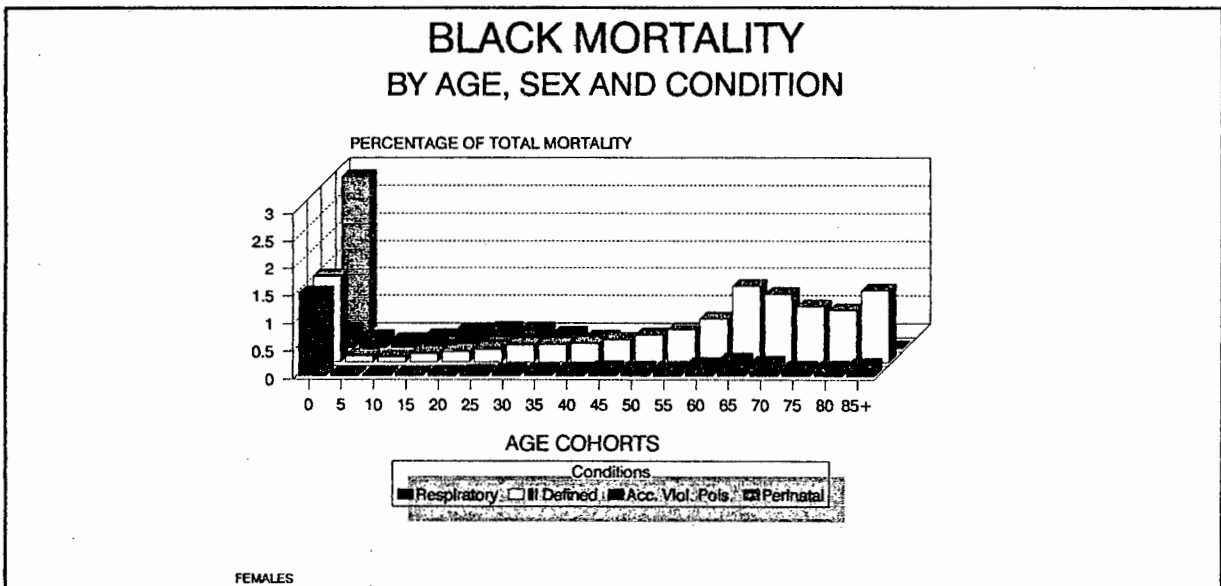


Figure 4.4



The mortality profile of blacks (figures 4.1 to 4.4) exhibit fairly distinctive characteristics. The mortality profiles of both males and females indicate that a very high percentage of total mortality occurs in the age cohort 0-4. Infectious and parasitic diseases are the most significant contributor in this age category, accounting for approximately 7% of the total mortality of blacks. This is largely due to the high incidence of, and consequently mortality from, essentially preventable and curable diseases, ie. Tuberculosis, Measles, and Tetanus. Infectious and parasitic diseases, although less life threatening in the older age cohorts, still play a major role in the mortality of these cohorts. Most of the mortality in the 0-4 age cohort occurs before the age of one. The high incidence of mortality during the perinatal, neonatal, and early childhood period is an indication of adverse socioeconomic conditions.

*"Perinatal mortality rates are influenced by the quality of perinatal care, the environment and the health status of the pregnant mother, and access to health care during the perinatal period. ... Post neonatal deaths are largely caused by social and environmental factors. Over the age of one year, children are at increased risk of dying of conditions related to the physical environment."*²

Diseases of the respiratory system (figures 4.1 and 4.3), ie. Pneumonia and Influenza, also play a major role in infant mortality, as do ill-defined conditions.

On the whole, the pattern of mortality reflected here suggests that the socioeconomic environment is most unsatisfactory for blacks. These adverse conditions are more extreme in the rural areas. In a survey conducted by the Regional Health Organization for Southern Africa for the Department of National Health and Population Development, it was found that approximately 2% of rural black children under the age

2 D. Bradshaw, H. Botha, G. Joubert, J.P. Pretorius, R. van Wyk, D. Yach, Review of South African Mortality, 1984, Medical Research Council, Technical Report No.1, 1987.

of five years in South Africa (excluding the 'independent' homelands) were wasted, 25% were stunted and 8% had low weight for age.³ The 1987 annual report of the Department of National Health and Population Development reported the results of a survey conducted among pre-school African children in Botshabelo in the Orange Free State. It was found that 15% of these children were under weight for age, with 36% showing significant growth retardation. In children under three years, 17% were undernourished and 41% showed retarded growth, which indicated that the nutritional status of this age group was lower than that for older children.⁴

The high mortality amongst children from an essentially curable disease such as measles can be ascribed to the following factors:⁵

- a Malnutrition: the risk of severe complications and death is significantly increased in the malnourished child.
- b Age: the younger the age of infection with measles, the greater the risk of complications. The highest case-fatality rates in South Africa have been found in children under one year of age.
- c Overcrowding: the rapid spread of the disease is facilitated by such conditions.

In studies carried out in the Ciskei it was found that:

Nutritional deficiencies "interact with and aggravate the communicable diseases which occur frequently among children: diseases like diarrhoea, measles, pneumonia and tuberculosis. ... Measles, for example, may cause undernourished children to have a mortality 400 times

3 South African Institute of Race Relations, Race Relations Survey 1988/89, Printed by Galvin and Sales, Cape Town, 1989, p.11.

4 South African Institute of Race Relations, p.12.

5 S. Fisher, "Measles and Poverty in Port Elizabeth", Second Carnegie Inquiry into Poverty and Development in Southern Africa, Paper No.172, 1984, p.1.

higher than that of their well nourished counterparts."⁶

It was further pointed out in this paper that studies revealed that Malnutrition diseases "... result not only from food deprivation (in a poverty economy), but in addition from social deprivation in a disorganized community."⁷

A study, which evaluated levels of household diet, perceptions of diet and food production, rural water supplies, and household incomes and expenditure of families in rural areas of the Transkei,⁸ found that 69% of families appear to be living below the poverty datum line. However, taking into account the cushioning effect of communal and kinship obligations, it was suggested that at least 40% of rural households live in a state of poverty.⁹

This study found that 74% of the drinking water which is obtained from dams, canals and streams, is vulnerable to contamination. On the whole the water was not boiled, either due to ignorance or lack of fuel, and the per capita use of water was well below the hygiene norm of 20-50 litres per day.¹⁰

The most significant condition resulting in death in older age cohorts is signs, symptoms and ill-defined conditions (figures 4.2 and 4.4). It is fairly disturbing that this category should be so large. It appears that medical attention was deficient or lacking during the illness prior to death. There is also the possibility of

6 Southern African Labour and Development Research Unit, Working Papers, Health and Health Services in the Ciskei, Community Health Research Project, SALDRU working paper No.54, Cape Town, 1983, p.38.

7 SALDRU p.37.

8 J.J. Bembridge, "Some Aspects of Household Diet and Family Income Problems in Transkei", South African Medical Journal, Vol.72, September, 1987, pp.425-428, p.425.

9 J.J. Bembridge, p.427.

10 J.J. Bembridge, p.427.

death certification problems.

The next most important cause of death is classified as accidents, poisonings and violence (figures 4.1 and 4.3). The major contributors to mortality in this category are motor vehicle related accidents, and deaths from violence. This type of mortality is clearly associated with an urbanized way of life. The high incidence of mortality due to violence is very disturbing and is an indication of high unemployment, and a generally unhealthy social environment in which urban blacks find themselves. It is interesting to note the low incidence of this type of mortality in female blacks.

Another disturbing sign of the ill-effects of urban life is the significance of diseases of the circulatory system (figures 4.1 and 4.3). Much of the mortality in this category is associated with hypertensive-related cases and various forms of heart disease. Mortality due to these causes is mostly the result of self-destructive habit-forming commodities such as alcohol and cigarettes. Although mortality in this category is clearly preventable without medical intervention, the social environment in which blacks have to live in the urban areas is not amenable to action of this sort. It should be pointed out that alcohol abuse would also contribute to the high mortality from accidents and violence.

Contributing to the health problems of blacks in urban areas are the general lack of adequate facilities (ie. water, electricity etc.) and housing. Quantification of the extent of the problem is extremely difficult, especially where housing is concerned.

By 1988 the total number of formal houses in the Pretoria, Witwatersrand and Vereeniging (PWV) black townships stood at around 352 900, of which 53.57% are rented. According to figures of families who are registered on the local authority's waiting list, there are approximately 600 000 homeless people in the

PWV townships.¹¹ Meanwhile, the urban foundation puts the population of people inhabiting the PWV area in informal housing at between 1 600 000 and 2 400 000.¹² In order to demonstrate the extent of the housing shortage, the following table shows statistics on the total population and the estimated quantity of people without formal housing in all the urban black townships in the PWV area.

Table 4.1

ESTIMATED QUANTIFICATION OF THE HOMELESS¹³ ¹⁴POPULATION IN THE URBAN TOWNSHIPS OF THE PWV¹⁵

Township	Population	Homeless	
Alexandra	120 000	21 120	(18%)
Atteridgeville	120 000	18 000	(15%)
Bekkersdal	49 450	27 275	(55%)
Boipatong	10 872	-	
Bophelong	11 271	-	
Botleng	12 000	11 850	(98%)
Daveyton	133 000	32 500	(24%)
Diepmeadow	261 426	-	
Dobsonville	86 100	-	
Duduza	30 450	30 000	(98%)
Ekangala	6 373	-	
Evaton	130 000	50 000	(38%)
Impumelelo	3 500	650	(19%)
Kagiso	73 060	9 730	(13%)
Katlehong	400 000	97 500	(12%)
Kwathema	143 123	45 250	(32%)
Mamelodi	320 000	92 500	(29%)
Mohlakeng	53 053	9 975	(19%)
Ratanda	22 000	10 500	(48%)
Refenkotso	1 668	-	
Refilwe	2 207	2 500	(113%)

11 H. Mashabela, Townships of the PWV, South African Institute of Race Relations Publication, Aurora Printers, Pretoria, 1988, p.11.

12 South African Institute of Race Relations, p.162.

13 "Homeless" here refers to those people without formal housing.

14 The population estimates are mostly based on official statistics. The estimates of people without formal housing are not always based on official statistics, and therefore some discrepancies do exist between the figures.

15 H. Mashabela, p.17-184.

Sebokeng	600 000	Unknown	
Sharpeville	46 143	-	
Soshangawe	96 400	40 000	(41%)
Soweto	1 474 000	1 150 000*	(78%)
Tembisa	200 000	25 000	(12%)
Thokoza	140 000	70 000	(59%)
Tsakane	75 000	34 000	(45%)
Vosloorus	112 500	11 325	(10%)
Wattville	35 000	15 000	(43%)
Zamdela	21 049	-	
Zithobeni	3 114	3 690	(118%)
TOTAL	4 792 759	1 808 365	(37%)

* This figure is based on the estimate of 230 000 informal structures multiplied by an average of 5 persons per structure.

In 1987 the following estimates (table 4.2) were produced showing the quantity of people living in informal settlements around the country.

Table 4.2

POPULATION LIVING IN INFORMAL STRUCTURES IN SOUTH AFRICA¹⁶

Region	Dep. of Const. Planning	Urban Foundation
Durban/Pinetown	3 353	1 700 000
Eastern Cape	11 560	200 000
KwanNdebele	N/A	350 000
PWV area	914 101	1 600 000 to 2 400 000
Western Cape	6 784	400 000
Winterveld	N/A	600 000

What can be seen in table 4.2 is the gross discrepancy between official and nonofficial estimates. If government has been using the official statistics for planning purposes, it would then become clear why severe backlogs still exist, as more accurate studies indicate. The sheer size of the problem ensures that by the year 2010 this will remain a major area of concern. Projections to the year 2000 estimate that the

housing shortage will affect a population of 3 413 000 by that stage.¹⁷ While the housing shortage, as well as the shortage in the provision of basic facilities persists, only a minimal contribution can be made to improving overall health conditions in the urban areas. However, despite the desperate conditions in which many blacks live in urban areas, some of the studies indicate that the situation is far worse in the rural areas with one estimate of infant-mortality rates at around 200 per 1000 live births.¹⁸ On the whole the mortality profile for blacks is similar to what one would expect to find in a low-income country.

17 South African Institute of Race Relations, p.190.

18 SALDRU p.35.

4.2.2 Whites

Figure 4.5

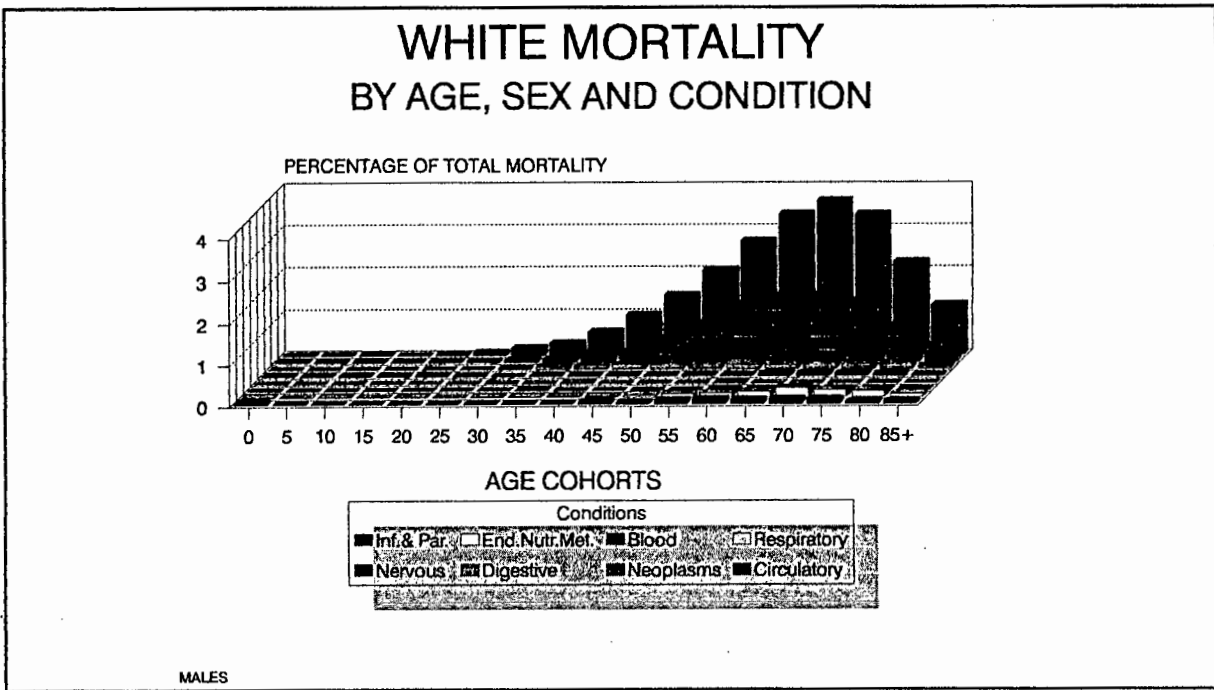


Figure 4.6

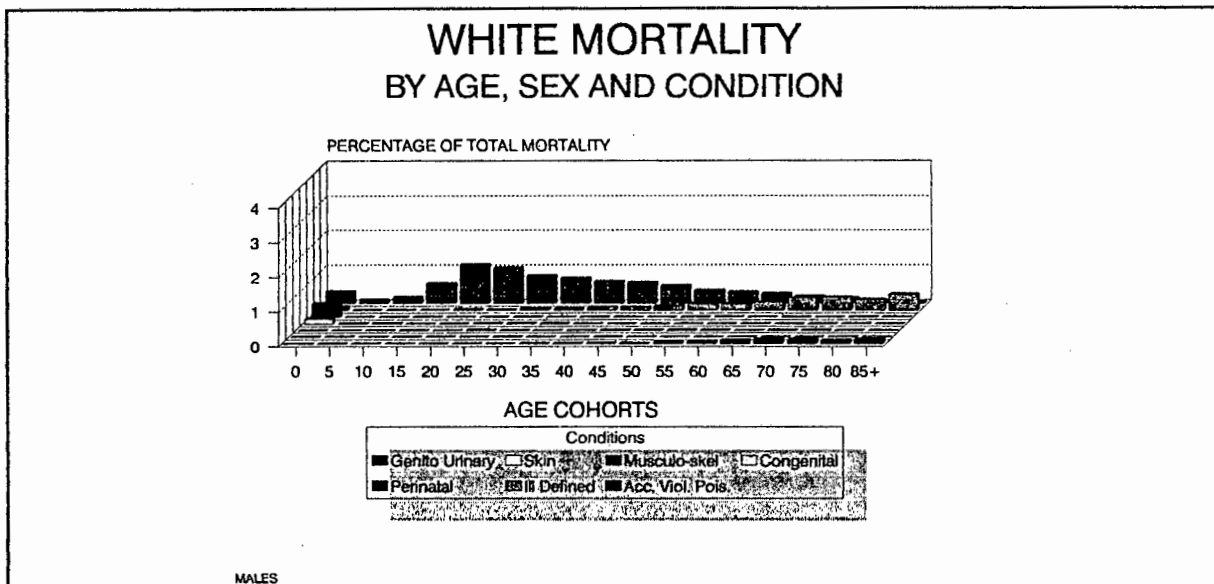


Figure 4.7

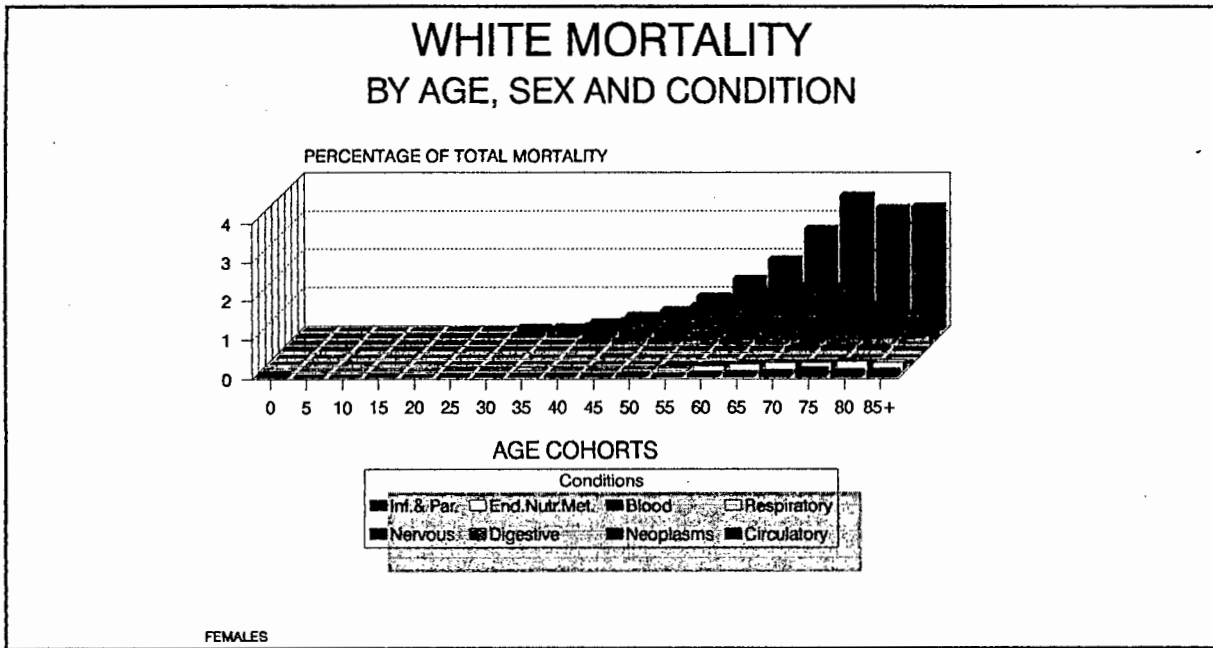
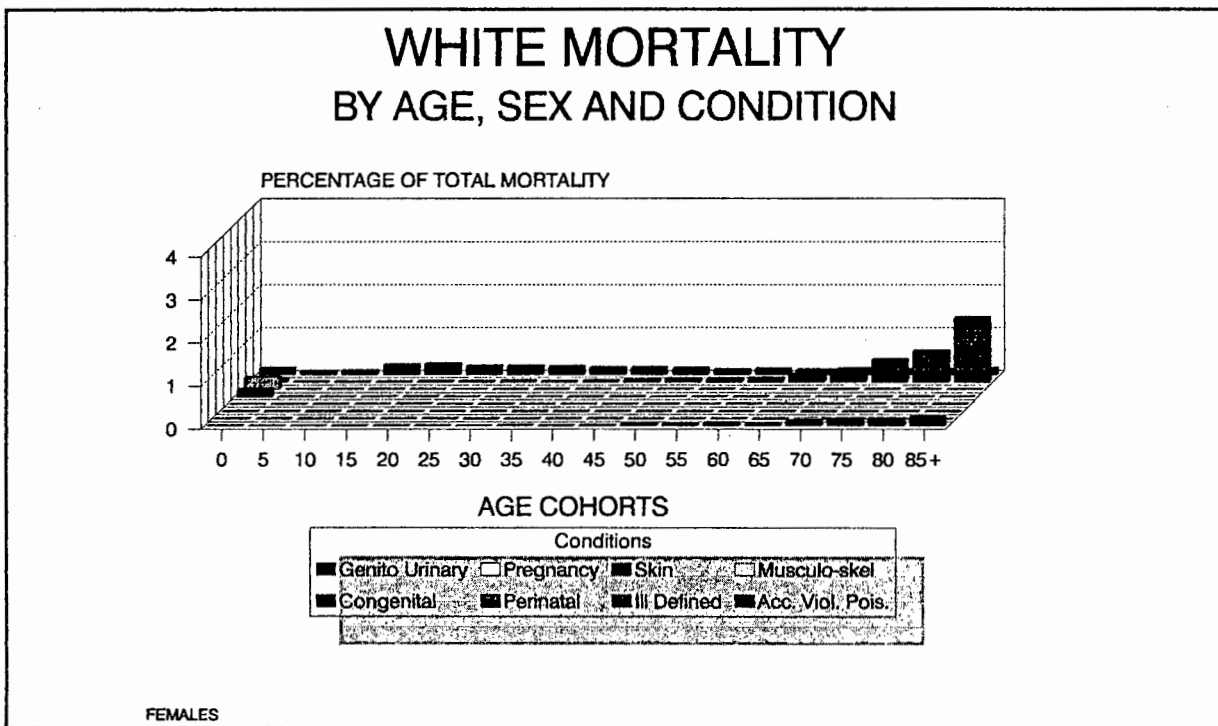


Figure 4.8



The mortality profile of the white population (figures 4.5 to 4.8) in South Africa appears consistent with the mortality profile that would be found in an industrialized country. Infant mortality is not a major contributor to total mortality, with infectious and parasitic diseases playing a very minor role. For the most part, mortality tends to be concentrated around the older age cohorts, unlike black mortality rates where mortality mostly tends to be associated with the younger age cohorts. The only condition that peaks in a younger age group is that classified as accidents, poisonings, and violence. As in blacks this is not a major cause of death of females. However, unlike the black population, most of the mortality under this heading is caused by motor vehicle accidents, and not through violence.

Neoplasms and diseases of the circulatory system are the most significant cause of death, with neoplasms (ie. cancer) the second most important contributor. As in the black population, diseases of the circulatory system can be associated with the over-consumption of self-destructive commodities, as can conditions listed under neoplasms (ie. cancer of the lung). Although non-medical preventive procedures could prove effective in reducing the burden these individuals place on medical services prior to death, and those that survive with reduced overall quality of life, such options are rarely investigated. Therefore the provision of expensive hospital facilities becomes the only available approach.

On the whole, the white mortality profile indicates the need for expensive hospital in-patient facilities, especially for the aged. It is clear from the graphs that whites exhibit a good health profile. Health problems associated with adverse socioeconomic living conditions have ceased to play a part in the mortality of whites. Many of the preventive procedures that would be necessary for the improvement of black health status are irrelevant for whites due, to their higher incomes, better nutrition and generally better living conditions.

4.2.3 Coloureds

Figure 4.9

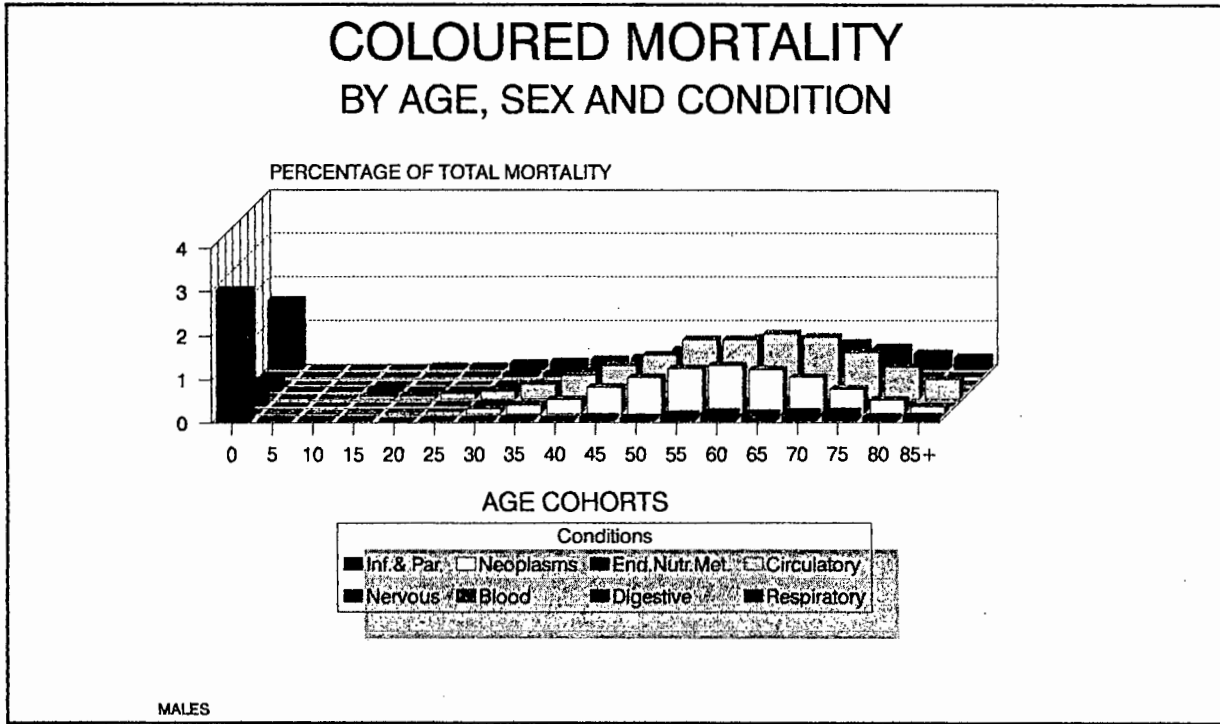


Figure 4.10

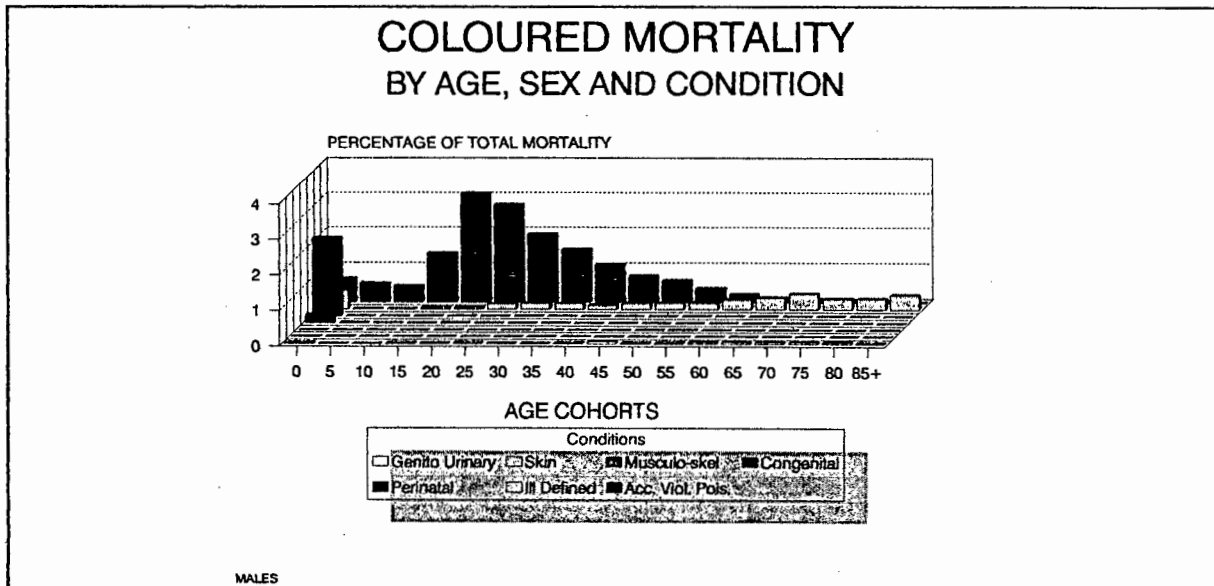


Figure 4.11

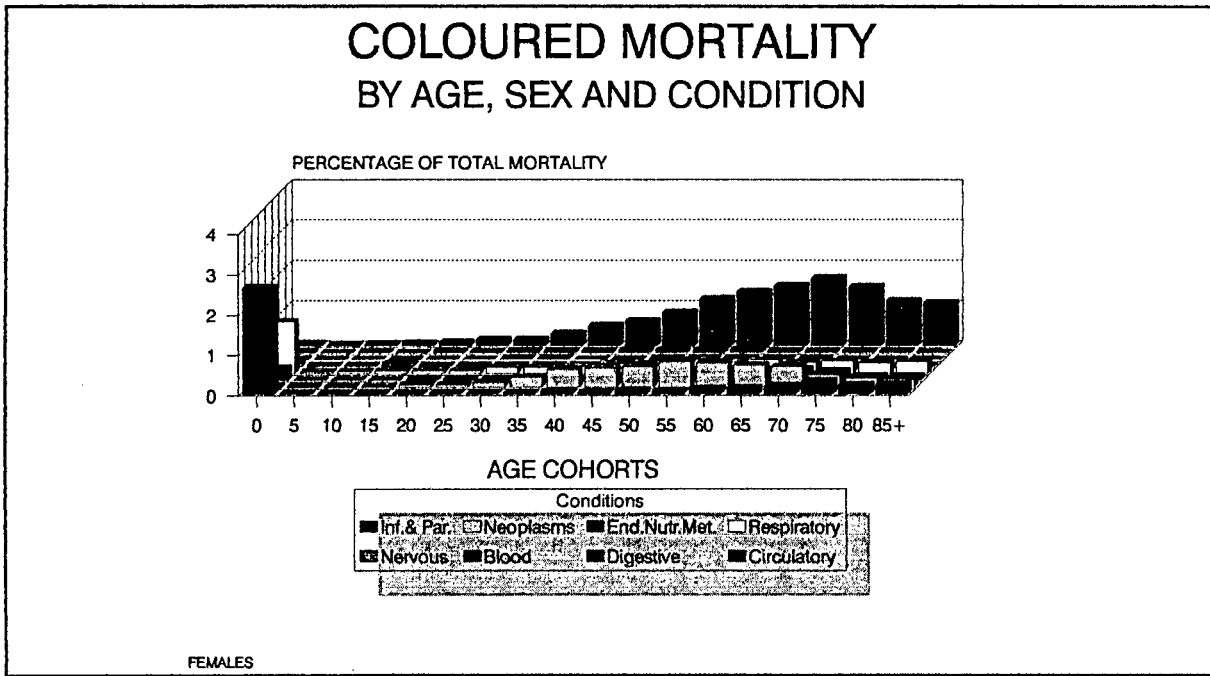
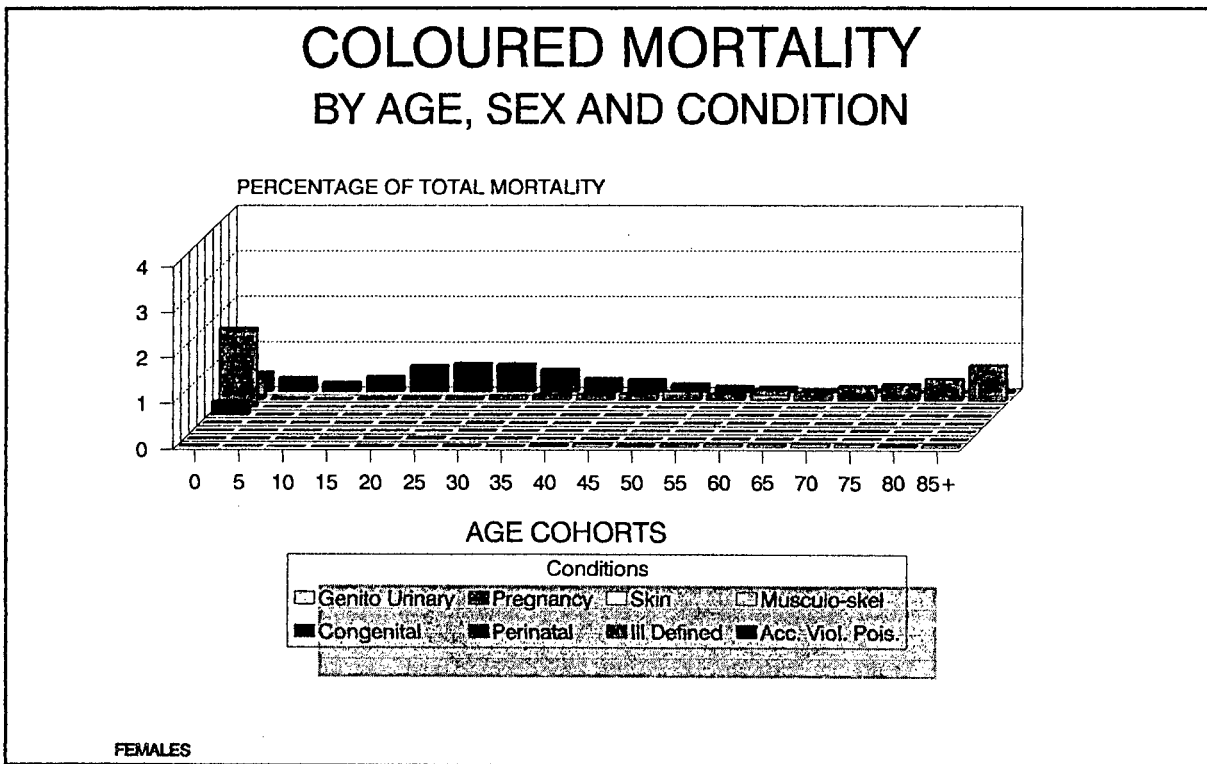


Figure 4.12



The coloured mortality profile (figures 4.9 to 4.12) appears to be mid-way between the white mortality profile and that for blacks. There are high levels of infant and child mortality co-existing with high levels of diseases of the circulatory system (figures 4.9 and 4.11) which are the most important overall cause of death. Neoplasms (4.9 and 4.11) are also important. To a certain extent this indicates that coloureds are more urbanized than blacks.¹⁹ Deaths from ill-defined causes (figures 4.10 and 4.12) also have a much lower representation here than in blacks, possibly indicating better access to medical services. Deaths from accidents, violence and poisoning (4.10 and 4.12) are prominent in young males (20-30 years), and as with blacks, much of the mortality in this category is due to acts of violence.

The high level of mortality in the 0-4 age group resulting from infectious and parasitic diseases, certain conditions arising in the perinatal period, and diseases of the respiratory system, indicates that the coloured population is still affected by an adverse socioeconomic environment. The low prevalence of breast-feeding, poor hygiene and nutrition as well as overcrowding can be regarded as major contributing factors. At present the incidences of perinatal and infectious causes of death among coloured children are similar to that for whites in 1929.²⁰

In a study by Bourne et al, on the spatial variation of mortality among white and coloured children under the age of five, it was found that whereas for whites there was no correlation between infant and 1-4 mortality, there was a significant correlation for coloured children. The conclusion reached was that age-specific mortality among coloured infants essentially forms a continuum with substantial mortality in

19 According to the 1985 census 77% of the "coloured" population live in urban areas. "Blacks" on the other hand are 39% urbanized in the RSA and 16% urbanized in the "self-governing" states. Central Statistical Services, Population census 1985, Age by Development Region, Statistical Region and District, RP/No 02-85-02, gov. printer.

20 D. E. Bourne, M. R. Rip, D.L. Woods, "Characteristics of Infant Mortality in the RSA 1929-1983, Part 2, Causes of Death Among White and Coloured Infants" South African Medical Journal, February, 1988, vol.73, pp.230-232.

the second and third years of life. The factors causing this are considered to be socioeconomic status and access to health care. For whites, however, the mortality rate has reached a low value in the first year of life.²¹

The coloured mortality profile therefore indicates that their health status is still far from optimal. On the whole there are characteristics associated with an urban life-style that both improve and act against health status. Diseases of the circulatory system as well as deaths due to violence, indicate the unhealthy social environment in which coloureds live in the urban areas. However, urbanization has reduced, to a certain extent, the influence of infectious and parasitic diseases and diseases of the respiratory system (figures 4.9 and 4.11) on mortality as a whole, and especially in infants. However, rural mortality from these causes is still fairly extensive.²²

21 D. E. Bourne, M.R. Rip, D.L. Woods, "The Spatial Variation of Mortality Among Children Under 5 Years of Age in South Africa", South African Medical Journal, June, 1989, vol.75, pp.565-570, p.567.

22 D. E. Bourne et al, pp.565-570.

4.2.4 Asians

Figure 4.13

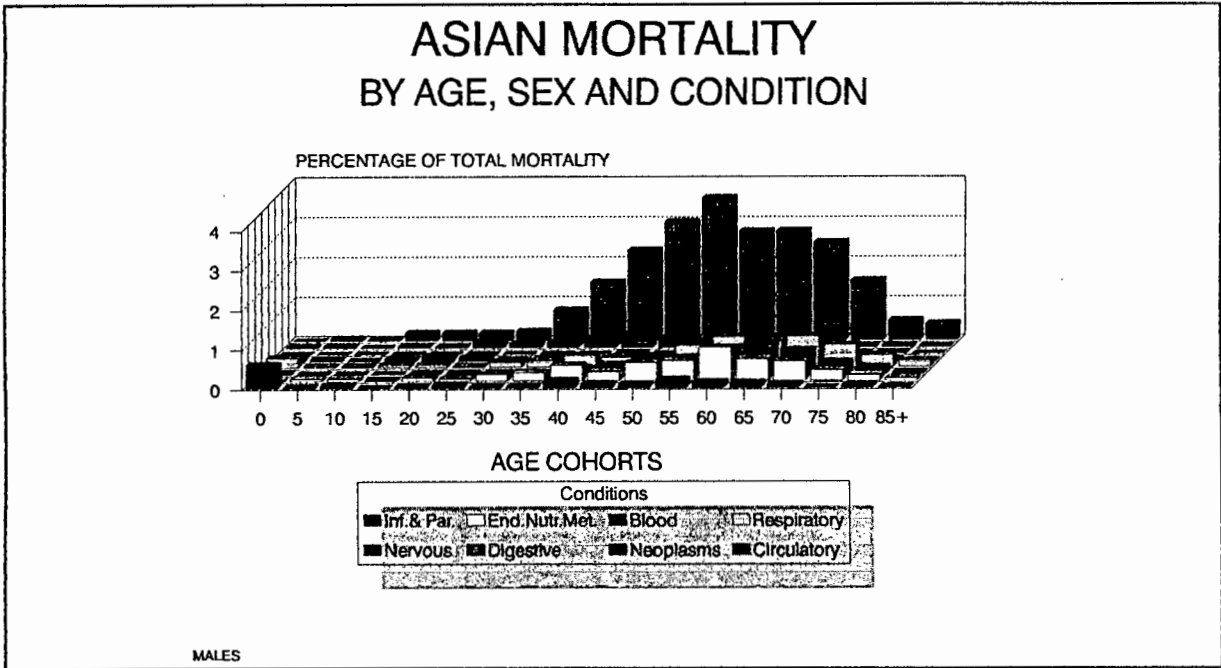


Figure 4.14

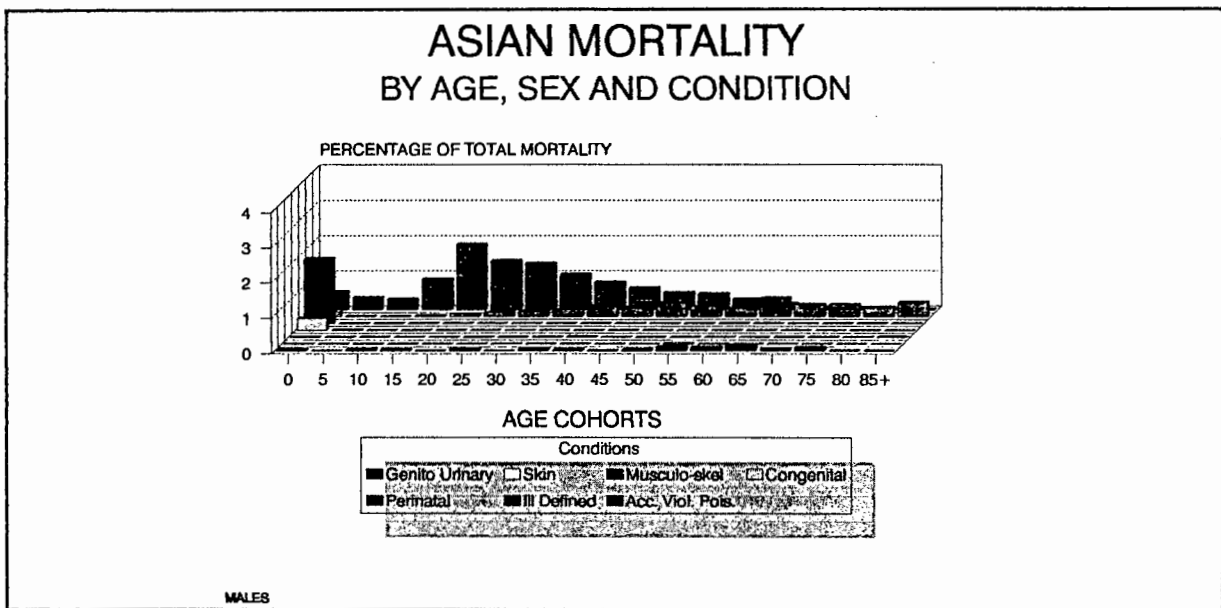


Figure 4.15

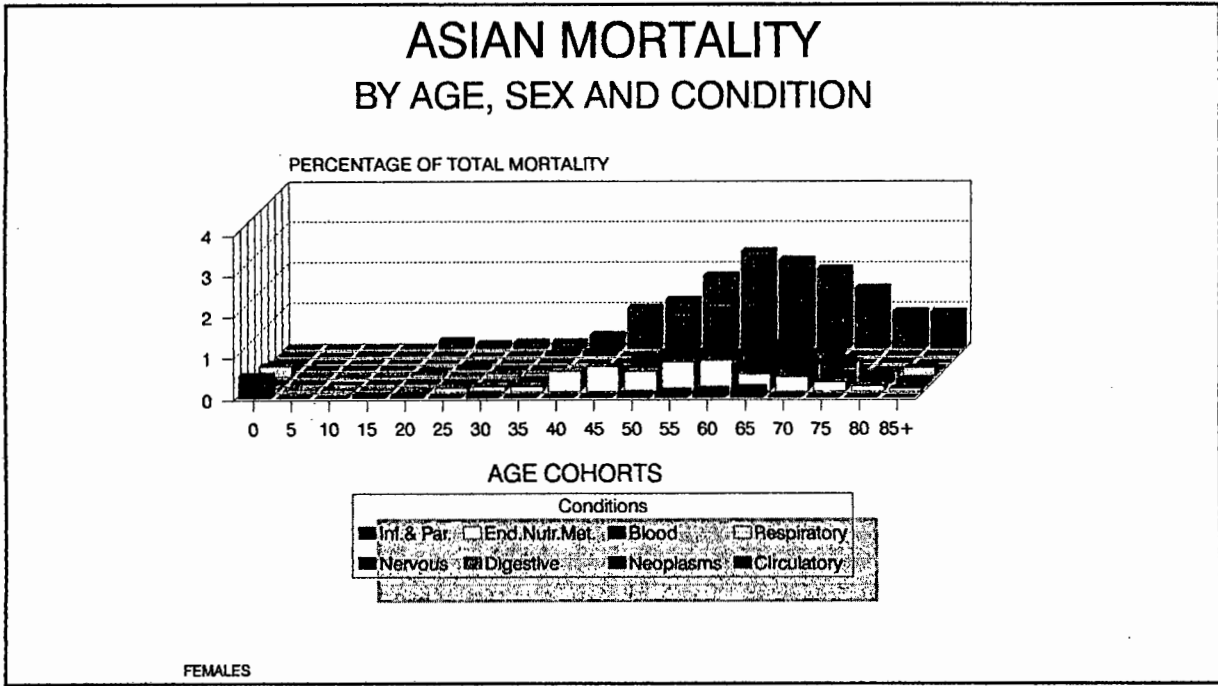
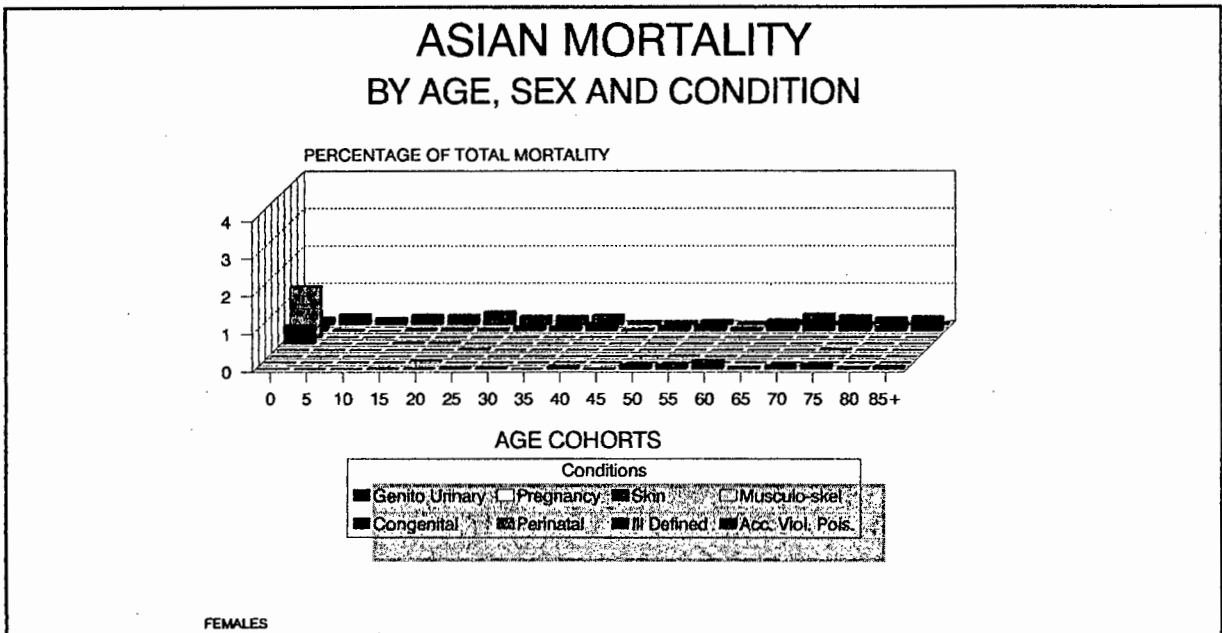


Figure 4.16



The Asian mortality profile (figure 4.13 to 4.16) largely resembles the profile of the white population. This can be attributed to the high socioeconomic status of the Asian population as a whole. The fact that the graphs are not smooth is a result of the small size of the overall population. Most of the comments that apply to whites apply to Asians as well, except for the observation that the profile is slightly worse for Asians, with most mortality occurring in younger age cohorts.

4.3 Conclusion

The overall health picture in South Africa is fairly disconcerting. With socioeconomic variables playing the role that they do, only negligible improvements can be expected in overall health status by the mid 1990s. Urbanization could play a part in both improving health in some respects, and harming it in others. It is essential that the urbanization process is complemented by a strategy both to accommodate people, and to provide them with basic needs, at least. The profiles of the black and coloured populations indicate that their health needs, in terms of effective interventions, are fundamentally different from those for whites and Asians. Strategies aimed at preventing illness, and intervening through improvements in the socioeconomic environment are far more crucial for blacks and coloureds than for the other two "race" categories.

On the whole, low-cost primary preventive procedures on a large scale could achieve more effective results than the provision of high-cost tertiary care, which fails to address the fundamental cause of the ill health. The requirements of whites and Asians, however, reflect a need for tertiary services, due to many of the socioeconomic causes of ill-health no longer being present in their communities.

The issue of whether one group should have priority over another need not be a normative one. What later chapters attempt to achieve is a methodology for evaluating the best approach for each level of

decision-making in the health sector. Priorities in terms of "who" gets "what" can be determined according to criteria that attempt to differentiate between specific needs and the best way to resolve those needs. This does not imply that any group of people, whether rich or poor, will be discriminated against. Essentially the approaches developed attempt to provide a methodology for determining what each socioeconomic group most needs, and providing it, subject to financial constraints.

CHAPTER FIVE

THE SOUTH AFRICAN PUBLIC HEALTH SECTOR

5.1 Introduction

This chapter provides some background statistics and comments on the prevailing distribution of health resources in South Africa. This information demonstrates the need for affirmative action towards improving the method of managing and distributing resources within the health sector. Chapter Six presents a methodology by which the maldistribution of these resources can be alleviated.

5.2 Structure of the Public Health Sector

The public-health system in South Africa is very complex. If all the "homelands" are included as part of South Africa, there are 14 ministries of health. These include the Department of National Health and Population Development (NHDP), the three Departments of Health and Welfare (ie. the "own affairs" departments created under the tricameral parliament system when the constitution of South Africa was altered in 1983), and the 10 "homeland" departments of health. Apart from these, another six central government departments are concerned with health in South Africa. They are: the Department of Development Planning (which finances the Provincial Authorities); the Department of Foreign Affairs (which supports the "independent homelands"); the Department of Development Aid (which finances the "self-governing homelands"); the Department of Defence; the Department of Police; the Department of Prisons; and the Department of Manpower.

The financial structure of the South African public-health system is derived from the existing institutional

structure. Public-health-care expenditure can be classified in terms of the following categories.¹

First-tier of government

- 1 Central Department of Health;
- 2 "Own affairs" Departments of Health;
- 3 Other central government departments providing additional health services;
- 4 "Homelands".

Second-tier of government (excluding "homelands")

- 1 Provincial Authorities.

Third-tier of government (excluding "homelands")

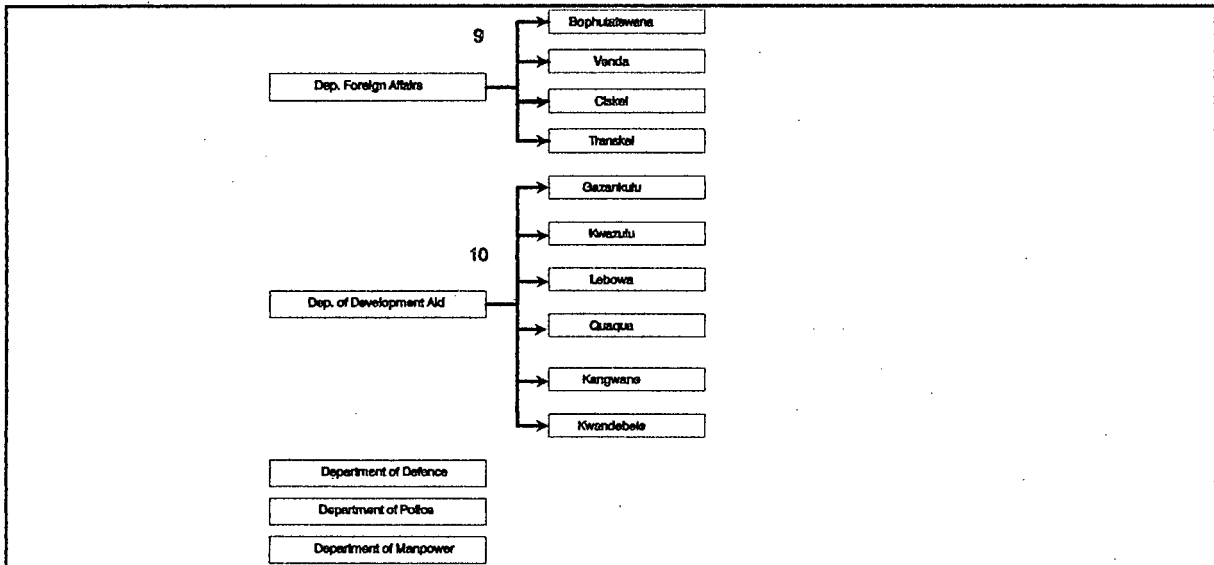
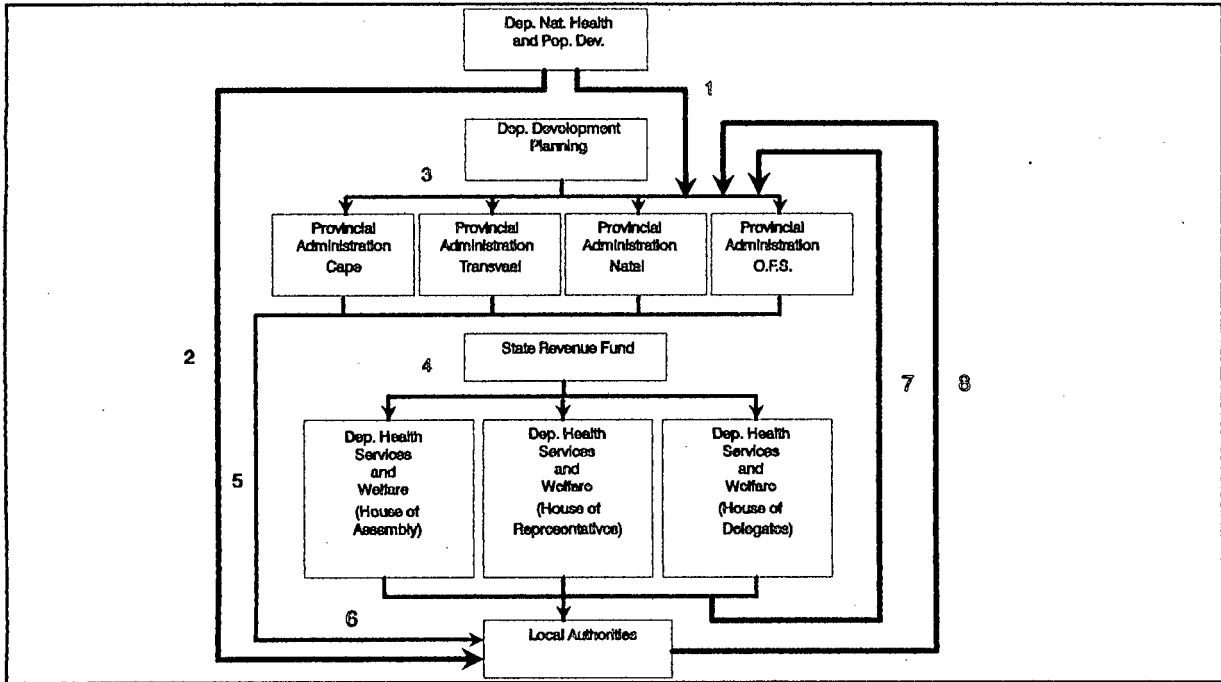
- 1 Local government.

Figure 5.1 indicates how the financing of the public-health sector takes place.

1 D. McIntyre, Health Care Expenditure in South Africa 1970 to 1990, Health Economics Unit, Department of Community Health, University of Cape Town Medical School, 1990, p.2.

Figure 5.1

Financial Structure of the South African Health Services



Key to figure 5.1²

- 1 Contributions to family planning services and for infectious disease in certain regions (in particular the Eastern Cape).
- 2 Contributions for the treatment of Tuberculosis patients, community health centres and mobile clinics, primary-health-care services, air-pollution control, public environmental services, nutrition services and family-planning services.
- 3 Budget for Provincial Administrations / second tier of Government is distributed through the Department of Development Planning.
- 4 The three "own affairs" administrations are funded by the State Revenue Fund.
- 5 Contributions from Provincial Administrations to Local government largely for ambulance services.
- 6 Contributions for mother and child care, care of the aged, infectious and communicable diseases, community health centres and mobile clinics and community based psychiatric care. (The extent of these transfers differ between the "own affairs" departments).
- 7 Contributions for general hospitalization and surgery for patients who are the statutory responsibility of the relevant "own affairs" department, out-patient psychiatric care and recreation for psychiatric patients.
- 8 Contributions from Local Government to Provincial Administrations largely for hospitalization of certain communicable disease patients.
- 9 The Department of Foreign Affairs provides financial aid to the "Independent Homelands". This aid amounted to approximately 44% of these countries' budgets in 1989/90.
- 10 The Department of Development Aid distributes financial aid to the "self-governing territories". In 1989/90, 76% of these territories' budgets were funded by this Department.

As can be seen from figure 5.1, the system of public-health care financing is complex, as is the general structure of the entire health system. Extricating accurate information on expenditure is difficult, and there

2D. McIntyre, p.33.

is no routine standardized publication of this information made available in South Africa. The fragmentation of South Africa's public-health services has become a major cause of concern. Since the introduction of the tricameral parliamentary system, creating the three "own affairs" departments at the first-tier level, public-health care expenditure has increased by 5.9%³ per annum. The relevant health administration expenditure has been increasing at an average rate of 8.5% per annum in real terms. Prior to the introduction of the tricameral parliament the average real rate of increase was 3.5% per annum.⁴

As the fragmentation is ethnically based, the indirect costs of operating such a system are considerable. People are often unable to use the nearest hospital or health facility, as they have to use the closest hospital of their own ethnic grouping.⁵ Using a racial criterion to determine the priorities and institutional basis for the public health system is wasteful and increases the costs of health care while decreasing its effectiveness. This fragmentation has also served to entrench the maldistribution of health resources in South Africa. Furthermore, the recording of statistics by the various departments is not coordinated or standardized, and of varying quality, severely limiting the ability of policy makers to plan adequately.^{6 7 8}

3 D. McIntyre, R. E. Dorrington, Trends in South African Health Care Expenditure 1971-1988, Paper presented at the conference "Containing costs in health care: Towards affordable health for all", Johannesburg, September, 1989, p.27.

4 D. McIntyre et al, p.27.

5 M. Price, "Health Care as an Instrument of Apartheid Policy in South Africa", Health Policy and Planning; Oxford University Press, 1986, 1, (2): pp.158-170, p.163.

6 D. McIntyre, 1990.

7 T. D. Wilson, Health services within Soweto, Carnegie conference paper No.170, 13-19 April 1984, Cape Town.

8 Personal communications with the Department of National Health and Population Development (Strategic Planning).

5.3 Maldistribution of Health Resources

Inequality of access to medical facilities and care is not merely the result of socioeconomic deprivation in South Africa. Two further aspects serve to make access more difficult for some people. The first, and most difficult to explain logically, particularly when the ethical nature of the service is taken into consideration, is a consequence of racial segregation or rather racial discrimination. The second is a consequence of medical care being provided on an ability to pay basis. Both of these aspects are now discussed in order to demonstrate the extent of the maldistribution of medical care in South Africa at present.

5.3.1 Geographical Distribution of Health-Care Spending

Table 5.1 provides an indication of the geographical distribution of public sector funds for health care in South Africa as a whole, for the financial year 1986/87.^{9 10}

9 The Provincial health care expenditure was obtained from D. McIntyre, p.66.

10 The "homeland" health care expenditure was obtained from estimates produced by the Central Economic Advisory Service, personal communication.

Table 5.1

THE GEOGRAPHICAL DISTRIBUTION OF PUBLIC HEALTH EXPENDITURE

	Expenditure 1986/87 (R 000)	Estimated Population (000)	Per capita Expenditure Rands
Natal	523 259	2 593	202
Cape	1 215 620	6 090	200
O.F.S	335 539	2 146	156
Transvaal	1 360 448	9 100	149
Ciskei	86 414	881	98
Gazankulu	42 542	606	70
Bophutatswana	91 664	1 480	62
Venda	23 365	446	52
QuaQua	20 923	221	95
Transkei	167 376	2 919	57
Kangwane	23 017	479	48
Kwazulu	214 829	4 565	47
Kwandebele	8 875	287	31
Lebowe	105 015	2 237	47

From table 5.1 it can be seen that the "homelands" are not regarded as a priority when it comes to health care financing. The average per capita expenditure in the Provinces is approximately three times the average per capita expenditure in the "homelands". However, the racial bias extends further than the geographical maldistribution of resources.

5.3.2 Racial Distribution of Health-Care Resources

Table 5.2¹¹ gives an indication of the disparity in the provision of publicly provided medical facilities on a racial basis.

11 Browne Commission, Eighth Interim Report of the Commission of Inquiry into Health Services, Interim Report on Hospitals and State Health Services, RP/No. 66/1986 gov. printer, p.13.

Table 5.2

SOUTH AFRICAN POPULATION AND BED DATA (EXCLUDING "HOMELANDS")

Area	Race Group	Population	No. of Beds/1000
Cape	White	1 264 040	5.20
	Non-white ¹²	3 827 320	2.59
	Total	5 091 360	3.24
Transvaal	White	2 362 060	4.05
	Non-white	5 988 440	2.10
	Total	8 350 500	2.65
Natal	White	561 860	4.09
	Non-white	2 114 480	4.81
	Total	2 676 340	4.66
Orange Free State	White	326 220	7.50
	Non-white	1 605 640	1.32
	Total	1 931 860	2.36
All Four Provinces	White	4 514 188	4.62
	Non-white	13 535 880	2.57
	Total	18 050 068	3.08

A glaring example of the maldistribution of resources can be seen in the Orange Free State where the the white population, which comprises approximately 17% of the total population of the Province, has access to more than 54% of all the in-patient hospital beds available. Non-whites, on the other hand, comprise 83% of the total population but have access to only 46% of the available hospital beds. The situation is shown to be even more inequitable according to table 5.3¹³.

12 "Non-white" refers to all race groups not classified as "white" in South Africa.

13 Browne Commission, p.85.

Table 5.3

AVERAGE PERCENTAGE BED OCCUPANCY IN SOUTH AFRICA

	Province	1979	1984
Whites	Transvaal	57.9	59.9
	Cape	59.3	58.7
	OFS	53.8	49.0
	Natal	57.6	51.1
Non-Whites	Transvaal	92.2	102.6
	Cape	107.6	102.6
	OFS	106.6	91.4
	Natal	78.7	74.8

The average percentage bed occupancy for whites in Provincial hospitals in 1984 was 49%, whereas for non-whites it was 91.4%. The total for all four Provinces shows a better distribution than in the Orange Free State, nevertheless there is a severe inequality of access to medical care, exacerbating the already difficult health status of South Africa's non-white population. Whites, who comprise approximately 25% of the population (within the Provinces), have access to 38% of all publicly provided beds in South Africa. Non-whites, on the other hand, comprise 75% of the total population and only have access to 62% of all available public sector beds. If it is noted that the deprived socioeconomic circumstances of the non-white population necessitates a greater need for health services than for whites, the present underprovision of these services, where they are needed most, aggravates an already difficult situation. A consequence of the inadequate provision of appropriate facilities is shown in table 5.4¹⁴.

Table 5.4

LENGTHS OF STAY IN INPATIENT FACILITIES (DAYS)

	Province	1979	1984
Whites	Transvaal	5.6	5.0
	Cape	7.4	6.8
	OFS	6.4	4.7
	Natal	6.8	6.4
Non-Whites	Transvaal	8.4	8.4
	Cape	8.9	7.5
	OFS	8.6	7.2
	Natal	7.8	7.1

Table 5.4 shows that non-whites tend to stay in hospital for longer periods than whites. The reasons for this are as follows:

- a Staff-patient ratios are inadequate where non-whites are concerned;
- b The overall health status of non-white patients is low necessitating longer recovery periods;
- c Inadequate outpatient and Primary Health Care facilities, which, if provided, would allow earlier discharge of patients, and allow illness to be detected at an earlier stage; and
- d Patient follow-up is inadequate usually on account of the poor education of non-white patients who fail to report for follow-up treatment. This results in hospitals having to detain patients beyond the stage in their illness where they could be released and treated on an out-patient basis, eg. tuberculosis treatment.

Thus, not only is there inadequate provision of hospital in-patient facilities for non-whites, but inadequate Primary- Health-Care facilities aggravate the problem by causing these patients to require longer occupation of expensive in-patient hospital facilities.

The Browne Commission¹⁵ report conceded the following with respect to the preceding observations:

"The Committee found that many of the conditions being treated in hospitals for Blacks could be considered as preventable, and that by increasing preventive health measures, the number of admissions, the bed occupancy, and the length of stay of Black patients, in particular, could be reduced significantly."

The second feature creating an inappropriate distribution of medical facilities is a consequence of the private sector health-care providers. Expenditure on private health-care facilities amounts to about 45%¹⁶ of total expenditure on health in South Africa. However, most of these facilities are used by white patients who belong to medical-aid schemes. Approximately 7% of the Black population are covered by medical-aid schemes¹⁷. McIntyre & Dorrington¹⁸ have calculated that annual per capita expenditure on blacks amounts to R 138 whereas for whites the figure is R 597. It was concluded that if these expenditures were to be expressed as percentages of Gross National Product (GNP), the amount spent on whites would be equivalent to around 13% of GNP, whereas for blacks it would be about 3% of GNP.

An additional problem with the private sector is that it charges a great deal for the little it provides. This is clearly shown by the fact that the public sector, which is responsible for around 74.6%¹⁹ of all the available beds, spends only 55% of total expenditure on health, whereas the private sector, which is

15 Browne Commission, p.157 par.49.

16 S. P. Taylor, J. M. L. Klopper, "South African Health Care Expenditure 1975-1984", South African Medical Journal, December, 1987, Vol.72, pp.802-804.

17 L. Hollis, from talk presented at the conference "Containing costs in health care: Towards affordable health for all", Johannesburg, September, 1989.

18 D. McIntyre et al, p.29.

19 Central Statistical Services, Census of Hospitals and Establishments for In-Patients (1983), RP/No.20-66-07, gov. printer.

responsible for approximately 25.4%²⁰ of all beds, is responsible for around 45% of the total expenditure on health. Furthermore, 60% of the total volume of drugs, which is distributed by the public sector, results in only 23% of the total spending on drugs in South Africa. In other words, 40% of the total volume of drugs, which is distributed by the private sector in South Africa, accounts for 77% of the total expenditure on drugs.²¹

Transferring facilities to the private sector, a policy which appears to have the support of the health authorities in South Africa, will only serve to further exacerbate the present maldistribution, with adequate medical care moving beyond the reach of most of the population.

5.4 Conclusion

The object of this chapter was to provide some indication of the general structure of the health-care system in South Africa. The discussion was primarily aimed at indicating the inadequate distribution of medical-care resources between the white and non-white population groups in South Africa. In order to redress adequately the historical racial and geographical bias that has developed in the distribution and technology selection of medical services, it is necessary to establish a sound and objective methodology whereby such a process can be facilitated. Chapter Six develops an approach to deal with some aspects of these problems.

20 Central Statistical Services, 1983.

21 P. Folb Head of the Department of Pharmacology, UCT Medical School, Paper presented at the conference "Containing costs in health care: Towards affordable health for all", Johannesburg, September, 1989.

CHAPTER SIX

THE GEOGRAPHICAL DISTRIBUTION OF SOUTH AFRICAN HEALTH RESOURCES

6.1 Introduction

The previous chapter dealt with the various distributive inequalities associated with the overall allocation of health-care facilities. Factors such as race, ability-to-pay, and geographical distribution of resources were important in determining access to medical care. The concern of this chapter is the determination of a methodology that will enable health authorities in South Africa to distribute new and redistribute existing resources on an objective and sound basis. However, the methodology that is presented here is limited in some respects. It only refers to the financial allocation of resources, and does not concern itself with the specific organizational issues or incentive structures within the health-care sector.

The methodology used in this chapter moves away from racially-based methods of distribution. South Africa's population is considered as a whole, with differentiation occurring according to defined variations in need on a geographical basis.

It is intended that the methodology be utilized in the form of a formula. At present in South Africa a formalized approach to the distribution of health resources is under consideration by the Department of National Health and Population Development. Such a formula has already been developed in various other countries, such as the United Kingdom, New Zealand and Portugal. Both the New Zealand and Portuguese formulae are based on the methodology and research used to establish the United Kingdom formula. The discussion in this chapter therefore centres around the formula established for the macro-distribution of resources within England and Wales.

The primary concern of such a formula is that it be sensitive to the specific medical needs defined within a geographical region relative to other geographical regions. The formula refers specifically to the distribution of resources within the public sector. At present the overall allocation of funds to health care is dependant upon an essentially arbitrary determination by Cabinet. This is not optimal, and hopefully such decision-making can be replaced by an approach similar to the one described in this chapter.

This chapter deals initially with issues relating to "need", ie. the need to redistribute medical-care resources, and the need versus demand arguments. Following these discussions, the various criteria that should be taken into account when evaluating target levels of service and relative resource allocation (on a geographical basis), are elaborated on. It should be borne in mind that this chapter does not attempt to establish a formula for resource allocation, but rather develops some of the issues that should be accounted for in its construction.

6.2 The Need to Redistribute Health Resources

The essence of the inequity, with regard to medical resources, that exists in South Africa, rests on the different levels of access to medical care experienced by people of different socioeconomic groups. It should be recognized that there are constraints placed on the accessibility of health facilities to people who are in the greatest need. Factors such as distance from the medical facility, queuing time, opportunity cost of missing work, poor quality of treatment or available facilities, educational levels of patients, and unsympathetic medical staff, could all play a role in reducing utilization rates of those people most in need of medical treatment. What is required to counter this tendency is a method of distributing resources according to the needs of a particular population. The object would be to ensure, on the macro level, the adequate distribution of funds for health spending, so that a particular authority will have finances available to improve the nature of the services being provided in accordance with the need of the population. Accessibility can hopefully be improved by the provision of appropriate care. Major reductions

in illness could be achieved, in lower socioeconomic communities, by providing services that are preventive in nature. The medical needs of communities that reside in better socioeconomic conditions, however, require very different services. It is therefore essential to ensure that the distribution of funding takes place in such a way as to account for both the level and nature of the need.

The Resource Allocation Working Party (RAWP)¹ report on the formula produced for England and Wales, made the following comments on the need for a re-allocation process in Britain:

"Supply of health facilities is, in England as elsewhere, also variable and very much influenced by history. The methods used to distribute financial resources by the National Health Services (NHS) have, since its inception tended to reflect the inertia built into the system by history. They have tended to increment the historic basis for the supply of real resources (eg. facilities and manpower) and, by responding comparatively slowly and marginally to changes in demography and morbidity, have also tended to perpetuate the historic situation."

As indicated in Chapter Five the distribution of health resources in South Africa is far from satisfactory. It is, therefore, essential to prevent the continuation and perpetuation of this historical bias. The difficulty is now how to define the need for health resources in an equitable manner.

6.3 Need versus Demand

This section concentrates on those issues that have led to the rejection of perceived demand as an indicator of medical need.

¹ Department of Health and Social Security, Sharing Resources for Health in England Report of the Resource Allocation Working Party, London, Her Majesty's Stationary Office, 1976, p.7.

As shown in Chapter Two, in many countries, medical care expenditure commands more and more of Gross National Product (GNP) every year. What is being experienced in these countries is a steady increase in the demand for medical goods and services, ie. a shift to the right of the demand curve with the supply curve either not moving or moving slowly to the right. This would not necessarily be a problem if the increased spending resulted in a visible increase in the health status of the affected countries. However, little of this increased spending appears to affect health status.²

The reasons for this increase in demand can be linked to the suppliers of medical goods and services. The initial demand for medical care expressed by the consumer is dependant upon his/her own perception of need for medical attention and this is qualified by the perceived price or cost of that medical attention. When the consumer arrives at the doctor he loses control over future demand for medical resources with respect to that visit. The doctor then examines the patient and recommends treatment. The patient has no means of establishing whether he/she is receiving the most cost-effective treatment, or whether it is necessary or not. Cost, therefore, ceases to play a determining role in the demand for medical goods and services leading to both cost and expenditure increases.

The potential for excessive increases in expenditure associated with this artificially-induced demand is largely in the hands of physicians. In the United States of America hospital services are the fastest growing component of medical care spending. In 1965 these services formed 32% of total medical care spending, while in 1978 it was 40%. Although physicians' services account for less than one fifth of total expenditure, they have been estimated to control 70% of total health spending, especially hospital services.³

2 Chapter Three supports this view. The regressions show very little relationship between health expenditure and life expectancy.

3 A. C. Enthoven, "The behaviour of health care agents: Provider behavior", Health, Economics, and Health Economics, eds. van der Gaag, Perlman, North Holland Publishing Company, Amsterdam, 1981, pp.173-188, p.188.

Attempts to identify what type of demand is being dealt with may prove illusory.

*"... a hospital may exhibit a comparatively high cost per day because its medical staff keeps patients in the hospital for a relatively short period of time. Or it may exhibit a comparatively high cost per stay, because its medical staff hospitalizes only very sick people, while another hospital may have low cost per stay because its physicians hospitalize people who could have been treated equally well on an outpatient basis. The behaviour of the physicians associated with the hospital exhibiting higher cost per day or stay may be associated with lower cost of health care. Similar remarks relate to the cost per doctor's office visit. Behaviour that produces low cost per day, stay or visit may or may not be consistent with behaviour that produces low cost per-capita. It all depends on the associated affects on use of service."*⁴

It is, therefore, the individual physician who is largely responsible for the nature of the health care demand. Specifically, the problem associated with the above mentioned cost increases is termed "supplier induced demand", "... where the supplier, in acting as agent for the consumer, brings about a level of consumption different from that which would have occurred if a fully informed consumer had been able to choose freely."⁵

However, it is asserted that a distinction should be made between "supplier-induced demand" and "overtreatment".

"Overtreatment relates to a technical judgement concerning the difference between the treatment deemed clinically necessary and that actually provided. It should be noted that

4 A. C. Enthoven, pp.173-188.

5 A. McGuire, J. Henderson, G. Mooney, The Economics of Health Care, Routledge & Kegan Paul, New York, 1988, p.162.

*overtreatment may exist even when supplier induced demand does not, if patients prefer what is technically considered to be too much treatment. Obversely supplier induced demand may occur with no overtreatment if patients prefer less than the technically necessary but are persuaded by the supplier/agent to consume more."*⁶

Supplier-induced demand is further described as an imperfect agency relationship, where the monopoly potential of the medical profession is linked to the hypothesis that: "... individual doctors aim to maximize a target income and the observation that per caput consumption of doctors' services tends to rise roughly in line with increases in the doctor/population ratio."⁷

In situations where consumers face zero prices and doctors receive salaries, as in the NHS in the United Kingdom, there can still be an incentive to induce demand. Usually in such situations physicians reduce demand. However increases in demand could be experienced as doctors may wish to satisfy other aspects of their utility functions, (ie. doctors associated with teaching functions) or if the doctor merely wishes to fill beds in order to secure a budget increase. Supplier inducement may also occur where doctors underestimate the costs of health care to the patient.⁸

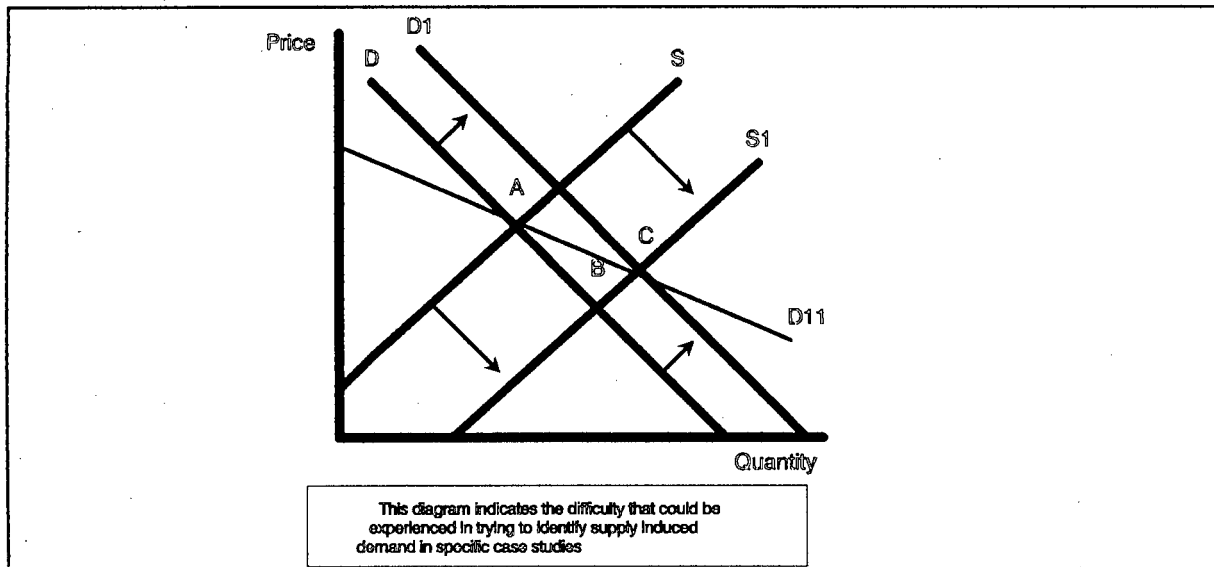
However, it is not easy to determine whether or not demand inducement is occurring where both the demand and supply curves are shifting. This can be explained graphically using figure 6.1.

6 A. McGuire et al, p.162.

7 A. McGuire et al, p.162.

8 A. McGuire et al, p.162.

Figure 6.1



The initial demand and supply curves are D and S respectively. If the supply of doctors' services increases to S1, the equilibrium point would move, assuming competitive conditions and no inducement, from A to B. Price would fall as quantity increased. With inducement, however, as supply increases, demand would also increase from D to D1. Although prices have fallen as quantity increased, the equilibrium point is at C and not B.⁹

Essentially the two possible shifts of the curve cannot be distinguished using data which are supposed to reflect equilibrium positions. If the data are consistent with the equilibrium moving from A to C, one of two explanations is possible. Either demand inducement is occurring, with a consequent shift in demand from D to D1, or assuming a competitive structure, the market may be represented by D11, rather than D1. This is then merely a shift in equilibrium under competitive conditions.¹⁰

9 A. McGuire et al, p.162.

10 A. McGuire et al, p.162.

For the above reasons accurately detecting demand inducement can prove difficult. Furthermore, it is possible that there are certain constraints on doctors exercising their monopoly power. Ethical codes and the entry of patients' welfare functions into the doctors' utility functions, could serve to limit extreme behaviour.

However, Cooper¹¹ makes the following observations in connection with supply induced demand:

"Collectively the profession appears to reassess its conception of need in line with actual levels of provision. Feldstein, for example, pointed out in 1967 that any attempt to allocate funds to hospital beds based on medical assessment of need for them as reflected in admission and waiting lists was likely, in practice, to have little or no meaning. Need seemingly tends to grow in line with provision, as doctors react to any expansion in supply by realigning their conception of need further along the continuum. In his study of 177 large acute hospitals, Feldstein found that both admissions and length of stay increased with bed availability and could discover no indication of a level of bed provision which could have fully satiated doctors' demands. Thus, like an iceberg, the more resources devoted to melting it, the more need seems to flow to the surface."

Sandier¹² makes the following comments which arose from studies of utilization in various countries:

"Our regional studies do not clarify the roles of supply of facilities and personnel in the rising consumption. The rate of growth of medical care utilization is not related to morbidity nor to

11 M. Cooper, "The Economics of Need: The Experience of the British Health Service", The Economics of Health and Medical Care, ed. M. Perlman, R & R Clark Ltd. Edinburgh, 1974, pp.89-107, p.94.

12 S. Sandier, The Health Care Cost Explosion, Which Way Now? ed. D. A. Erlich, Hans Huber Publishers, Bern, 1975, p.118.

the level of utilization but it is strongly related and perhaps can we consider that it is explained by the opportunities that physicians and hospitals have to increase their production."

The above observations, as well as those mentioned in earlier chapters, suggest the inappropriate nature of demand as an indicator of medical need. Demand can be manipulated upwards and downwards in accordance with the incentives generated within the medical care industry, whether run by the private or public sectors. Demand inducement is, however, not the only argument against the use of demand as a need indicator. For instance, hospital utilization rates (which can be taken as a reflection of demand) in Kwazulu tend to be quite low for blacks, at around 66.8%. This is largely due to a shortage of doctors and nursing staff. Hospitals in Kwazulu are estimated to be understaffed, in terms of nurses, by at least 47.8%.¹³

Natal hospitals, on the other hand, have very high utilization rates in hospitals treating black patients. However, the high utilization rates in Natal appear to be caused by Kwazulu residents seeking medical care in Natal rather than in Kwazulu. The reason for this situation is the inferior provision of medical services within Kwazulu, causing potential patients in that area to make the perfectly rational decision to seek care elsewhere. However, the Natal health authorities could claim, and do claim,¹⁴ that they should receive extra funding to cope with the high cross-border flow to their hospitals. If this rationale were to be acted on, Natal would be able to increase its supply of services relative to Kwazulu, reinforcing the cross-border flow of patients. This would result in further applications for increased funds in Natal. However, what is really needed is an increase in the supply of good quality medical services in Kwazulu, reducing the

13 Unpublished Report on the Kwazulu Health Services obtained from the Central Economic Advisory Service.

14 This remark is based on a personal discussion with the health representative from the Natal Provincial Authority, at a meeting of the Health Matters Advisory Committee, Sub committee on finance, 19 October 1989.

cross-border flow.¹⁵

Using a "dirty" indicator, such as utilization rates, to indicate need for medical goods and services will lead to decisions being made that are the exact opposite of those that should be made. The better method for evaluating the need for medical goods and services, is to use an indicator that is totally unrelated to observed demand. A final element suggesting the inappropriate nature of demand indicators for the purposes of defining need can be illustrated in the following analogy.

Assume that demand for emergency services are calculated only by taking into account the amount of telephone calls received to make use of those services. Imagine that an earthquake occurs causing a severe disruption in the telephone service. If demand continued to be calculated in the same way, the indication would be that there was a drop in the need for emergency services over that period. A government official "analyzing" the "available" statistics may conclude that severe earthquakes reduce the need for emergency services. With regard to medical services, if certain race groups, or socioeconomic classes, have limited access to that care, the observed utilization rates of the available facilities will not be a true reflection of their need for access.

The RAWP report¹⁶ recommended the removal of case-load indicators from the need-based formula in use in England and Wales using the following explanation:

"In our Interim Report we relied upon Regional in-patient and out-patient case-loads as an indicator of relative need over and above that arising from the age/sex structure of the population. We realized that this had serious imperfections. Whilst numbers of cases clearly

15 Note that this cross-border flow does not refer to the people leaving Kwazulu to permanently reside in Natal, but instead to those people who normally reside in Kwazulu and merely travel to Natal for medical treatment.

16 DHSS, 1976, par.2.6, p.15.

reflect need, they do so in terms of the available supply of services. Caseloads fail both to distinguish between degrees of need and to assess the extent to which need is unmet through lack of facilities. Waiting lists as one indicator of need are also known to have questionable reliability. Moreover, there is ample evidence to support the view that the level of supply has a significant influence on the level of demand. Need must, therefore, be measured by an indicator that is far less dominated by supply."

On the basis of the extensive evidence suggesting that demand indicators do not reflect the actual need of the population, other possible indicators, reflecting need for medical services are proposed in the following sections. These suggestions are primarily concerned with in-patient hospital services, although some of the discussion is relevant to the other service categories.

6.4 Weighting Population for Age and Sex

The size of the population must clearly be the basic determinant of need. However, per-capita needs are not identical. It is necessary to adjust the base or crude population to take these other characteristics into account. The RAWP report recommended as follows:¹⁷

"... the elderly (men and women aged 65 and over) form about 14% of the total population, yet they occupy more than half the non-psychiatric hospital beds (excluding maternity). Women have different needs from men, and children too are heavy users of health care facilities. Similarly, patterns of morbidity are different between the sexes at different ages. Thus the age/sex make-up of the population needs to be taken into account as well as its size."

Each population that is examined should therefore be broken down according to age and sex. However,

17 DHSS, 1976, par.1.7, pp.8-9.

in order to determine the relative burden a particular age cohort of a particular sex is going to have on medical facilities, if the facilities existed, some indicator of expected utilization is required. The method used in RAWP was to establish national bed utilization rates by age and sex, and then to use these rates to estimate expected utilization, by age and sex on a regional basis, of in-patient services. It should be pointed out at this stage that these utilization rates can be established through the collection of national data and sample studies. These rates are used to weight a regional population to reflect age and sex composition. Variations in need for in-patient services according to socioeconomic status are not taken into account by this indicator.

As national hospital in-patient utilization rates have not been determined in South Africa, the following rates (table 6.1) for the United Kingdom (U.K.) illustrate how the rates tend to differ according age and sex. In a study by McIntyre et al it was found that the difference in national hospital in-patient utilization rates between countries as diverse as India and the United States of America made less than 1% of a difference to their weighted population using both rates. It is therefore possible that the U.K. rates can be used as a valid proxy within the South African context.

Table 6.1

NATIONAL HOSPITAL IN-PATIENT UTILIZATION RATES (UNITED KINGDOM) BY AGE GROUP FOR EACH SEX IN TERMS OF NON-PSYCHIATRIC BED DAYS PER 1000 POPULATION¹⁸.

Age Groups	Males	Females
0 - 4	958	726
5 - 14	420	318
15 - 19	405	990
20 - 24	443	1 616
25 - 34	480	1 428
35 - 44	607	939
45 - 64	1 406	1 170
65 - 74	3 284	3 000

Although the above rates reflect the relative needs of a first-world population, they indicate characteristics general to all populations. These common features are:

- a The higher utilization requirements of children aged 0-4;
- b The generally higher utilization requirements of females from the ages of 15 to 44 (ie. child-bearing age);
- c The very high utilization of these services by the aged.

Table (6.2) shows the various regional populations in South Africa, adjusted by age, sex, and expected utilization (based on the UK rates).

Table 6.2

POPULATION ADJUSTED FOR BED UTILIZATION BY AGE AND SEX

Province/State	Crude Population ¹⁹ (1985)	Population Weighted for Age and Sex
Cape	5 591 067	6 091 491
Natal	2 145 018	2 218 259
Transvaal	7 531 513	7 501 915
O F S	2 193 403	2 256 160
Kwazulu	3 746 835	3 698 261
Kangwane	392 782	366 223
Quaqua	181 559	181 198
Gazankulu	496 943	462 870
Lebowa	1 835 586	1 822 531
Kwandebele	235 855	227 468

In order to weight the population, the United Kingdom national bed utilization rates by age and sex were combined with the base population by age and sex in each region. For convenience (ie. data deficiencies) Transkei, Bophuthatswana, Venda and Ciskei (the TBVC states) were omitted. The population breakdown used is that for the year 1985. The population figures adjusted to reflect age and sex composition are then summed for each region.

As the above weighted population only reflects information concerning usage by age and sex, further indicators must be obtained in order to account for variations in socioeconomic living conditions. In this respect an indicator of relative morbidity would be required.

19 Central Statistical Services Population Census, Age by Development Region (1985), Statistical Region and District, RP/ No. 02-85-02, gov. printer.

6.5 Weighting Population for Morbidity

Age, sex and socioeconomic environment can all play a role in the morbidity of a population. A particular environment affects particular age/sex cohorts differently. For instance, living conditions at Alexandra Township in Johannesburg may cause higher-than-normal levels of preventable diseases in the age cohort 0-4, whereas the age cohort 34-39 may be suffering higher than normal levels of hypertension due to excess consumption of alcohol and cigarettes. Age cohort 20-24 for males may exhibit higher levels of morbidity as a result of high levels of violence experienced in that environment. What is needed, therefore, is an indicator that will reflect abnormal levels of morbidity within certain populations on a regional basis. Estimating morbidity within a particular geographical region is difficult. The only cost-effective method is to find an easily obtainable proxy. The following four indicators are all possible candidates for consideration:

- 1 Case-load;
- 2 Potential Years of Life Lost (PYLL);
- 3 Health Status Indices; and
- 4 Mortality Rates.

6.5.1 Case-Load

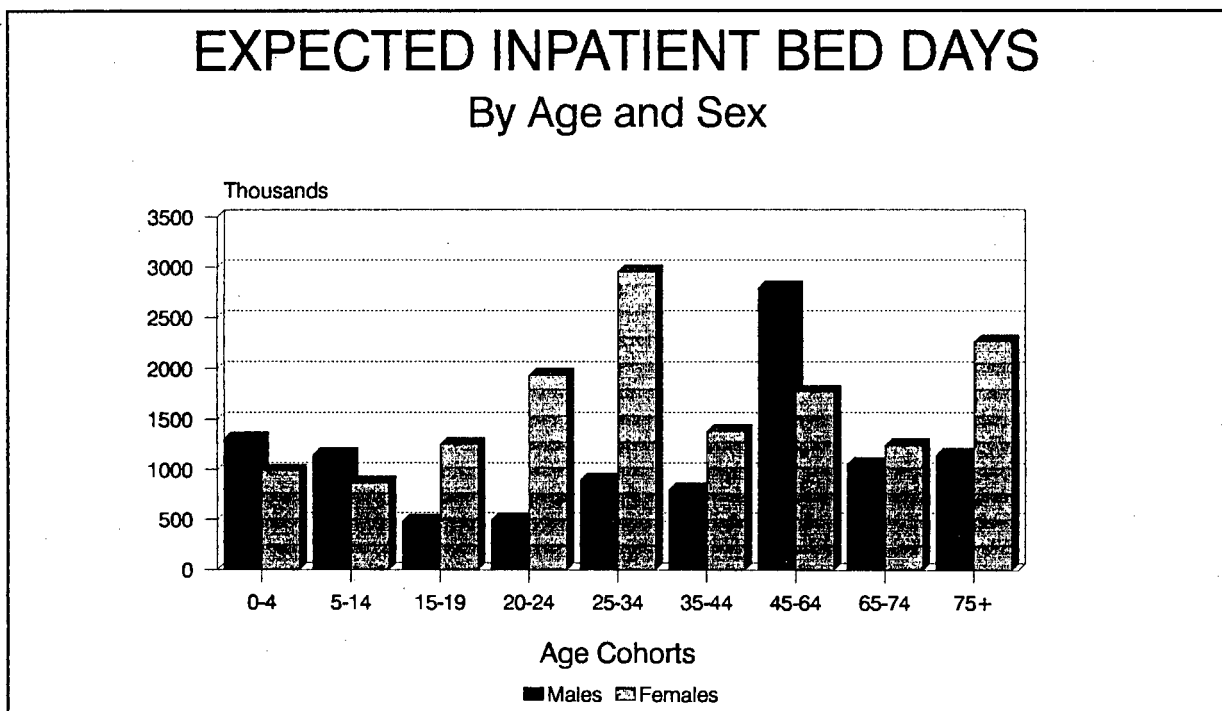
Case-load has already been discussed in the section on need versus demand. This figure tends to be strongly influenced by factors unrelated to morbidity levels. Therefore this would not be a very useful indicator of need.

6.5.2 Potential Years of Life Lost

This indicator is calculated by estimating how many years a patient has lost by dying at a particular age.

As only in-patient facilities are being discussed at this stage, the following comments do not necessarily relate to the possible use of PYLL with respect to other service categories. According to the in-patient utilization rates used earlier to calculate expected bed days in South Africa, the following more complete breakdown of expected bed utilization can be produced.

Figure 6.2



On the basis of the information given in figure 6.2, the aged (over the age of 65) are expected to utilize, on average, 23% of all bed days (not excluding maternity bed days). This is a significant percentage. The use of PYLL to calculate need would give the aged less of a weighting than would be given to younger age cohorts. This contradicts the purpose of a need-based formula which attempts to calculate the expected burden a particular age/sex cohort should place on in-patient facilities. The essential question to ask is: do PYLL reflect morbidity? That is do PYLL reflect the level of illness prevalent within a particular age/sex

cohort? The answer is no. PYLL are primarily used as outcome indicators (as discussed in greater detail in Chapters Seven and Eight) with a bias towards age cohorts with longer to live. Therefore, PYLL should be rejected as a measure indicating morbidity for the purposes of evaluating need for in-patient facilities.

6.5.3 Health Status Indices (HSI)

Health-status indices are discussed in depth in the chapter on cost-effectiveness analysis. Theoretically such an index would provide an indication of health status by grading individuals or communities on a continuum that includes various states of health, ranging from death to healthy. Such an approach is the most sound of all the options, as morbidity levels would be accurately reflected. However, such information could only be compiled through the use of complex questionnaires. Illiterate or poorly-educated people would have extreme difficulty in providing the information required. Furthermore, people living in South Africa's rural and underdeveloped areas could not easily be reached. It would only be possible to do sample studies which would be expensive to administer. Although the creation of and use of HSIs on a small scale would be useful in checking the other proxies that are used, the costs involved militate against its use as a general indicator of need.

6.5.4 Mortality Rates

With available technology, the only cost-effective method of estimating morbidity is to use mortality data. The reliability of mortality as a proxy for morbidity rests on whether mortality accurately reflects the expected burden a particular group of people will place on medical services. In reality the relationship between mortality and morbidity is not perfect as the RAWP report points out:

"Many of the commonest conditions including some which lead to death place relatively little demand

on health care services. Others require expensive care, perhaps over a long period."²⁰

A further problem is concerned with the use of crude rates to reflect regional variations in morbidity, as these fail to account for differing mortality rates by age and sex.

The solution to both the above-mentioned problems, and the method that is focused on in this chapter, rests on the following criteria. If the variation in mortality from the norm for the country, by age, sex and region, can be quantified, these can be used to identify those populations experiencing higher-than-normal morbidity due to socioeconomic conditions. If these unusual variations are then used to weight the base population, by age and sex, those cohorts that have a mortality rate higher than the norm will have their populations weighted upwards, whereas those with mortality rates below the norm will have their populations weighted down. The base population will be weighted in such a way that, after the population has been adjusted, each cohort in each region will have a mortality rate that is the same as the norm for that age/sex cohort for the entire country. The intention would be to combine these Standardized Mortality Ratios (SMRs), ie. the ratio of mortality by age, sex and region to that expected if the national mortality rate by age and sex were applied, with expected national bed utilization which is just the base population weighted for expected bed utilization.

This solves the problem associated with the crude rates, but does not solve the problem of different kinds of illness resulting in different levels of usage. The solution may lie in refining the data further. Both national expected-bed utilization and SMRs could be collated and processed according to age, sex, region and condition. The term condition refers to the specific classification of the ailment, resulting either in death or in the period of hospitalization of a particular person. Due to many different countries developing their own classifications of these diseases or ailments, an internationally accepted list of specific conditions has been drawn up in order to standardize the collation of sickness and mortality data. The list is termed

the International Classification of Diseases (ICD), the 17 chapter headings of which are as follows:²¹

- 1 Infective and parasitic diseases;
- 2 Malignant, benign, lymphatic, haematopoietic and unspecified neoplasms;
- 3 Endocrine, nutritional and metabolic diseases;
- 4 Diseases of blood and blood-forming organs;
- 5 Mental disorders;
- 6 Diseases of the nervous system, eye, ear and mastoid process;
- 7 Rheumatic fever, hypertensive and heart diseases, and diseases of the peripheral circulatory system;
- 8 Diseases of the respiratory system;
- 9 Diseases of the digestive system;
- 10 Diseases of the urinary system, male genital disorders and diseases of breast and female genital system;
- 11 Conditions of pregnancy, childbirth and puerperium;
- 12 Diseases of skin and subcutaneous tissue;
- 13 Diseases of musculoskeletal system and connective tissue;
- 14 Congenital anomalies;
- 15 Certain causes of perinatal morbidity;
- 16 Symptoms and ill-defined conditions;
- 17 Other injuries and reactions, fractures, dislocations and sprains;

Thus SMRs by age, sex, condition and region, combined with expected national bed utilization by age, sex and condition, should provide a useful picture of the burden that a regional population may place on in-patient facilities.

21 Central Statistical Services, Deaths of Blacks (1985), RP/No. 07-05-08, gov. printer.

However, the use of mortality rates to measure need for medical services has been subject to some criticism. It is suggested that death rates are more a reflection of social class than of the availability of health services. Furthermore, it is estimated that given the full application of current medical diagnosis in the United Kingdom a mere 6% mortality reduction would be achieved.²²

The use of SMRs in the RAWP formula are also criticized on the basis that their usage embodies a tendency to over-compensate for an inadequate socioeconomic environment. In other words, the distribution of in-patient services that is implied in the use of SMRs, is an inaccurate reflection of the need for medical resources, as opposed to social goods and services. In commenting on the RAWP and Scottish Health Authorities Revenue Equalization (SHARE)²³ formula, Forwell²⁴ remarks:

"RAWP and SHARE assume there is a clear distinction between the needs of social services and health services and that only the latter is the responsibility of the NHS. ... the underlying hazard in trying to distinguish between social and medical needs is that this might, by implication, define the present social responsibilities of the NHS as somehow being a burden which ought to be shed now, without regard to continuing unresolved defects in social services. It would be wrong, in theory, to allocate NHS resources on the principle that these should reflect social conditions as distinct from medical conditions (although these and the SMR in part are consequences of social conditions). In particular it should not be the objective of health service revenue allocations to compensate for the absence of social welfare or for failure in other fields of social services."

22 A. Barr, R. F. L. Logan, "Policy Alternatives for Resource Allocation", The Lancet, May 7, 1977, pp.994-996, p.995.

23 SHARE is the Scottish equivalent of RAWP. Both essentially use the same criteria to measure relative need.

24 G. D. Forwell, "Scottish Health Authorities Revenue Equalization, a Community Medicine View of the SHARE Report." Scottish Medical Journal, 1978, Vol.23, pp.207-212, p.207.

Forster²⁵, using regression analysis to examine the correlations between the results of the General Health Survey (GHS)²⁶ for 1972 and 1973 in England and Wales, criticized the relationship between SMRs and Morbidity. The results were as follows:

Mortality and acute sickness0,49 (not significant)

Mortality and bed sickness0,31 (not significant)

Mortality and chronic sickness0,73 (p < 0.05)

Forster drew the following conclusions:

"I have shown that there is no significant correlation between standardized mortality and acute sickness, and sickness causing work or school absence as measured by the GHS. Although there is a significant correlation between mortality and chronic sickness, the exacerbations of chronic sickness are in fact included in the acute sickness indicator. ... It seems reasonable to doubt, therefore, whether standardized mortality accurately measures relative need for health care. If SMRs are to be used in the calculation of revenue allocations to individual regions, as recommended by RAWP, the evidence from this study suggests that errors may occur in the distribution of funds."

Bevan and Mays²⁷ criticize Forster's study, suggesting that the conclusions drawn appear exaggerated as the results obtained could be used to support the use of SMRs as a morbidity indicator.

25 D. P. Forster, "Mortality, morbidity, and Resource Allocation" The Lancet, May 7, 1977, pp.997-998, p.997.

26 The General Health Survey is a routine health study using a sample population in the United Kingdom.

27 N. Mays, G. Bevan, Resource Allocation in the Health Service. A Review of the Methods of the Resource Allocation Working Party (RAWP), Bedford Square Press / NCVO, 1987, p.46.

"While 'acute' bed, and 'off-work' sickness were not found to be correlated with SMRs, these measures are likely to include episodes of illness which are not serious and which may not require any professional health service intervention. By contrast 'chronic' sickness which was associated with mortality, would seem far more likely to include severe, long-standing illness and disability which would warrant extended use of health services. Chronic sickness is responsible for far more use of National Health Service resources than acute sickness."

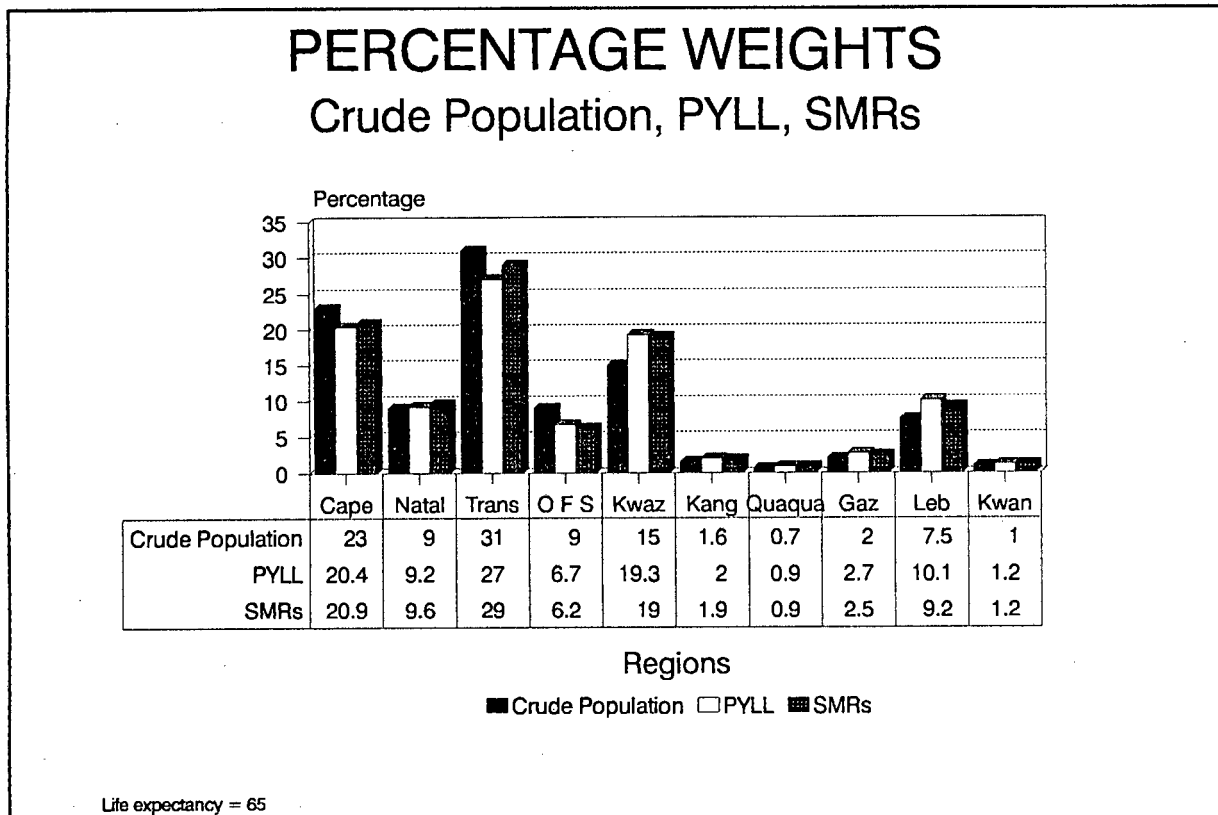
Despite the fact that SMRs do not accurately reflect the distribution of need for medical services, and embody a defect in that such services will be provided where other forms of social goods and services should have priority, this indicator remains the only viable method of defining regional morbidity. It is interesting to note that the use of PYLL may have a tendency to over-compensate, to a greater extent than SMRs, for socioeconomic deprivation.

The following graph illustrates the different weightings that would result if either SMRs or PYLL were used to weight the crude population. Both the SMRs and the PYLL are calculated by age and sex using 1985 mortality²⁸ and population data. A single life expectancy was used for the entire population²⁹. Figure 6.3 illustrates that both PYLL and SMRs indicate that need is higher, in the self-governing states, than would be indicated by crude population. In the Provinces the situation is the reverse except in Natal. If PYLL were used as the basis for redistribution, they would suggest a more extreme weighting than if SMRs were used.

28 As in the following section, the mortality rates in the self-governing states for blacks were adjusted upwards using the mortality rates for blacks in the Provinces.

29 A uniform life expectancy is used so as not to bias lower socioeconomic groups. It is therefore assumed that given optimal conditions every human being born is viable for 65 years.

Figure 6.3



The reason for the greater weighting, as already explained, is a consequence of the bias toward younger demographic profiles, which have higher mortality in younger age cohorts. Insofar as the SMRs will already result in some degree of inappropriate redistribution, with respect to in-patient services, PYLL would only make this situation worse.

6.6 Using the Formula to Indicate Relative Need for In-Patient Services on a Geographical Basis

Taking the above criteria into account, it is possible to derive a formula by which the geographical variation in need can be calculated for in-patient services. In this section the formula is used to indicate the variation in levels of relative need by region, in South Africa, for in-patient services, and then compared

to expenditure data. The equation used is the one developed by the Resource Allocation Working Party for England and Wales and is as follows:³⁰

**A NEED-BASED FORMULA
FOR DEFINING VARIATIONS IN NEED FOR
IN-PATIENT SERVICES ON A REGIONAL BASIS**

$$\left[\sum_i \sum_j \sum_k \frac{R_{jk} \cdot NB_{ijk} \cdot SMR_{jk}}{NP_{jk}} \right] \left[\sum_r \sum_j \sum_k RP_{rjk} \right]$$

RP: Regional Population
NB: National Bed Utilization
NP: National Population

i: Age
j: Sex
k: Condition
r: Region

The first bracket relates to the multiplication of the factors to produce expected bed-days for each Region as a fraction of the total. The second bracket is the national population which is used to convert the fraction to a weighted population³¹. This equation was not followed exactly as condition specific utilization rates could not be obtained. Thus all the references to condition in the equation should be ignored for the purposes of this exercise.

The mortality rates and crude population figures were obtained from the Central Statistical Services publications for 1985. The mortality rates for blacks in the self-governing states were adjusted upwards due to the significant undercounting evident in these regions. In fact, the general unreliability of mortality data collection for blacks is such that Central Statistical Services does not produce Life Tables for South African blacks.³² In order to partially correct these figures the mortality rates for blacks in the Provinces,

30 DHSS, 1976, p.101.

31 DHSS, 1976, p.101.

32 Central Statistical Services, South African Life Tables 1979-81: Whites (1985), gov. printer.

where it was assumed the data collection process would be more reliable, were taken as a proxy for the mortality rates of blacks in the self-governing states.³³ This may still involve a significant undercount as mortality of blacks in the homelands has been found to be higher than that of blacks living in the Provinces.

In working out the SMRs, blacks, whites, coloureds and Asians were evaluated separately, by age and sex, before being grouped together to give one single-need weighted population figure for each region in South Africa. Table 6.3 shows the results of weighting the populations by means of the above equation.

Table 6.3

POPULATIONS WEIGHTED FOR BED UTILIZATION AND SMRS COMPARED TO TOTAL HEALTH EXPENDITURE BY REGION.

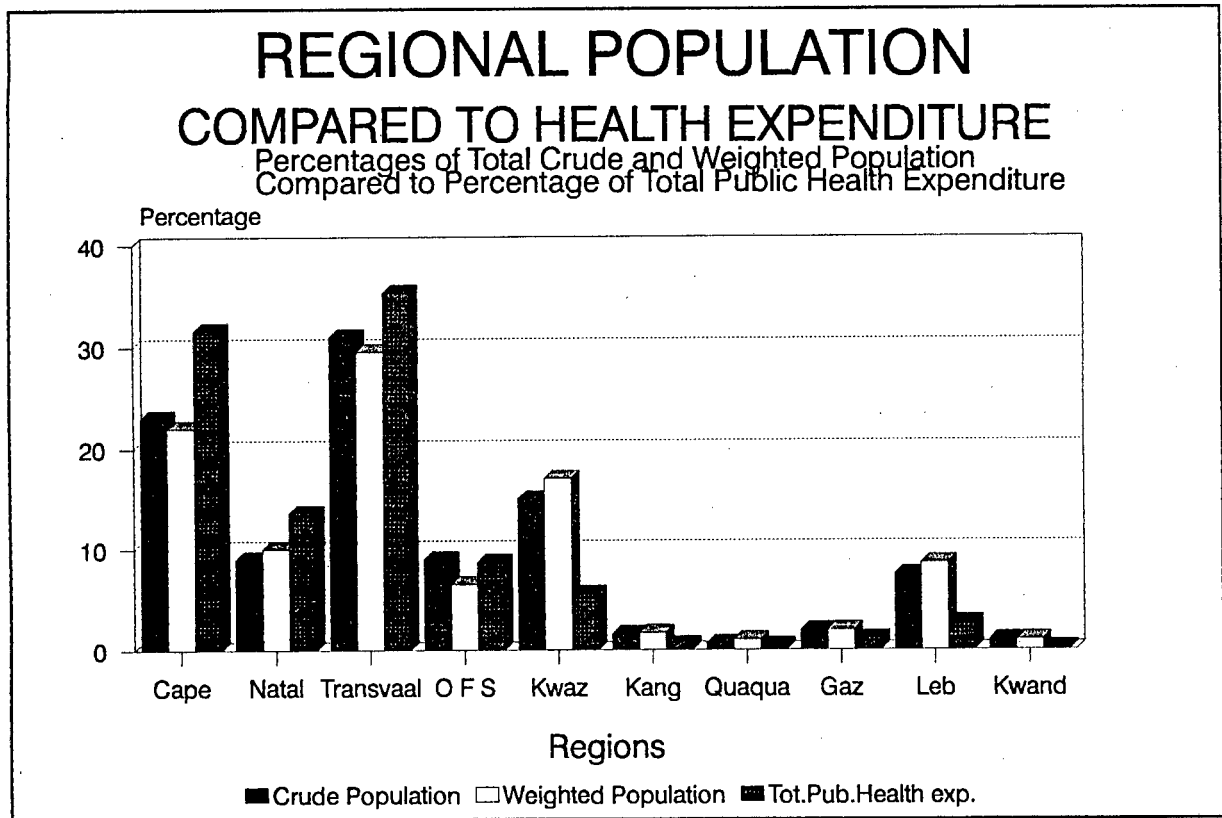
Province	Crude Population (1985)	Weighted Population	Public Expenditure on Health ³⁴ (1986/87) (R 000)
Cape	5 591 067	5 478 322	1 215.6
Natal	2 145 018	2 383 788	523.3
Transvaal	7 531 513	7 203 969	1 360.4
Orange Free State	2 193 403	1 559 527	335.5
Kwazulu	3 746 835	4 212 353	214.8
Kangwane	392 782	420 528	23.0
QuaQua	181 559	206 058	20.9
Gazankulu	496 943	534 991	42.5
Lebowe	1 835 586	2 092 918	105.0
Kwandebele	235 855	258 102	8.9
TOTAL	24 350 561	24 350 561	3 349.2

Figure 6.4 shows the above figures as percentages of the totals.

33 The TBVC states were excluded for the sake of convenience.

34 See section 5.3.1 table 5.1.

Figure 6.4



The table and the graph indicate a clear disparity in the broad distribution of health expenditure in South Africa. Even if crude population was the primary indicator, the maldistribution remains significant. In most of self-governing states, the need-based formula adjusted their base populations upwards, while in three of the Provinces the populations were adjusted downwards. As far as the distribution of expenditure is concerned, all the Provinces receive funding disproportionately, in terms of both crude and weighted population, when compared with the homelands.

It must be remembered, however, that the expenditure data presented here reflects expenditure on all health services, and not just in-patient services. Nevertheless, as in-patient services constitute the bulk of health expenditure, it is a reasonable indicator of the flow of funds. Private-sector expenditure on health

is not included in the above figures. If they were included, the bias towards the Provinces, in terms of expenditure, would be greater, particularly when taking into account that private expenditure on health constitutes 45% of total health expenditure in South Africa (chapter 5). Very little of this is spent in the self-governing states.

6.7 Further Aspects of the Formula Approach

6.7.1 Service Categories

In-patient services are not the only category of service for which a need-based formula should be derived. Out-patient services, primary health care services, ambulance services, etc. should all be assessed in terms of appropriate criteria, specific to their functions, so that relative need on a regional basis can be determined for the purpose of allocating resources efficiently. The RAWP formula weights the regional populations in England and Wales with regard to seven service categories: non-psychiatric inpatients; all-day patients and outpatients; community health services; and family practitioner committee administration. Seven different population figures are produced for each region and a weighted average of these is taken to determine regional revenue and capital targets. The weighting given to each of the seven service groups was derived from past expenditure breakdowns.³⁵

Table 6.4 shows the proportions of revenue expenditure allocated to each service category, in the RAWP formula.

Table 6.4

ESTIMATED PROPORTIONS OF REVENUE EXPENDITURE ON EACH SERVICE CATEGORY (UNITED KINGDOM)³⁶

Non-psychiatric inpatients	55.9%
Day and outpatients	13.4%
Community health	8.8%
Ambulances	3.5%
Mental illness inpatients	12.2%
Mental handicap inpatients	5.7%
FPC administration	0.5%

Based on the methodology used in RAWP, it may be useful to categorize South Africa's health sector similarly, and to develop an appropriate formula for each category. A possible point of departure from the RAWP formula would be to calculate the relative weightings between service categories according to criteria other than historical spending patterns. For instance, it may be more appropriate to construct and finance several Primary Health Care Clinics in a region characterized by low socioeconomic living conditions, rather than build and finance a single high-cost in-patient facility. In a region characterized by low socioeconomic living conditions, a method should be found to put a greater weighting on those facilities that offer primary preventive rather than tertiary preventive (ie. curative) services. Thus, not only should there be a formula evaluating need on a regional basis, but there should also be a means of weighting service categories.

There are so many departments in South Africa dealing with functions which overlap that the distribution of funds according to this formula may prove difficult. However, once the service categories have been

adequately defined, the distribution of funds between different health authorities within the same geographical region could be calculated according to the service categories or the proportion of them for which they are responsible.

In summary therefore, it is essential that adequate service categories be defined, and that a separate formula be created for each of them. Furthermore, criteria should be developed to determine the relative importance of each service category to the others. Finally, these service categories should provide the basis on which different authorities within the same geographical region will be financed.

6.7.2 Targeting Resource Needs

Achieving spatial redistribution and distribution goals must take into account the overall level, as well as the nature (ie. preventive or curative) of the medical care required. Sharing the existing resources equally could negatively influence the entire system, as well-financed regions are not necessarily over-funded. Target levels of facilities, by service category, based on the formula, could prove a useful criterion for allocating funds. This could imply, however, that total government expenditure on health would have to increase in real terms in order to achieve equity.

Target levels of service provision, on a regional basis, should define the eventual levels of required expenditure in each region. In other words, the need of a particular region determines the target level of services, and the target level of services determines the required expenditure. If a region is above target, services can be reduced, and if below, they can be raised. Suggestions as to how this adjustment process could be implemented are discussed in section 6.8.

Achieving target levels of resource provision is a concept developed by the RAWP,³⁷ and was implemented

37 DHSS, 1976.

in 1976. This method of resource allocation was upheld in the RAWP review of 1986.³⁸ In the original RAWP it was decided to allow the actual regions themselves discretion as to how the allocated funds were to be spent in achieving particular targets. This was also upheld in the 1986 RAWP review.³⁹

6.7.3 Differential Costs

The costs of constructing and staffing various facilities could differ significantly between developed and under-developed regions. In outer-lying, under-developed regions, an income incentive may be necessary to attract highly qualified staff. Furthermore, the costs of constructing and maintaining facilities in areas with limited infrastructure and skilled manpower could be considerably higher than in developed areas. Therefore, the distribution of funds, based on the formula, must take into account these differences in what the same quantity of money can buy in different regions.

6.8 Ideas on the Practical Application of the Formula Approach

Defining how redistribution should take place, and at what pace is probably the most important aspect of the formula. If the obvious need of certain geographical regions relative to others has not stopped the misallocation of resources in the past, why should a more rigorous definition of need make any difference in the present? Criteria need to be developed suggesting how the information processed in the formula should be acted on. Furthermore, as the allocation of funds for current expenditure requirements is closely related to the existence of capital stock, the redistribution process must account for both capital and current expenditure. In other words, it will not be possible to alter the levels of current expenditure, unless the level of capital stock is altered first.

38 Department of Health and Social Services, Review of the RAWP Formula, Report by the NHS Management Board, 1986.

39 DHSS, 1986, p.8.

Another important aspect to consider is that cross border flows of people can be accounted for in this part of the formula. If one particular geographical region is providing extensive medical services for a neighbouring region, there is no sense in moving vast sums of money into the under-provided area at the cost of the over-funded area in a relatively short period of time. The adjustment process must occur at such a pace that as the improved facilities in the under-funded area reduce the cross-border flows, the excess capacity is reduced in the over-funded area. A suggestion regarding the method of redistribution is now presented, based on the points already mentioned.

The first aspect to note is that existing capital stock acts as a kind of money base with a multiplier effect: the larger (smaller) the capital stock the larger (smaller) the recurrent expenditure. It should be possible to determine a fairly constant relationship between capital stock and recurrent expenditure⁴⁰ (both by service category). If a certain level of capital stock already exists in a particular area, redistribution of current expenditure away from this area without reducing the capital stock, could result in an overall decline in the efficiency and effectiveness of the services being provided. Therefore, the problem is not the speed at which redistribution of funds for recurrent expenditure should take place, but rather the speed at which the existing capital stock is re-arranged. This re-arrangement will be determined by the following factors:

- a the level of capital expenditure;
- b the rate of natural depletion of aging capital stock; and
- c the ability of the various departments either to shut down facilities or otherwise dispose of them.

The most convenient method of reducing existing levels of state funded capital stock is to turn facilities over to the private sector for use. Ownership of the facilities should remain vested in the public sector, but the stock could be leased, or merely allowed to compete with the private sector for patients.

40 For a discussion on the various issues relating to recurrent expenditure problems see C. Waddington, and M. Thomas, "Recurrent Costs in the Health Sectors of Developing Countries", International Journal of Health Planning and Management, 1988, Vol.3, pp.151-166.

A further method is to establish Health Maintenance Organizations (HMOs)⁴¹ that are state-owned, but funded externally ie. through pre-payment. These methods of disposing of capital stock will retain a useful degree of flexibility. Furthermore, allowing the public sector to compete with the private sector for paying patients will create the unusual anomaly of making the private medical-care sector more efficient and effective.

To summarize: it is suggested that the proposed redistribution of revenue expenditure, according to the need- based formula, should occur at a rate consistent with the redistribution of capital stock. The obvious implication of this suggestion is, therefore, that the distribution of capital stock and capital expenditure should be dictated by the formula in such a way that the distribution of revenue expenditure reaches the targeted levels. The speed at which this redistribution takes place will be determined by the rate at which capital stock can be created in under-funded areas, and the rate at which capital stock can be depleted in over-funded areas.

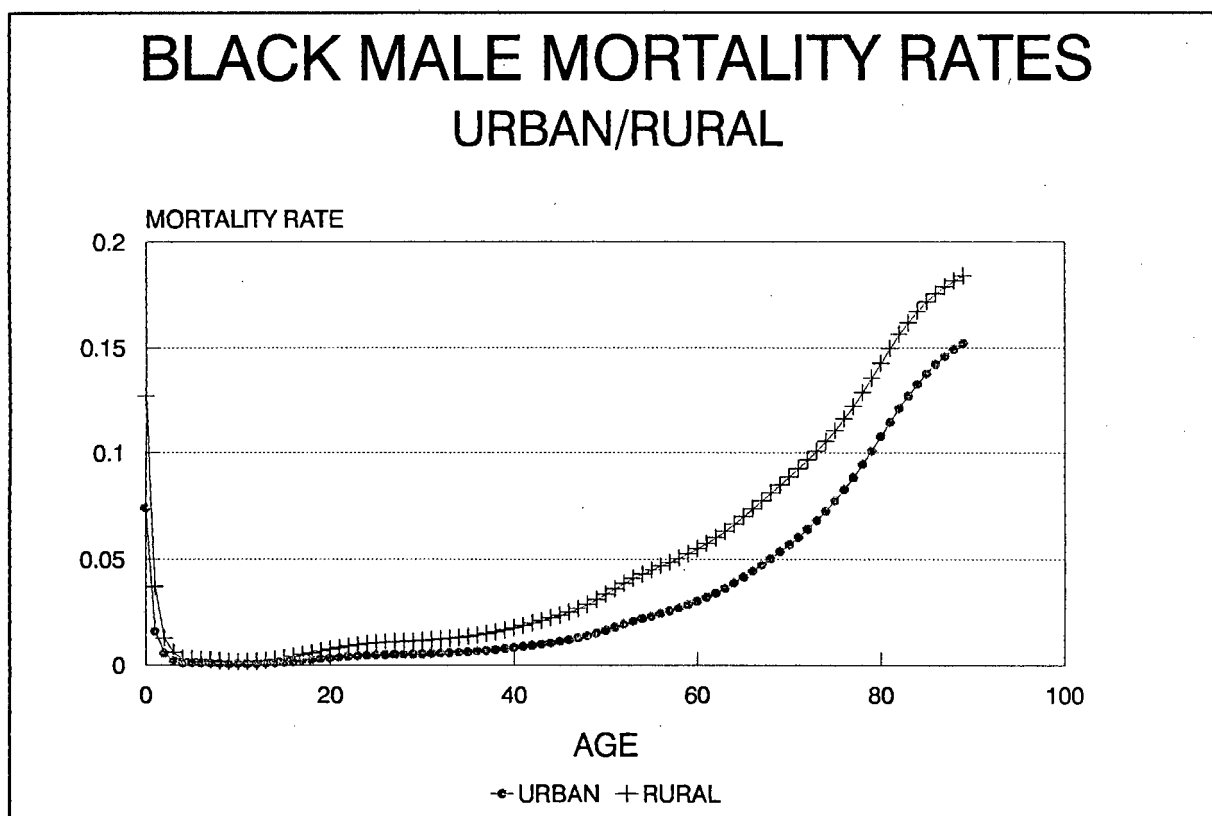
6.9 Regions

The basic units for relative evaluation used in this chapter have been large, ie. at the Provincial and homeland level. As the method of distributing finances from Central Government to the various Departments is done on a similar basis, the methodology presented is relevant to this broad distribution. However, more specific regional variations in need for health services could prove to be useful in defining, fairly specifically, how and where to spend. The variation in need between rural and urban areas could then become an issue dealt with in an objective manner according to acceptable criteria.

41 A Health Maintenance Organization (HMO) is a medical facility that is financed on a pre-payment (similar to a premium) basis by members. The members theoretically never pay anything over-and-above their pre-payment for treatment received. The difference between belonging to an HMO and a medical aid, is that in an HMO the third party payer and provider are one and the same, whereas with a medical aid scheme, the third party payer and the provider are separate.

The following two figures (6.5 and 6.6) showing the variations in mortality rates from age 0 to 85, and comparing the mortality rates of blacks living in rural areas to those of blacks in urban areas, indicate the strong possibility that morbidity differences between these areas are significant^{42 43}.

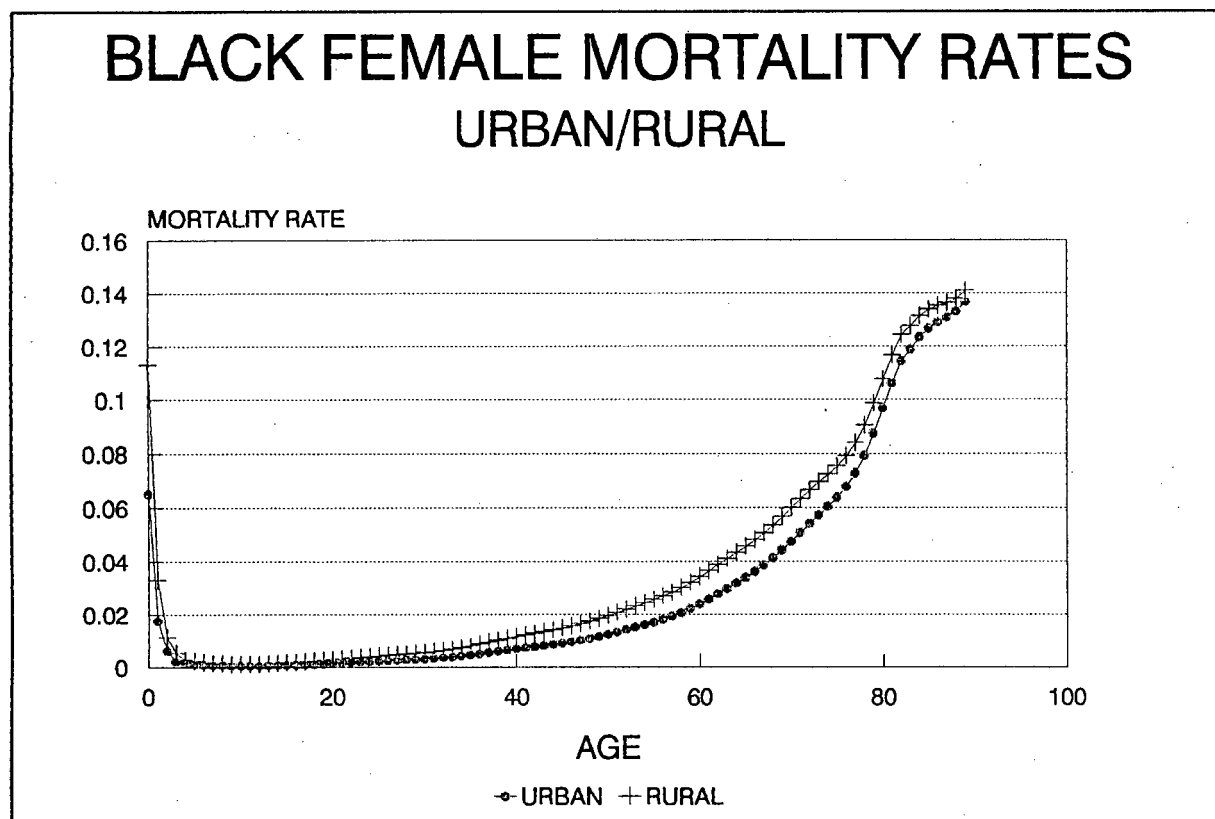
Figure 6.5



42 R. E. Dorrington, "African Morbidity Rates - An Initial Estimate" Transactions of the Actuarial Society of South Africa, ed. R. E. Dorrington, C. J. B. Greeff, 1987, Vol. (vii) part 1, pp.169-197.

43 The data used to produce figures 6.5 and 6.6 were corrected as far as possible through the use of coloured life tables as life tables for blacks are not available.

Figure 6.6



However, the optimal size of region for use in revenue allocation needs to be determined. The use of magisterial districts may be too small⁴⁴. Statistical regions may also be too small as far as revenue distribution is concerned but, in their case, application of the formula may still prove useful in determining priorities. In terms of defining the revenue distribution more specifically, it is possible that Planning Regions could prove to be a useful basis. However, it should be taken into account that too explicit an instruction on the disposal of revenue expenditure, emanating from Central Government, could antagonize

44 From personal communication with D. E. Bourne, UCT Medical School, Department of Community Health. According to his research, the determination of SMRs on a magisterial district level is so affected by "noise" and irregular methods of data collection, as to render them useless for the purposes of effectively defining relative need.

local administrators and even cause inefficiency.⁴⁵

6.10 Cross-Border Flows

The extent to which one regional authority in South Africa provides services for another's population cannot be quantified at present due to the complexity of the present institutional set-up. Although some aspects of this problem are dealt with in section 6.8, it is necessary to elaborate further.

The reasons for cross-boundary flows are as follows:

- 1 Inadequate spatial distribution of medical resources, ie. the product of necessity and the result of difficulty of access⁴⁶; and
- 2 Administrative boundaries that are not consistent with the population served by medical facilities within those boundaries. These could be termed flows of convenience and can be regarded as rational.⁴⁷

The flows resulting from (1) above are covered in section 6.8. The flows resulting from (2), however, cannot be compensated by resource redistribution, as this is not the problem. The fact that the administrative boundary does not reflect the catchment area has to be compensated for in the formula. If it is not, serious contradictions could arise.

45 The level of specification is affected by the accuracy of the data. In the present instance the data is not sufficiently accurate to define anything more than a general indication of the state of affairs. It should therefore be up to the administrators, on the spot, to interpret how to spend the available funds, given more specific information.

46 N. Mays et al, p.17.

47 N. Mays et al, p.17.

The RAWP formula, in use in England and Wales, makes a distinction between the population within an authority's boundary and the population served by the medical facilities. On the basis of the adjustment made to the formula in this regard, the regional authorities can more accurately target their resource requirements. The method of calculating cross-boundary flows was based on "historic (net) cross-boundary flows costed at average specialty costs".⁴⁸

Although there has been no criticism of the principle of compensating the various regional authorities for cross-boundary flows in the UK, the method by which it is achieved has been criticised. The basis of the criticism is not discussed here as it may not be relevant to the South African situation.⁴⁹

What is of interest, however, is the alternative method for compensating cross-boundary flows referred to as cross-charging. Cross-charging involves the creation of an internal market within the public-health sector, between regional authorities, whereby a regional authority is charged if residents under their jurisdiction receive treatment at another regional authority's medical facility. This raises many possibilities. In Britain the following suggestions have been made:

"Whereas under RAWP, District managers have no control over their residents' use of services outside the District, cross-border charging gives Districts the power to choose where their residents may be treated, since Districts will only pay for the treatment of their residents in other health authorities when this has been the subject of a specific contract. Since the contract is likely to give full details of the services to be provided at agreed costs and volumes, all cross-boundary flows for non-emergency cases will be controllable, predictable and

48 DHSS, 1986, p.22.

49 For a discussion of the problems with the RAWP methodology in this context the reader is referred to the RAWP review, DHSS, 1986, pp.22-26.

*accounted for, thus obviating the need for the estimation of catchment populations.*⁵⁰

In South Africa a RAWP type formula could be used to allocate resources according to need (for in-patients) while at the same time allowing regional authorities to have the ability to control residents' use of resources outside of the region. The prices charged could be decided on by the region receiving patients from outside its jurisdiction. This method would allow the possibility of price competition within the public health sector. Supplier-induced demand (described in section 6.3) would be limited in this instance as doctors would be purchasing services on behalf of patients.

It is possible that a method, such as the one described above, could be used to redress the problem of spatial inequity, to a degree, and eliminate the problem of natural cross boundary flows. In Britain the problems with cross-boundary flows are mainly at sub-regional level. In South Africa this is a major consideration even at Provincial and homeland level. However, this is mainly caused by the spatial maldistribution of resources rather than natural flows. The problems of spatial distribution at a sub-regional level can be expected to be significant and should be addressed in the formulation of a RAWP type formula for South Africa. At the present time a methodological approach as put forward in this chapter would not be able to achieve many of its goals as the institutional set-up is too complex and cumbersome. Nevertheless, without the formula, the prevailing scenario could still use the internal-market concept.

6.11 Conclusion

The concern of this chapter has been the assessment of a methodology for redistributing existing, and distributing new, health resources on a need basis. The creation and utilization of a formula has been discussed as well as the various elements that should be included in it. The criteria of need for in-patient services have largely been discussed by way of example, especially as expenditure for this service category

50 N. Mays et al, p.104.

constitutes the bulk of the present health budget. However, outpatient, day care, and primary-health-care facilities could all be evaluated by the use of similar criteria as set out for in-patient facilities. In this connection it has been suggested that PYLL be used instead of SMRs in defining need for these service categories⁵¹. This will positively discriminate in favour of younger age cohorts with high mortality rates in youth. The logic behind the use of PYLL in this regard is that in South Africa lower socioeconomic groups tend to require more preventive type services, and more primary medical care than other groups. SMRs would still be useful, but would not stress socioeconomic deprivation to the same degree.

Two complications not yet mentioned, are:

- a How to deal with teaching facilities; and
- b How to account for the private sector within the formula.

Teaching hospitals tend to be the biggest expenditure item on the Provincial health budgets in South Africa.⁵² Teaching hospitals, however, provide a service for the country as a whole and not merely for the region in which it is situated. Excluding teaching hospitals from a formula that attempts to define regional disparities in need, and correct them, will result in any reallocation becoming, possibly, insignificant. The following comments present a way of dealing with this problem.

It is important to distinguish between the two functions of a teaching facility:

- a It provides medical goods and services; and

51 Suggested by D. McIntyre and S. Taylor UCT Medical School, Department of Community Health.

52 For instance in the Cape Province just two teaching hospitals, Groote Schuur and Tygerberg, account for 40% of the total budget allocated for that Province. S. P. Taylor, J. M. L. Klopper, "South African Health Care Expenditure 1975-1984" South African Medical Journal, December, 1987, Vol.72, pp.802-804, p.803.

b It trains medical personnel.

The training of medical staff provides a service that responds to the need of the country as a whole and not of a particular region. Therefore, insofar as a hospital requires greater funding to finance this teaching, the allocation of expenditure on this basis should not be determined by an inter-regional need-based formula. However, insofar as a hospital, teaching facility or not, provides medical goods and services for the public, this must be taken into account within the inter-regional need-based formula.

Populations served by private medical facilities should be excluded from the formula. Possibly a method that could be applied here, is to exclude all people who are included within some sort of medical aid scheme from the base population. All those members of medical aid schemes who are subsequently compelled to use public facilities would then be required to refund the medical facility in full.

Finally, what such a formula would hope to achieve, is an objective method of distributing resources on a need basis in such a way that appropriate services are provided where they are needed. By such a method regional inequalities in access to medical care can be reduced over time. This chapter establishes a methodology for breaking the grasp that historical spending patterns exercise over public- health financing. The next two chapters deal with a methodology for establishing the most effective interventions on the micro-level.

CHAPTER SEVEN

COST-EFFECTIVENESS ANALYSIS IN THE HEALTH SECTOR

7.1 Introduction

Cost-effectiveness analysis (CEA) has been developed as a tool for economic decision-making in situations where a choice has to be made between at least two alternatives for attaining a particular objective. This chapter is primarily concerned with a description of CEA and its relationship with health planning and decision-making. The technique has much to offer in creating a sound basis for micro-economic decision-making in the health sector.

An important aspect concerning health planning, especially in a developing country such as South Africa, is the necessity not only for an even geographical spread of resources, but also for those resources to be provided in accordance with actual need. The complexity of the health sector contributes to the extreme difficulty central planning authorities have in making strategic micro-economic decisions. In other words, although a hospital or clinic may be involved in decision-making that is micro in nature, the results of those decisions usually force the hands of central planners. Micro decisions, to a large extent, determine macro decisions. It is essential, therefore, that, as far as possible, the micro-economic actions of the health sector reflect the actual need of a community, rather than perceived demand. For example, the fact that a clinic sees many cases of Tuberculosis reflects a need for primary preventive interventions although the perceived demand is for tertiary preventive (ie. curative) interventions.

Cost-effectiveness analysis would hope to identify the most appropriate intervention, given the specifics of the community or population that is being evaluated. In this way the most effective intervention can be determined according to criteria of need.

7.2 Cost Benefit Analysis and Cost-Effectiveness Analysis

7.2.1 Cost Benefit Analysis

Cost benefit analysis (CBA) simply involves the calculation of both the advantages and disadvantages of implementing a particular project or programme. If the benefits exceed the costs then the project or programme becomes viable. Usually the costs and benefits are evaluated using a common unit of value, usually money.

7.2.2 Cost-Effectiveness Analysis

Cost-effectiveness analysis (CEA) attempts to evaluate the best option between various alternatives. The following is a definition offered by Lee:¹

CEA "aims to ascertain either the programme capable of achieving certain specified goals at the lowest possible cost, or the programme which will maximize the benefit to be gained from a defined budget. In either case one of the parameters is fixed and the objective is simply to maximize the benefit or to minimize the cost."

It is important to note that the findings may in some cases indicate that a rise in spending may be required in order to improve the cost-effectiveness ratio. Furthermore, the output (ie. the benefit aspect) of the health programme is not expressed in terms of money, or a common numeraire with the cost element.

1 K. Lee, A. Mills, "The Role of Economists and Economics in Health Service Planning: A General Overview", Economics and Health Planning, ed. K. Lee, Biddles Ltd., Great Britain, 1979, p.166.

7.3 "Output" or "Outcome" Measurement in CEA

The output of the health sector can be easily determined, ie. number of patients treated, number of injections administered, etc.. However, as the medical field is far from being an exact science, output, expressed in this way, is not a sufficient criterion on which to evaluate the efficiency or effectiveness of medical procedures. What is required is a measure that reflects the actual consequences of the procedure on the patient.

Deliberations along these lines have led to the identification of outcome as opposed to output measures. An outcome measure, theoretically, reflects the actual success, or benefit, which a particular medical facility or procedure has had in curing or preventing ailments. A further disadvantage with output measures relates to their being applicable only to certain procedures. For instance the number of coronary bypass operations performed cannot be compared to the number of injections administered. The benefit measure should be comparable to the measures used in other interventions.

It is, furthermore, not sufficient that this measure merely evaluate the quantitative aspect of outcome, ie. life-years extended. A qualitative aspect is crucial to the accurate determination of outcome. The extra benefit provided in terms of life years may not be of great value if the quality of those life years is poor.

The following three aspects can therefore be regarded as necessary ingredients for an effective outcome measure:

- 1 It should be non-procedure specific;
- 2 It should indicate life-years extended; and
- 3 It should indicate the quality of those extended life years.

Essentially (1) and (2) are both incorporated in a measure such as Potential Years of Life Saved (PYLS). This measure can be determined fairly easily in consultation with medical experts. The utilization of estimations of survival rates, published in the established literature, as well as statistical sampling of existing case evaluations can also be used. The qualitative aspect, however, is the most difficult aspect to determine, and has become the focus of much attention over the past 15 years. Section 7.4 therefore covers some of the views and knowledge that have evolved over the years in this area.

7.3.1 Quality of Life²

The primary concern of work in this field of study has been focused on the development and application of a Health Status Index (HSI).^{3 4 5} This appears to be the only approach by which outcome valuation can be standardized and objective. Although HSIs have, as yet, not fully achieved either of these goals, they have gone a long way to establishing a sound methodology.

7.3.1.a Health Status Indices

The principle behind the utilization of an HSI is to obtain an accurate, objective and consistent

-
2. It can be argued that when quality-of-life measures are brought into account what is actually being measured is the utility of patients in terms of their health status. It could therefore be suggested that what is being performed is a cost-utility analysis. However, it should be noted that utility evaluation merely contributes to the level of specification of an outcome measure, and in no way detracts from the original objectives or methodology of CEA. Therefore, no further mention is made in this thesis of cost-utility analysis as the concept is contained within the definition of CEA.
 3. R. Rossér, P. Kind, "A Scale of Valuations of States of Illness: is There a Social Consensus?" International Journal of Epidemiology, 1978, Vol.7, No.4, pp.347-358.
 4. H. Sintonen, "An Approach to Measuring and Valuing Health States" Social Science Medicine, 1981, Vol.15C, pp.55-56.
 5. D. L. Sackett, L. W. Chambers, A. S. MacPherson, C. H. Goldsmith, R. G. Mcauley, "The Development and Application of Indices of Health: General Methods and a Summary of Results", American Journal of Public Health, May, 1977, Vol.67, No.5, pp.423-428.

methodological approach to valuing the quality of life. A description of an HSI developed by Torrance⁶ is as follows:

"The central concept underlying this approach is that at any point in time an individual occupies one of a continuum of possible health states ranging from good health to death. For estimation purposes the continuum is divided into discrete health states, $j=1\dots n$. Each state is assigned a value U_j on a utility scale ranging from 0 to 1. Health interventions alter the probability of occupying different health states. The effectiveness of an intervention is measured in terms of the increment in health status units, defined as the discounted sum over all years of the number of days changed from state j to k in each year multiplied by the utility weight of state j ."

The information required to assess where a particular patient would be situated on this index can be determined through consultation with medical practitioners. However, a major difficulty with HSIs involves the construction of the criteria to be included within the index, and the order in which the various possibilities should be ranked.

The assistance of medical practitioners would be required to identify those functional levels that a particular patient would be in, and the duration of being classified at a particular level. The defining of a particular functional level and the contributing factors that will enable a medical practitioner to decide which level adequately explains the condition of a patient, is illustrated by the following HSI (table 7.1).⁷ Here health states are defined, each representing a particular combination of descriptive statements on three

6 G. W. Torrance, "Health Status Index Models: a Simplified Mathematical View", Mgt. Sci., 1976, vol.22, pp.990-1001.

7 J. W. Bush, M. M. Chen, D. L. Patrick, "Health status indices in cost-effectiveness: analysis of a PKU program", Health Status Indices, ed. Berg, R. L. Hospital Research and Educational Trust, Chicago, 1973, pp.172-208.

dimensions: mobility, physical activity, and social activity. Two sets of individuals were asked to assign weights to different health states. Death was given a value of 0.000 and "symptom free" a value of 1.000. From the results it can be seen that there is disagreement between consultants' and students' rankings. Thus a crucial factor affecting the construction of HSIs is that individuals' values with respect to health status vary, and that some form of social consensus must be established before an HSI can be successful.

Table 7.1

PREFERENCE RATINGS FOR FUNCTION LEVELS⁸

Function Level	Mobility	Physical	Social activity	Preference rating	
				Consultants'	Students'
L30	Travelled freely (No symptom/problem complex)	Walked freely	Performed major and other activities	1.000	1.000
L29	Travelled freely (Symptom/problem complex present)	Walked freely	Performed major and other activities	0.845	0.804
L28	Travelled freely	Walked freely	Performed major but limited in other activities	0.805	0.689
L28	Travelled freely	Walked freely	Performed self-care but not major activity	0.580	0.648
L24	Travelled with difficulty	Walked freely	Performed major activity with limitations	0.610	0.536
L17	Confined to house	Walked freely	Performed self-care but not major activity	0.435	0.594
L10	In hospital	Walked freely	Required assistance with self-care	0.273	0.505
L9	In hospital	In bed or chair	Performed self-care but not major activity	0.290	0.534
L8	In hospital	In bed or chair	Required assistance with self-care	0.186	0.436
L7	In hospital	Walked freely	Performed self-care but not major activity	0.293	0.528
L3	In hospital	Walked freely	Required assistance with self-care	0.165	0.440
L2	In hospital	In bed or chair	Performed self-care but not major activity	0.205	0.440
L0				0.000	0.000

Kind⁹ suggests that, due to the paucity of instruments designed for measuring quality of life a system of shadow or proxy measurements has to be developed. Kind¹⁰ suggests the following possibilities.

Employment status	Leisure activity
Days of disability	Dependence on others
Anxiety	Dependence on drugs
Depression	Fatigue
Mood	Hospitalization
Sexual function	Problems with sleep
Relief of symptoms	Intellectual function

However, the link between these factors and quality of life is a matter of conjecture.

"Whilst the researcher or clinician may intuitively feel that there is a positive correlation between them, there is precious little hard evidence. Neither is there evidence about how they are perceived by providers or consumers of health care."¹¹

A single proxy index is weak in that movements from one state to another are assumed to be improvements in health status, when in many instances this assumption may not be valid. In order to correct for this short-coming many different indicators, reflecting different areas of activity and functioning should be used. However, it is difficult to evaluate how improvements of one indicator and deteriorations in another, affect overall benefit. Essentially what can be done is to use different instruments reduced to a single index score, using a weighting system which permits results on one dimension to be traded off against those on other dimensions.¹²

9 P. Kind, The Design and Construction of Quality of Life Measures, Discussion Paper 43, Centre for Health Economics, Health Economics Consortium, University of York, 1989.

10 P. Kind, p.4.

11 P. Kind, p.3.

12 P. Kind, p.5.

Sintonen¹³ viewed the qualitative side of health as consisting of three major components: perceived health, psycho-physical functioning, and social functioning. Social functioning could be divided into two components: social participation and ability to work. Psycho-physical functioning was broken down into 12 dimensions. They were as follows: Perceived health, seeing, breathing, sleeping, communication, moving, eating, intellectual functioning, incontinence, working, hearing, and social participation.

Rosser and Kind¹⁴ developed a classification of health states combining eight categories of disability and four of distress. Various groups of people were interviewed to determine how they valued the relative undesirability of the combinations of states. Six groups of subjects with different experiences of illness were interviewed. The resulting valuation matrix can be seen in table 7.2a and 7.2b.

Table 7.2a

TRANSFORMED VALUATIONS FOR 29 HEALTH STATES¹⁵

DISTRESS

DISABILITY	A	B	C	D
I	1.000	0.995	0.990	0.967
II	0.990	0.986	0.973	0.932
III	0.980	0.972	0.956	0.912
IV	0.964	0.956	0.942	0.870
V	0.946	0.935	0.900	0.700
VI	0.875	0.845	0.680	0.000
VII	0.677	0.564	0.000	-1.486
VIII	-1.028	-	-	-

13 H. Sintonen, pp.55-56.

14 R. Rosser et al, pp.348-349.

15 P. Kind, R. Rosser, A. Williams, "Valuation of Quality of Life: Some Psychometric Evidence", Value of Life and Safety, ed. M. W. Jones-Lee, 1982, pp.159-170.

Table 7.2b

CLASSIFICATION OF STATES OF ILLNESS**Disability**

- I** No disability
- II** Slight social disability
- III** Severe social disability
- IV** Choice of work or performance at work very severely limited. Housewives and old people able to do light housework only but unable to go shopping
- V** Unable to undertake any paid employment. Unable to continue any education. Old people confined to home except for escorted outings and short walks and unable to do shopping. Housewives able to perform a few simple tasks.
- VI** Confined to chair or wheel chair or able to move around in the home only with support from an assistant.
- VII** Confined to bed.
- VIII** Unconscious.

Distress

- A** No distress
- B** Mild distress
- C** Moderate distress
- D** Severe distress

The HSI shown in table 7.2a and 7.2b has been used in several cost-effectiveness studies.^{16 17} An interesting

¹⁶ C. Gudex, Qualys and their use by the health service, University of York, Centre for Economics, Discussion Paper 20, October, 1986.

feature of Rosser and Kind's index is the recognition of states with values less than 0, ie. where life is not worth living. This concept has been the subject of some debate and is discussed in the following section. Although such an HSI is not a perfect indicator of health status, it can be regarded as an essential requirement for CEA studies in health. If such an index is developed in South Africa, CEA studies can become a very useful tool.

7.3.1.b Valuation of Death as an Outcome

Not valuing death as the worst state creates a few problems. Firstly, if death is not the worst state, should the worst state be valued at 0.000.

"It could be argued that the opposite of 'pain-free ability to conduct normal activities', is, 'in very severe pain and totally unable to conduct normal activities'. It may be further asserted that this is worse than being unconscious (ie. in no pain but totally unable to conduct normal activities), and perhaps even worse than being dead."¹⁸

Sintonen suggests that there is actually not a problem with a utility value of death greater than 0, so long as death is regarded as the worst state. Difficulty arises when the quantitative and qualitative aspects of life are compared, as there is an implicit assumption in the quantitative valuation that the utility value of death is equal to zero. This aspect could be solved technically (as was done in the Rosser and Kind HSI shown in table 7.2) by assigning death, in the qualitative index, the value of 0.000, and transforming the values of other levels so as to retain their existing distance relationships.¹⁹ Death as a state and death as an event

17 A. H. Williams, "Economics of Coronary Artery Bypass Grafts", British Medical Journal, 1985, Vol.291, pp.326-329.

18 A. C. Williams, "Welfare Economics and Health Status Measurement", Health, Economics, and Health Economics, North Holland Publishing Company, Amsterdam, (1981), pp.271-282, p.274.

19 H. Sintonen, p.60.

would be distinguished and yet remain comparable.

7.3.2 Duration of Health States

The "duration" referred to in the title to this section has nothing to do with the quantitative aspects of life, but rather to the weighting an individual would assign to a particular health state depending upon the time that individual would spend in that health state. For example, on the Rosser and Kind index (table 7.2) an individual may weight the combination "severe distress", coupled with "confined to bed", quite differently if the period was half an hour compared to several years. In fact, this aspect of valuing health states has led to the development of an alternative to the HSI. This is referred to as the Time Trade-off Technique (TTO).

The TTO approach is an equivalence technique which involves subjects judging how many years in a state of full health are equivalent to a given number of years in a described imperfect state.²⁰

"The subject is offered two alternatives - alternative 1: state i for time t (life expectancy for an individual with the chronic condition) followed by death; and alternative 2: healthy for time $x < t$ followed by death. Time x is varied until the respondent is indifferent between the two alternatives, at which point the required preference value for state i is given by $h_i = x/t$."²¹

Sintonen²² in deriving an HSI, however, felt that it was not possible to derive values for all possible health

20 M. Buxton, J. Ashby, "The Time Trade-off Approach to Health State Valuation", Measuring Health: A Practical Approach, ed. G. Teeling-Smith, John Wiley & Sons Ltd., 1988, pp.69-87, p.69.

21 G. W. Torrance, "Measurement of health state utilities for economic appraisal: a review", Journal of Health Economics, 1986, vol.5, pp.1-30, p.23.

22 H. Sintonen, p.60.

states and all possible durations.

The "duration of health states are in practice more or less uncertain. It was felt that the values to be assigned should reflect this uncertainty and, therefore, no reference was made to potential duration. In consequence, subjects had to define their attitude to uncertainty and the resulting mean values can be claimed to reflect the distribution of this attitude among subjects."

As both the HSI approach and the TTO approach are in the process of being developed it is beyond the scope of the present requirements of this thesis to make any relative assessment of the various approaches. It is sufficient for the purposes of this chapter merely to establish the existence and importance of such work. Both the HSI and TTO approaches are worthwhile attempts at quantifying the qualitative aspects of life. The more advanced such approaches become the greater will be the accuracy of outcome measurement. It is also important to make medical practitioners more aware of the need for such assessments, and to establish a basis for the routine generation of statistics that can be used for accurate outcome evaluation.

7.3.3 Relative Valuation of a Life Year

One issue that would never fail to cause concern is the question of whether quality adjusted life years can be regarded as a numeraire commodity. In other words, can a programme that gives ten quality adjusted life years to one person be regarded the same as a programme that gives one quality adjusted life year each to ten people?

The answer to this question lies in the fact that quality adjusted life years are used, and not merely life years extended. Due to the weighting of life years for quality, five actual years of life for one person may only equal one quality adjusted life year extended. Intrinsicly, therefore, the use of quality adjusted life

years establish a uniform value system for life years gained. Furthermore, the use of a discount rate to reflect the myopia involved in any decision involved with time enhances the numeraire quality of a such a valuation system.

7.3.4 Discounting Life Years

The principle of calculating life years cannot exclude the preference that most people have for favouring those life years that can be consumed now rather than later. However, the problem is not whether a discount rate exists or not, but rather, what rate to use.

There are essentially two view-points to this issue. The first argues that each individual applies his/her own valuation system between present and future periods, and that these valuations vary significantly. Somebody who smokes is revealing a different attitude to future (distant) life years than somebody who only consumes health foods. What this view-point would suggest is that a uniform discount rate should not be applied, as a society discount rate may be significantly different from individual discount rates. Therefore, no discount rate can be applied.

The second alternative, suggested by Williams²³, is based on the premise that health can be transferred into wealth, and vice versa, at any point in time.

Since "'wealth' is ideally allocated through time with reference to the rate of social time preference, then it would be inconsistent to apply a different rate of discount to 'health' from that being applied to 'wealth'."

Williams' view is the more consistent. To apply no discount rate is equivalent to applying a social discount rate of zero. Clearly this is further from reality than the discount rate applied to wealth. Essentially,

23 A. H. Williams, 1985, p.277.

therefore, the same discount rate that is applied to the cost side of a CEA would be applied to the outcome side.

7.4 Costs and Cost-Effectiveness Analysis

7.4.1 Indirect Costs

A feature of economic evaluations such as CBA and CEA when performed by the public sector, is the inclusion of indirect as well as direct costs. In many instances the indirect costs considerably exceed the direct costs of particular interventions. For example, in a study by Rice²⁴ the costs of various disease groups were calculated based on total health care expenditure in the United States of America in 1963, and on projections of best productivity value for 1963 and all future years for the populations concerned. It was found that the total cost of the grouping, "Diseases of the Circulatory System" was \$20,9 billion. Of this amount only \$2,2 billion could be attributed to the direct costs of the diseases.

It is intuitively clear that social costs and values are of major importance when establishing the soundness of a particular public sector expenditure. However, applying the principles of social cost and social utility is no easy matter, in terms of both definition and practical application. One difficulty arises concerning the source of price information used to value social costs. If price information is obtained from the private sector it must be determined whether these prices represent genuine social opportunity costs, or whether they are strongly biased by the distribution of income. Williams²⁵ takes the view that "practicality" should replace "idealism":

24 Cited in A. Crystal, A. W. Brewster, "Cost Benefit and Cost-Effectiveness Analysis in the Health Field: and Introduction", *Inquiry*, 1966, vol.III, No.4, pp.3-13, p.6.

25 A. H. Williams, 1985, p.279.

"... it is in societies' interests to give some precedence in matters of health care to those members of the society whose activities others value most highly. Not to do so seems perversely masochistic."

Thus, in this respect, it may be concluded provisionally that, where social values have to be estimated, the value that most closely approximates the opportunity cost to society should be used. This is rather vague, but in essence there is no course for the planner other than to use that information which is both feasible and obtainable.

A further issue is to evaluate which indirect costs to take into account. Factors such as: travel costs of patients, stress, lost work hours, etc., can all be relevant, depending on the interventions being analyzed. However, one of the most significant factors that should be included in the evaluation of indirect costs associated with health care is the cost of death, or rather the value of life. This aspect is, therefore, discussed in the following section.

7.4.1.a The Value of Life

One of the most difficult indirect costs to attribute a social value to is that of life itself. Culyer²⁶ mentions three methods of assigning a value to human life:

- 1 The social decisions approach;
- 2 The human capital approach; and
- 3 The value of risk avoidance approach.

26 A. J. Culyer, "Assessing Cost-effectiveness" Resources for Health, Technology Assessment for Policy Making, Praeger Publishers, Printed in the USA, 1982, pp.107-120, p.115.

Approach (1) refers to the value society places on a life, inferred by various public sector decisions about life-saving expenditure. However, this method is circular in its logic, ie. society must already have attributed a value to lives lost before obtaining a value by this method. On the whole, this method results in a very broad range of valuations. Card and Mooney²⁷, for example, have reviewed the value of life implications of past policy decisions in Britain, and found values ranging between 50 pounds and 20 million pounds.

Approach (2) is based on a valuation related to the earnings of an individual. This would be expressed in terms of the discounted present value of expected future earnings (gross of income tax) allowing for expected unemployment.²⁸ Rice²⁹ used the following criteria when utilizing this approach in an actual study,

"This method of derivation must consider life expectancy for different age and sex groups, changing patterns of earning at successive ages, varying labour force participation rates, imputed values for housewives services, and the appropriate discount rate to convert a stream of costs or benefits into present worth."

This method is usually objected to on moral grounds as it assumes that a person's social value is related solely to his/her contribution to the national output. The result of utilizing such a method is that the young and the old are discriminated against.

Approach (3) values a life on the basis of how much a society is prepared to pay in order to reduce the risk

27 Cited in A. Ludbrook, "A cost-effectiveness analysis of the treatment of chronic renal failure", Applied Economics, 1981, Vol.13, pp.337-350, p.339.

28 A. J. Culyer, p.115.

29 A. Crystaletal, p.4.

of a statistical death.

*"Suppose there is a population of 100 individuals at risk. Suppose also there is a procedure by which the risk that anyone of them will die prematurely can be reduced from 2 to 1 percent, so that if the procedure is adopted one premature death is expected to be avoided. Now ascertain ... the value of each of the 100 places on the reduction in risk and, since the procedure is expected to save one life, that sum is also the value of the statistical life saved. Plainly this can, in principle, exceed the whole wealth of even the wealthiest of the 100 individuals at risk."*³⁰

There are several problems with this approach as well. They can be listed as follows:

- a Individuals are not always aware of the probability of dying, or that they are at risk;
- b The value obtained could be strongly biased by the wealth of the people at risk; and
- c The methodology is complex.

Of the three methods the human capital approach is most often preferred in CBA studies. Despite its flaws it presents the most accessible methodology. However, a great deal of research is required in this area to improve the acceptability of valuing life in economic evaluations of this nature.

In the following chapter, a fourth alternative is used, based on the human capital approach. A value is attributed to life according to the annual expected consumption of a particular income group. This method discriminates against poor people. However, this factor did not affect the validity of the study, as only one income group was assumed.

30 A. J. Culyer, p.115.

7.4.2 Discounting Costs

The social rate of discount is an important concept in evaluating inter-temporal transformations in consumption. The reasoning behind the necessity for a discount rate can be based on the following two propositions:³¹

- 1 Society prefers the present to the future.
- 2 Future generations are likely to have higher levels of consumption. If the principle of diminishing marginal utility of consumption operates, then the utility gains to future generations from a gain in consumption will be less than the utility to the present generation from the same gain. Hence the future gain should be discounted.

The need to discount is not controversial. However, the rate that is to be used can be. A suggestion is to use market interest rates. This method could however involve an element of risk which is not present in public sector investments. The community may also take a different view when acting collectively than if they were to act as individuals. It could therefore be stated that social time preference will be lower than the observed market rates. Dasgupta and Pearce³² make the following observations:

"Will market rates reflect social time preference? Certainly the existence of positive interest rates suggest that people do discount future expenditures compared to present ones. If no allowance is made for risk, the relevant market rate would be the one which reflects the individuals' willingness to make risk-free loans, ie. the government borrowing rate. The use of this rate for social discounting has considerable attractions. Firstly, it is easily observed

31 Dasgupta, Pearce, Cost Benefit Analysis, Richard Clay, (The Chaucer Press) Ltd, Bungay, Suffolk, 1972, p.137.

32 Dasgupta et al, p.137.

(though some decision has to be reached about whether it is the long or short rate which is used). Secondly, it appears to meet the requirement generally made in private investment decisions that the rate of return should exceed the cost of capital to the firm. Since the cost of capital to the government is the borrowing rate, the analogy is met."

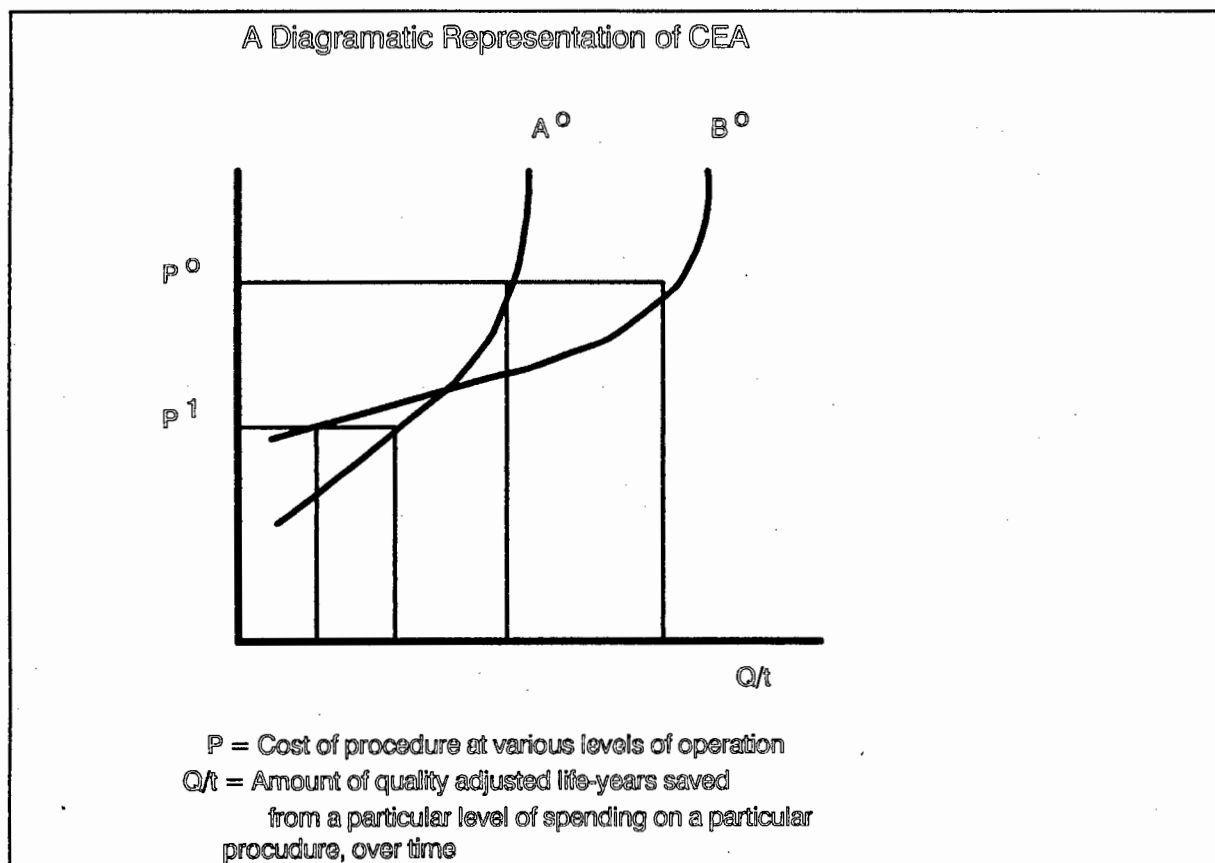
The above arguments concerning the cost of capital are appropriate for CBA studies where a single programme is evaluated through time to completion. As CEA evaluations refer to projects of an on-going nature, discount rates involving the cost of capital may not be completely appropriate.

As much of the literature refers to discount rates for CBA, the issue of how to evaluate expenditure through time with respect to CEA is not resolved. The question of which discount rate to use is resolved, to a degree, by the usage of official discount rates for public sector projects. Nevertheless, regardless of which theory is adhered to, any study should involve a sensitivity analysis where a range of discount rates is applied.

7.5 Conclusion

The value of CEA to health planning is derived from its ability to provide an objective and valid basis for differentiating between various procedures for dealing with the same problems. Given a particular context or environment in which certain procedures are being utilized, CEA will enable the most appropriate intervention to be emphasized, until such time as the context or the environment alters. Figure 7.1 illustrates what a CEA would be able to show.

Figure 7.1



The curves A° and B° represent two alternative procedures for treating the same disease. From the way that the curves have been drawn, at higher levels of expenditure (eg. at P°) B° is more cost-effective than A° . At lower levels of expenditure (eg. P^1) the reverse is true. This could be attributed to the type of inputs being used. A° could be a curative procedure, which at lower levels of expenditure proves to be more cost-effective than the preventive procedure. If the money is available to spend more, procedure B° proves to be the more cost-effective and saves many more lives. If there is a change in the exogenous variables (eg. a change in socioeconomic conditions) that could affect the risk of falling ill of this particular ailment, there would be a shift in the curves. If the risk factor was reduced this could shift B° to the left while A° could move to the right. In other words, the curative option could improve in effectiveness, while the preventive option would lose much of its usefulness.

Chapter eight involves an actual cost-effectiveness study and indicates the kind of information such research could provide for planning purposes. The study does not include the use of sophisticated outcome measures such as quality adjusted life years. Nevertheless, it does provide a useful comparison of alternative medical procedures despite the use of simpler measures of outcome and cost. The intention of including this explanation of CEA and the study in Chapter Eight, is to illustrate the potential that exists for solving resource allocation problems in the health sector. The technique of CEA as applied in the field of health resource allocation is becoming highly specialized and sophisticated. It is possible that the utilization of such methods of analysis could result in the saving of much money and many lives.

CHAPTER EIGHT

A COST-EFFECTIVENESS ANALYSIS OF A POTENTIAL COMPREHENSIVE NATIONAL SCREENING PROGRAMME FOR CERVICAL CANCER IN SOUTH AFRICA, USING A COMPUTER SIMULATION

8.1 Introduction

8.1.1 Objective of Chapter

The primary objective of this chapter is to demonstrate the potential usefulness of cost-effectiveness analysis for micro-economic decision-making in the health sector.

8.1.2 Motivation for Study

The choice for analysis of cervical cancer in the black female population was made for the following reasons:

- a It is a major cause of mortality among the black female population in South Africa. In 1987 total mortality for the disease stood at 869.¹ Using the population projection estimates described in section 8.7.1.4 and black female mortality by age cohort for 1985², total mortality for the disease is expected to rise to around 1005 by the year 1993.
- b The disease is extremely expensive to treat in its later stages.

1 Central Statistical Services, Deaths of Blacks 1987, RP/No 03-10-01(TBVC states excluded), gov. printer.

2 Central Statistical Services, Deaths of Blacks 1985, RP/No 07-05-08 (TBVC states excluded), gov. printer.

- c The disease is very amenable to early detection and cure.
- d The disease is more prevalent in lower Socioeconomic populations, and for this reason is not a major cause of death for whites, coloureds and Asians.

8.1.3 Description of Study

The study compares two scenarios. Scenario one (S1) is an estimate of the total incidence of the various stages of the disease, and the total mortality in black females, assuming that no screening takes place. Scenario two A (S2a) is an estimate of the incidence of the various stages of the disease and mortality of black females, if a comprehensive screening programme for cervical cancer were already in place in South Africa. Thus in both scenarios the same time period is examined. S2a assumes a 100% compliancy rate.³ In order to check the sensitivity of the compliancy rate, a third scenario, scenario two B (S2b), is included, in which an identical screening scenario as S2a is used. The only difference is that a 60% compliancy rate is used. As S2b is the more achievable of the two screening scenarios the hypotheses are tested in relation to this scenario. The study excludes the independent homelands.

8.1.4 Hypotheses

- H0** S2b is more cost-effective than S1.
- H1** S2b is less cost-effective than S1.

³ The compliancy rate referred to here is the percentage of female black women of the total population, who allow themselves to be subjected to screening. This aspect is a major hindrance to the effectiveness of such a screening programme, but could be overcome through educational and advertising campaigns.

8.1.5 Structure of Chapter

The chapter is set out as follows:

- a The disease is described and how it progresses to death in females.
- b The various possible causes of the disease are discussed.
- c The possible treatment regimes are discussed, as well as the associated direct costs.
- d The findings of various studies and cost-effectiveness analyses are presented and commented on.
- e The South African situation, as it affects black females, is elaborated on.
- f The characteristics and assumptions of the computer model designed to simulate the various scenarios are discussed.
- g The methodology is described.
- h The results of the various simulations are presented and discussed.

8.2 Cervical Cancer

The following is a brief description of the disease and the various stages through which it passes to death.⁴

The stages are described in order, from the least to the most serious.

8.2.1 Cervical Intraepithelial Neoplasia (CIN) I, II, III

Cervical cancer begins in the squamous epithelium separating the ectocervix from the endocervix. The CIN I, II and III stages are all pre-invasive.

4 D. Llewellyn-Jones, Fundamentals of Obstetric Gynaecology, 4th ed., Vol.2, faber & faber, London, 1986, pp.181-192.

a CIN I (or dysplasia of low degree)

This stage is regarded as being insignificant and reversible, and rarely progresses to further stages.

b CIN II (or moderate dysplasia)

In this stage abnormal cells are distinct and resemble cancerous cells. There is an approximate 10% progression from this stage to the next.

c CIN III (or severe dysplasia and Carcinoma-in-situ)

In this stage the epithelial lining of the cervix shows all the indications of malignant tissue, but has not invaded the cervix. Approximately 15% to 30% of cases progress to invasive carcinoma. There is some doubt as to the actual progression rates to invasive cancer.⁵ In this study the rate that is assumed is placed quite high, at 50% for all age groups.⁶ There were no studies found that could provide information on age specific progression rates. This high progression rate is justified by the very high incidence and mortality rates reported for black females, in specific studies.^{7 8}

5 Yu Shun Zhang, A.B. Miller, G.J. Sherman, "Optimizing the Age, Number of Tests, and Test Intervals for Cervical Screening in Canada", Journal of Epidemiology and Community Health, 1982, Vol.36, pp.1-10, p.2

6 G. Leiman, "Cervical Cancer Screening in a Johannesburg Family Planning Centre", South African Medical Journal, 1976, vol.50, pp.611-615.

7 M. C. Gordon, "Carcinoma of the Cervix, A Tragic Disease in South Africa", South African Medical Journal, 1982, Vol.61., pp.819-822.

8 G. Leiman, "Project Screen Soweto - a Planned Cervical Cancer Screening Programme in a High-Risk Population", South African Journal of Epidemiology and Infection, 1987, Vol.2, pp.61-68, p.62.

8.2.2 Invasive Cervical Carcinoma

The following are the stages through which a female with cervical cancer will pass if not treated.⁹

International staging of cancer of the cervix:

Stage 1	:	Confined to cervix.
1A	:	Early Stromal Invasion.
1B	:	Occult Ca-Frank invasion recognised histologically but not clinically and all other cancers limited to the cervix.
2	:	Carcinoma extending beyond the cervix but has not extended onto the pelvic wall; involves the vagina but not the lower one-third.
2A	:	No obvious parametrial involvement.
2B	:	Obvious parametrial involvement.
3	:	Carcinoma extending onto the pelvic wall (on rectal examination, there is no free space between the tumor and the pelvic wall). The tumor involves the lower one-third of the vagina. All cases with a hydronephrosis or non- functioning kidney.
3A	:	No extension onto the pelvic wall.
3B	:	Extension onto the pelvic wall and/or non-functioning kidney.
4	:	Extension outside the pelvis/reproductive tract.
4A	:	Involvement of mucosa of bladder/rectum.
4B	:	Distant metastases of disease outside the true pelvis.

9 S.B. Gusberg, H.M. Shingleton, "Diagnosis and Principles of Treatment of Cancer of the Cervix", Female Genital Cancer, eds. S.B. Gusberg, H.M. Shingleton, G. Deppe, Churchill Livingstone, New York, 1988, pp.275-296, p.285.

The following are the approximate 5 year survival rates with treatment:¹⁰

Stage 0 and 1A.....	100%
Stage 1B.....	90%
Stage 1C.....	70%
Stage 2.....	50-60%
Stage 3.....	30-40%
Stage 4.....	5%

The overall mortality from invasive cancer with treatment is around 50% after 5 years.¹¹ However, various South African studies appear to have found a much faster progression to mortality amongst black females, of approximately 50% in 1.8 years.¹² Although this faster rate can be attributed to late presentation, Walker¹³ suggests another possibility on the basis of studies carried out in the United States of America. In "a series of cancer patients matched approximately for stage of disease and also for appropriate treatment the survival rate of the poor was lower than that for the higher class". This issue is discussed further in section 8.3.

8.3 Causes of Cervical Cancer

Although it is not certain what actually causes cervical cancer, there are certain factors that are thought

10 S.B. Gusberg et al, p.287

11 A.R.P. Walker, Review Article, "Cancer of the Cervix", South African Medical Journal, 1985, Vol.68, pp.316-320, p.319.

12 A.R.P. Walker, p.319.

13 A.R.P. Walker, p.319.

to greatly increase the risk of getting this disease. The primary risk factors are as follows:¹⁴

- 1 Onset of sexuality before the age of 17;
- 2 Multiple sexual consorts;
- 3 Herpesvirus hominus type 2 transmitted during coitus by the male donor to the female at risk; and
- 4 Human papilloma virus (HPV) infection transmitted during coitus.¹⁵

However, there is much debate concerning other possible influences. In a review of international studies on cervical cancer, Walker¹⁶ concludes that ethnic affiliation, age at first menarche, age at first marriage, age at first pregnancy, age at birth of first child, mean number of children, and use of oral contraceptives, appear to have no association with cervical cancer. Nevertheless, as indicated earlier, class factors appear to play a role. It is suggested that the association between class and cervical cancer may have more to do with working-class men bringing home carcinogens from the work-place rather than "bringing home herpes or wart virus from their paramours".¹⁷

A study by Andrews et al¹⁸ differs from some of Walker's conclusions. It was found that of 115 cases of "surface dysplasia", ie. severe dysplasia, 8 could be attributed to promiscuity, 28 to birth control pills, and in 67 cases both factors were associated. It is also suggested that poor hygiene and basic proteins of the sperm head may be contributory factors. The study makes the following statement and recommendations:

14 A.R.P. Walker, p.317.

15 S. Markowitz, G. Leiman, K.A. Margolius, "Human papilloma virus and cervical intra-epithelial neoplasia in an African population", The Southern African Journal of Epidemiology and Infection, 1986, Vol.1, pp.65-69.

16 S. Markowitz et al, p.317.

17 S. Markowitz et al, p.318.

18 F.J. Andrews, J.J. Linehan, D.H. Melcher, "Cervical Cancer in Younger Women", The Lancet, October 7, 1978, pp.776-778.

*"Cervical cancer does not begin in women at some vague mid-procreative term, but is a concomitant of early sexual activity. ... Until the aetiology of the disease is established, no lower age limit should be set for cervical screening; and all young women should be entitled and encouraged to have cervical cytology tests."*¹⁹

The class aspect was noted in a study by Rosenwaike²⁰ on cancer mortality among Puerto Rican-born residents in New York City, where it was found that there were much higher mortality rates for recently settled Puerto Ricans, compared with the white population. The age-adjusted rate was three times greater than that for white non-Puerto Rican females. Furthermore, it has been found that New York City's Puerto Rican population is "occupationally, educationally, and in terms of income ... at the bottom of the urban stratification heap".²¹ Socioeconomic status and high incidence and mortality from cervical cancer can therefore be regarded as being strongly linked. There could be a link between socioeconomic status and personal hygiene.

In a study by Lyng²² in Denmark, it was suggested that the low incidence of cervical cancer among Jewish women, and the link between the socioeconomic gradient, indicates that sexual hygiene is an important factor in the risk of cervical cancer.

"Of importance for the declining trend in Denmark may therefore be the considerable improvement in housing conditions during the 1960s. In 1960 some 45% of dwellings had a

19 F.J. Andrews et al, p.776.

20 I. Rosenwaike, "Cancer Mortality Among Puerto Rican-Born Residents in New York City" American Journal of Epidemiology, 1984, vol.119, No.2, pp.177-185.

21 I. Rosenwaike, p.180.

22 E. Lyng, "Regional Trends in the Incidence of Cervical Cancer in Denmark in Relation to local Smear Taking Activity", International Journal of Epidemiology, 1983, Vol.12, No.4, pp.405-413, p.408.

*private bath in contrast to 73% in 1970.*²³

Harris²⁴, agrees with this view, suggesting that lack of penile hygiene in the male could explain the variable incidence rates between different racial groups. Associated with this factor is the strong possibility that circumcision could play a role reducing the risk of cervical cancer in females.

Although there is still doubt as to the actual determinants of cervical cancer, it is nevertheless possible to determine high-risk categories of females. The reasons for high incidence within specific communities and racial groups suggests a non-random element increasing the exposure of this category of females to the disease. There are many different factors that could play a role, from psychological²⁵ to the immunosuppressant actions of smoking. However, the association of the disease with socioeconomic status is a notable factor. The maintenance of personal hygiene is extremely difficult for much of South Africa's black population, due to the lack of housing and associated facilities. It is possible that this element alone accounts for the high incidence of the disease amongst black females in South Africa.

8.4 Treatment and Prevention

The one positive feature about cervical cancer is the length of time it takes for the disease to move from the pre-invasive stage to the invasive stage, and then to the incurable stage. It can take around 20 years for the disease to progress to the invasive stage.²⁶ Richart and Barron²⁸ estimated a 80% probability for severe

23 E. Lyng, p.408-412.

24 R.J.C. Harris, Cancer, The Nature of the Problem, 3rd ed. Penguin Books, Made and Printed in Great Britain by Cox & Wyman Ltd, 1976, p.49.

25 For a discussion on these aspects Peter Lambly provides some interesting observations in his book, The Psychology of Cancer, Great Britain, Futura Publications, 1987, p.114-122.

26 A.I. Spriggs, O.A.N. Husain, "Cervical Smears", British Medical Journal, Vol.1, June, 1977, pp.1516-1518, p.1516.

dysplasia to progress to carcinoma-in-situ in 10 years.

On the whole this disease is very amenable to early detection and cure. However, it is possible, although difficult to prove where regular screening is absent, that in a certain proportion of cases progression may be so rapid that death could not be prevented by any screening programme.²⁹

The following discussion briefly sets out the necessary procedures required at each stage of the disease, followed by rough estimates of costs (in South Africa).

8.4.1 Screening

During the early stages of the disease (ie. CIN II and III) it is possible to detect abnormalities of the epithelium of the cervix through the use of a simple test called the Papanicolaou Smear (pap. smear). A positive pap smear is an indication that further investigation is necessary. It is not taken as exclusive evidence of pre-invasive cancer of the cervix. A follow-up involves either a colposcopically directed biopsy, or a cone biopsy.³⁰ The former is preferred as this leads to fewer possible complications. If the patient is found to be suffering from either CIN II or III, treatment is surgical, with conization, cryosurgery, or hysterectomy (if patient no longer wishes to bear children). If the disease is treated at this stage, the patient has a 100% possibility of survival. After treatment, follow-up smears and examinations are required to check that the disease has been removed.

27 F.J. Andrews, p.778.

28 R.M. Richart, B.A. Barron, "A follow-up study of patients with cervical dysplasias", American Journal of obstetric Gynecology, 1969, vol.105, pp.386-393.

29 Yu Shun Zhang et al, p.9.

30 M. C. Gordon, p.821.

8.4.2 Treatment of Invasive Cervical Carcinoma

Once the disease has invaded, it becomes difficult to detect using the pap smear. However, in the invasive stages it is normally possible for the physician to visually detect the presence of the disease which is not the case in the pre-invasive stages. The major problem with this disease, and many other neoplasms, is that it is asymptomatic in the curable stages. By the time symptoms are apparent to the sufferer, it is often too late to save the patient. Stages III and IV are regarded as incurable³¹, although patients can survive with treatment for longer than five years.

For stages IA, IB, and IIA, surgery is still regarded as an option. For stage IA a total hysterectomy is performed or, if the patient wishes to preserve her child-bearing capacity, conization. For stages IB and IIA, a radical hysterectomy with lymphadenectomy can be performed, but five and ten year survival rates are not much improved over the other alternative which is radiotherapy. For all later stages of the disease radiation therapy is the only treatment.

The following are the costs in 1984 prices of the various procedures.³²

Cost to State of one Cervical Smear	-	R5
Approximate cost of treating one Pre-malignant lesion	-	R500
Treatment of Invasive Cervical Carcinoma, one patient	-	R20 000

31 M. C. Gordon, p.820.

32 S. Emdon, U. Gerard, R. Jones, "Knowledge About and Utilization of Facilities for Cervical Smears Among Black Women in Johannesburg", South African Medical Journal, 1984, Vol.65, pp.289-290.

8.5 Cost-Effectiveness of Screening (various findings)

The evidence that screening programmes are successful in reducing mortality from cervical cancer appears convincing. This section discusses the evidence that screening programmes do reduce mortality and are justified on an economic basis.

In a study of mortality from cervical cancer in Britain it was found that, except for women under 35, cervical cancer is on the decline in Britain. In Scotland the two areas with established screening programmes have lower mortality rates for the disease than the rest of the country.³³

Another study in Britain on the costs of detecting and treating cervical cancer in North-East Scotland, concluded:

*If "... it is assumed that early detection does abort the cause of this disease, the figures suggest that the running of an efficient screening service in North-East Scotland is financially justified."*³⁴

A cost-effectiveness study, performed in Denmark, in Frederiksberg Borough, found that:

the "... modified economic model of the study has yielded a reliable estimate of some basic features of cytological screening. The estimated price per avoided premature death seems to indicate strongly that cytological screening for cancer is justified from an economic point of

33 J.E. Macgregor, S. Teper, "Mortality From Carcinoma of the Cervix Uteri in Britain" The Lancet, October, 1978, pp.774-776.

34 J.B. Thorn, J.E. Macgregor, E.M. Russell, K. Swanson, "Costs of Detecting and Treating Cancer of the Uterine Cervix in North-East Scotland in 1971", The Lancet, March, 1975, pp.674-676, p.675.

view.³⁵

It was found in the United States of America that the incidence and mortality from invasive cervical cancer decreased by 59% and 50% respectively from 1947 to 1970, while the incidence of carcinoma-in-situ actually increased over the same period. A negative correlation was found between mortality rates and the rates of cytological screening during the 24-year period. However, it was suggested the relationship may not be causal due to evidence of a decline prior to the initiation of wide-spread screening.³⁶ On the other hand it was found in Iceland that mortality rates were not declining prior to the introduction of a screening programme; but subsequent to the introduction of such a programme, significant reductions have been achieved.³⁷ In Finland it was found that the mortality rate had a significant negative correlation with the intensity of screening.³⁸

Theoretically, if the disease takes between 10 and 15 years to reach the invasive stage, just one examination within this period will be sufficient to prevent a case from progressing. However, not all cases progress at the same rate. As some will be missed, the frequency of screening is an important consideration. The age to begin and disperse with screening, can significantly affect the cost-effectiveness of the programme. It is often suggested that screening should no longer take place after the age of 70. Spriggs and Husain, however, suggest that there is sufficient doubt about the extent of new preinvasive disease in women aged 70 and over for them not to be excluded from a screening programme.³⁹

-
- 35 C. Gad, E. Ostergaard, "Population Screening for Cervical Carcinoma in Frederiksberg Borough: Economic Considerations", Danish Medical Bulletin, 1977, Vol.24, No.2, pp.60-65, pp.64-65.
- 36 Yu Shun Zhang et al, p.1.
- 37 Yu Shun Zhang et al, p.1.
- 38 Yu Shun Zhang et al, p.1.
- 39 A.I. Spriggs et al, p.1516.

8.6 Cervical Cancer Among South African Black Females

The computer model used in this study predicts that for the period 2000 to 2004 mortality from cervical cancer will increase to an average of 1145 per year, from an average of 689 per year for the period 1985 to 1989. The model also predicts that the average annual incidence of invasive cervical cancer will increase from 2054 in the period 1985 to 1989, to around 3474 in the period 2000 to 2004. These increases will place a severe strain on existing curative resources unless some preventive measures are implemented.

The late presentation of many black patients is a major problem. In 1980 in the Witwatersrand 87.9% of all black patients first presented in stages II, III, or IV, 43% being in stage III, and 16.9% in stage IV. At Groote Schuur, 55% of all cervical cancer patients presented in stages III or IV, while at Tygerberg 46.8% first presented in stages III or IV. It should be noted that stages III and IV are regarded as incurable.⁴⁰

A further serious problem that makes the application of screening less effective is the inability of many hospitals to follow-up women who have tested positive. This is a consequence of black women not realizing the significance of testing positive, and therefore not responding to notifications to this effect. Poor education and inadequate staff-patient contact are also to blame. The following problems with respect to follow-up were encountered in a failed attempt to implement a screening programme in Soweto.

"The pilot studies highlighted inadequacies of follow-up and patient tracing which severely compromised final treatment rates. These inadequacies appeared to be essentially outside the control of the purely administrative role of the laboratory based service and included (1) incorrect addresses recorded for patients; and (2) a barrier at the point of colposcopy which discouraged 21/70 (30%) patients from all four pilot studies who did undergo colposcopic evaluation, from returning for definitive treatment. The impression gained by follow-up staff

was that this barrier was due to poor patient comprehension of an asymptomatic disease, reluctance to undergo hysterectomy and inadequate patient-doctor rapport at the time of colposcopy."⁴¹

In one of these pilot studies, pick-up rates of CIN II and III were unprecedented in world literature, with yields of around 41/1000 detected from the screening of women at Primary Health Care clinics.⁴²

On the whole it is very difficult to deal with this disease in a piece-meal fashion. The above-mentioned pilot studies failed to have any appreciable effect on the overall incidence of invasive cervical Carcinoma, and mortality from the disease. Problems with persuading patients to submit to initial screening, as well as inadequate follow-up, contributed to low overall coverage of the targeted population and therefore the limited effect of the programme. It is essential, if a screening programme is to be implemented, that coverage be extensive, or the expense of the programme will be incurred without any appreciable drop in the incidence of invasive cervical Carcinoma. Overall costs will be raised without any benefits.

8.7 The Model

This model compares the life histories of selected cohorts for the black female population of South Africa. The natural history of the disease is estimated through the use of transition rates between five different states. These are: normal, severe⁴³ dysplasia, carcinoma-in-situ, invasive cervical cancer, and death of cervical cancer.

41 G. Leiman, 1987, p.65.

42 G. Leiman, 1987, p.65.

43 Although the present convention is to classify both severe dysplasia and carcinoma-in-situ together, with moderate dysplasia preceding this stage, severe dysplasia is separately defined in this study. There is some difficulty in accurately distinguishing between stages. Severe dysplasia in this study should be regarded as the earliest stage of the disease deemed worthy of treatment.

The procedures were compared as follows:

- a No screening⁴⁴. In other words this would reflect the present situation in South Africa.
- b Screening at 5 yearly intervals, starting at ages 20-24 and continuing to ages 80-84.⁴⁵ A 100% compliancy rate was assumed, in order to assess the maximum potential of such a programme.
- c Screening as in (b) above except at a 60% compliancy rate.

8.7.1 Assumptions

This study uses various assumptions concerning progression rates, population estimates, costs and discount rates.

8.7.1.a Progression Rates

With regard to the progression rates from one stage of the disease to the next the following were used:

- a A 10%⁴⁶ progression rate from severe dysplasia to carsinoma-in-situ in five years.
- b A 50%⁴⁷ progression rate from carsinoma-in-situ to invasive cervical cancer in five years.

44 Screening here refers to the use of the Papanicolaou smear.

45 Screening to ages 80-84 was continued due to the fairly significant mortality from the disease into the 85+ age cohort.

46 D. Llewellyn-Jones, p.185, Although this figure is considerably lower than that given by Richart & Barron (see section 8.4) it is used because it is the lowest possible progression rate from severe dysplasia to CIS. This progression rate therefore provides the maximum possible estimate of the incidence of severe dysplasia likely.

47 D. Llewellyn-Jones, p.185.

- c The progression from invasive cancer to death was estimated using the actual statistics available on mortality by age cohort⁴⁸, and the incidence of invasive cervical cancer by age cohort.⁴⁹ It was assumed, for example, that the progression rate to death of a particular age cohort could be represented by taking the actual incidence rate of invasive cervical cancer from 1986 for that cohort and comparing it to the estimated mortality in that same cohort five years on. It was therefore assumed that, if patients who had invasive cervical cancer in 1986 were still alive in 1991, then they were regarded as having been cured. This produces what could be termed rather conservative progression rates to death. The progression rates to death also vary according to age cohort.
- d The total span of the disease is assumed to be fifteen years.⁵⁰

8.7.1.b Costs

The costs of the treatment are based on the figures mentioned in section 8.4.2.. These costs are expressed in terms of 1984 prices. Thus, all costs through time are in constant 1984 prices. There are, however, some difficulties involved in using these figures. They are as follows:

- a They do not take into account treatment through time, and therefore no discount rate has been

48 Central Statistical Services, Deaths of Blacks 1985, RP/No 07-05-08 (TBVC states excluded), gov. printer.

49 From personal communication with South African Institute of Medical Research, Johannesburg, 1989.

50 Although the exact speed of progression differs from person to person, this period was chosen as a reasonable generalization based on the literature reviewed.

applied to the figure to make it represent a present value of the costs through time.

- b Although the study involves a national strategy with regard to screening facilities, economies of scale have not been taken into account. In other words, with the enlargement and possible centralization of various activities the average costs could drop significantly. Therefore, the use of a constant price in this respect is not accurate, and involves a possible overestimate of the cost.

Unfortunately these figures cannot be regarded as exact. Due to the paucity of adequate cost studies done in the field of health economics these figures must remain as the best available estimates. Ideally, a major study, investigating the direct costs involved, should have been performed taking into account the criteria mentioned in chapter seven. However, this would have been beyond the scope and intention of this study.

The costs of an educational programme to change the attitudes of black females with regard to the benefits of the pap smear test have not been included. It is extremely important that an educational and advertising campaign accompany a screening programme of this nature. As the expense of such educational and advertising campaigns will not be constant through time, it was not included in the model. However, the amount of money allocated to advertising would depend on whether the expected benefits of achieving a particular compliancy and follow-up rate are worth the expense. What this model indicates are the expected benefits at particular compliancy rates. From these results a decision could be made on the value of the extra expenses required to achieve a particular outcome.

The only indirect cost included is that of death. The value of a life is assumed to be R396 per year.⁵¹

51 A figure for the value of a life was obtained from the Central Economic Advisory Service through personal communication. The value of life chosen is based on the estimated annual consumption (in 1984 prices) of the lowest income group in South Africa. As it is consumption based it does not bias those age groups not actively involved in the economic production process.

8.7.1.c Screening Programme

The following assumptions have been made on the basis of the preceding literature survey:

- a Screening begins from the age cohort 20-24 and continues every five years until death;
- b All those screened and discovered to have either severe dysplasia or carcinoma-in-situ are cured and do not progress;
- c Seven percent false negatives are assumed to occur;^{52 53}
- d There are two compliancy rates chosen for examination; one of 100% the other of 60%. The rate of 60% was chosen because it was assumed that, with a reasonable amount of education and advertising, this would be the minimum compliance rate that could be expected. In the case of the 60% compliance rate, it is assumed that 40% of the people with either severe dysplasia or carcinoma-in-situ are not detected.

Regression rates from carcinoma-in-situ back to dysplasia, and from dysplasia back to normal have not been included in the model, as it was felt that they will not affect:

52 D. Llewellyn-Jones, p.186.

53 According to a review of the literature by Schikendantz from 1950 to 1978, the majority of studies had results fluctuating between 5 and 25% false negatives. His own analysis suggested a figure between 10 and 15%. G. A. Schikendantz, "Exfoliative Cytology in Cervical Cancer. Role and Accuracy", Proceedings of the 12th International Cancer Congress, Advances in Medical Oncology, Research and Education, Gen. eds. A. Canonico, O. Estevez, R. Chacon, S. Barge, Gynecological Cancer Vol.III, ed. N. Thatcher, Pergamon Press, Oxford, (1978), pp.163-173, p.166. However, Lele and Piver suggest that false negatives can be reduced to below 5% through improved techniques. S.B. Lele, M.S. Piver, "Cervical Dysplasia and Cervical in Situ Carcinoma, C.I.N.", Manual of Gynecologic Oncology and Gynecology, ed. M. S. Piver, Printed U.S.A., 1989, p.64.

- a the amount of cases being detected in a particular period; and
- b the chosen progression rates.

8.7.1.d Population Estimates

The base population estimates on which all estimates of the incidence and mortality are based, are taken from corrected population figures by Sadie.⁵⁴ However, as these figures only gave population data at ten year intervals, it was necessary to make rough estimates of the population for intervening and future years in order to create sufficient information for the computer programme to process. Sadie only projected the population as far forward as the year 2005. This was extended to 2019 by taking the growth rate of each population cohort in the last two years estimated, ie. 2004 and 2005, and projecting it forward to 2019 at these rates. Although these figures are acknowledged as being rough, they were sufficient for their purpose.

8.7.1.e Discount Rate

The rate chosen to discount both the indirect costs and the Potential Years of Life Lost and Saved is 8%.⁵⁵ In order to check the sensitivity of the discount rate the above results were compared with the results obtained where no discount rate had been applied. Although the totals varied quite considerably, the relative magnitudes between the three scenarios remained virtually unchanged. Therefore, only the results discounted at 8% are discussed in the study.

54 J. L. Sadie, A Reconstruction and Projection of Demographic Movements in the RSA and TBVC Countries, Bureau of Market Research, University of South Africa, Pretoria, Report No.148, 1988.

55 This figure was obtained from the Central Economic Advisory Service through a personal communication. Although no official public sector discount rate exists at the moment, the one recommended in the text represents the most likely prospect for future Public Sector Cost Benefit studies carried out in South Africa.

8.7.1.f Life Expectancy

Two life expectancies were used for comparative purposes. They are 85 and 65 years. The results are only shown for a life expectancy of 65, however, as the relative results were not significantly altered by using different life expectancies.

8.7.2 Methodology

8.7.2.a Scenarios

This study compares two options over the same period of time. The curative option is an estimate of what will happen, or is happening, given no changes in present policy. It is assumed that the present levels of screening are so low that they have no effect on the incidence of invasive cervical cancer and mortality from the disease. The preventive procedures (ie. screening programmes) assume the pre-existence of the programme, and thus the simulation answers the question: what would the direct and indirect costs, and mortality have been, had this programme always been in existence? The first four age cohorts are ignored, as the incidence and mortality are insignificant within these age groups.

8.7.2.b Outcome Measures

The outcome measure is Potential Years of Life Saved (PYLS). PYLS for S1a are calculated by working out how many patients with invasive cervical cancer did not die as a result of treatment. It is assumed here that all patients who develop invasive cervical cancer would die if not treated. Thus PYLS would be calculated by adding all the potential years of life patients with invasive cervical cancer would have lost if they had died five years hence. The PYLS are discounted at 8% per year.

Potential Years of Life Lost (PYLL) are calculated by working out the amount of life years a particular individual lost by dying at a particular age. Thus the remaining life years for each individual within a particular age cohort could be calculated by giving the average remaining life years for individuals in that cohort, ie. someone dying in the age cohort 0-4 would have lost 82.5 potential years of life if life expectancy were 85. The PYLL are discounted at 8% per year.

The PYLS for S2a and S2b (the screening scenarios) are worked out by finding the difference between the $PYLL_u$ (= PYLL unscreened) and the $PYLL_s$ (= PYLL screened), and then adding this to the $PYLS_u$.

8.7.2.c Direct Costs

The costs for the various individual procedures are as follows:

The cost to the state of one Cervical Smear	=	R5
The cost of treating one pre-malignant lesion	=	R500
Treatment cost of one case of invasive CC	=	R20 000

All patients with invasive cervical cancer, whether they live or not, are regarded as costing R20 000 each. Those who are screened and tested positive are assumed to receive follow-up costing an additional R500 per individual. It is assumed, furthermore, that there is a 100% cure rate from early detection and treatment at this stage. This simulation does not include an adjustment for follow-up rates (ie. those cases that test positive but who do not return for further treatment), as this again depends on the publicity given to the programme.

8.7.2.d Indirect costs

The value of life is assumed at R396 per year (see footnote 51). This figure is multiplied by the amount of life years lost, discounted at 8% per year. The discounting procedure is used here to ensure consistency with the calculation of PYLS.

8.7.3 Discussion of Results

Although the computer model can evaluate any 20-year period between 1950 and 2010, the period 1990 to 1994 is chosen for evaluation here. The detailed results for the 20-year period 1985 to 2004, coupled with a detailed explanation of the model are contained in Appendices B and C. All the results presented here are in terms of 5-year periods. The following is an explanation of the results.

8.7.3.a Mortality

In figure 8.1 the mortality resulting from the different interventions is indicated according to age cohort. The actual mortality from cervical cancer in the option without screening (S1)⁵⁶ is significantly greater than that of the screening options S2a (at 100% compliancy) and S2b (at 60% compliancy). Total mortality in S1 comes to 4065 (table 8.1) compared to 1228 and 119 in S2b and S2a respectively. The difference in mortality between S2a and S2b indicates the benefits to be gained from maximizing the intensity of the screening option. It is clear from this figure that the greatest gains to be made from prevention in terms of total mortality are in the age cohorts 35-39 to 65-69.

56 As S1 represents the existing situation in South Africa it should be noted that some screening does take place. Therefore, the term "without screening" merely refers to the situation where screening is very limited in its coverage.

Figure 8.1

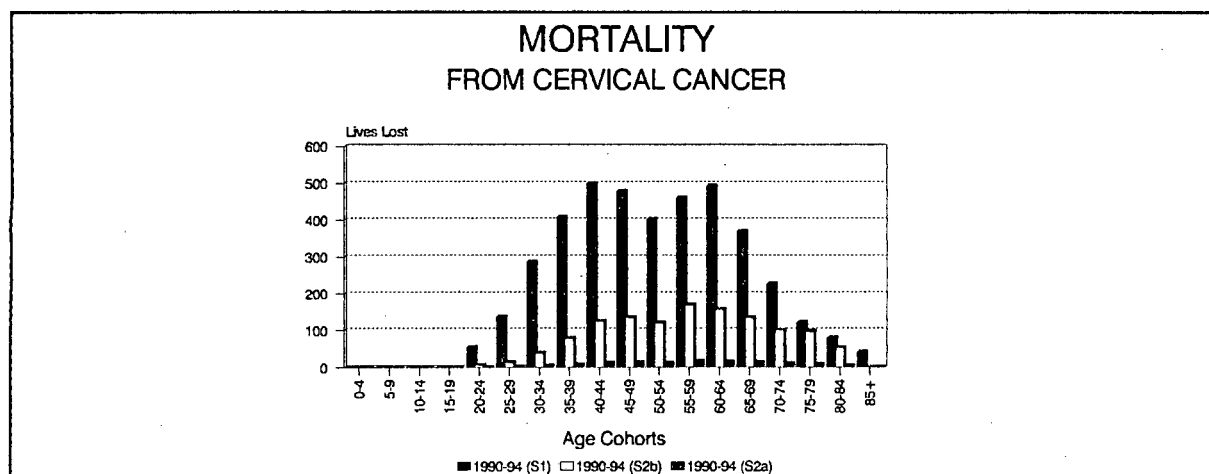


Table 8.1

MORTALITY FROM CERVICAL CANCER (FOR 1990-94)

Age	S1	S2a	S2b
0 - 4	0.0	0.0	0.0
5 - 9	0.0	0.0	0.0
10 - 14	0.0	0.0	0.0
15 - 19	0.0	0.0	0.0
20 - 24	54.9	0.5	5.7
25 - 29	136.8	1.4	14.5
30 - 34	287.2	3.8	39.1
35 - 39	410.0	7.6	78.6
40 - 44	499.0	12.0	123.9
45 - 49	478.5	13.0	134.4
50 - 54	403.9	11.6	120.5
55 - 59	461.0	16.2	168.2
60 - 64	493.4	15.1	157.2
65 - 69	370.0	12.9	134.0
70 - 74	227.3	9.7	100.4
75 - 79	122.0	9.4	97.0
80 - 84	79.6	5.2	54.3
85+	41.2	0.0	0.0
Total	4 065.0	118.5	1 228.0

8.7.3.b Incidence of Invasive Cervical Cancer

The model estimates that the incidence of invasive cervical cancer can be greatly reduced with both

programmes S2a and S2b, as can be seen in figure 8.2. There is a significant difference between S2a and S2b, again indicating the gains that can be made from intensifying the screening programmes. The total incidence for S1 is 12 181 (table 8.2), with S2b and S2a getting totals of 3892 and 375 respectively. At this stage the treatment involves much discomfort for the patient.

Figure 8.2

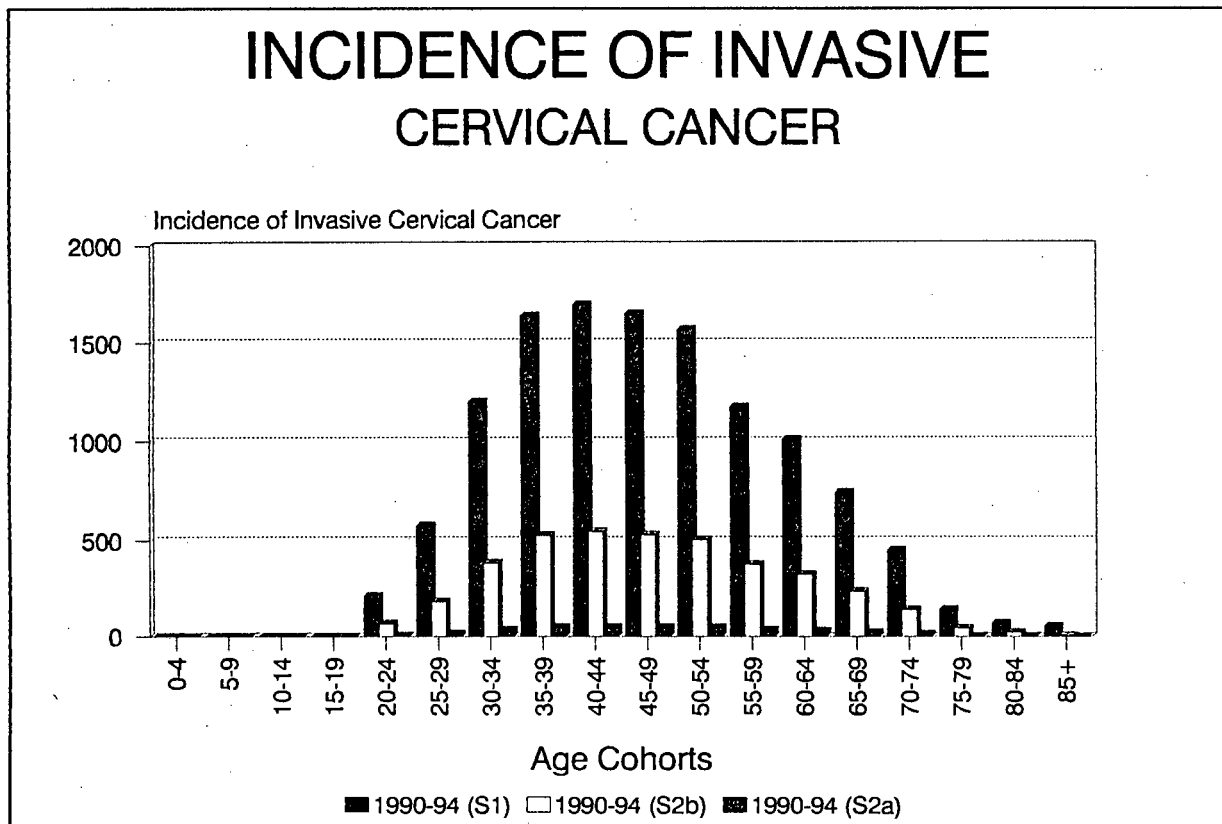


Table 8.2

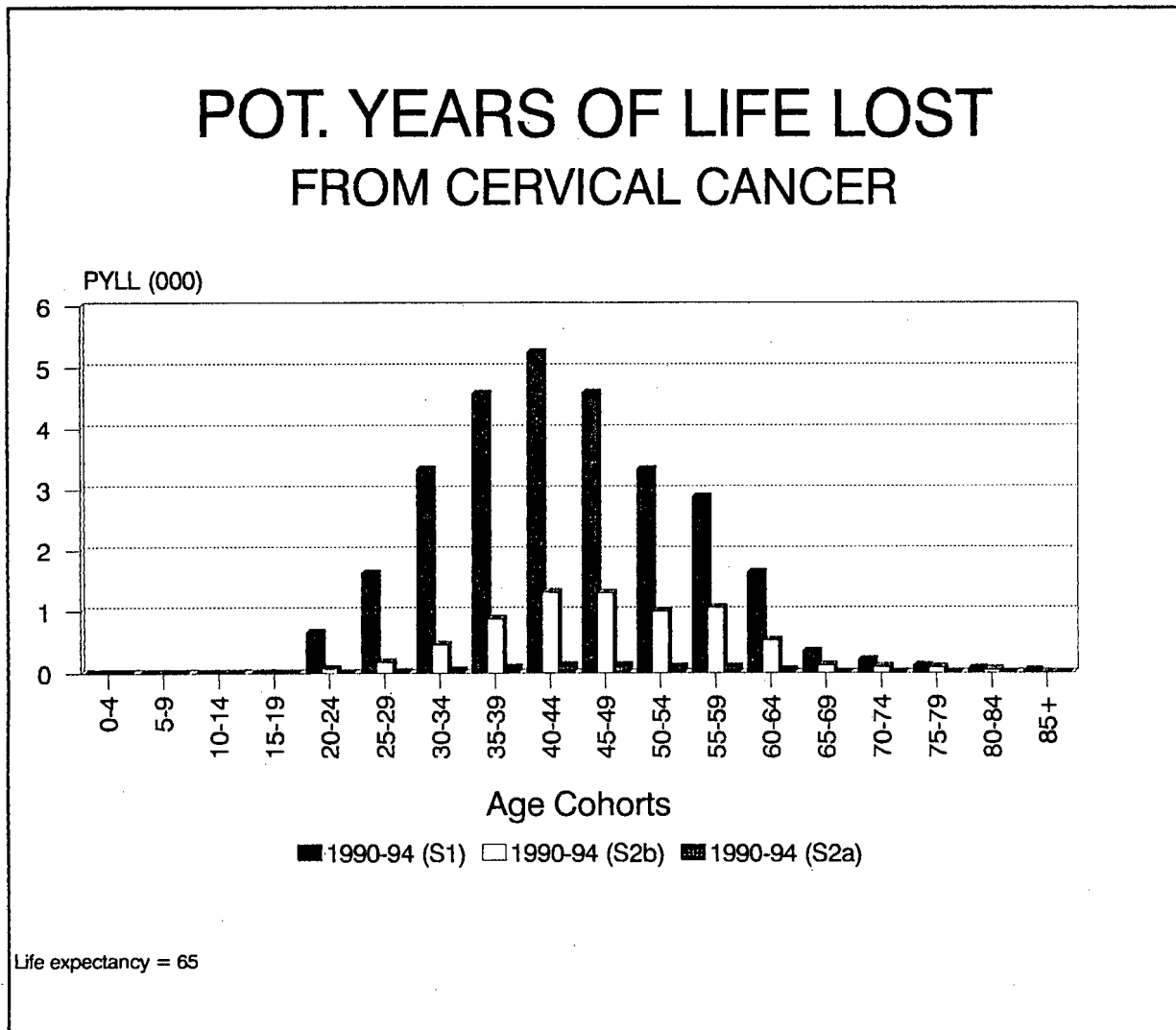
THE INCIDENCE OF INVASIVE CERVICAL CANCER (FOR 1990-94)

Age	S1	S2a	S2b
0 - 4	0.0	0.0	0.0
5 - 9	0.0	0.0	0.0
10 - 14	0.0	0.0	0.0
15 - 19	0.0	0.0	0.0
20 - 24	213.9	6.6	68.6
25 - 29	574.6	17.8	184.2
30 - 34	1 199.6	37.1	384.5
35 - 39	1 635.9	50.6	524.4
40 - 44	1 696.0	52.5	543.6
45 - 49	1 647.5	51.0	528.1
50 - 54	1 566.3	48.5	502.1
55 - 59	1 172.3	36.3	375.8
60 - 64	1 010.5	31.3	323.9
65 - 69	740.7	22.9	237.4
70 - 74	449.9	13.9	144.2
75 - 79	145.0	4.5	46.5
80 - 84	72.2	2.2	23.2
85+	56.6	0.5	5.4
Total	12 80.8	375.6	3 891.9

8.7.3.c Potential Years of Life Lost

The measure Potential Years of Life Lost (PYLL) is a method of prioritizing mortality, by quantifying the extent of life lost by dying at a particular age. Figure 8.3 illustrates the gains to be made, in terms of PYLL, of introducing a comprehensive screening programme in South Africa for black females. Both S2a (screening with 100% compliancy) and S2b (screening with 60% compliancy) are a significant improvement over S1 (existing scenario) in all age cohorts from 20-24 through to 85+.

Figure 8.3



The results used in figure 8.3 are shown below in table 8.3. In terms of total PYLL S1 has 28 664 compared to 684 for S2a, and 7087 for S2b. Thus the PYLL for S1 are approximately 42 times that of S2a, and 4 times that of S2b. However, the results are better explained with reference to the amount of life-years saved. This is shown in section 8.7.3.d.

Table 8.3

THE POTENTIAL YEARS OF LIFE LOST FROM CERVICAL CANCER (FOR 1990-94)

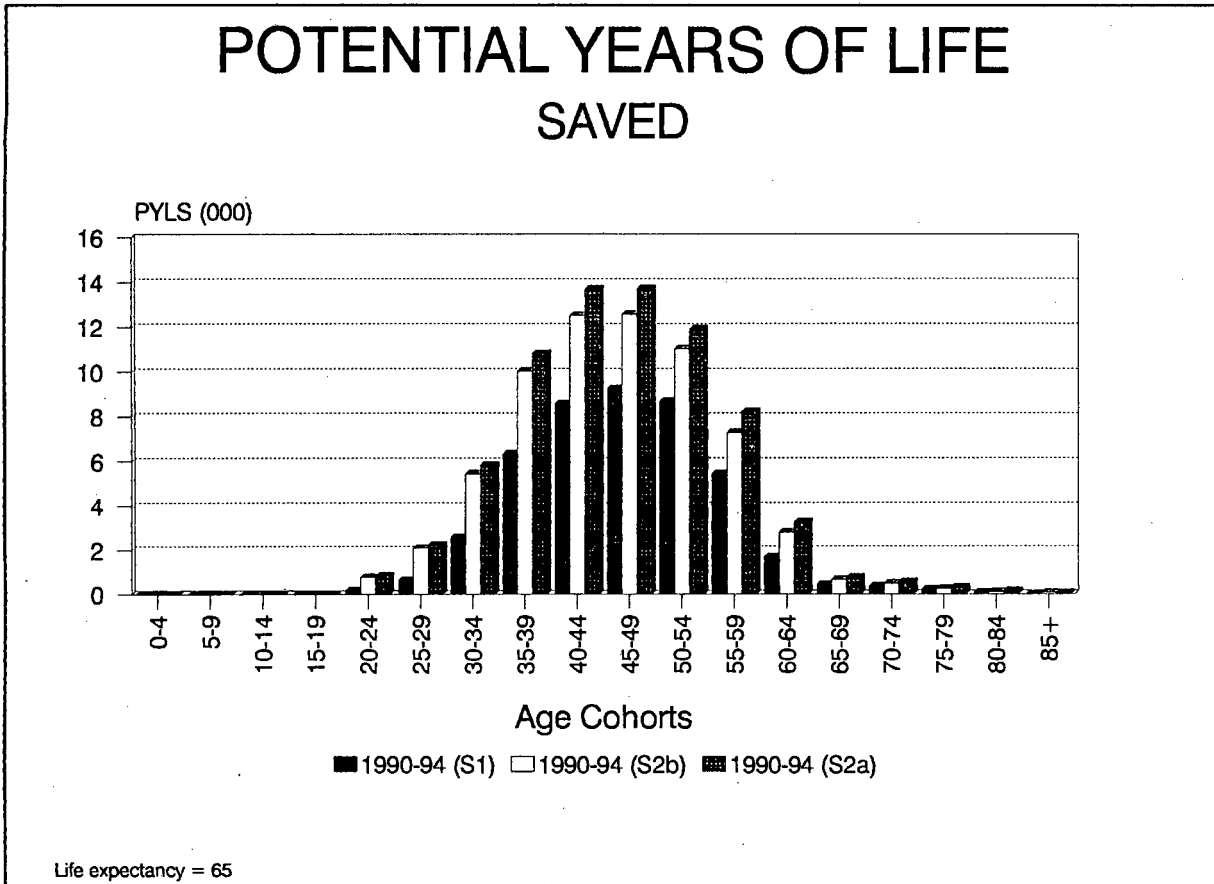
Life expectancy = 65 Discount rate = 8%

Age	S1	S2a	S2b
0 - 4	0.0	0.0	0.0
5 - 9	0.0	0.0	0.0
10 - 14	0.0	0.0	0.0
15 - 19	0.0	0.0	0.0
20 - 24	663.4	6.6	68.3
25 - 29	1 625.2	16.6	172.2
30 - 34	3 328.3	43.8	453.4
35 - 39	4 576.0	84.6	876.9
40 - 44	5 254.4	125.9	1 304.4
45 - 49	4 595.2	124.5	1 290.3
50 - 54	3 329.9	95.9	993.3
55 - 59	2 879.7	101.4	1 050.8
60 - 64	1 634.1	50.3	520.6
65 - 69	342.5	12.0	124.1
70 - 74	210.5	9.0	92.9
75 - 79	113.0	8.7	89.8
80 - 84	73.8	4.8	50.2
85+	38.1	0.0	0.0
Total	28 664.0	684.1	7 087.4

8.7.3.d Potential Years of Life Saved

Figure 8.4 indicates the estimated Potential Years of Life Saved (PYLS) by the three scenarios. Both S2a and S2b are shown to be superior to S1, in terms of PYLS throughout all age cohorts. S2a (screening with 100% compliancy) saves the most lives of the three scenarios.

Figure 8.4



In table 8.4 the actual values used in figure 8.4 are shown. The total PYLS for S1, S2a and S2b respectively are 44 488, 72 467 and 66 064. This represents an improvement in PYLS of 27 979 for S2a over S1, and 21 576 for S2b over S1. The total improvement in PYLS of S2a over S2b is only 6403. Therefore, if a comprehensive screening programme is introduced with only a 60% compliancy rate and no change in the environment, approximately 75% to 80% of the maximum PYLS can be achieved.

Table 8.4

THE POTENTIAL YEARS OF LIFE SAVED (FOR 1990-94)

Life Expectancy = 65 Discount rate = 8%

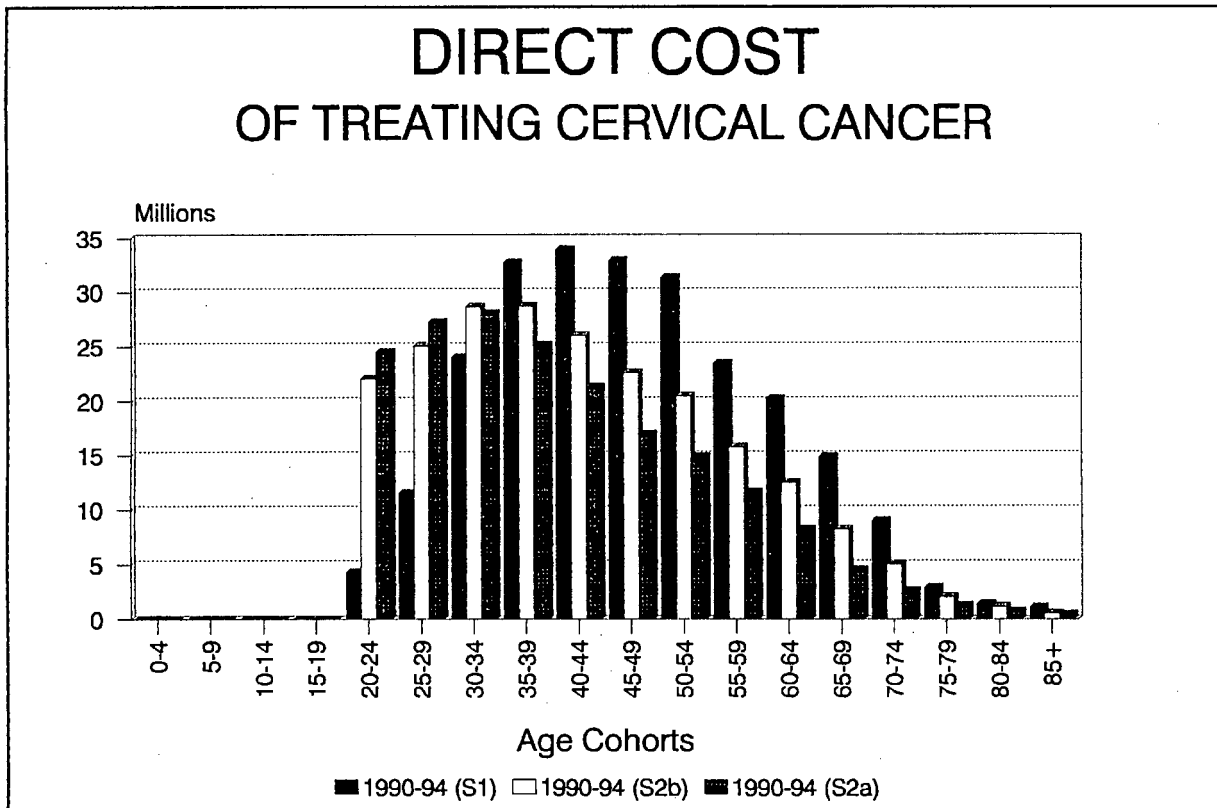
Age	S1	S2a	S2b
0 - 4	0.0	0.0	0.0
5 - 9	0.0	0.0	0.0
10 - 14	0.0	0.0	0.0
15 - 19	0.0	0.0	0.0
20 - 24	167.0	823.8	762.1
25 - 29	630.5	2 239.1	2 083.5
30 - 34	2 579.0	5 863.5	5 453.8
35 - 39	6337.2	10 828.5	10 036.2
40 - 44	8 574.1	13 702.5	12 524.1
45 - 49	9 265.1	13 735.7	12 570.0
50 - 54	8 687.0	11 921.0	11 023.6
55 - 59	5 456.7	8 235.0	7 285.7
60 - 64	1 692.2	3 276.1	2 805.8
65 - 69	432.5	763.1	651.0
70 - 74	363.4	564.9	481.0
75 - 79	220.1	324.4	243.2
80 - 84	68.4	137.3	91.9
85+	14.3	52.4	52.4
Total	44 488.0	72 467.0	66 064.0

8.7.3.e Direct Costs

In figure 8.5 the direct costs for the three scenarios are indicated. It is interesting to note that the direct costs of S2a and S2b, in the age cohorts 20-24, 25-29, and 30-34, are all greater than the direct costs of S1. However, from the age cohort 35-39 onward, the direct costs of S1 exceed S2a and S2b. This may indicate the possibility that screening should start at an older age cohort. It is also interesting to note that initially S2a is the most expensive option (ie. for cohorts 20-24 and 25-29). However, in the age cohort 30-34 S2b becomes the most expensive of the three scenarios followed by S2a. By the age cohort 35-39 the order of the three options is the reverse of the age cohort 20-24. From there on the order does not change. Thus, the more comprehensive the screening programme from the age cohort 35-39 onward, the

cheaper it becomes. This is primarily due to the drop in the incidence of invasive cervical cancer which is very expensive to treat.

Figure 8.5



The total cost of S1 (table 8.5) for the period 1990 to 1994 is estimated at R243 616 288, with an average of R48 723 257 per year (1984 prices). The total cost of S2a for the same period is R187 347 495, or an average of R37 469 499 per year. This involves an estimated saving in total costs of R56 268 793 over S1 for the five-year period, or R11 253 759 per year on average for the five-year period. The total cost of S2b is estimated at R218 559 089 or an annual average saving of R5 011 440. It is important to note that both S2a and S2b cost significantly less than S1 in total. Although many direct costs have not been included, such as spending on education and advertising, the indication is, that if between 60 and 100% of the

population are screened, these costs could actually be financed out of savings.

Table 8.5

THE DIRECT COST OF TREATING CERVICAL CANCER (FOR 1990-94)

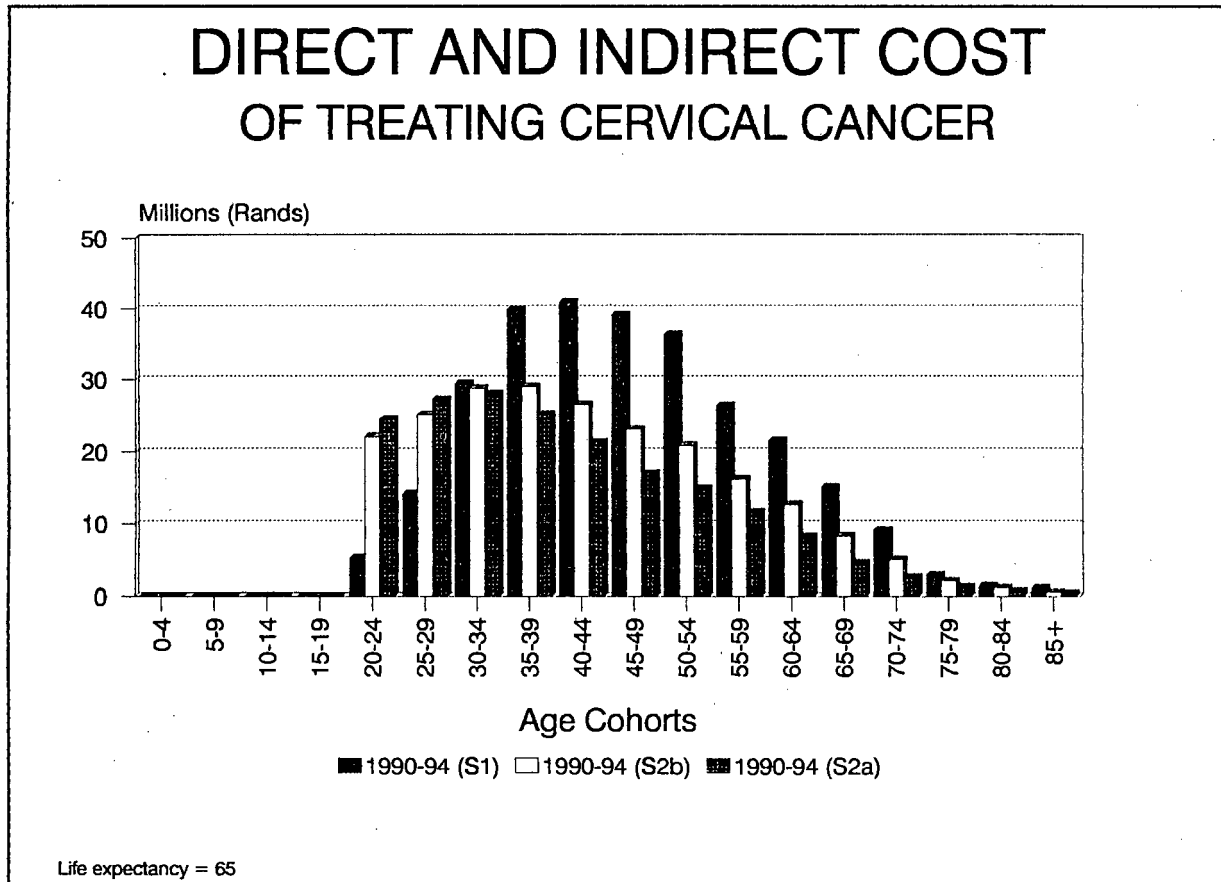
(1984 Rands)

Age	S1	S2a	S2b
0 - 4	0	0	0
5 - 9	0	0	0
10 - 14	0	0	0
15 - 19	0	0	0
20 - 24	4 278 853	24 477 689	22 011 921
25 - 29	11 492 472	27 213 000	25 013 647
30 - 34	23 991 244	28 022 802	28 632 731
35 - 39	32 717 846	25 139 969	28 666 302
40 - 44	33 917 335	21 300 026	25 979 388
45 - 49	32 950 117	16910 074	22 578 025
50 - 54	31 325 487	14 826 505	20 444 415
55 - 59	23 445 910	11 577 373	15 736 335
60 - 64	20 209 075	8 223 999	12 428 081
65 - 69	14 813 219	4 501 180	8 228 373
70 - 74	8 997 906	2 614 384	5 000 795
75 - 79	2 900 074	1 292 244	2 075 099
80 - 84	1 444 744	754 636	1 171 989
85+	1 132 002	493 610	591 984
Total	243 616 288	187 347 495	218 559 089
Weights	38	29	33

8.7.3.f Direct and Indirect Costs

The inclusion of indirect costs (ie. the cost of death) alters the costs in favour of S2a and S2b. In figure 8.6 it can be seen that by the age cohort 30-34 S1 is already more costly than both S2a and S2b unlike the situation when only direct costs are considered (figure 8.5).

Figure 8.6



The relative magnitudes of the costs of S1, S2a and S2b change from 38, 27 and 33 respectively (table 8.5), to 41, 27 and 32 (table 8.6). This is also an indication of the shift to the screening options when the cost of death is included in the calculation of total costs. It should be noted that the lowest possible value of life is used. If a higher value of life were used, this would further increase the cost of S1 relative to S2a and S2b.

Table 8.6

THE DIRECT AND INDIRECT COST OF TREATING CERVICAL CANCER (FOR 1990-94)

Life expectancy = 65 Discount rate = 8% (1984 Rands)

Age	S1	S2a	S2b
0 - 4	0	0	0
5 - 9	0	0	0
10 - 14	0	0	0
15 - 19	0	0	0
20 - 24	5 302 039	24 480 301	22 038 984
25 - 29	14 195 455	27 219 580	25 081 820
30 - 34	29 495 347	28 040 134	28 812 285
35 - 39	3 9946 410	25 173 488	29 013 565
40 - 44	40 988 062	21 349 885	26 495 926
45 - 49	39 215 623	16 959 395	23 088 991
50 - 54	36 438 930	14 864 473	20 837 766
55 - 59	26 345 897	11 617 538	16 152 449
60 - 64	21 534 388	8 243 898	12 634 241
65 - 69	15 084 794	4 505 922	8 277 507
70 - 74	9 162 867	2 617 936	5 037 600
75 - 79	2 953 242	1 295 678	2 110 679
80 - 84	1 471 230	756 556	1 191 882
85+	1 152 755	493 610	591 984
Total	283 287 046	187 618 401	221 365 685
Weights	41	27	32

8.7.3.g Cost-Effectiveness

The cost-effectiveness, ie. the cost⁵⁷ per PYLS, of the three scenarios is depicted graphically in figure 8.7, by age cohort. Both S2a and S2b are shown to be more cost-effective than S1 in all age cohorts into old age.⁵⁸ On the whole, S2a is shown to be the most cost-effective, as can be seen in table 8.7, with a value of R2589 per life year saved, compared to R3350 for S2b, and R6367 for S1. This study, therefore, indicates that if S2a were achievable it would be more than 2.5 times more cost-effective than S1, and S2b

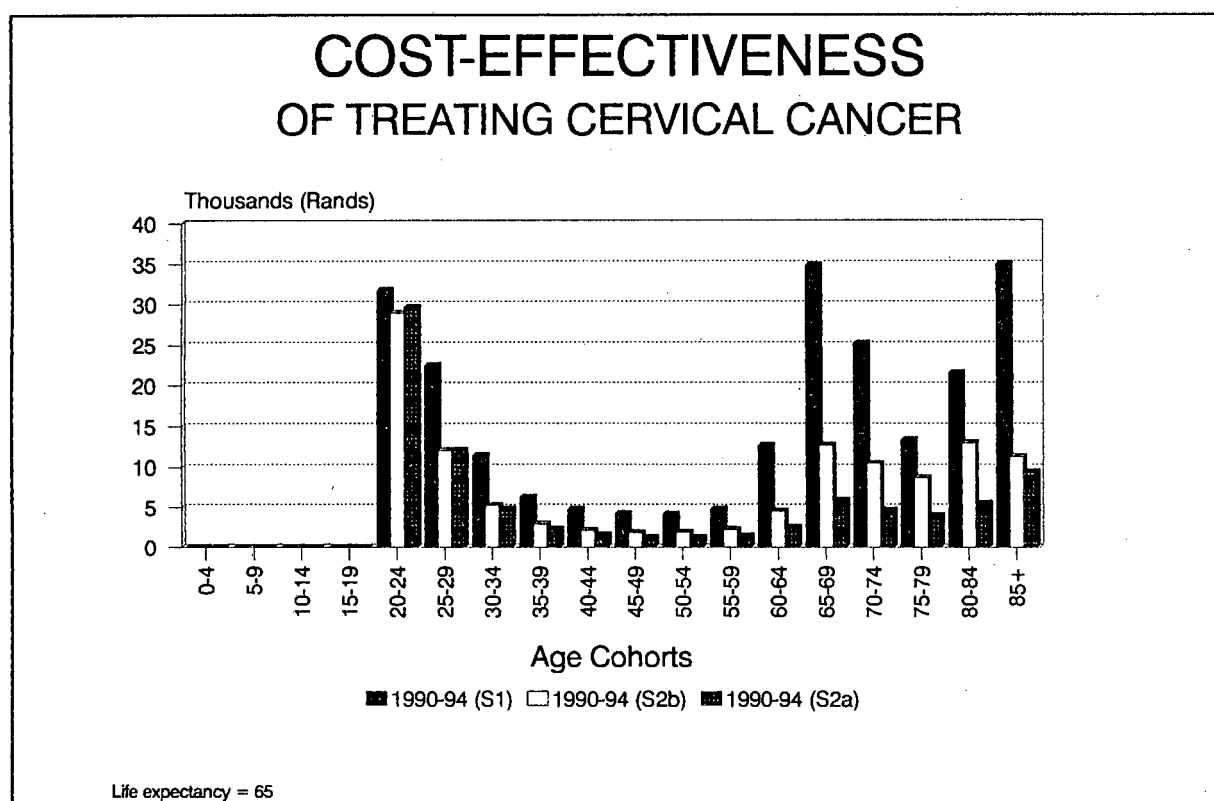
57 The costs used were both the direct and indirect costs combined.

58 The cost-effectiveness figure for S1's age cohort 85+ has been put at R35 000 to avoid distorting the graph. The actual figure is R80 719 as can be seen in table 8.7.

would be 2 times more cost-effective than S1.

In terms of direct costs only S1 has a cost-effectiveness value of R5476 compared to R2585 for S2a and R3308 for S2b.

Figure 8.7



If the programme were adequately prepared on a national basis, it would not be unrealistic to expect to achieve at least a 60% compliancy rate.⁵⁹ Over the years as the black female population becomes more aware of the need to utilize such a service, much higher compliancy rates could be expected.

Thus, S1 not only saves fewer lives, but costs more, in terms of both direct and indirect costs. The

59 This view was confirmed by Dr. G. Leiman who is head of the cytology unit at the South African Institute of Medical Research.

evidence provided in this study suggests the need for a re-evaluation of the present policy on the treatment of cervical cancer.

Table 8.7

THE COST PER POTENTIAL YEAR OF LIFE SAVED (FOR 1990-94)
 Life expectancy = 65 Discount rate = 8% (1984 Rands)

Age	S1	S2a	S2b
0 - 4	0.0	0.0	0.0
5 - 9	0.0	0.0	0.0
10 - 14	0.0	0.0	0.0
15 - 19	0.0	0.0	0.0
20 - 24	31 746.1	29 715.9	28 920.0
25 - 29	22 515.3	12 156.7	12 038.2
30 - 34	11 436.9	4 782.2	5 282.9
35 - 39	6 303.5	2 324.7	2 890.9
40 - 44	4 780.5	1 558.1	2 115.6
45 - 49	4 232.6	1 234.7	1 836.8
50 - 54	4 194.6	1 246.9	1 890.3
55 - 59	4 828.2	1 410.7	2 217.0
60 - 64	12 725.3	2 516.4	4 503.0
65 - 69	34 876.9	5905.1	12 715.9
70 - 74	2 5212.3	4 634.1	10 474.2
75 - 79	13 417.2	3 993.8	8 677.2
80 - 84	21 517.9	5 511.1	12 970.5
85+	80 719.7	9 418.7	11 295.8
Total	6 367.8	2 589.0	3 350.8

8.3.7.h Overall Comment

In terms of the results presented, the indication is that the null hypothesis (H0) should be accepted, ie. that,

S2b is more cost-effective than S1. This result can be further qualified as follows:

- a S2b is more cost-effective than S1.
- b S2b involves lower overall direct costs than S1.
- c S2b involves lower overall direct and indirect costs than S1.

8.8 Conclusion

The results suggest that a national screening programme for cervical cancer, specifically focused on black women, would be justified in terms of lives saved, total costs, and cost-effectiveness. Although this model assumed the pre-existence of the two screening programmes, ie. setting-up costs, and the lag between beginning the programme and achieving results has not been evaluated, the gains to be made appear significant enough to justify the possibility of setting up such a programme.

What this also illustrates is the contradiction that often occurs in the "fire-fighting" approach to health adopted by many countries including South Africa. The rationing of medical goods and services usually leads to curative procedures getting priority over primary preventive procedures. This eventually results in greater incidence and mortality from illness, as well as a greater overall cost, due to the high cost of curative procedures. This form of rationing, therefore, increases both the suffering and the cost. It is essential that cost-effectiveness analyses, of this nature, be used to determine the micro priorities within the health sector.

CHAPTER NINE

CONCLUSION

9.1 Results and Findings

9.1.1 Need-Based Resource Allocation on the Macro Level

The existing distribution of health-care resources in South Africa would in all probability have occurred, regardless of apartheid. This statement unfortunately cannot be tested against reality. However, the evidence presented in Chapters Two and Six (section 6.3) shows that the health sector has an internal dynamic unrelated to the needs of the broader populace of any country or community. Unless a degree of central planning exists, much of the potential of the health sector will be left unrealized. If the premise is accepted that the South African government takes responsibility for promoting the health of all South Africans,¹ allowing the health sector to determine objectives consistent with its internal dynamic will contradict that responsibility. The notion established in Chapter Two is that a high degree of central control of the health sector is required in order to achieve health goals consistent with the needs of an entire population. However, the success of this responsibility depends on whether the health sector can provide health care according to criteria of need. In other words, need must be quantifiable.

Chapter Six discusses the development of a methodology for distributing resources according to quantifiable need. The methodology is based on the Resource Allocation Working Party (RAWP) formula

1 The preamble to the Health Act, No.63 of 1977, reads as follows:

"To provide for measures for the promotion of the health of the inhabitants of the Republic; to that end to provide for the rendering of health services; to define the duties, powers and responsibilities of the several authorities which render health services in the Republic; to provide for measures for the co-ordination of such health services..."

developed in the United Kingdom. This formula has been in operation since 1976 and has achieved much success, in terms of its objectives.² However, it is recognized that using an approach designed for an industrialized country may not be appropriate for the South African situation. Nevertheless, the discussion in Chapter Six aims to illustrate the applicability of a RAWP type methodology to the South African situation, as well as motivating the need for one. The analysis provided was not complete, as such a task was beyond the scope of this thesis.

The important issues raised in Chapter Six were as follows:

- 1 The importance of focusing on aspects of need, rather than demand was discussed;
- 2 A methodology for quantifying need was developed;
- 3 Target levels of resource provision based on a need-based formula should be established as the basis for the distribution and redistribution process;
- 4 The nature of capital and recurrent expenditure in the redistribution process was elaborated upon, ie. the redistribution of capital stock should be the primary determinant of the redistribution of financial resources;
- 5 The importance of a more appropriate regional basis, in order to facilitate the resource allocation process, was discussed; and
- 6 Cross-border flows were shown to be important for the resource allocation process through:
 - i the more appropriate distribution of medical facilities; and
 - ii through the method of cross-charging.

2 When the formula was introduced, Regions were on average over eight percent away from their RAWP targets, with a range from 11 percent below to 15 percent above. Now 11 of the 14 are within three percent. (White Paper on the National Health Service in the United Kingdom, Working for Patients, Her Majesty's Stationary Office, London, 1989, p.31.)

9.1.2 Nature of Health Status

The socioeconomic environment is the most important factor determining the health status of any community or population. Chapter Three indicated the importance of such variables as education, nutrition and levels of urbanization in explaining life-expectancy in a cross-sectional study of low- and upper-middle income countries. However, variables such as population per doctor and population per nurse had little influence on health status. This result places health services in context, rather than diminishing their importance. The purpose of medical facilities (in most cases) is to prevent ill people from progressing to later stages of an illness, or to death. They minimize the damage, given particular circumstances. Where medical facilities provide primary preventive services, such as inoculations against a particular disease, in a low-income population, specific diseases may be prevented, but illness remains prevalent and severe (see Chapter Four section 4.2.1).

The mortality profiles of blacks and coloureds in South Africa (Chapter Four) indicates the existence of a severe health problem in this country. It is likely that the underlying causes of this problem stem from the socioeconomic environment in which most blacks and coloureds live. It is unlikely, however, that this situation will be alleviated in the next ten years (see the projections for the housing backlog to the year 2000, Chapter Four, section 4.2.1). It is, therefore, important for the health-care system to adapt to the particular needs of the populace. Unfortunately, achieving this objective is not easy. Need is very rarely expressed in a form where the most appropriate and cost-effective procedures can be discerned (see Chapter Six, section 6.2), and so it is necessary to evaluate medical and non-medical interventions in terms of their ability to achieve certain health objectives.

9.1.3 Cost-Effectiveness Analysis

Cost-effectiveness analysis (CEA) is a tool used to facilitate decision-making in situations where a choice has to be made between at least two alternatives to attain a particular objective. One problem in using CEA in the health sector is the requirement for an outcome measure. For this purpose much work has centred around the creation of a health-status index and related measures used to weight the life years extended as a result of utilizing a particular medical procedure. The development of such measures, as well as the utilization of the technique of CEA, could prevent considerable wastage and inefficiency in the health sector.

The cervical cancer study (Chapter Eight) provides an actual case study as an example of the useful nature of CEA. The hypothesis tested, and confirmed, was that cervical cancer screening of the entire black female population at five year intervals (from the age of 20) would be more cost effective than the scenario where no comprehensive screening takes place. Although this study cannot be regarded as conclusive, there is a strong indication that cervical cancer screening on a large scale for black females is justified, and deserves further attention by the health authorities in South Africa. It should also be noted that the potential gains to be made from more appropriate medical intervention (ie. the papanicolaou smear) are achieved in spite of the existing socioeconomic environment.

9.2 Areas for Further Research

During the process of researching this thesis many areas deserving of further research in the field of health economics came to light. The following are some suggested topics:

- 1 The pharmaceutical industry in South Africa is characterized by high prices in the private sector, and considerably lower prices in the public sector. It is also characterized by multinational firms

(apparently) making super-normal profits, while retail pharmacists struggle to make normal profits. Furthermore, the introduction of the dispensing doctor has served to make matters more complicated. It is evident that this industry needs to be evaluated in some depth.

- 2 With the desegregation of the health-care industry in South Africa, much work needs to be done in order to eliminate the fragmentation that occurred as a consequence of racially-oriented policies. A coordinated health service needs to be developed. In this respect, consideration should be given to the development of a national health insurance scheme for centralized financing of medical services.
- 3 The involvement of the private sector in the provision of health services causes cost and expenditure increases in spending on health care. Due to the size of South Africa's private health sector, an analysis of how best to manage its excesses is required.
- 4 The exact nature of delivering lower level health services for South Africa's underdeveloped sections of the population needs to be evaluated. It is not sufficient merely to be aware that primary-health-care facilities should be provided. What proportion of the health budget should be spent on primary health care, and other levels of health care? What should South Africa's total health bill come to? All these issues should be evaluated as scientifically as possible.
- 5 Aids is going to have a significant impact on South Africa's economy. The interventions developed to minimize the impact of the disease need to be evaluated for effectiveness. A cost-effectiveness analysis is essential in this regard for defining the best path to follow in combatting this disease.
- 6 A study needs to be done to determine a more scientific approach to the setting of salaries and

wages of health-care personnel. ³The preamble to the Health Act, No.63 of 1977, reads as follows:

"To provide for measures for the promotion of the health of the inhabitants of the Republic; to that end to provide for the rendering of health services; to define the duties, powers and responsibilities of the several authorities which render health services in the Republic; to provide for measures for the co-ordination of such health services..."

3 When the formula was introduced, Regions were on average over eight percent away from their RAWP targets, with a range from 11 percent below to 15 percent above. Now 11 of the 14 are within three percent. White Paper on the National Health Service in the United Kingdom, Working for Patients, Her Majesty's Stationary Office, London, January, 1989, p.31.

APPENDIX A

The following is an explanation of the data used for the study produced in chapter three, and their reliability. The data was obtained from the World Bank¹ and the information provided in this appendix was obtained from the same source.

1 Population per doctor and nurse for the year 1981 (POPDOC and POPNUR)

The estimates of population per physician and nursing person are derived from World Health Organization (WHO) data and have been slightly revised to take account of more recent estimates of population. For a few countries the information shown relates to a year later than 1981. The figure for physicians normally refers to the total number of registered practitioners in the country. Nursing persons include graduate, practical, assistant, and auxiliary nurses; the inclusion of auxiliary nurses provides more realistic estimates of available nursing care. Because definitions of doctors and nursing personnel vary - and because the data shown are for a variety of years, generally not more than three years distant from those specified - the data for these two indicators are not strictly comparable across countries.

2 Central Government Expenditure

The data on central government finance ... are from the IMF Government Finance Statistics Yearbook, 1987 and IMF data files. The accounts of each country are reported using the system of common definitions found in the IMF Manual on Government Finance Statistics (1987). The shares of total expenditure and revenue by category are calculated from series in national currencies. Because of differences in coverage of available data, the individual components of central government expenditure and

1 World Bank, 1988, World Development Report 1988, Published by Oxford University Press

current revenue shown in these tables may not be strictly comparable across all economies.

Moreover, inadequate statistical coverage of state, provincial, and local governments, dictated by the use of central government data, may seriously understate or distort the statistical portrayal of the allocation of resources for various purposes, especially in countries where lower levels of government have considerable autonomy and are responsible for many economic and social services. In addition, central government can mean either one of two accounting concepts: consolidated or budgetary. For most countries, central government finance data have been consolidated into one overall account, but for others only the budgetary central government accounts are available. Since all central government units are not included in the budgetary accounts, the overall picture of central government activities is incomplete.

It must be emphasised that the data presented, especially for education and health, are not comparable across countries for the above and other reasons. In many economies private health services represent the major component of total expenditure but may be financed by lower levels of government. Great caution should therefore be exercised in using data for cross-country comparisons.

2A Government Health Expenditure (GOVHTH)

Health expenditure covers public expenditure on hospitals, medical and dental centres, and clinics with a major medical component, and on family and preventive care. Also included is expenditure on the general administration and regulation of relevant government departments, hospitals and clinics, health and sanitation, and national health and medical insurance schemes; and on research and development.

2B Government Education Expenditure (GOVED)

Education comprises expenditure on the provision, management, inspection, and support of pre-primary, primary, and secondary schools; of universities and colleges; and of vocational, technical, and other training institutions by central governments. Also included is expenditure on the general administration and regulation of the educational system; on research into its objectives, organisation, administration and regulation of the education system; on research into its objectives, organisation, administration, and methods; and on such subsidiary services as transport, school meals, and school medical services.

2C Government Expenditure on Housing and Community Services (GOVSOC)

Housing and community amenities and social security and welfare cover, public expenditure on housing, such as income related schemes; on the provision and support of housing and slum clearance activities, on community development; and on sanitary services. They also cover public expenditure on compensation to the sick and temporarily disabled for loss of income; on payments to the elderly, the permanently disabled, and the unemployed; and on family maternity, and child allowances. They also include the cost of welfare services, such as care of the aged, the disabled, and children; as well as the cost of general administration, regulation and research associated with social security and welfare services.

3 Total Consumption of Medical Care and Education (MEDEXP and EDUEXP)

The variables MEDEXP (total expenditure on medical-care) and EDUEXP (total expenditure on education) are calculated differently from GOVHTH and GOVED. MEDEXP and EDUEXP are calculated as shares of total household expenditure. These estimates include both public and private expenditure.

Estimating the structure of consumption is one of the weakest aspects of national accounting in developing countries. The structure is estimated through household expenditure surveys and similar sampling techniques, and shares any bias of the technique. For example, some countries limit surveys to urban areas or, even more narrowly, to capital cities.

4 Private Medical Care and Education Expenditure (CONMED and CONED)

Private expenditure of medical care and education was worked out by subtracting the public expenditure from the total expenditure on these items. The information used came from the data described in sections 2 and 3 of this appendix. It should be noted that the reliability of this data is suspect.

5 Daily Calorie Intake (CALOR)

The daily calorie supply per capita is calculated by dividing the calorie equivalent of the food supplies in an economy by the population. Food supplies comprise domestic production, imports less exports, and changes in stocks; they exclude animal feed, seeds for use in agriculture, and food lost in processing and distribution.

6 Infant Mortality Rate (IMR)

The infant mortality rate is the number of infants who die before reaching one year of age, per thousand live births, in a given year. The data are from a variety of UN sources - "Infant Mortality: World Estimates and Projects, 1950-2025" in the Population Bulletin (1983), recent issues of Demographic Yearbook, and Population and Vital Statistics Report - as well as from the World Bank.

7 Urbanisation (URBPOP and URB85)

The data on urban population as a percentage of total population are from the UN publication *The Prospects of World Urbanisation, Revised as of 1984-85, 1987*, supplemented by data from various issues of the *Demographic Yearbook*, and from the World Bank.

The growth rates of urban population are calculated from the World Bank's population estimates of urban population shares are calculated from the sources cited above.

Because the estimates are based on different national definitions of what is urban, cross-country comparisons should be interpreted with caution. Data on urban population are from population censuses, which are conducted at only five or even ten-year intervals. The summary measures for urban population as a percentage of total population are calculated from country percentages weighted by each country's share in the aggregate population.

8 Education (EDUPM, EDUPF, EDUPT, EDUSM, EDUSF, EDUST and EDUTT)

The data refer to a variety of years, generally not more than two years distant from those specified, and are mostly from Unesco. However, disaggregated figures for males and females sometimes refer to a year earlier than that for overall totals.

The data on primary school enrollments are estimates of children of all ages enrolled in primary school. Figures are expressed as the ratio of pupils to the population of school-age children. While many countries consider primary school age to be 6 to 11 years, others do not. The differences in country practices in the ages and duration of schooling are reflected in the ratios given. For some countries with universal primary education the gross enrollment ratios may exceed 100 percent, because some pupils are younger or older

than the country's standard primary school age. The data on secondary school enrollments are calculated in the same manner, but again the definition of secondary school age differs among countries. It is most commonly considered to be 12 to 17 years. Late entry of more mature students as well as repetition and the phenomenon of bunching in final grades can influence these ratios.

The tertiary enrollment ratio is calculated by dividing the number of pupils enrolled in all post-secondary schools and universities by the population, age 20 to 24. Pupils attending vocational schools, adult education programs, two-year community colleges, and distance education centers (primarily correspondence courses) are included. The distribution of pupils across different types of institutions varies among countries. The youth population, that is 20 to 24, is used as the denominator since it represents an average tertiary level cohort. While in higher income countries, youths aged 18 to 19 may be enrolled in a tertiary institution (and are included in the numerator), in developing and in many industrialised countries, many people older than 25 years are also enrolled in such an institution. These data and definitions come from Unesco.

The summary measures are country enrollment rates weighted by each country's share in the aggregate population.

9 Gross National Product (GNPPC and GNPGR)

Gross National Product (GNP) measures the total domestic and foreign output claimed by residents and is calculated without making deductions for depreciation.

GNP per capita figures are calculated according to the World Bank Atlas method. The Bank recognises that perfect cross-country comparability of GNP per capita estimates cannot be achieved. Beyond the classic, strictly intractable, index number problem, two obstacles stand in the way of adequate comparability. One

concerns GNP and population estimates themselves. There are differences in national accounting and demographic reporting systems and in the coverage and reliability of underlying statistical information. The other relates to the conversion of GNP data, expressed in different national currencies, to a common numeraire - conventionally the US dollar - to compare them across countries.

As a consequence of these shortcomings the Atlas conversion factor has been developed to improve the estimates. It is described as follows:

The Atlas conversion factor for any year is the average of the exchange rate for that year, and the exchange rates for the two preceding years, which have been adjusted for differences in relative inflation between the country and the United States. This three-year average smoothes out fluctuations in prices and exchange rates for each country. The resulting GNP in US dollars is divided by the midyear population estimate for the latest year to derive per capita GNP.

10 The Crude Birth and Fertility Rate (BIRTH and FERT)

The crude birth rate indicates the number of live births occurring per thousand population in a year.

The total fertility rate represents the number of children that would be born to a woman, if she were to live to the end of her childbearing years and bear children at each age in accordance with prevailing age-specific fertility rates.

The estimates are based on data from the U.N. Population Division and from World Bank sources.

11 Life Expectancy at Birth (LE)

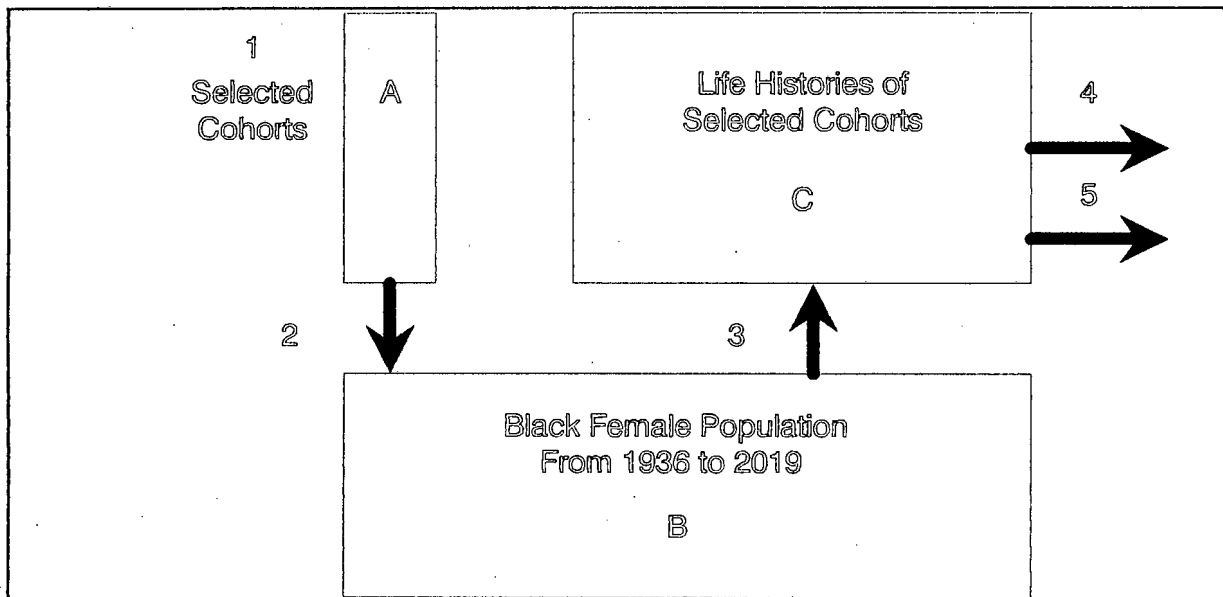
Life expectancy at birth indicates the number of years a newborn infant would live if patterns of mortality prevailing for all people at the time of its birth were to stay the same throughout its life. Data are from the UN Population Division, supplemented by World Bank estimates.

APPENDIX B

This appendix explains how the computer model determines the various results. A flow-chart is used to facilitate the explanation. Those aspects of the model that require elaboration are numbered on the flow-chart and commented on in more depth below the chart. Part One of the explanation describes how the model predicts the incidences, total costs, and cost-effectiveness of S1 (the scenario consistent with no comprehensive screening programme). Part Two explains how this information is used to predict the incidences, total costs, and cost-effectiveness of S2a and S2b (the scenarios consistent with a comprehensive screening programme). S2a and S2b are worked out according to identical procedures in the model, and are consequently discussed as one in this appendix. Both are referred to as S2. Printouts of what various parts of the flow-chart actually represent are provided in Appendix C. Each part of the flow-chart in this appendix is associated with a letter which is used to identify it in Appendix C.

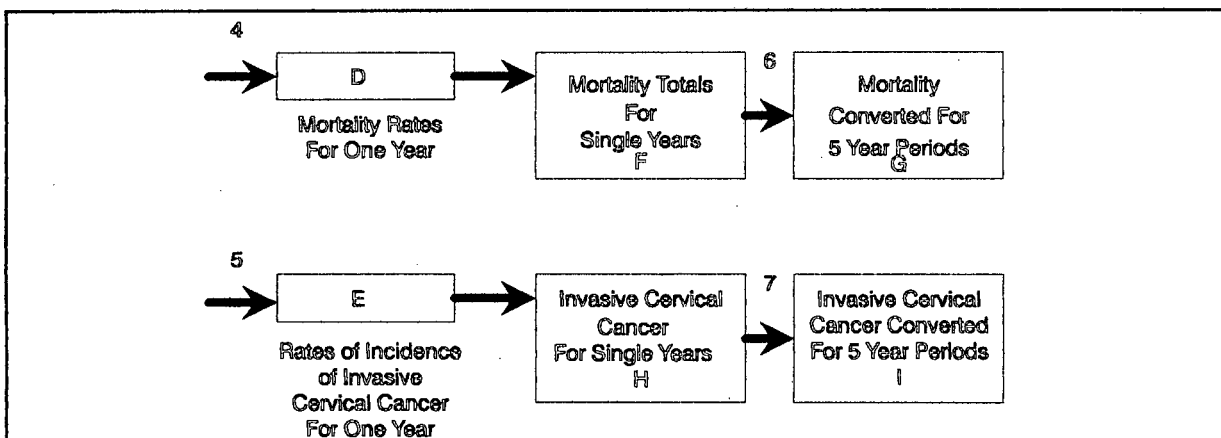
PART ONE (the determination of cost per potential year of life saved for S1)

Figure 1



- 1 The computer model begins with the selection of a date by the operator. The programme searches the population database for the population cohorts that existed, exist, or are predicted to exist, associated with the selected date. The cohorts are arranged according to five-year age categories (ie. 0-4, 5-9, 10-14, etc.) for the particular year chosen. (From Figure 1)
- 2 The programme now searches the population database for the complete life-histories of the selected age cohorts, from birth to death. (From Figure 1)
- 3 The life histories of the selected cohorts are arranged according to five-year age categories. An age cohort here refers to the entire population born in a particular five year period. Thus if the entire black, female, population is represented for a particular year in five year age categories, each five-year category (from 0-4 to 85+) is referred to as an age cohort. However, when a life history is referred to, each five year age category (from 0-4 to 85+) represents a single age cohort at different periods in the total lifetime of that cohort. The various stages in the life histories of a particular age cohort are from now on referred to as age "groups", to avoid confusion with age "cohorts". (From Figure 1)

Figure 2



4-5 The mortality rates from and incidence rates of invasive cervical cancer (ICC) by age category (obtained from actual data) are multiplied by the age groups by age category in order to obtain estimated incidence and mortality from the disease throughout the life-times of the selected cohorts. As the rates used were obtained from actual data for mortality in a single year, the initial multiplication of the rates with the population groups does not produce the total expected mortality from and incidence of the disease for the five-year period. (From Figure 2)

6-7 The mortality and incidence of cervical cancer is adjusted to five-year periods by the following method. The total mortality and incidence of invasive cervical cancer is known for the first year of a particular age group, and for the first year of the following age group five years hence. The following equations indicate how this information is used to estimate the total incidence and mortality from the disease in the intervening years:

$$((yr6 - yr1) / 4) + yr1 = yr2 \quad (1)$$

$$((yr6 - yr2) / 3) + yr2 = yr3 \quad (2)$$

$$((yr6 - yr3) / 2) + yr3 = yr4 \quad (3)$$

$$((yr6 - yr4) + yr3 = yr5 \quad (4)$$

Where: yr_x = mortality and incidence of ICC in period x

yr_1 = first year of age group

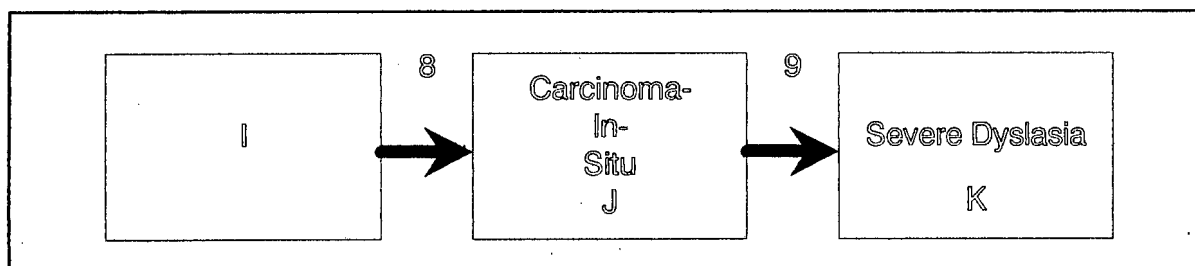
yr_6 = first year of following age group

(Please note, year one of each age group is known)

This rough method provides an adequate estimation of the incidence and mortality in the intervening years when the decline in the age group size is unknown, as is the associated mortality

and incidence rates for those age group sizes. (From Figure 2)

Figure 3

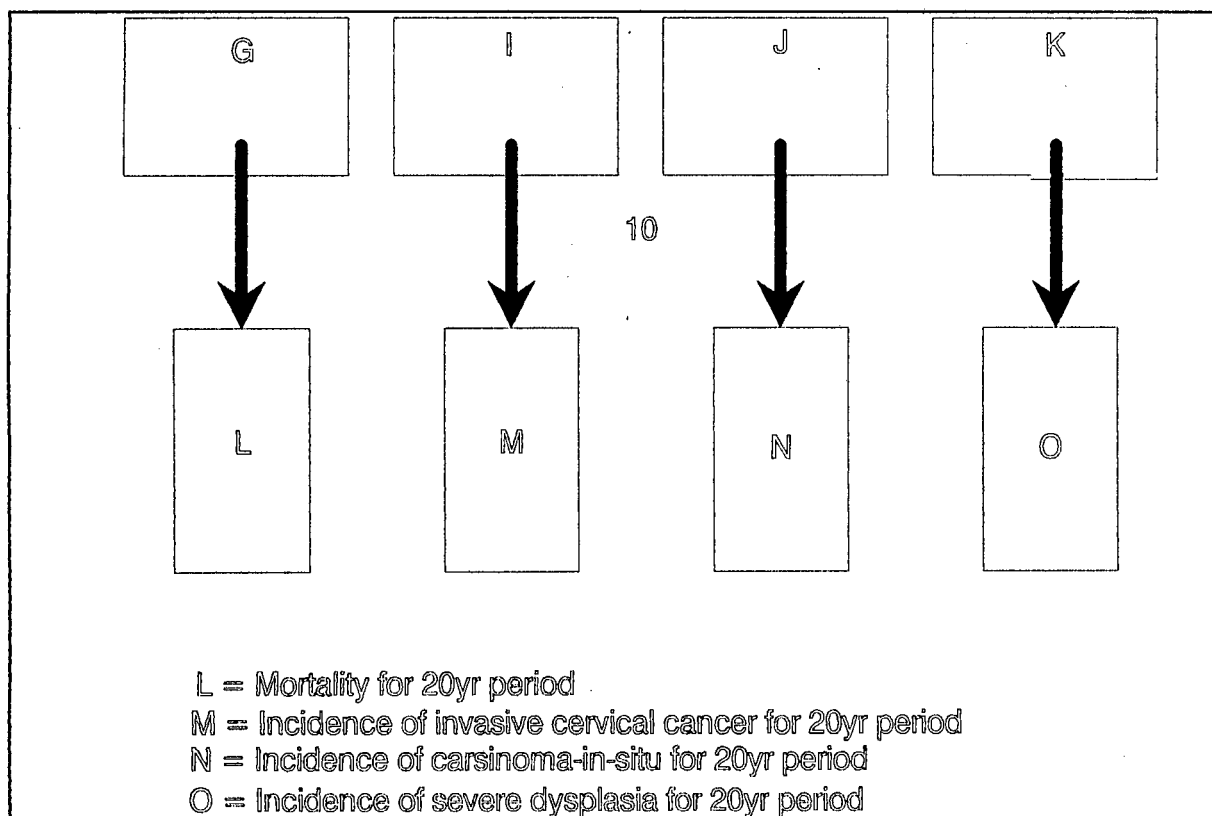


8 The incidence of carcinoma-in-situ (CIS) and severe dysplasia are now estimated using the estimated incidence of invasive cervical cancer, based on actual data, and the progression rates from dysplasia to CIS, and from CIS to invasive cervical cancer, provided in the literature. A straight 50% progression from CIS to invasive cervical cancer in five years is used. Thus an older age group with X_a estimated incidence of invasive cervical cancer ("a" refers to period), is associated with an incidence of $2X_a-5$ CIS in the age group five years previously. By this method the total incidence of CIS in each age group of the life history of an age cohort can be estimated. (From Figure 3)

9 As the incidence of CIS has been estimated, and as a progression rate of 10% from severe dysplasia to CIS is provided by the literature¹, it is possible to estimate the incidence of severe dysplasia. If an older age group in the life history of a particular cohort has an incidence of CIS of X_a , it is associated with an incidence of severe dysplasia of $(X_a-5) \times 10$ in the age group five years previously. (From Figure 3)

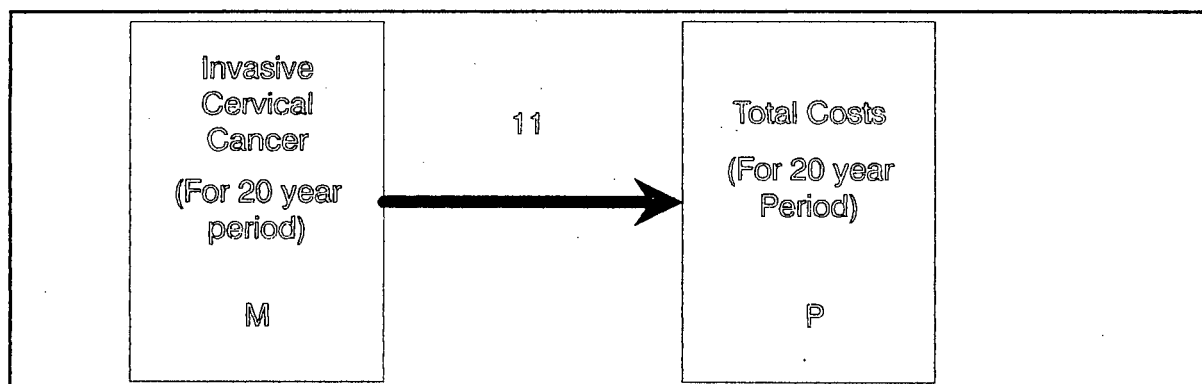
1 See footnote 46, p.195.

Figure 4



10 In order to express the estimates in terms of a particular 20-year period, starting from the selected date, those age groups of each cohort (ie. their mortality and incidence rates that would be in existence during the 20-year period) are now presented. The rest of the analysis in Part One takes place with the data expressed in this form. (From Figure 4)

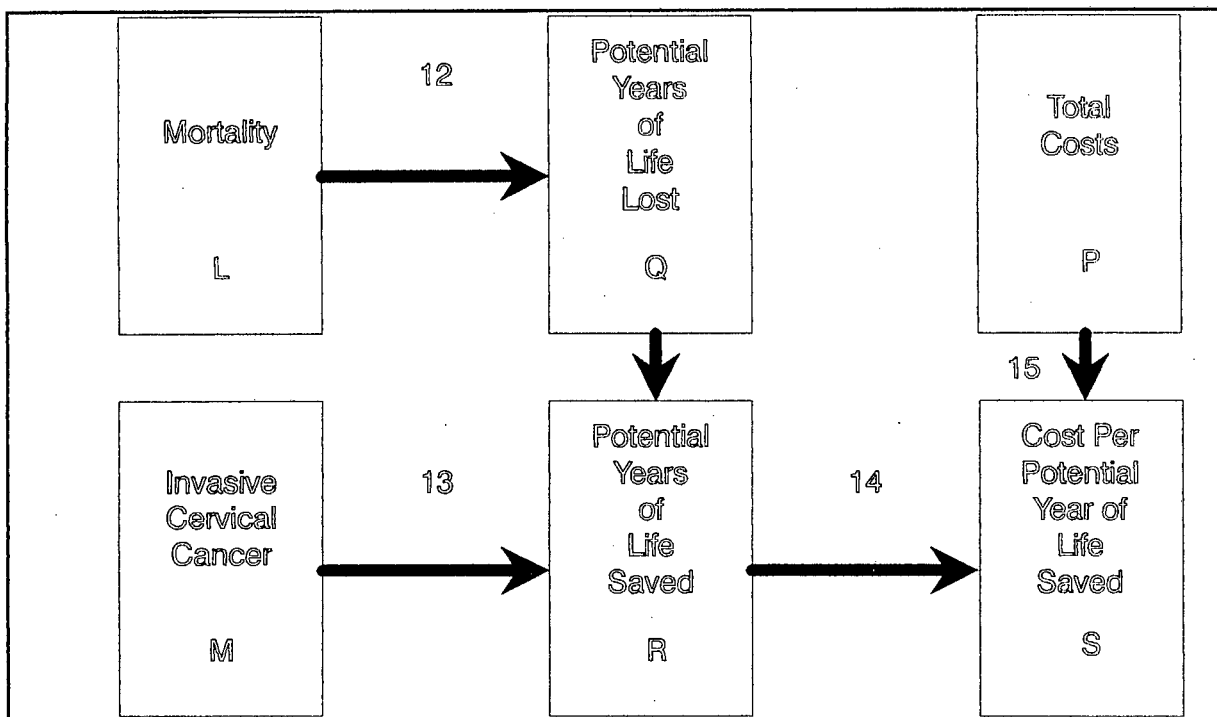
Figure 5



- 11 In S1 it is assumed that all cases of invasive cancer are treated. The total costs are estimated at R20 000 multiplied by the incidence of invasive cervical cancer. This provides a rough indication of the total costs per age cohort for the period under consideration.

The indirect costs are calculated by estimating the value of a life lost by dying at a particular age. This is done by attributing a value to a life year lost (assumed at R396 per year) and multiplying it by the amount of life years lost discounted at 8% per year. A life expectancy of 65 is used. All deaths occurring after 65 are assumed to lose only a single life year. (From Figure 5)

Figure 6



- 12 The potential years of life lost (PYLL) are those life years that the patient could have lived had she survived to 65 years of age. (From Figure 6)
- 13 The potential years of life saved (PYLS) are calculated by working out how many patients with

invasive cervical cancer did not die, as a consequence of the treatment received. It is assumed that all patients who develop invasive cervical cancer would die if not treated. Thus, PYLS are calculated by adding together all the potential years of life a patient would have lost if she had died five years hence. Again, these figures are based on a life-expectancy of 65 years for the whole population. (From Figure 6)

14-15 Using the figures obtained for potential years of life saved, and the total costs, by age cohort, an estimate can be obtained for the cost per potential year of life saved. The equation is as follows:

$$Ca = TCa/PYLSa$$

Where: Ca = cost per PYLS by age cohort

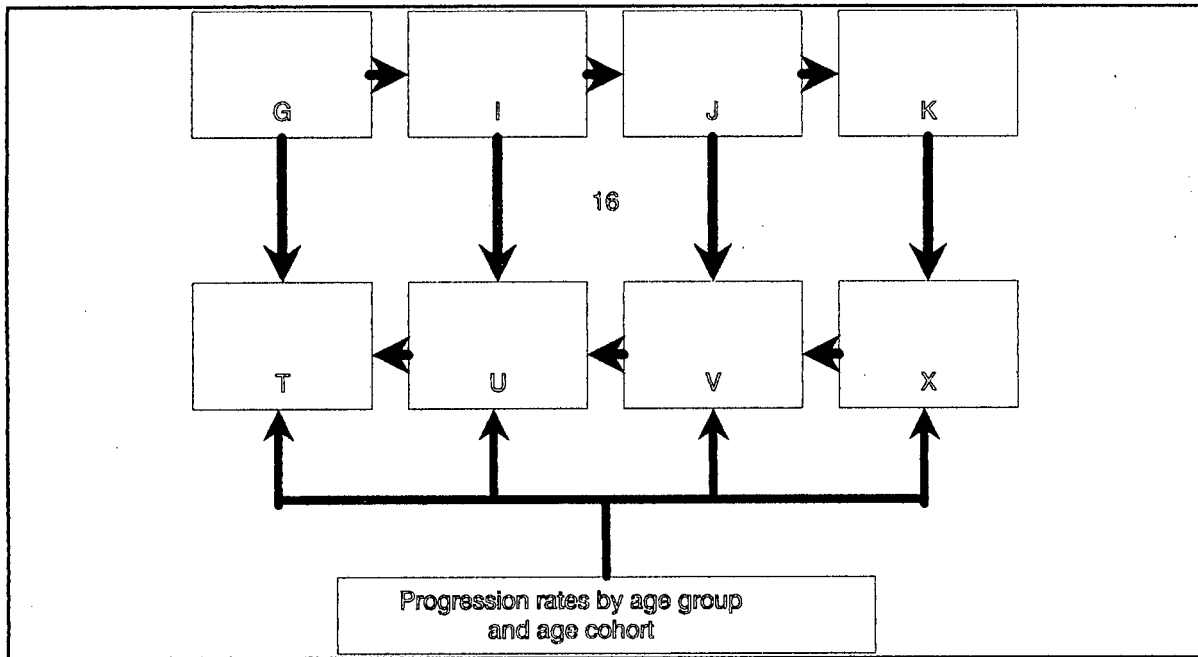
TCa = total cost by age cohort

PYLSa = PYLS by age cohort

A figure is also obtained for overall cost per PYLS for a particular time period, ignoring the age cohort breakdown. (From Figure 6)

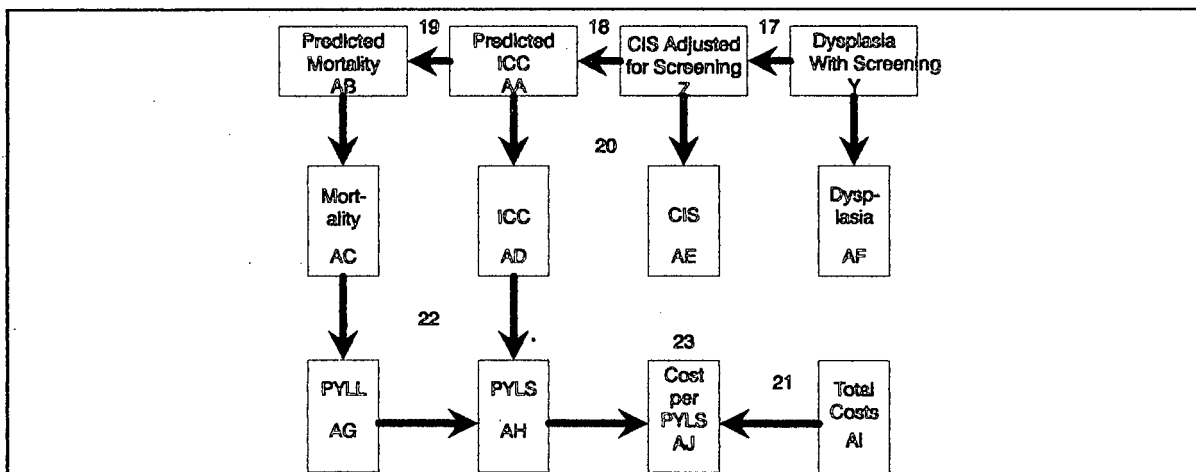
PART TWO (the determination of cost per PYLS for S2a and S2b)

Figure 7



16 The progression rates from severe dysplasia to mortality for the life histories of the selected cohorts are obtained from Part One of the computer programme. (From Figure 7)

Figure 8



- 17 With the comprehensive screening programme many of the women with severe dysplasia are assumed to be detected and the progress of the disease halted. However, those that are not detected progress to CIS. The following equation shows how the amount of cases that progress to CIS is worked out:

$$\text{CISa+5} = \text{Ia} \times 0.07 \times 0.1 (+\text{Ia} \times 0.4) \quad (1)$$

Where: CISa+5	=	the quantity of women with severe dysplasia who progress to CIS
a	=	age group
Ia	=	Incidence of severe dysplasia by age group
0.07	=	the percentage of severe dysplasias that are missed by the pap smear test due to error
0.1	=	the 10% progression rate to CIS
(+Ia x 0.4)	=	this is only for S2b where 40% of the women do not present themselves for screening

As severe dysplasia is the first stage of possible detection, and due to the assumption that screening of the total population takes place over a five-year period, some cases are assumed to be missed because they were screened within the five-year period before developing severe dysplasia. It is therefore assumed that in year one, of the five-year period, only 1/5 of all the incidences of dysplasia that are going to come into existence in the five-year period are apparent. However, it is not known which 1/5. Assuming a random sample of the population is selected, only 1/5 of 1/5 incidences would be detected in year one. In year two 1/5 of 1/4 incidences would be detected using the above logic. The amount of cases progressing to CIS would therefore be increased by the following amount:

$$(1/5 \times 1/5) + (1/5 \times 1/4) + (1/5 \times 1/3) + (1/5 \times 1/2) + 1/5 = 0.4 \text{ (rounded off)}$$

(From Figure 8)

- 18 The amount of women who progress to invasive cervical cancer from CIS, by age group, are also reduced by the comprehensive screening programme. All the cases of CIS detected are assumed to be completely cured after treatment. The amount of cases progressing from CIS to invasive cervical cancer can be determined using the following equation:

$$ICCa+5 = CISa \times 0.07 \times 0.5 \text{ (+CISa} \times 0.4)$$

- Where: ICCa+5 = incidence of invasive cervical cancer as a result of progression from CIS, in age group five years on
- CISa = incidence of CIS by age group
- 0.07 = the percentage of CIS that is missed by the pap smear test due to error
- 0.5 = the progression rate to invasive cervical cancer of undetected CIS
- (+CISa x 0.4) = this is only for S2b where 40% of all women do not present themselves for screening

(From Figure 8)

- 19 All the women that progress to invasive cervical cancer are treated using curative medical procedures, already discussed in section 8.4.2. The progression rates to mortality by age group were determined by the computer programme using actual data from Part One of the model. These rates are used to determine the progression to mortality for S2. (From Figure 8)

20 In order to express the estimates in terms of a particular 20-year period, starting from the originally selected date, those age groups of each cohort (ie. their mortality and incidence rates) that would be in existence during the selected 20-year period are now presented. The rest of the analysis in Part Two takes place with the data expressed in this form. (From Figure 8)

21 The total costs are calculated as follows:

- a the entire black female population is screened at a cost of R5 per person;
- b those who are screened and tested positive are assumed to receive follow-up costing R500 per individual; and
- c all patients with invasive cervical cancer are assumed to receive treatment costing R20 000 each.
- d the cost per life lost, discounted by 8% over the life years the person would have been expected to live (life expectancy = 65). (From Figure 8)

22 The PYLL are calculated in the same way as in Part One. The PYLS for S2 are calculated by finding the difference between the PYLL(S1) and the PYLL(S2), and then adding this to PYLS(S1). The difference between PYLL(S1) and PYLL(S2) represents the additional amount of life years that would be saved, over and above those that would already have been saved in S1. (From Figure 8)

23 The cost per potential year of life saved for S2 is worked out using the same method as for S1. (From Figure 8)

APPENDIX C

(L) MORTALITY FROM CERVICAL CANCER WITHOUT SCREENING

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	48.7	54.9	62.8	73.0
25-29	121.3	136.8	154.3	176.9
30-34	234.1	287.2	324.2	366.4
35-39	329.2	410.0	504.1	570.4
40-44	424.6	499.0	623.2	764.6
45-49	423.5	478.4	564.0	706.5
50-54	343.9	403.9	458.3	542.8
55-59	394.7	460.9	544.1	620.8
60-64	408.5	493.3	580.0	689.1
65-69	309.7	369.9	450.7	534.6
70-74	181.5	227.3	274.5	338.5
75-79	129.1	122.0	154.0	187.8
80-84	62.8	79.6	79.0	97.2
85+	35.0	41.1	51.0	54.0
Total	3447.3	4064.3	4824.9	5723.2
Total for entire 20 year period				18060.5

**(AC) MORTALITY FROM CERVICAL CANCER WITH SCREENING
60% compliancy**

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	5.0	5.6	6.4	7.5
25-29	12.8	14.4	16.3	18.6
30-34	31.9	39.1	44.1	49.7
35-39	63.1	78.5	96.3	108.8
40-44	105.6	123.8	154.3	189.6
45-49	119.4	134.3	157.8	196.9
50-54	103.0	120.4	136.2	160.7
55-59	144.9	168.2	197.5	224.1
60-64	131.0	157.1	183.5	216.6
65-69	113.3	134.0	161.6	189.8
70-74	80.8	100.3	120.1	146.6
75-79	103.7	97.0	121.3	146.3
80-84	42.9	54.2	52.4	65.1
85+	0.0	0.0	0.0	0.0
Total	1058.0	1227.6	1448.3	1720.9
Total for entire 20 year period				5454.9

(M) INCIDENCE OF INVASIVE CERVICAL CANCER WITHOUT SCREENING

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	189.9	213.9	244.8	284.5
25-29	509.8	574.6	648.4	743.0
30-34	978.0	1199.6	1354.2	1530.3
35-39	1313.4	1635.9	2010.9	2275.2
40-44	1443.2	1695.9	2117.6	2599.2
45-49	1457.6	1647.5	1942.9	2434.9
50-54	1334.5	1566.3	1776.4	2102.8
55-59	1004.3	1172.3	1383.3	1577.6
60-64	837.0	1010.5	1187.5	1410.6
65-69	619.8	740.7	903.0	1071.7
70-74	359.7	449.9	542.7	668.2
75-79	153.5	145.0	183.1	223.2
80-84	56.6	72.2	70.0	87.6
85+	14.5	16.9	21.0	22.3
Total	10271.9	12141.1	14385.8	17031.2
Total for entire 20 year period				53830.0

**(AD) INCIDENCE OF INVASIVE CERVICAL CANCER WITH SCREENING
60% compliancy**

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	60.9	68.6	78.5	91.2
25-29	163.4	184.2	207.8	238.2
30-34	313.5	384.5	434.1	490.5
35-39	421.0	524.4	644.6	729.3
40-44	462.6	543.6	678.8	833.2
45-49	467.2	528.1	622.8	780.5
50-54	427.8	502.1	569.4	674.0
55-59	321.9	375.8	443.4	505.7
60-64	268.3	323.9	380.6	452.1
65-69	198.7	237.4	289.4	343.5
70-74	115.3	144.2	174.0	214.2
75-79	49.2	46.5	58.7	71.6
80-84	18.1	23.2	22.4	28.1
85+	4.6	5.4	6.8	7.1
Total	3292.6	3891.7	4611.2	5459.2
Total for entire 20 year period				17254.7

(N) INCIDENCE OF CARCINOMA-IN-SITU WITHOUT SCREENING

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	1149.2	1296.8	1486.0	6028.8
25-29	2399.1	2708.4	3060.6	3473.3
30-34	3271.8	4021.9	4550.3	5135.5
35-39	3391.7	4235.1	5198.5	6028.8
40-44	3295.0	3885.8	4869.8	5969.8
45-49	3132.5	3552.9	4205.6	5170.9
50-54	2344.6	2766.5	3155.2	3682.0
55-59	2020.9	2375.0	2821.1	3263.8
60-64	1481.3	1806.0	2143.4	2549.4
65-69	899.8	1085.4	1336.4	1603.5
70-74	290.0	366.2	446.5	550.8
75-79	144.5	140.0	175.3	212.8
80-84	34.0	42.1	44.6	52.5
85+	0.0	0.0	0.0	0.0
Total	23854.5	28282.0	33493.4	43721.9
Total for entire 20 year period				129351.8

**(AE) INCIDENCE OF CARCINOMA-IN-SITU WITH SCREENING
60% compliancy**

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	415.4	468.7	537.1	615.7
25-29	867.2	979.0	1106.3	1255.5
30-34	1182.6	1453.7	1644.8	1856.3
35-39	1226.0	1530.8	1879.0	2179.2
40-44	1191.0	1404.6	1760.2	2157.9
45-49	1132.3	1284.2	1520.2	1869.1
50-54	847.5	1000.0	1140.5	1330.9
55-59	730.5	858.5	1019.7	1179.7
60-64	535.4	652.8	774.8	921.5
65-69	325.2	392.3	483.1	579.6
70-74	104.8	132.4	161.4	199.1
75-79	52.2	50.6	63.3	76.9
80-84	12.3	15.2	16.1	19.0
85+	0.0	0.0	0.0	0.0
Total	8622.5	10222.8	12106.5	14240.2
Total for entire 20 year period				45192.0

(O) INCIDENCE OF SEVERE DYSPLASIA WITHOUT SCREENING

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	27084.1	30606.4	34733.4	39282.5
25-29	40218.7	45503.4	51355.4	57781.4
30-34	42351.2	51984.9	60288.3	69729.5
35-39	38858.1	48697.7	59698.4	72565.4
40-44	35528.6	42056.1	51708.6	62753.9
45-49	27665.1	31552.4	36819.6	42683.9
50-54	23750.1	28211.2	32637.6	37595.8
55-59	18059.6	21433.9	25493.6	30087.0
60-64	10853.8	13364.0	16034.6	19051.9
65-69	3662.0	4465.0	5508.1	6702.2
70-74	1400.0	1752.5	2128.0	2551.1
75-79	421.4	446.1	524.8	611.1
80-84	0.0	0.0	0.0	0.0
85+	0.0	0.0	0.0	0.0

Total 269852.6 320073.6 376930.4 441395.7

Total for entire 20 year period 1408252.3

**(AF) INCIDENCE OF SEVERE DYSPLASIA WITH SCREENING
60% compliancy**

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	8612.7	9732.8	11045.2	12491.8
25-29	12789.5	14470.1	16331.0	18374.5
30-34	13467.7	16531.2	19171.7	22174.0
35-39	12356.9	15485.9	18984.1	23075.8
40-44	11298.1	13373.8	16443.3	19955.7
45-49	8797.5	10033.7	11708.6	13573.5
50-54	7552.5	8971.2	10378.8	11955.5
55-59	5743.0	6816.0	8107.0	9567.7
60-64	3451.5	4249.7	5099.0	6058.5
65-69	1164.5	1419.9	1751.6	2131.3
70-74	445.2	557.3	676.7	811.2
75-79	134.0	141.9	166.9	194.3
80-84	0.0	0.0	0.0	0.0
85+	0.0	0.0	0.0	0.0

Total 85813.1 101783.4 119863.9 140363.8

Total for entire 20 year period 447824.2

**(Q) THE POTENTIAL YEARS OF LIFE LOST FROM CERVICAL CANCER
WITHOUT SCREENING**

Discount Rate = 8%

Life Expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	588.8	663.4	759.1	882.1
25-29	1441.9	1625.2	1833.8	2101.4
30-34	2713.6	3328.3	3757.4	4246.1
35-39	3673.8	4576.0	5625.4	6364.8
40-44	4470.7	5254.4	6562.5	8050.9
45-49	4067.2	4595.2	5416.5	6785.3
50-54	2835.8	3329.9	3778.6	4475.1
55-59	2466.0	2879.7	3399.4	3878.7
60-64	1353.2	1634.1	1921.1	2282.6
65-69	286.8	342.5	417.3	495.0
70-74	168.1	210.5	254.2	313.5
75-79	119.6	113.0	142.7	173.9
80-84	58.2	73.8	73.1	90.1
85+	32.5	38.1	47.3	50.1
Total	24276.2	28664.0	33988.5	40189.6

Total for entire 20 year period 127118.3

**(AG) POTENTIAL YEARS OF LIFE LOST FROM CERVICAL CANCER WITH
SCREENING**

60% compliancy

Discount rate = 8%

Life Expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	60.7	68.3	78.1	90.7
25-29	52.8	172.2	194.0	221.9
30-34	370.1	453.4	511.1	576.7
35-39	705.1	876.9	1075.6	1214.3
40-44	1112.5	1304.4	1625.1	1996.5
45-49	1146.7	1290.3	1515.5	1891.5
50-54	849.6	993.3	1123.3	1325.4
55-59	905.5	1050.8	1233.8	1399.9
60-64	434.2	520.6	607.9	717.5
65-69	104.9	124.1	149.7	175.8
70-74	74.8	92.9	111.2	135.8
75-79	96.1	89.8	112.4	135.5
80-84	39.7	50.2	48.6	60.3
85+	0.0	0.0	0.0	0.0
Total	6052.9	7087.4	8386.1	9942.0

Total for entire 20 year period 31468.4

(R) POTENTIAL YEARS OF LIFE SAVED WITHOUT SCREENING

Discount rate = 8%

Life-expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	148.4	167.0	190.3	220.2
25-29	560.7	630.5	707.6	806.7
30-34	2107.9	2579.0	2900.7	3266.7
35-39	5100.9	6337.2	7759.8	8745.9
40-44	7326.0	8574.1	10661.4	13121.7
45-49	8245.5	9265.1	10869.9	13550.9
50-54	7447.8	8687.0	9803.8	11542.7
55-59	4721.1	5456.7	6384.9	7218.5
60-64	1422.0	1692.2	1961.7	2298.9
65-69	368.3	432.5	518.3	604.5
70-74	294.5	363.4	431.6	522.6
75-79	236.7	220.1	273.9	328.6
80-84	53.4	68.4	61.1	79.5
85+	11.8	14.3	19.6	14.7
Total	38044.9	44487.5	52544.6	62322.0

Total for entire 20 year period 197399.0

(AH) POTENTIAL YEARS OF LIFE SAVED WITH SCREENING

60% compliancy

Discount rate = 8%

Life Expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	676.6	762.1	871.3	1011.6
25-29	1849.7	2083.5	2347.4	2686.1
30-34	4451.4	5453.8	6147.1	6936.1
35-39	8069.6	10036.2	12309.6	13896.4
40-44	10684.2	12524.1	15598.9	19176.1
45-49	11165.9	12570.0	14771.0	18444.7
50-54	9434.0	11023.6	12459.1	14692.4
55-59	6281.6	7285.7	8550.5	9697.2
60-64	2341.0	2805.8	3274.9	3864.0
65-69	550.2	651.0	785.9	923.7
70-74	387.7	481.0	574.6	700.3
75-79	260.2	243.2	304.2	367.0
80-84	71.9	91.9	85.7	109.2
85+	44.3	52.4	66.9	64.8
Total	56268.3	66064.2	78146.9	92569.6

Total for entire 20 year period 293049.0

(S) THE COST PER POTENTIAL YEAR OF LIFE SAVED FOR CERVICAL
CANCER WITHOUT SCREENING

Including Direct and Indirect Costs

Life-expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	31706.3	31746.1	31883.4	32020.1
25-29	22463.7	22515.3	22638.0	22753.9
30-34	11408.7	11436.9	11479.1	11518.8
35-39	6287.4	6303.5	6328.0	6352.3
40-44	4761.4	4780.5	4800.5	4787.6
45-49	4207.8	4232.6	4254.6	4277.0
50-54	4168.6	4194.6	4215.5	4238.3
55-59	4780.7	4828.2	4868.8	4911.7
60-64	12544.9	12725.3	12900.7	13076.3
65-69	34272.1	34876.9	35483.9	36105.7
70-74	24881.1	25212.3	25611.7	26040.1
75-79	13208.6	13417.2	13614.8	13839.0
80-84	21591.5	21517.9	23327.6	22457.8
85+	24944.1	80719.7	159589.7	497031.8
Average	6281.5	6367.8	6420.5	6464.8
Average for entire 20 year period				6395.8

(AJ) THE COST PER POTENTIAL YEAR OF LIFE SAVED FOR CERVICAL
CANCER WITH SCREENING

Including Direct and Indirect Costs

60% compliancy

Life-expectancy = 65

Age cohorts	1985/89	1990/94	1995/99	2000/04
20-24	28863.3	28920.0	28804.0	28946.4
25-29	12020.6	12038.2	12036.0	12061.2
30-34	5309.8	5282.9	5327.7	5377.4
35-39	2895.7	2890.9	2880.7	2947.2
40-44	2107.6	2115.6	2101.9	2088.1
45-49	1822.4	1836.8	1836.2	1805.5
50-54	1881.8	1890.3	1902.6	1897.4
55-59	2209.3	2217.0	2218.5	2253.3
60-64	4482.9	4503.0	4529.0	4547.7
65-69	12581.9	12715.9	12802.4	12867.7
70-74	10454.7	10474.2	10528.9	10611.7
75-79	8404.4	8677.2	8643.9	8723.9
80-84	13144.9	12970.5	13565.5	13312.4
85+	11197.9	11295.8	10904.7	11373.1
Average	3347.6	3350.8	3322.0	3297.9
Average for entire 20 year period				3325.8

BIBLIOGRAPHY

Albert, A. (1981) "Estimated Cervical Cancer Disease State Incidence and Transition Rates", J.N.C.I., Vol.67, No.3, pp.571-576

Anderson, O. W. (1989) "Issues in the Health Services of the United States", Success and Crisis in National Health Systems, ed. Field M.G., McKinlay, J. B. Routledge, New York, London, pp.49-69

Andrews, F.J. Linehan, J.J. Melcher, D.H. (1978, October 7) "Cervical Cancer in Younger Women", The Lancet, pp.776-778

Barr, A. Logan, R. F. L. (1977, May 7) "Policy Alternatives for Resource Allocation", The Lancet, pp.994-996

Bembridge, J. J. (1987, September) "Some Aspects of Household Diet and Family Income Problems in Transkei", South African Medical Journal, Vol.72, pp.425-428

Bourne, D. E. Rip, M. R. Woods, D. L. (1988, February) ", Characteristics of Infant Mortality in the RSA 1929-1983, Part 2, Causes of Death Among White and Coloured Infants" South African Medical Journal, vol.73 pp.230- 232

Bourne, D. E. Rip, M. R. Woods, D. L. (1989, June) "The Spatial Variation of Mortality Among Children Under 5 Years of Age in South Africa", South African Medical Journal, vol.75, pp.565-570

Bourne, D. E. (1989, September) "Nomenclature in a Pigmentocracy - a Scientists Dilemma", South African Medical Journal, vol.76

Bradshaw, D. Botha, H. Joubert, G. Pretorius, J. P. van Wyk, R. Yach, D. (1987), Review of South African Mortality (1984), Medical Research Council, Technical Report No.1

Buxton, M. Ashby, J. (1988) "The Time Trade-off Approach to Health State Valuation", Measuring Health: A Practical Approach, ed. Teeling-Smith, G. John Wiley & Sons Ltd, pp.69-87

Bush, J. W. Chen, M. M. Patrick, D. L. (1973) "Health status indices in cost-effectiveness: analysis of a PKU program", Health Status Indices, ed. Berg, R. L. Hospital Research and Educational Trust, Chicago, pp.172-208

Central statistical services:

Census of Hospitals and Establishments for In-Patients (1983), RP/No.20-66-07, Pretoria, Government Printer

Deaths of Blacks 1985, RP/No 07-05-08, Pretoria, Government Printer

Deaths of Blacks 1986, RP/No 03-10-01, Pretoria, Government Printer

Deaths of Blacks 1987, RP/No 03-10-01, Pretoria, Government Printer

Deaths of Whites, Coloureds and Asians, RP/No 03-09-01, Pretoria, Government Printer

Population census 1985, Age by Development Region, Statistical Region and District (1985),

RP/No 02-85-02, Pretoria, Government Printer

Population Census, Age by Development Region, Statistical Region and District (1985), RP/No.02-85-02, Pretoria, Government Printer

South African Life Tables 1979-81: Whites (1987), RP/No.02-06-04, Pretoria, Government Printer

Cooper, M. (1974) "The Economics of Need: The Experience of the British Health Service", The Economics of Health and Medical Care, ed Perlman, M. Printed by R. & R. Clark Ltd. Edinburgh, pp.89-107

Creese, A.L. Darby, S.C. Palmer, S.R. Patrick, D.L. (1978, November 18) "NHS Priorities and RAWP", British Medical Journal, pp.1446-1447

Crystal, A. Brewster, A. W. (1966) "Cost Benefit and Cost-Effectiveness Analysis in the Health Field: and Introduction", Inquiry, vol.III, No.4, pp.3-13

Culyer, A. J. Cullis, J. G., (1974), "Private Patients in NHS Hospitals: Waiting Lists and Subsidies", The Economics of Health and Medical Care, ed. Perlman, M., Printed in Great Britain by R. & R. Clark Ltd., Edinburgh

Culyer, A. J., (1982), "Assessing Cost-effectiveness", Resources for Health, Technology Assessment for Policy Making, Praeger Publishers, Printed in the USA, pp.107-120

Dasgupta, Pearce, (1972), Cost Benefit Analysis, Richard Clay, The Chaucer Press Ltd., Bungay, Suffolk

de Ferranti, D., (1985), Paying for Health Services in Developing Countries, an Overview, The World Bank, Washington D.C., USA, World Bank Staff Working Papers, No.721

Department of Health and Social Services, (1986), Review of the RAWP Formula, Report by the NHS Management Board.

Department of Health and Social Security, (1976), Sharing Resources for Health in England Report of the Resource Allocation Working Party, London, Her Majesty's Stationary Office

Do Rosario Giraldes, M. (1988) "The Equity Principle in the Allocation of Health Care Expenditure on Primary Health Care Services in Portugal: The Human Capital Approach", International Journal of Health Planning and Management, Vol.3, pp.167-183

Dorrington, R. E. (1987) "African Morbidity Rates - An Initial Estimate" Transactions of the Actuarial Society of South Africa eds. Dorrington, R. E. Greeff, C. J. B. Vol.(vii) part 1, pp.169-197

Detsky, A. S. (1978) The Economic Evaluations of National Health Policy, Ballinger Pub. Co., Cambridge Massachusetts

Eddy, D. M. McGivney, W. (1988, March) "The Value of Mammography Screening in Women Under Age 50 years", J.A.M.A., Vol.259, No.10, pp.1512-1519

Eighth Interim Report of the Commission of Inquiry into Health Services, Interim Report on Hospitals and State Health Services, RP/No. 66/1986, Pretoria, Government Printer

Emdon S, Gerard U, Jones R. (1984) "Knowledge About and Utilization of Facilities for Cervical Smears Among -Black Women in Johannesburg", South African Medical Journal, Vol.65, pp.289-290.

Enthoven, A. C. (1981) "The Behaviour of Health Care Agents: Provider Behaviour", Health, Economics and Health Economics, ed. van der Gaag, Perlman, North Holland Publishing Company, Amsterdam, New York, Oxford, pp.173-188

Evans, R. G. Lomas, J. Barer, M. L. Labelle, R. J. Fooks, C. Stoddart, G. L. Anderson, G. M. Feeny, D. Gafni, A. Torrance, G. W. Tholl, W. G. (1989, March 2) "Controlling Health Expenditures - The Canadian Reality" New England Journal of Medicine, special article, Vol.320, No.9, pp.571-577

Findley, S. Carr, A. C. (1978) Patterns of Urban-Rural Fertility Differentials in Developing Countries, General Electric Company, Tempo Centre for Advanced Studies, California Washington D. C.

Fisher, S. (1984) "Measles and Poverty in Port Elizabeth", Second Carnegie Inquiry into Poverty and Development in Southern Africa, Paper No.172

Folb, P. (1989, September) Head of the Department of Pharmacology, UCT Medical School, Paper presented at the conference Containing costs in health care: Towards affordable health for all, Johannesburg

Forster, D. P. ,(1977, May 7), "Mortality, morbidity, and Resource Allocation", The Lancet, pp.997-998

Forwell, G. D.(1978) "Scottish Health Authorities Revenue Equalization, a Community Medicine View of the SHARE Report." Scottish Medical Journal, Vol.23, pp.207-212

Fuchs, V. R. (1974) "Some Economic Aspects of Mortality in Developed Countries" The Economics of Health and Medical Care, ed. Perlman, M., Proceedings of a Conference Held by the International Economic Association at Tokyo. Macmillan, Great Britain, pp.174-205

Gad, C. Ostergaard, E. (1977) "Population Screening for Cervical Carcinoma in Frederiksberg Borough: Economic Considerations" Danish Medical Bulletin, Vol.24, No.2: pp.60-65

Gordon, M. C. (1982) "Carcinoma of the Cervix, A Tragic Disease in South Africa" South African Medical Journal, Vol.61, pp.819-822

Grosse, R. N. Perry, H. (1983) "Correlates of Life Expectancy in Less Developed Countries", Human Capital and Development, eds. Salkever, Sirageldin, Sorkin, JAI press inc., London England

Gudex, C. (1986,October) Qualys and their use by the health service, University of York, Centre for Economics, Discussion Paper 20

Gusberg, S. B. Shingleton, H. M. (1988) "Diagnosis and Principles of Treatment of Cancer of the Cervix", Female Genital Cancer, eds. Gusberg, S. B. Shingleton, H. M. Deppe, G. Churchill Livingstone, New York, Edinburgh, London, Melbourne, pp.275-296

Harris R.J.C., (1976) Cancer, The Nature of the Problem, 3rd ed. Penguin Books, Made and Printed in Great Britain by Cox & Wyman Ltd

Herber, B. P. Richard, D. (1975) Modern Public Finance, 3rd ed. Irwin, inc., Homewood, Illinois 60430

Hollis, L. (1989, September) (Medscheme) Unpublished paper presented at the conference Containing

costs in health care: Towards affordable health for all, Johannesburg

Kahn, H. Berelowitz, L. Boroda, A. Gardner, S. Gordon, B. Manoim, T. Mirman, R. Smith, R. Leiman, G. (1987, October) "Knowledge of and Attitudes Towards Cancer of the Cervix in Johannesburg Women", Southern African Journal of Epidemiology and Infection, Vol.3, pp.89-92

Kenney, K. (1980) Architect of Apartheid H F Verwoerd - an Appraisal, National Book Printers, Cape Town

Kind, P. Rosser, R. Williams, A. (1982) "Valuation of Quality of Life: Some Psychometric Evidence", Value of Life and Safety, ed. Jones-Lee M. W., pp.159-170

Kind, P. (1989) The Design and Construction of Quality of Life Measures, Discussion Paper 43, Centre for Health Economics, Health Economics Consortium, University of York

Lambley, P. (1987), The Psychology of Cancer, Futura Publications, Great Britain

Lee, K. Mills, A (1979) "The Role of Economists and Economics in Health Service Planning: A General Overview", Economics and Health Planning, ed. Lee, K. Biddles Ltd, Great Britain, Guildford, Surrey

Lehman, P. Gutzwiller, F. Martin, J. F. (1989) "The Swiss Health System: the paradox of ungovernability and efficacy", Success and Crisis in National Health Systems, ed. Field M.G., McKinlay, J.B., Routledge, New York, London, pp.25-48

Leiman, G. (1976) "Cervical Cancer Screening in a Johannesburg Family Planning Centre", South African Medical Journal, vol.50, pp.611-615

Leiman, G. (1987) "Project Screen Soweto - a Planned Cervical Cancer Screening Programme in a High-Risk Population", South African Journal of Epidemiology and Infection, Vol.2, pp.61-68

Lele, S. B. Piver, M. S. (1989) "Cervical Dysplasia and Cervical in Situ Carcinoma,C.I.N.", Manual of Gynecologic Oncology and Gynecology,ed. Piver, M. S. Printed U.S.A.

Llewellyn-Jones, D. (1986) Fundamentals of Obstetric Gynaecology, 4th ed., Vol.2, faber & faber, London, Boston, pp.181-192.

Ludbrook, A. (1981) "A cost-effectiveness analysis of the treatment of chronic renal failure", Applied Economics, Vol.13, pp.337-350

Lynge, E. (1983) "Regional Trends in the Incidence of Cervical Cancer in Denmark in Relation to local Smear Taking Activity",International Journal of Epidemiology, Vol.12, No.4, pp.405-413

Macgregor, J.E. Teper, S. (1978, October) "Mortality From Carcinoma of the Cervix Uteri in Britain" The Lancet, pp.774-776

Markowitz, S. Leiman, G. Margolius, K.A. (1986) "Human papilloma virus and cervical intra-epithelial neoplasia in an African population", The Southern African Journal of Epidemiology and Infection, Vol.1, pp.65-69

Mashabela, H. (1988) Townships of the PWV, South African Institute of Race Relations Publication, Aurora Printers, Pretoria

Mays, N. Bevan, G. (1987) Resource Allocation in the Health Service. A Review of the Methods of the Resource Allocation Working Party (RAWP), Bedford Square Press / NCVO

McGuire, A. Henderson, J. Mooney, G. (1988) The Economics of Health Care, Routledge & Kegan Paul, London & New York.

McIntyre, D. Dorrington, R. E. (1989, September) "Trends in South African Health Care Expenditure 1971 - 1988", Paper presented at the conference, Containing costs in health care: Towards affordable health for all, Johannesburg

McIntyre, D. (1990) Health Care Expenditure in South Africa 1970 to 1990, Health Economics Unit, Department of Community Health, University of Cape Town Medical School

Miller, A. B. (1981) "An Evaluation of Population Screening for Cervical Cancer", Advances in Clinical Cytology, eds. Koss, L. G. Coleman, D. V. Butterworths, pp.64-94

Osmond, C. (1985) "Using Age Period and Cohort Models to Estimate Future Mortality Rates", International Epidemiological Association, Vol.14, No.1, pp.124-129

Murray, C.J.L (1987) "A Critical Review of International Mortality Data", Soc. Sci. Med., Vol.25, No.7, pp.773-781.

Parkin, D. M. Moss, S. M. "An Evaluation of Screening Policies for Cervical Cancer in England and Wales Using a Computer Simulation Model", Journal of Epidemiology and Community Health, Vol.40, pp.143-153

Platzky, L. Walker, C. (1985) The Surplus People. Forced Removals in South Africa, Raven Press, Johannesburg

Prescott, N. Warford, J. Economic Appraisal in the Health Sector, World Bank Reprint Series: No.322, Reprinted with permission from Lee, K. and Mills, A. (1983) The Economics of Health in Developing Countries, Oxford U.N. Press. pp.127-45

Price, M. (1986) Health Care as an Instrument of Apartheid Policy in South Africa. Health Policy and Planning 1986; Vol.1 (2): pp.158-170, Oxford University Press

Repetto, R. (1979) Economic Equality in Developing Countries, John Hopkins University Press, Baltimore and London

Richart, R.M. Barron, B.A.(1969) "A follow-up study of patients with cervical dysplasia", American Journal of Obstetric Gynecology, vol.105, pp.386-393

Rosenwaive, I. (1984) "Cancer Mortality Among Puerto Rican-Born Residents in New York City" American Journal of Epidemiology, vol.119, No.2, pp.177-185

Rosser, R. Kind, P. (1978), "A Scale of Valuations of States of Illness: is There a Social Consensus?" International Journal of Epidemiology, Vol.7, No.4, pp.347-358

Rutton, F. (1980) "Invited discussion" Griffiths, A. Bankowski, Z. Economics and Health Policy, XIIIth CIOMS Round Table Conference, Published jointly by the Council for International Organization of Medical Services and the Sandoz Institute for Health and Socio-economic Studies, Geneva, pp.266-269

Sackett, D. L. Chambers, L. W. MacPherson, A. S. Goldsmith, C. H. Mcauley, R. G. (1977, May) "The Development and Application of Indices of Health: General Methods and a Summary of Results", American Journal of Public Health, Vol.67, No.5, pp.423-428

Sadie, J. L. (1988) A Reconstruction and Projection of Demographic Movements in the RSA and TBVC Countries, Bureau of Market Research, University of South Africa, Pretoria, Report No.148

Sandier, S.(1975) in The Health Care Cost Explosion. Which Way Now? ed. Erlich, D. A., Based on a symposium in Geneva organized by the Henry Dunant Institute, Hans Huber Publishers, Switzerland

Schickendantz, G. A. (1978) "Exfoliative Cytology in Cervical Cancer. Role and Accuracy", Proceedings of the 12th International Cancer Congress. Advances in Medical Oncology. Research and Education, Gen. eds. Canonico, A. Estevez, O. Chacon, R. Barge, S. Gynecological Cancer Vol.III, ed. Thatcher, N. Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt, pp.163-173

Scrimshaw, S. (1981) "Infant Mortality and Behaviour in the Regulation of Family Size", Fertility Decline in Less Developed Countries, ed. Eberstadt. N, Praeger, USA

Sintonen, H. (1981) "An Approach to Measuring and Valuing Health States", Social Science Medicine, Vol.15C, pp.55-56

Smith, A. The Wealth of Nations, Book 4 and 5, pp.182-422, William Berton Publisher, London, Encyclopaedia Britannica.

South African Institute of Race Relations, (1989) Race Relations Survey 1988/89, Printed by Galvin and Sales, Cape Town

Southern African Labour and Development Research Unit, (1983) Working Papers, Health and Health Services in the Ciskei, Community Health Research Project, SALDRU working paper No.54, Cape Town

Spriggs, A.I. Husain, O.A.N. (1977, June) "Cervical Smears", British Medical Journal, Vol.1, pp.1516-1518

Taylor, S. P. Klopper, J. M. L. (1987, December) "South African Health Care Expenditure 1975-1984", South African Medical Journal, Vol.72, pp.802-804

Thorn, J.B. Macgregor, J.E. Russell, E.M. Swanson, K. (1975, March) "Costs of Detecting and Treating Cancer of the Uterine Cervix in North-East Scotland in 1971", The Lancet, pp.674-676

Torrance, G. W. (1976) "Health Status Index Models: a Simplified Mathematical View", Mgt. Sci., Vol.22, pp.990-1001

Torrance, G. W. (1986) "Measurement of health state utilities for economic appraisal: a review", Journal of Health Economics, vol.5, pp.1-30

Vogel, R. J. Greenspan, N. T. (1983) "Health Planning in the Sudan", Human Capital and Development, ed. Salkever, Sirageldin, Sorkin, vol 3, JAI press inc. London England, pp.271-309

Waddington, C. Thomas, M. (1988) "Recurrent Costs in the Health Sectors of Developing Countries", International Journal of Health Planning and Management, Vol.3, pp.151-166

Walker, A.R.P. (1985) Review Article, "Cancer of the Cervix", South African Medical Journal, Vol.68,

pp.316-320

Wagstaff, A. (1986) "The Demand for Health, Some New Empirical Evidence", Journal of Health Economics, Vol.5, pp.195-233

Webber, M. (1983) "Health Sector Planning and National Development Planning", The Economics of Developing countries, eds. Lee, K. Mills, A. Oxford University Press

Wells, S. Klees, S. (1980) Health Economics and Development, Praeger Publishers (USA)

West, M. E. Boonzaier, E. A. (1989, September) "Population Groups, Politics and Medical Science", South African Medical Journal, vol.76, pp.185-186.

White Paper (1989) Working for Patients, Her Majesty's Stationary Office, London

Williams, A. C. (1981) "Welfare Economics and Health Status Measurement", Health, Economics, and Health Economics, eds. van der Gaag, J. Perlman, M. North Holland Publishing Company, Amsterdam, New York, Oxford, pp.271-282

Williams, A. C. (1985) "Economics of Coronary Artery Bypass Grafts", British Medical Journal, Vol.291, pp.326-329

Wilson, T. D. (1984) Health services within Soweto, Carnegie conference paper No.170, 13-19 April 1984, Cape Town.

Wolfson, M. C. (1989, August) A System of Health Statistics. Toward a New Conceptual Framework for Integrating Health Data, A paper prepared for the International Association for Research in Income and Wealth, 21st Conference

World Bank (June 14, 1989), Korea Health Insurance and the Health Sector, Country Department 11, Asia Regional Office, Document of the World Bank, RP/No 7412-K0

World Bank Policy Study (1987) Financing Health Services in Developing Countries, an Agenda for Reform

World Bank (1988) World Development Report 1988, Oxford University Press

Zhang, Y. S. Miller, A.B. Sherman, G.J. (1982) "Optimizing the Age, Number of Tests, and Test Intervals for Cervical Screening in Canada", Journal of Epidemiology and Community Health, Vol.36, pp.1-10