HEALTH INFORMATICS IN DEVELOPING COUNTRIES:

An Analysis and Two African Case Studies

Mary Ekundayo Lucretia Forster

Department of Information Systems
London School of Economics and Political Science

UMI Number: U050308

Dissertation Publishing

THESIS
F
6798

x210923775

ABSTRACT

This thesis relates informatics to the problems of health and medicine experienced in less developed countries. It evaluates the potential of health informatics and investigates the issues that constrain successful implementations. This serves as a basis for establishing a generic description of viable computer applications in the developing world.

The thesis contains two case studies from sub-Saharan Africa. The first, undertaken in The Gambia, is based on a computer-assisted data collection system used in a longitudinal child health survey. The second, undertaken in Kenya, relates to a medical decision-aid system used in an out-patient clinic of a district hospital. In each case, an outline is given of the background to the application domain, and an analysis is made of some comparable prior systems that have been developed and evaluated. The two case studies provide interesting investigatory comparisons since both systems are used by health personnel with little computer experience, and exploit some state-of-the-art technologies despite the identified constraints that exist in developing countries. The context, system design, methods, and results of each case are described. A generalised evaluation approach is proposed and is used to summarise the case study findings.

The evaluation framework employed includes components related to functional and human perspectives as well as the anticipated benefits to the health care system. The thesis concludes by suggesting some guidelines for the design and evaluation of future health information systems.

TABLE OF CONTENTS

CHAPTER 1
HEALTH INFORMATICS FOR DEVELOPING COUNTRIES

CHAPTER 4

THE FIELD DATA COLLECTION SYSTEM:

DEVELOPMENT AND EVALUATION

CHAPTER 5

EXPERT SYSTEM ON TROPICAL DISEASES:

DEVELOPMENT AND EVALUATION

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1

HEALTH INFORMATICS FOR DEVELOPING COUNTRIES

1.1   INTRODUCTION

There is worldwide consensus that health should be a priority in national development strategies. This has resulted in a reorientation of policy to place a particular importance on the strengthening of health services in developing countries, where only a minority of the population has access to a high level of health care. The World Health Organisation's (WHO) Health For All strategy, for example, states that social development can no longer be regarded as distinct from economic development. This is because the progressive improvements attained by the development process can only be achieved and enjoyed by a healthy population [World Health Organisation, 1978]. This view has been supported by other authors on development. In defining development for all, Harrison [1983] asserts that development must aim for an eradication of absolute poverty, hunger and want. Streeten [1981] advocates a change from traditional economic viewpoints of development, towards regarding health as a basic need which should be viewed as a national priority.

A large difference exists between the standards of health provision, and the capabilities to provide the required services, in the industrialised countries and those of the developing world. This is reflected in regional statistics on basic health indicators and in the amount of resources available for health. In the poorest countries, life expectancy at birth is on average 22 years lower, and the infant mortality rate is 10 times higher, than that in industrialised countries. An average of one nursing person to 3,130 people in developing countries is contrasted with 1 to 130 people in industrialised countries; expenditure on health as

a percentage of total public spending is also low (3.4% compared with 12.5%), and decreasing [The World Bank, 1989b].

Nonetheless, there is wide agreement that major improvements are possible, even within the constraints that exist, if resources are better targeted and managed [The World Bank, 1989a]. Information technology (IT), in particular, is increasingly seen as a valuable tool for better resource management in health [Wilson *et al.*, 1988; World Health Organisation, 1987]. It must however be recognised that advocating IT in developing countries can be a contentious position, and some do not view it as an appropriate option for developing countries in general.

Much of the early resistance to IT adoption was fuelled by doubts about the creation of further unemployment and technological dependence [Sardar, 1988; United Nations Industrial Development Organisation, 1983; Bhatnagar and Bjorn-Andersen, 1990; Vasquez and Zimmerman, 1987]. Many governments in developing countries responded by imposing controls on imports of computer hardware. Sometimes, as was the case in India, discussion with trade unions about the use of the technology was a legal requirement. In some African countries the punitive import tariffs still in place are a legacy of this period.

This initial unease has been countered by two trends. Firstly, a recognition that modern information processing and communications can generate new service sectors within developing countries. Although IT does result in some unemployment in traditional sectors, it tends to displace labour, producing a demand for new types of skills [United Nations Industrial Development Organisation, 1983; Cole, 1986]. The second trend is the emergence and increasing role of the newly industrialising countries which have been able to develop a significant IT industry. The production of computer technology is no longer the exclusive province of the industrialised world, and countries like Brazil, Singapore and India now have the ability to implement a range of applications on a national basis [Schware, 1987; The Courier, 1989]. In short, in the 1990's it is widely seen that IT can be harnessed for development, and need no longer be regarded principally as a threat [Shires, 1982; 1983a].

Problems of implementation of the technology in developing countries, however, do arise in many instances, as well as difficulties in selecting the right type of applications [Schware and Trembour, 1985]. For developing countries, IT offers a wide range of application options [Association internationale futuribles, 1985]. The difficulty lies in choosing applications that can have a high impact on health care delivery. Intuitively, it may be easier to justify the use of IT for normal administrative purposes (such as nurse rotas, and book-keeping) than that of more innovative possibilities (for example satellite communications in operation theatres). For intermediate applications, which combine a sense of the pragmatic as well as the innovative, an informed decision is more difficult. Given the scarcity of resources, it is imperative that implementations are designed for maximum impact. In the tradition of appropriate or intermediate technology, IT applications should also be selectively adopted to ensure their sustainability.

The problems of applying IT to health are double-edged. They arise from the lack of resources, including technical skill, to implement the technology, as well as the inability of many health systems to absorb the technology into their organisational structures. It is still a debatable issue as to what type of information systems can best be applied and why. This question cannot be adequately answered solely by theoretical debate, but also requires real-life evidence. Studies that illustrate successes or failures, and identify factors responsible for outcomes are invaluable as pointers to what types of systems can be most effective. However, few such studies exist, and this thesis counters this deficiency by providing two real-life experiences.

Practical investigations of computer applications in health will guide towards more sensible approaches to choosing future applications. The procedure adopted in this thesis is to develop systems that reflect local constraints, and then to carry out evaluations to investigate their effectiveness. Based on the evaluation outcomes, guidelines for implementing health information systems are made which, in turn, can contribute to the general discussion on computers for developing countries, and the identification of promising areas for future development.

1.2   RESEARCH CONTEXT

The two key concepts used widely in this thesis, *developing countries* and *health informatics*, are introduced here. Both need clarification because a variety of meanings has been attached to each by different researchers. Each term is discussed, and the view taken by this thesis is presented, defining more strictly the sense in which these concepts are used.

1.2.1   DEVELOPING COUNTRIES

The concept of underdevelopment is barely two centuries old, and results mainly from the immense surge in economic activity that occurred during the industrial revolution in Europe, creating a marked difference in the economic wealth of nations [Pearson, 1970]. Sardar [1988] claims it is also linked to the colonial histories of many developing nations, and may be regarded as a consequence of old colonial structures of power and economic control. Today the distinction between developed and developing countries is generally made in a purely economic sense between countries that have developed a strong economic base, and have a self-sustaining growth, and those that are still largely dependent [Pearson, 1970].

Under this definition, the majority of countries in Africa, Asia and Latin America are classified as developing. Researchers who disagree with this viewpoint tend to object to the use of the word "developing", explaining that at different times in world history, some parts of the developing world were leaders in innovation when parts of the developed world were lagging behind [Sardar, 1988]. Although this stream of thought has spawned a number of other terminologies - Third World, the South, and so on - all of them accept that a group of countries can be identified which are economically poorer than another group of richer countries, and therefore poses distinct challenges.

These developing countries share certain common characteristics [Todaro, 1982]. Although they have 70% of the world's population, they control just 20% of its income. These countries are often well endowed with natural resources (minerals,

water, etc.) but are rarely able to develop an indigenous capacity for exploiting such resources. They have a very small industrial base; most exports are primary products subjected to a minimum amount of processing. Capital and skilled manpower resources are scarce. Economies tend to be agrarian, and government income is largely dependent on the world price of (usually) one major commodity. Lacking the capacity to produce many of the consumer goods required (or desired) by their populations, developing countries import a large proportion of their needs from the industrialised nations. Population growth is high, and there is increasing pressure on basic public facilities (such as schools and hospitals) concentrated in the urban areas, which now have to cater for the stream of population migrating from the rural areas. Governments in developing countries also have additional constraints on their spending in the face of growing external debt burdens.

Generalised descriptions, such as given above, cover more particularly the poorest developing countries, classified by the World Bank as low income. While some of these characteristics may well describe aspects of other developing countries, it by no means does so comprehensively. In addition to the poorest of the poor, the World Bank also classifies developing countries into lower middle income, upper middle income, and high income countries. It must be noted that the classification criterion of Gross National Product (GNP) per capita used is not entirely satisfactory, because, although income remains a major determinant of the ability to provide for basic needs, the extent to which economic benefits have been distributed remains significantly different. In some countries with GNP per capita levels comparable to poorer European countries, increases in national income are relatively recent, and their impact in raising the general living standards is still low with much inequality remaining [Conyers and Hill, 1982]. For example, Gabon, as a result of its income from oil is classified as middle income. However, average life expectancy is still 52 years, below the middle income country average of 67, and infant mortality at 105 per 1000 live births is double the middle income country average.

As well as differences in income within developing countries, there also exists a diversity in population size, population growth rate, percentage of rural population, adult literacy rate, skilled manpower level, industrial activity, climate and

ecological endowment (Figure 1.1). Thus, different countries are equipped with varying levels of resources, both capital and labour, in their quest for national development [Gupta, 1987].

In this thesis, a definition of developing countries is used as follows: those countries with a large proportion of their population living under absolute poverty conditions, with restricted access to basic needs such as food, shelter, clothing, health and education services. This is a severe definition but serves to provide a clear indication of the background against which the studies reported in later chapters are to be seen and evaluated.

1.2.2   HEALTH INFORMATICS

A few definitions of health informatics exist, but these are generally inadequate for the purposes of this thesis primarily because they fail to distinguish the concept of health from that of medicine. This distinction is important in this work because different conceptual approaches breed different ways of dealing with the technology, and in a developing country setting the relative balance of health versus medicine is crucial to effective development.

> "Health informatics is *not only* the computer applications in health but also is the body of knowledge acquired by applying the gamut of methodologies and techniques that may lead to computer support of a health activity, particularly support to coping with its related information" [Mandil, 1983].

Mandil also goes on to say that health informatics is *not only* medical informatics. His definition covers not only computer applications in the direct delivery of health care, but is more wide reaching, also including information support activities. Mandil seems to incorporate medical informatics into health informatics, with the former presented as a subset of the latter. No strong dichotomy between applications in medicine and health is established. Instead, his view is that medical informatics is of a more restricted scope than health informatics. As examples of present uses of health informatics he includes statistics and epidemiology, budgeting, hospital management, telematics, and bio-medical instrumentation.

SELECTED CHARACTERISTICS	BANGLADESH	CHILE	GABON
World Bank Classification (1989 report)	Low Income	Lower middle Income	Upper middle Income
World Bank GNP ranking (1989 report)	5	67	86
GNP per capita ($) (1987)	160	1310	2,700
Average growth rate of GNP (%) (1987)	0.3	0.2	1.1
Debt (millions of $) (1987)	8,851	15,536	1,605
Debt as % of GNP (1987)	50.6	89.4	52.5
Area ('000 square kilometres)	144	757	268
Population (millions)	106.1	12.5	1.1
Population growth rate (%) (1987-2000)	2.4	1.4	2.6
Population per physician (1984)	6,730	1,230	2,790
Population per nursing person (1984)	8,980	370	270
Life expectancy at birth (years) (1987)	51	72	52
Urbanisation (% of total population) (1987)	13	85	43

FIGURE 1.1 : The diversity within developing countries

Gremy's [1983] attempt to define medical informatics is quite similar.

> "... the meaning I personally give to the expression medical informatics ... cannot be limited to the use of computing machinery in medical public health, and biomedical research. It covers the whole set of sciences, methods and techniques - including, of course computer sciences - for handling a specific kind of information: medical information."

In addition, his definition supports the use of both terms interchangeably, as he later goes on to say:

> "The first and main concern of health informatics has dealt with hospital and curative medicine, specially in its most advanced technical aspects. Of course, it is possible to point out many exceptions, ..., but it remains clear that primary medical care, and public health have been neglected for a long time."

He criticises applications in industrialised countries for being biased towards curative rather than preventive aspects, illustrating his case by listing several types of activities that attract most attention - information systems for medical action, clinical research and hospital organisation.

This lack of uniformity in definition can be partially attributed to the infancy of this new branch of science, as illustrated by the relative novelty of societies for medical informatics. The British Medical Informatics Society, for example, was formed in 1986. Its aims of bringing together people from all professions supports Duncan's [1983] claim that this emerging discipline draws from and is related to other fields of study such as informatics, health science and sociology.

Perhaps this lack of clarity can be resolved by examining the relationship between medical and health sciences. Traditionally, the distinction between the two lies in their different emphases, being highly individualistic in medicine and with a more community based perspective in health. If medicine is defined as the science of *restoring health*, it could be argued that, although the two disciplines are closely linked, health science, by being concerned with disease causes, prevention and cure, at a more global level than in medicine, is more strongly biased towards a community and developmental role in *sustaining health*. It covers, for example, education, provision of clean water and sanitation services, immunisation and control of communicable diseases.
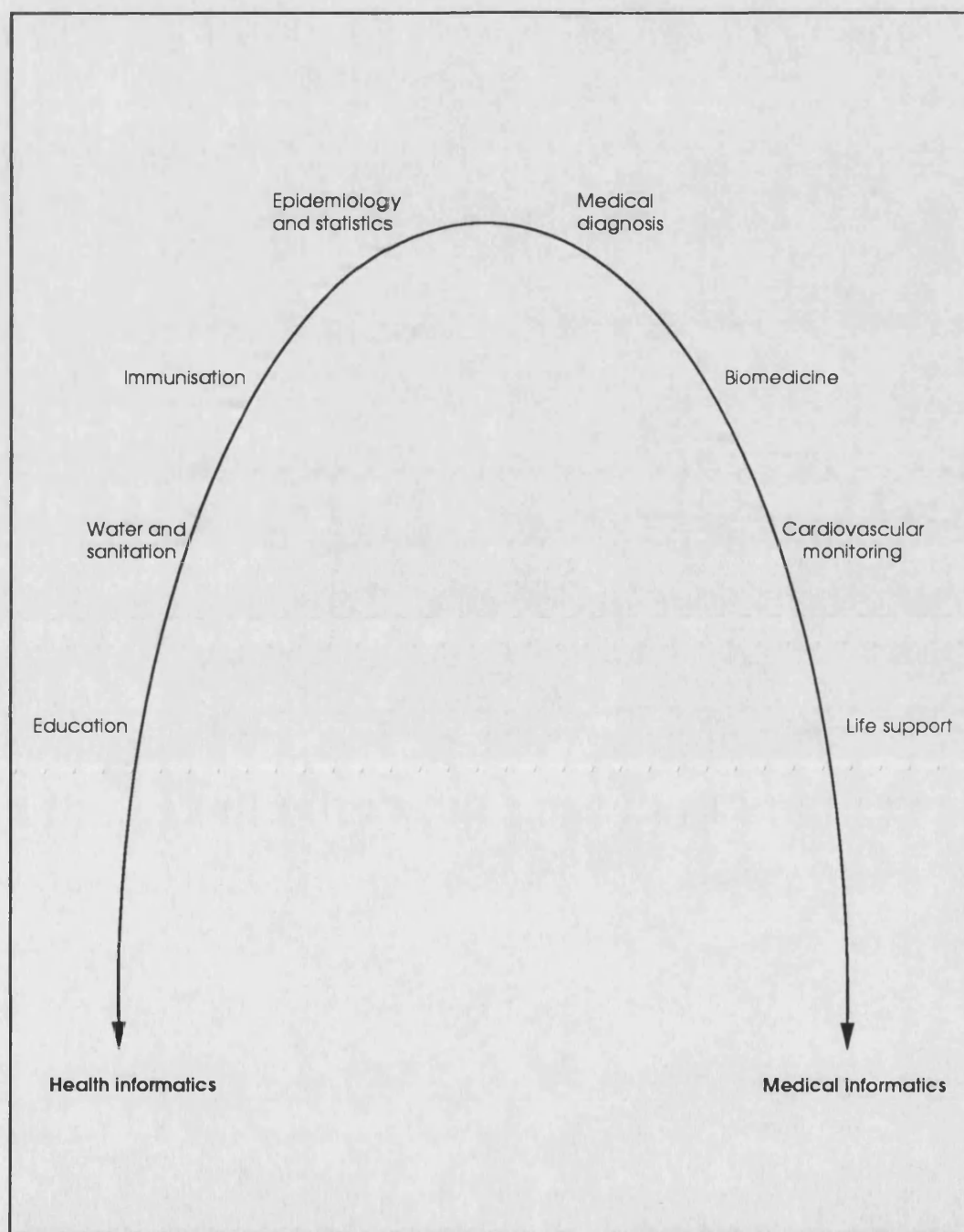
FIGURE 1.2 : The health / medical informatics spectrum

From the body of literature on IT applied to medicine and health that has built up, it is generally possible to discern grouping around one of the two application area clusters. Some papers tend to involve the use of technology in the diagnosis and cure of individual patients, in cancer therapy for example, these are focused on medical informatics. Others focus more often on health policy formation and administration, health education, and basic treatment services; these are more clearly health informatics. A spectrum of applications is identifiable between the two extremes of medical informatics and health informatics (Figure 1.2). Attempts to delineate the two will always be based on arbitrary distinctions as there is much overlap between the two interconnected areas, for example, as in the use of computers in basic clinical research. Nonetheless, this thesis can be placed firmly in the health informatics domain.

A working definition of health informatics would exclude intricate engineering miracles and advanced clinical procedures. Rather, it focuses primarily on the use of computers in the handling of information relating to disease distribution, transmission and causes, and the mechanisms for the control as well as diagnosis and cure of disease at the primary health care level.

1.3   PROBLEMS IN THE HEALTH SYSTEMS OF DEVELOPING COUNTRIES

The diversity within developing countries is reflected in the state of their health systems and the extent of health coverage. The ability of a government to provide health facilities for its population is partially dependent on its income, and therefore would mirror closely the classification of a country into income groups [Agbalajobi, 1983]. But even within a given country, it generally also holds true that the provision of health services is not evenly distributed but is concentrated in urban areas and that the poor in peri-urban and rural areas often lack access to facilities or cannot afford them. The problems in trying to meet this need are similar across the range of developing countries and provide a further context for investigating IT, and particularly its use as a distributional technology.

Evans *et al.* [1981] and Vallbona [1983] present two interesting perspectives on health systems. Evans *et al.* consider the present systems in a historical context, describing the development of health systems as stages in an evolutionary process. Vallbona presents a structural description, identifying different levels of care in a health system.

Evans *et al.* [1981] describe stages in the evolution of health systems, the developed countries having progressed through the stage which developing countries are currently at. They link the pattern of diseases to the type of care provided at each stage (Figure 1.3). The first stage was dominated by infectious diseases, whose spread was linked to poverty, malnutrition and poor personal hygiene. The introduction of public health measures by the provision of sanitation, immunisation and clean water resulted in a decline in child mortality and a subsequent rise in life expectancy.

The second stage was characterised by chronic diseases such as cardiac diseases and cancer. Personal health services became more important as emphasis shifted from public health services, to the provision of increasingly more specialised curative treatment in hospitals. However, the use of sophisticated and expensive methods did not made as striking an impact as expected on the health status of the population.

In the third stage, more attention is paid on personal health care and increasing awareness of ways to avoid patterns of behaviour that lead to disease, in seeking protection from environmental hazards, acknowledging the effect of changes in family and social conditions and so on. There is renewed interest in preventive and protective measures, and an added emphasis on individual and community responsibility.

Developing countries have to cope with all three stages of evolution simultaneously, whereas developed countries have evolved through to the third stage. The majority of populations in developing countries are rural. Although they have the most pressing and most easily preventable diseases, the existence of a vocal and more affluent urban population, with disease patterns similar to that of developed countries, tends to divert resources away from the simple and effective measures
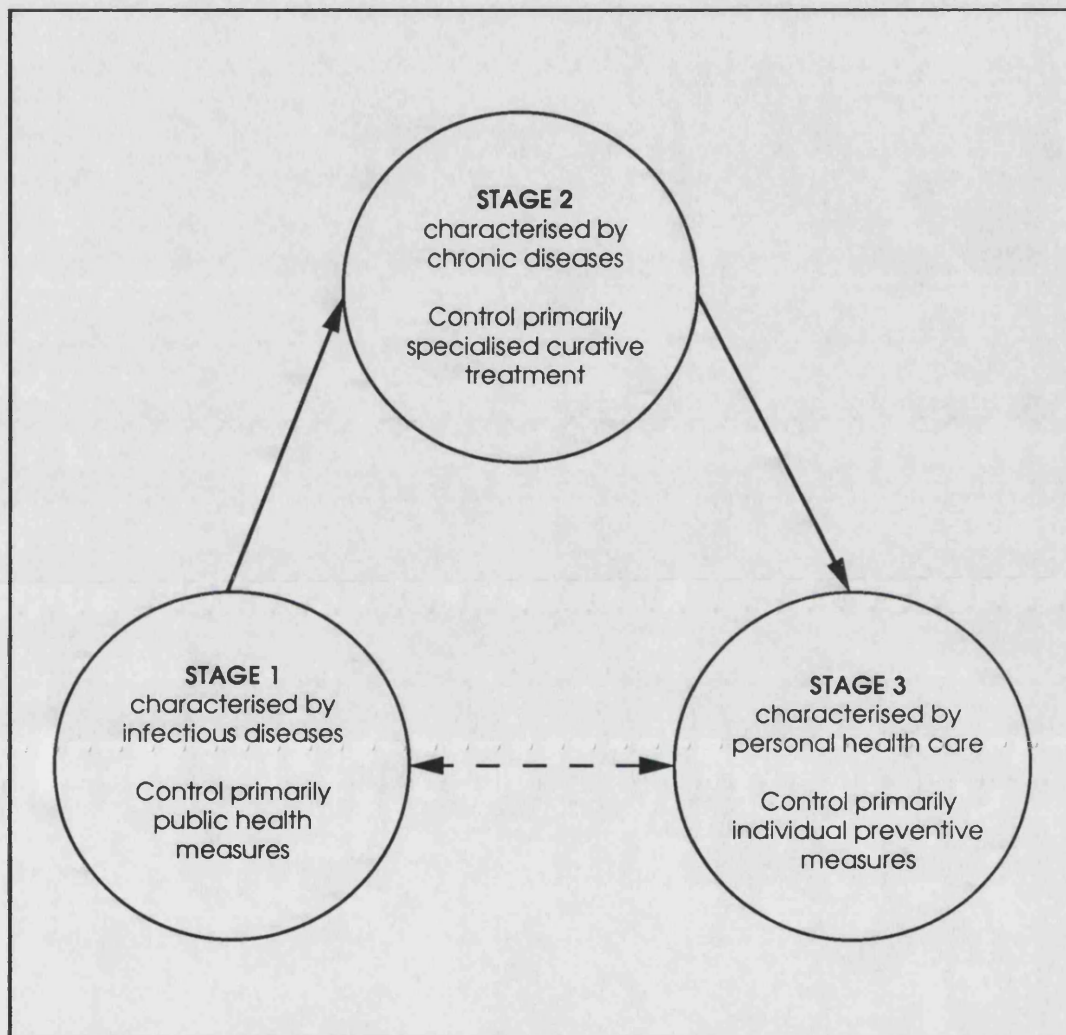
FIGURE 1.3 : A historical representation of the health system (based on Evans et al.)

needed in rural areas. The prevalent pattern of disease in rural areas is characterised by infectious diseases, its spread exacerbated by poverty and the lack of hygiene and sanitation. Most of these problems could be improved by public health measures, and indeed, the life expectancy of populations in developing countries provided with basic health facilities has increased. The pressure put on services by larger and increasing populations will dampen this effect, and many suggest that the situation can only improve further if family planning is introduced as a control measure [Harrison, 1990].

Vallbona [1983] presents a different model. He describes structural differences in health care provision in terms of the different care levels at which health services operate (Figure 1.4). At the primary care level, primary health care workers and doctors deal with common problems at health centres. More serious problems are treated by specialists at community hospitals; this is the secondary level. Super-specialists deal with complex and critical problems at the tertiary level either in regional or national centres. A pyramidical illustration represents primary care at the base and caring for most cases, and tertiary care at the top, dealing with complex cases. Figure 1.4 includes an additional level, a self-care level below the primary level, at which the patient or family copes with limited problems within the home.

The facilities that exist in many developing countries for providing health care at the higher levels are limited. Developing countries have low doctor and nurse ratios, and the sophisticated technology that is required at such levels is expensive. The distribution of services is concentrated around urban areas, which attract most doctors and contain most clinics and hospitals. Provision at the secondary level is usually available, although only in urban areas. A few scattered clinics typify the services available to the majority of rural dwellers, where self administration of medicines and the use of traditional healers outside of the formal health structure is more common. Vallbona also points out that there are many differences in the management and delivery styles in health systems within developing countries. He describes health systems as being at varying positions on spectrums of variables, so that there exist numerous managerial combinations and options. Systems range from highly regulated to fragmented and unintegrated; either balancing preventive and curative aspects or
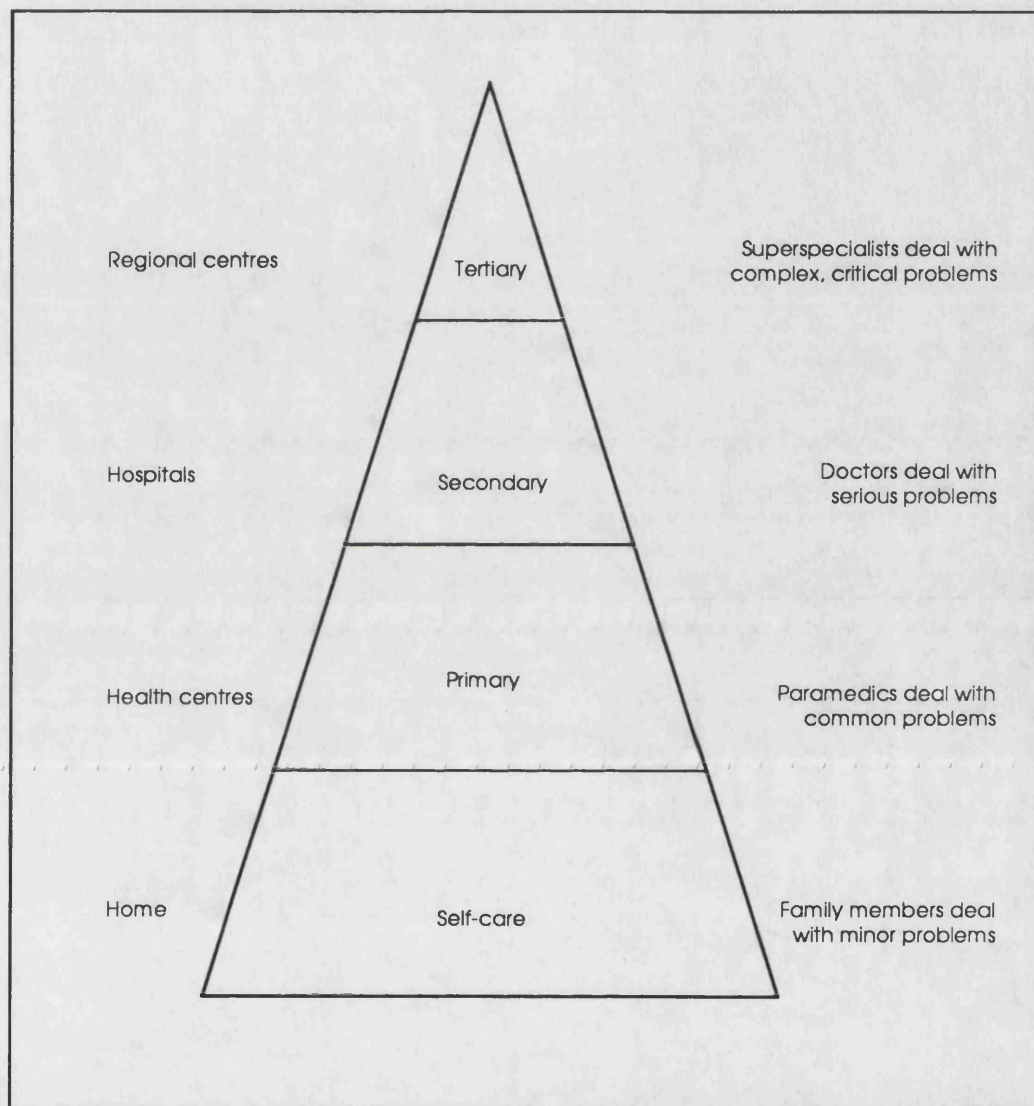
*FIGURE 1.4 : A structural representation of the health system
(based on Vallbona)*

biased towards one ofthe two; with a varying degree of technological sophistica-
tion; and with a variable proportion of private care available in addition to a
publicly funded system.

These two approaches to describing health systems help us in two ways. Evans
et al. explain how the current health systems of industrialised countries evolved,
and shows that the greatest impact was made at the first evolutionary stage.
Developing countries are now faced with choices from the wide range of medical
knowledge and experience available today, in a situation different from the past
when the limits of knowledge served to constrain the development of curative and
preventive medical practices. It is clear that their best choice, at present, is not to
imitate industrialised countries. Vallbona's structural descriptions paint a similar
picture in a cross-sectional form, presenting the different types of care available at
any moment in time. This view provides an immediate focus for health planning,
and illustrates the level of resources that must be directed towards different health
service strata, with primary health care needing the most attention and resources
to achieve the greatest impact within developing country populations.

This basic argument was recognised at the Alma Ata Conference in 1978, where
the WHO proposed a strategy for helping governments achieve a minimum
standard of health care throughout a nation [World Health Organisation, 1978].
universallysulting Primary Health Care approach has since been adopted by various
nations throughout the world. It aims to provide essential health care that is
universally accessible and acceptable with community support and participation.
It is regarded as the means of raising and levelling an unequal distribution of
health services. This approach supports Vallbona's classification, by advocating
primary health care as the means of increasing health coverage. It confirms his
placement of primary health at the base of the health system pyramid, serving the
basic needs of the majority of a population.

Several recurring key ideas in the declaration - affordability, self-reliance, co-
ordination - indicate the intended emphasis. The system should provide care at
affordable costs, involve community and individual self-reliance for quick
implementation and wider coverage, and should be considered as part of a
national development policy co-ordinated with other levels of the health system

and other sectors. By stressing that social values and economic conditions of countries and communities vary, the WHO allows scope for diversity in implementing a national health strategy and policy.

Evans *et al.* [1981] discuss several factors that they consider as obstacles to achieving the WHO goals.

1. *The uneven distribution of health services.* Large segments of the rural population are unreached because facilities are concentrated in urban areas. Low national doctor and nurse to population ratios disguise the large disparities between densities of personnel based in urban and rural areas. Personnel training has been biased towards procedures and techniques in developed countries, resulting in the production of doctors who need sophisticated facilities for their work. To redress this balance and to increase broad coverage, less skilled paramedical staff from rural communities will have to be trained in inexpensive health care provision. They will need access to a support system for distributing drugs and supplies, and supervision to ensure that they deliver a maximum amount of primary care. This approach would divert pressure from the stretched secondary and tertiary care levels whose skilled staff could be devoted to more specialist care provision.

2. *The lack of appropriate technology.* Although technology is available for immunisation against many infectious diseases, little research has been directed towards developing cheap technologies for serious diseases frequent in the developing world. As many medical technologies are developed in the industrialised world, they are often expensive and complex, with delicate mechanisms that require skilled use and maintenance. Scientific knowledge on the biology, vector transfer and control of common tropical diseases is poor. The scientific capabilities of developing countries are limited and disease research is concentrated in industrialised countries, whose preoccupations lie in other areas such as cancer and cardiac diseases.

3. *Pharmaceutical policies to control the purchase and distribution of drugs.* In many cultures patients expect to receive drugs when consulting medical personnel. The excessive use of drugs, which are mostly imported, places high demands on foreign exchange and diverts funds that could be used in other aspects of health care delivery. At present an average of 40 to 60% of total health expenditure is on drugs (in contrast to 15 to 20% spent in developed countries).

4. *Ineffective management of health resources.* The organisation of a complex health delivery system involves the co-ordination of different elements at various delivery levels. It also requires co-ordination between sectors that complement community development to reduce duplication and conflict between the objectives of the various programs. It therefore includes the management of education, budgets, supply services, and personnel. The most obvious deficiency is in administration at the district or local level, where managerial capability is very low. This link in the delivery chain is vital to maintain contact with communities and to respond to their evolving needs. Effective management depends on decision making based on information about health needs, but planning staff are often not trained to demand such information and the available information system is usually inadequate.

5. *Poverty.* Low GNPs and mass poverty restrict the amount of funds that can be dedicated to health services, which has to compete at the governmental level with other development concerns. Growing populations place an additional demand on the already strained services.

6. *The financing of health services.* The percentage of GNP spent on health in developing countries is less than a third of that spent by developed countries. Between $1 and $19 per capita is spent in developing countries (poorest countries to middle income ones) compared with up to $1000 per capita in the United States and Europe. To finance deficits in the health budget, governments in developing countries need to investigate different sources of funding and not depend solely on foreign input.

This raises several interesting questions. What are the prospects for increased public expenditures on health, and to what extent are improvements in health dependent on economic progress? Can existing resources be used more effectively? Within these constraints is *Health for All* through Primary Health Care a feasible objective? Although this thesis does not investigate these issues in great detail, it concentrates on one important aspect: how the use of information technology can help to improve the delivery and management of health systems in developing countries.

1.4   THE POTENTIAL OF HEALTH INFORMATICS

The use of micro-electronics and computers in the health systems of industrialised countries has improved aspects of their medical and health delivery systems [BOSTID, 1986; Fernandez Perez de Talens *et al.*, 1983]. With the special problems that developing countries face, how could the potential of informatics be maximised? Some are optimistic, thus a report of a workshop on Management Information Systems and Microcomputers in Health Care suggests that the use of microcomputers is the only means of achieving the aims of the WHO Strategy of *Health for All by the Year 2000* [Wilson *et al.*, 1988].

With a move towards small-scale and decentralised organisation, computers now offer a wealth of opportunities for increasing efficiency in various sectors in developing countries, including government delivered public services. Computers are capable of performing information handling and storage tasks with greater efficiency, speed and reliability than humans. Recent trends in micro-electronics have resulted in more powerful, more reliable and faster processors at a lower cost, smaller and more portable computers that are more flexible, have lower energy-consumption, larger storage devices, and increasingly more friendly interfaces that can be used by non-technical people [Guy and Arnold, 1987; Schware and Trembour, 1985; Forester, 1975].

The use of computers has permeated most sectors in developed countries and the so-called information technology revolution has also influenced aspects of medical

care in a number of ways. Meindl [1985] summarises the many applications of microcomputers to medicine and includes:

(a)   *medical research* - in modelling biological systems, simulating disease processes, physiological mechanisms and pharmacological interactions, and in the collection and analysis of data in biomedical research;

(b)   *medical data collection* - in standardising records of patient histories, allowing storage and retrieval of past consultations, and in tracking disease progression and treatment response;

(c)   *medical decision-making* - using decision analysis, clinical judgement or problem-solving paradigms to give advice under conditions of uncertainty;

(d)   *computed tomography* - to produce more detailed and precise images than can be obtained with x-rays;

(e)   *ultrasonic imaging* - to produce an image of soft tissues such as a foetus, heart or muscle using radar signals;

(f)   *clinical laboratories* - mechanical and electronic equipment are interfaced with computers to reduce response time, errors and costs;

(g)   *interpretation of data from diagnostic tests* - the computer refers abnormal results from a large volume of data, e.g. in electrocardiography;

(h)   *monitoring of critically ill patients* - to provide prompt and accurate information about patients in intensive care on a continuous basis; and

(i)   *aids for the handicapped* - to increase the functional abilities and job opportunities for the handicapped by using the sensory, memory and graphic capabilities of microelectronic chips.

In considering how computers could be used in the provision of medical care, Meindl's listing is clearly a developed world view, unsuitable for approaching the problem of improving health delivery to the majority of populations in developing countries at the primary health care level. The first three application areas in his listing, covering medical research, medical data collection and medical decision-making, could equally be assigned a health role. But even for these, the focus will have to be modified to accommodate a health emphasis. For example, health research to help understand what a balanced diet should contain using locally available food; health data collection on nutritional status, to assess the impact of some new intervention and guide policy formulation; health decision-making in deciding how to allocate drugs and schedule their distribution. This has not been

the primary focus of health information system implementations in industrialised countries; rather, there has been an unfortunate bias towards curative rather than community based systems [Gremy, 1983].

Clearly, because their needs and objectives are different from that of the industrialised world, not all of the applications of interest in the developed world are directly useful in helping developing countries solve their most urgent health problems. Information technology will therefore have a greater impact if used to improve the delivery of basic treatments and preventive care rather than expensive curative care.

Agbalajobi [1983] suggests that developing countries are at present none the worse as a result of the slow penetration (or absence of) informatics, because the current demand for management information is low and automated hospital-wide information systems are a luxury. But he stresses that as the health care delivery system grows, effective management becomes a non-trivial exercise and informatics then becomes invaluable. The problems of health systems in developing countries discussed earlier indicate that, irrespective of the size of the health system, informatics can have a role in implementing well-directed and targeted policies. But with the circumstances and financial resources of developing countries being very different from those of developed countries, their needs and priorities also differ. Developing countries seek inexpensive technological applications that can be adapted to their conditions.

There have been various suggestions about the types of systems that could be used in developing countries. Systems can be designed to support the managerial process for national health development and in the formulation, programming, budgeting, planning, implementation, evaluation and replanning of health programmes [World Health Organisation, 1981b]. Statistical data would be required to assess the health status of different communities, allocate resources, monitor socio-economic indicators, and analyse Primary Health Care coverage. Other more ambitious systems could be designed for medical diagnostic aids, establishing national databases, biomedical laboratories and instrumentation and for use in education. Potential users would include policy makers at the highest political and administrative levels, programme managers, health care personnel,

supporting technical and administrative staff, research and training personnel and physicians [Uemera, 1983; Mandil, 1983; Salamon, 1983].

Not all of these suggestions are either feasible or practical in view of the formidable problems which exist, mainly because the capacity of present health systems to absorb these new methods and technologies is very limited [Vallbona, 1983]. As Boafo [1987] stresses, the acquisition and utilisation of information technology should be determined by authentic needs of development, and not simply by easy access to the technology. Developing countries should therefore resist the temptation to follow the rapidly accelerating rhythm of IT development.

The problem in health systems of developing countries could be simply restated as: to meet the health needs of the majority of the population as quickly as possible. The challenge is to use currently available information technologies to devise applications essentially focused on health, directed towards the community level.

1.5   KEY IMPLEMENTATION ISSUES OF HEALTH INFORMATICS IN DEVELOPING COUNTRIES

To identify possible uses of IT in health is easy compared to considering the problems of implementing sustainable applications. Various issues are influential in assessing the sustainability of health informatics applications in the longer term, which apply to the technology itself as well as to the health context. Both the technological and health aspects are discussed in this section under the various issues introduced.

It is important to consider firstly, whether the introduction of computers will reinforce the technological dependency of developing countries on industrialised ones, without any significant gain in the development process. This also applies in the area of health as many developing countries continue to import the majority of their drug requirements. Next it is necessary to consider how the infrastructural deficiencies of developing countries will determine the viability of a computer-based information system implementation. The requirement of skilled

manpower to plan, operate, and maintain the systems is often not available, but is it possible to generate these? A major constraint in all plans, ambitious or otherwise, is that of capital, often foreign linked and with limitations on use. Ultimately, however, the best laid plans are of no use if the people who should implement them feel threatened by a proposed system; the socio-cultural aspects must be taken into account before and during system implementation.

1.5.1 TECHNOLOGICAL DEPENDENCY

The use of advanced technology in developing countries in the past has created a dependency relationship, characterised by an increasing reliance on outside sources for machines, equipment, inputs and both managerial and technical personnel [United Nations Industrial Development Organisation, 1984]. Stewart [1977] attributes this to the historical development of inappropriate technologies. Most of the technological innovations in the past 150 years have been in the developed countries and have incorporated techniques that were judged efficient in that society, and which are dependent on already available and older technologies. For instance, some common assumptions are that the technology will be operated in a temperate climate, that as much energy as is required will be available, that certain managerial and organisational methods exist, and that levels of investment and labour costs are high.

Computer technology, as a product of these industrialised societies could be assessed as inappropriate in a similar sense. Developing countries have low per capita incomes and do not have high levels of capital at their disposal. In addition to this, there is a shortage of highly skilled labour although there is an excess of unskilled labour. To avoid unemployment, they require technology that utilises high proportions of unskilled labour. This could be described as a pay-off situation, where it has to be decided whether the benefits of IT adoption outweigh the possible disadvantages of not doing so. The question arises as to whether the advantages can be maximised and the disadvantages minimised in the interests of development. What makes IT appropriate? How can it be used to address essentially developing country problems? What should be done to avoid dependency?

Proponents of intermediate technology promote the use of technologies that are relatively labour-intensive and increase the productivity of traditional methods, yet remain less costly than the most advanced technologies of industrialised countries [Elliott, 1986]. A related proposal has been made, suggesting the blending of old and new methods, utilising the potential of IT for increasing the efficiency of small-scale and decentralised organisation in a selective fashion, without sacrificing large amounts of capital or reducing employment levels [Cole, 1986; Lent, 1987].

Others argue that the mere utilisation of IT will not break dependency and that an accompanying policy of production is necessary, because the ability to adapt technologies for local use can only evolve from some indigenous production experience [United Nations Industrial Development Organisation, 1981]. This could be seen as an essential factor in overcoming the persistent and cyclical nature of dependency and for producing additional IT requirements such as peripherals and consumables. A further angle on this debate is the view that the problem lies not with the technology, but with the policies that determine its use - whether laissez faire or regulated [Zimmerman, 1990; El Sayed Noor, 1984]. Each country must decide, on the basis of their resources, what their integrated approach to IT utilisation for development will be.

At this point, it might be asked if the problem of technological dependency in the sense of continuing reliance on industrialised countries need be a permanent characteristic of all developing countries. If technological exchange can and does occur among the industrialised countries, where there is an increasing level of specialisation, a similar arrangement might emerge among the developing countries. The resource-rich ones could transfer the results of their labour, with a more pertinent applicability, to the poorer ones for their mutual benefit. Such a solution will be found by pursuing applications with a redirected focus, dissimilar in approach from the aims of research and development in industrialised countries. The potential of IT must be exploited in a selective way, using non-imitative applications to increase efficiency in targeted areas.

Concern has also been expressed about the dependency of most developing countries for drugs and vaccines for their health services. Few countries have

viable drug industries to supply the local market, and most pharmaceutical products are imported. Much has been written about the unethical practices of many multinational pharmaceutical companies when operating in the developing world [Sardar, 1988]. In an attempt to help stem the overwhelming supply of similar drugs with a bewildering array of names and supposed benefits, the WHO constructed a list of essential drugs [World Health Organisation, 1980]. However, as Evans *et al.* [1981] explain, implementations of health programmes usually lead to a rapid increase in the consumption of drugs, and ways must be found of reducing this to more sensible levels. In pursuit of less dependence on drug imports from industrialised countries, regional co-operation by sharing resources and research costs can help a determined group of countries to design and implement stronger pharmaceutical policies [World Health Organisation, 1978].

1.5.2 INFRASTRUCTURAL DEFICIENCIES

In the developing world, much of the basic physical infrastructure that is taken for granted in developed countries is either non-existent or poor [Wofsey and Dickie, 1971; Bennett and Kalman, 1981; Baark, 1986; Bhatnagar and Bjorn-Andersen, 1990]. Electricity supplies are often sporadic and inconsistent, the variations being extremely damaging for computer equipment, which have to be protected from power surges and failures by stabilisers, uninterruptible power supplies, or generators [Vuister, 1988]. Modern telecommunications equipment that transmit error-free digital data may not exist, limiting the range of possibilities of system configuration, so that decentralisation by distributed networks cannot be implemented. Climatic factors are different, and controlled air-conditioned environments are often required to protect against high temperatures and high humidity levels [Byass, 1989a]. As a result of this, installation costs may be considerably more than the mere acquisition cost. It may be necessary to purchase an additional set of equipment as backup in case of failure, to minimise disruption of essential work procedures.

The existence of a national scientific infrastructure could direct research effort towards developing suitable - small, cheap, self-maintaining - systems for such environments. As far as hardware is concerned, new methods and media for data

storage will have to be found to overcome, for example, the possibility of disks melting, and disk drives that attract fungus [Byass, 1987]. More research on alternative ways of powering systems is needed, on solar power, longer-life batteries, and on systems that require less power or can automatically regulate their power intake. Other types of communications have to be pursued, for example, by radio or satellite transmission.

The cost of software development has increased in relation to the cost of hardware because it is a labour-intensive and highly skilled exercise. Acquisition from abroad is an expensive business, particularly if the software still requires further adaptation for local use. For example, in many now independent states, the languages introduced during the colonial era have remained as official languages. The common assumption that all potential users have an adequate grasp of this foreign language may not hold. Locally adapted user interfaces are needed that could use suitably designed and locally understood icons that are action-oriented.

Many other aspects of support infrastructure are also relevant. Training facilities are generally poor, requiring staff, software, literature, equipment and other facilities that are not easily available locally. The supply of computer peripherals and consumables, for example furniture, paper, disks and ribbons, will rarely be provided through local production, most will be imported. Computer suppliers may be poorly represented, and even when they do exist, often cannot offer independent advice, being only equipped to market and support the products of a single parent manufacturer. Maintenance facilities that guarantee a fast response to equipment failure are rare. Often there are regional manufacturer representatives who are responsible for a large area, and because of their mobility, they are often difficult to contact, resulting in a longer down-time of computer equipment than is necessary.

In health systems within developing countries, there are similarly these three types of problems. In terms of physical infrastructure, there are not enough physical buildings that can be assigned a strongly identifiable role as health centres. The scientific and professional structure to support research and refine practice using locally available materials, especially in that of herbal remedies is poor, with a few exceptions in China, Kenya, and Ghana. Support in terms of

timely distribution of drugs and equipment can be partly attributed to transportation difficulties, but also suffers from a lack of managerial skill. This extends also to a lack of supervision and continuous in-service education of health workers [Evans *et al.*, 1981].

1.5.3 SKILL AND PERSONNEL AVAILABILITY

One of the obvious ways of helping to break dependency and of improving the local infrastructure is to educate and train nationals to perform the necessary tasks. In most developing countries, in spite of a surplus of unskilled labour, there is a lack of skilled personnel in all fields, including IT.

Expertise in IT is needed at all levels, to include technical and managerial staff, as well as users. Technical staff are needed to advise on equipment, and to install, operate and maintain systems. Their skills have to be linked with those of managerial staff who make decisions about work procedures and are more capable of devising economically-sound strategies. There are often problems in combining the two, with technical staff working to strict technical constraints but not aware of management problems, and with managers on the other hand operating with a similar managerial bias. For organisational planning, an ideal would be hybrid personnel skilled in both technical and managerial matters who could liaise with other employees, equipped with an awareness of the dual constraints.

The level of computer literacy within the population is low, and, unlike industrialised countries, widespread familiarity with this new technology does not exist. This is because high income levels in industrialised countries encourages home computer ownership, and consequently increases general computer awareness. New computerised information systems may be introduced in organisations where none of the staff have previously used or perhaps even seen a computer. In such contexts, it is important that familiarity with computers is also built up amongst this semi-skilled group.

In the short-term, foreign consultants may be brought in to supervise the initial stages of information technology implementation. This is of course expensive, and unless they are experienced and carefully trained, it is unlikely that they will appreciate and take the full organisational context into account when acting in their advisory role. Instead they may impose certain imported procedures that may be unsuitable. As these posts are temporary, systems may be implemented that are not sustainable after the expatriates leave. The use of foreign staff is invaluable initially, but it can contribute to the establishment of an indigenous IT capability only if skills are transferred [Boye *et al.*, 1988].

If nationals are trained, there is a possibility that once qualified, they may leave for more lucrative posts overseas unless they are given enough incentives to stay - the 'brain drain' problem. Within the health care system, the problem could also be extended to internal personnel movement, with staff leaving the public sector to work in more attractive private sector positions.

A significant problem in health care delivery is a lack of managerial skill, the weakest links in the managerial process being at district and local levels. To help alleviate the problem of too few staff in the rural areas, it has often been suggested that rural health workers be recruited and trained to deal with the most common and easily cured diseases [Gowers, 1987]. China and Peru are examples of countries that have dealt with this by training "barefoot doctors" [Harrison, 1983].

1.5.4 CAPITAL AND RESOURCE MANAGEMENT

Extra costs are generally incurred in developing countries because of geophysical conditions, so that backup systems, air conditioners, and uninterruptible power supplies have to be added on to the cost of acquisition [Delapierre and Zimmerman, 1985]. Costs are also increased because of a lack of indigenous resources - hardware and spare parts have to be imported, software packages purchased at international prices, expensive foreign expertise is required to customise software, and maintenance is difficult [Baark, 1986; Bhatnagar and Bjorn-Andersen, 1990]. The lack of some basic infrastructure, such as telecommunication networks and

training facilities, increases installation costs if IT users have to provide for this themselves.

The international market tends to work against developing countries, so that strongly agrarian economies have to cope with the vagaries of fluctuating world prices for their products. The all too important foreign exchange required for imports also suffers likewise. Not being able to accurately predict income hampers long-term expenditure planning. In the case of Nigeria, for example, the government overspent in the days of the oil boom, borrowed too much in expectation of future earnings, and was then left with dwindling incomes and a debt burden. Many countries suffer from a similar lack of continuity of access to capital.

Governments may choose to finance a project by seeking foreign creditors, but loans often come with strings attached [Boafo, 1987]. The unwise lending in the 1970's has led to the current international debt crisis. Alternatively, they may seek aid. As has been mentioned by many authors, foreign aid affects choice [Kluzer, 1990]. Foreign governments may prefer to fund certain types of projects, may decide to fund only parts of a project, or may require a donor content. The need to seek foreign financial assistance may result in little command over monetary resources, and a lack of continuity in planning.

Health has been given a low priority in the past, more focus being given to sectors that seemed more directly linked with economic growth. This is changing to some extent, but the percentage of GNP spent on health in developing countries is still far short of that of industrialised countries. In addition current World Bank policy, as part of their Structural Adjustment Programmes [The World Bank, 1989a], is against heavy government subsidy of health services, and encourage patient payment for health care. Part of the pro-fee argument is that payment increases the perceived value of services. This new element has been introduced in several countries already, including The Gambia and Kenya, as either user fees or contributions. The amount patients have to pay for health services is partly determined by how much the government can spend. With many rural populations already living under poverty conditions, paying for health becomes an additional burden on their household income.

The crisis over finances influences what governments can devote to promoting the use of informatics, and also in improving health service provision.

1.5.5   SOCIO-CULTURAL ASPECTS

Any technology must take local ways of thinking and behaving into account; it must try to cause 'soft' cultural changes which can easily be absorbed by the local community [Baumer, 1980]. For example, notions of rationality and efficiency may differ, affecting the ways in which skills are accepted, modified and embedded in a society [Boye, 1988]. Language, attitudes towards time, and the meaning of authority may also differ [Robey *et al.*, 1990]. In addition to national cultural traits, organisational culture has also been distinguished as a separate issue [Bell and Wood-Harper, 1990]. If, for example, people are used to working with informal, non-standardised methods of data processing, IT implementation becomes more difficult [Baark, 1986].

Indeed, a computer-based information system cannot be divorced from its social environment. Walsham *et al.* [1990] go as far as to say that an information system should be regarded as embedded in a social system. There are a growing number of researchers who have pointed out that organisational and social aspects may damp out the intended effect of technical innovations. Keen [1981] and Dickson [1970] discuss several potential hurdles. System designers are often surprised to find that the processes by which individuals make decisions in their work do not necessarily approximate to the rational ideal. Information systems often increase formalisation, requiring unfamiliar techniques which may be seen as threatening or unneeded by the user. The change that occurs is best if based on the users' felt need and motivation, and if implementation is incremental and in small stages, with precise objectives. If management is given data that allows them to more closely monitor subordinates' and therefore study their 'productivity', there will be obvious resistance on the part of the users. Resistance may be expressed by inertia or sabotage, completely undermining an otherwise technically efficient system.

In many countries in the developing world, a parallel culture of traditional medical practices exists alongside the western system. Trust is an essential element in treatment, and traditional medicine incorporates more cultural content than imported styles. Patients using both systems often believe that certain types of illness are more easily treated with Western style medicine than others [Harrison, 1983]. Often, also, patients seeing health workers expect to be given drugs or injections, believing their cure to be impossible without at least one [Evans *et al.*, 1981]. The influence of traditional culture may still be strong, but this can be over-dramatised by authors like Binet [1983], who seems to view Africa in a time warp, ignoring trends in an emerging modern society in which both co-exist. However, it is clear that new methods of delivering care must be accepted by patients before they can be useful.

Many trained health providers in developing countries have either come from the urban areas, or become accustomed to urban life during training, so they are reluctant to work under more difficult conditions in the rural areas. This is an example of a situation where a social pattern, as suggested by Boye *et al.* [1988], can have an effect on the implementation of a health policy.

1.5.6. *TOWARDS SUSTAINABLE INFORMATION TECHNOLOGY IMPLEMENTATIONS IN HEALTH*

These 5 issues, of vital importance when considering the implementation of sustainable applications, are very much interrelated (Figure 1.5). The lack of an indigenous scientific research and development capacity can be described as a sign of technological dependency and can equally be described as an infrastructural deficiency. The dependence on foreign expertise and funding is also a characteristic of technological dependency. Infrastructural deficiencies can be partially attributed to a lack of capital and resources. Skill and personnel availability can also be linked to social expectations. Yet this multi-dimensional perspective provides a basis for discussing positive steps that developing countries can take towards overcoming existing constraints.
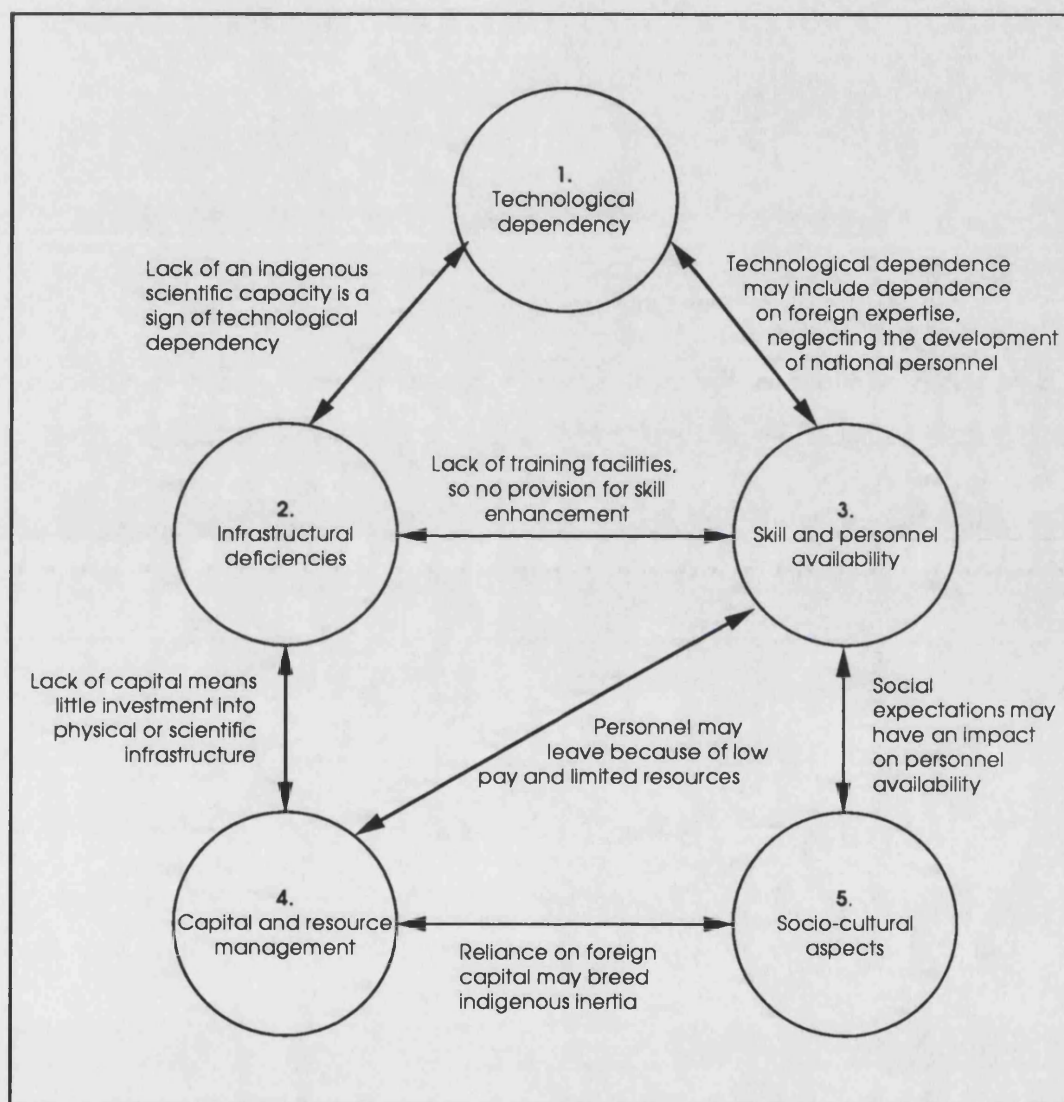
FIGURE 1.5 : Some interrelationships between the key implementation issues of health informatics in developing countries

Success in implementing IT applications and improving health services might breed yet more dependency, requiring more imports of computer technology and drugs. Several options are possible for developing countries, including that of producing for their own needs and developing an indigenous scientific capability [United Nations Industrial Development Organisation, 1981; 1983]. Industrial policies could also be refocused, so that developing country governments encourage more trade amongst themselves [Vasquez and Zimmerman, 1987]. Regional co-operation could also include pooling resources and establishing institutions for training and research.

Most developing countries have a surplus of skilled labour, but some, like India, have an educated class whose skills are often underutilised. In such circumstances, a viable software industry is a possibility that could be pursued with government support, expressed in terms of deregulation, enforcing legal copyright procedures, active financing of the sector, and so on [Kalman, 1981]. Similar government sanctioned encouragement is also needed to provide infrastructure which, if not available, tends to increase the overheads of organisations wishing to use IT, telecommunications and training facilities, for example. To help disseminate knowledge, professionals and users could establish computer societies and nursing organisations, to provide a forum for the exchange of ideas.

Government incentives could stem the flow of skilled professionals either overseas as an external drift of expertise, or into the private sector as internal movement. This should be expressed not only in pecuniary terms, but also in terms of career structure, continuing training, and management appreciation. Governments should not only seek to increase the absolute numbers of staff, but also ensure that personnel of different levels of skill are all put to the best use. A cascaded skill ladder can extend from managers, to people within communities who are trained in some required skills, such as rural health workers. In some instances, foreign expertise will be needed in the short-term, but ultimately, it will be best to encourage the development of local managerial skill alongside [Madu, 1989].

There is a possibility that, armed with skilled negotiators and a well thought out plan, developing countries could exploit foreign aid to their advantage. This might mean the staggering of policy implementation, approaching donors more

likely to fund particular sectors of the overall project. Another tactic might be to approach a number of different donors for the same project, and select that offering the best all-round alternative.

Ways of overcoming social difficulties in developed countries include the participative approach, favouring user involvement in the design and implementation process. One of the difficulties of using this approach in a developing country is that users are less likely to have had previous exposure to IT, and therefore are less able to contribute in the early stages, having little conception of what the technology might achieve. A modified delayed approach which identifies a suitable application area, designs and implements a prototype which could then be used and criticised by the users, is probably more appropriate. More training opportunities should be created for rural people who are more likely to return to work within their communities. Governments should recognise that there is a role for traditional medicines and practitioners to be drawn into the public health system.

In the short-term, it is unlikely that many of these constraints will be overcome to a great extent. For the moment, the technology that is implemented must function despite existing limitations. This thesis will therefore concentrate on relatively low-cost, community oriented, sustainable, non-imitative applications that are sensitive to the levels of skill available and to acceptability by those affected by its use.

1.6   RESEARCH APPROACH

Before launching into the main body of this thesis, it is appropriate to pause here to more precisely formulate the research framework. The need for this research is put into the context of the preceding discussion. The reasons for choosing computer-assisted data collection and medical decision-aid systems as the two application areas of interest are given. The research methods used are introduced, and the section concludes with an overview of the major research contributions.

1.6.1   THE NEED FOR THIS RESEARCH

> "The technological maturity of a nation is a key influence on the choice of the type of IT that can be productively utilised in that nation." [Mohan *et al.*, 1990]

> "Further research is required to analyse how the benefits of the applications of the microcomputer as they are currently defined can trickle down to the farm and village level." [BOSTID, 1986]

> Emphasis should be on "the development and use of small scale, locally controlled systems for solving locally defined problems and achieving progress in locally defined terms." [Cooney, 1986]

The authors quoted above imply that the real pay-off from the use of computers in developing countries will come from applications of direct importance to the population, in sectors with well-understood and defined needs. Health has been identified as one of the basic needs required by a large proportion of populations in developing countries. Preceding sections have discussed the needs and constraints of providing adequate care, in particular the need for better management. This provides the basic motivation for introducing IT in the field of health, to improve overall health care delivery by enabling better resource utilisation.

Although it may be said that the right types of information systems are not available, it is unclear what these 'right' types are. There is very little evaluatory material available to allow identification of systems with high potential. Some systems have been developed but little published material exists to explain why they fail or succeed. One of the principal ambitions of this thesis is to help to establish an understanding of success and failure by a systematic evaluation of health information systems.

1.6.2   THE RESEARCH STUDIES

The case studies introduced in this thesis are therefore used as a basis for developing the required evaluation framework. The two applications have been chosen because of their potential as means of improving problematic areas: the

dearth of reliable data, and support for paramedical personnel. Generally, there is a lack of data of sufficient quality for health management purposes in developing countries, and this results in an inability to measure improvement brought about by policy implementations [The World Bank, 1989b; Wilson *et al.*, 1988]. There is also a need for human resource development, to encourage, train and support health personnel of all types, in particular at the paramedical level [Kaul and Kwong, 1988; World Health Organisation, 1978].

Hence, this research focuses on computer-assisted data collection systems and medical decision-aid systems. Each of these has been suggested by various researchers as directly meeting health needs; in the case of the first by enabling fast access to accurate data, and in the case of the second by disseminating medical expertise to areas where it is lacking. Both approaches have been proposed to be used at the lower levels of the health system as self-contained applications that can be operated by staff with little previous exposure to computer technology. Examples and prototypes of both types of systems have been created and implemented in developing countries, but none are routinely operational and it is unclear why this is so. The two represent an interesting divergence in the level of controversy they have attracted. In developed countries the notion of machine expertise has been attacked vociferously and few medical decision-aid systems are in routine use. Computer-assisted data collection systems, on the other hand, offer a safe and basic technological application.

1.6.3 RESEARCH METHOD

The basic format of this research has been to implement computer-based health information systems and to observe their impact in the environments for which they were designed. In this thesis, evaluation methods are constructed around monitoring the technical and performance characteristics of the systems, the attitudes of the health personnel and those of the public affected by the systems. In addition, the training requirements and the impact of the systems on working methods are investigated. The aims of this thesis can be summarised as:

(a)   To develop and implement two health information systems: a computer-assisted data collection system and a medical decision-aid system.

(b)   To evaluate them in context, considering training needs, user acceptability and effectiveness, and to consider the wider and longer term implications of such systems for primary health care.

(c)   To contribute to the future development of health informatics applications by suggesting criteria for successful identification and evaluation of projects.

The thesis has examined the literature concerning overall development issues, the use of information technology in developing countries, the state of health systems, and the expected benefit of information technology in improving their effectiveness. By covering so wide a range of topics, there is a danger of superficiality. This cannot be entirely avoided because there are few research precedents straddling all of these aspects, and this is an area of diverse and fragmentary literature. However, by taking an information system perspective, the thesis attempts to consolidate knowledge and provide a fruitful direction for more focused research. This overview, together with a description of the social and economic context of the countries within which the studies took place, provided a good footing for the field research.

The strongest element has been that of research case studies in which the researcher both undertakes the innovation, and observes and records reactions. This is not a 'pure' research methodology, but one that is widely recognised as valid in the complex and rapidly changing information system environment. Such an approach is often referred to as 'action research', "... applicable to an examination of human activity systems carried out through the process of attempting to solve problems" [Checkland, 1984]. A number of techniques were used, including personal interviews, the use of observation methods, and opinion surveys.

At the same time, where appropriate, more classical research designs have been used, in which before and after comparisons are made and validated by statistical means, consideration being given to recommended approaches to the validation

of computer-assisted data collection and medical decision-aid systems. Thus the research does provide some quantitative results alongside a more holistic analysis.

The research has been an evolutionary process of successive hypothesis formulation and refinement (Figure 1.6). The initial set of issues discussed earlier in this chapter generated some ideas of what systems are desirable and feasible. The design and development stage modified the idea of what systems are achievable. An evaluation approach was formulated as a means of assessing system utility. The case study evaluation stage showed the real-life aspects and what is practicable. The analysis and retrospective assessment stage produced guidelines for future developers as well as hypotheses for future researchers.

Overall, there are several distinctive characteristics of the research approach described. The author's role as systems' developer, as well as social science researcher and evaluator, requires careful balance. To counteract the potential bias in reporting, the original data is summarised to retain its original integrity, and the methods used are made explicit so other researchers can evaluate them. The very specific nature of case studies is another feature. Each has unique qualities which provide interesting and illuminating snapshots of a *particular* environment, and consequently generalisations from these are problematic. When made, therefore, they will be explained in the light of the issues discussed in this chapter, relating the narrower context to the wider issues. Finally, part of the aim of the research was to develop systems that *work*, and construct software that can be used in practice. Thus, underlying the next few chapters is a strong technical basis which must be recognised and considered by those reading this work. Ultimately, in order to get the systems described into action, a certain single-mindedness was required.

1.6.4   RESEARCH CONTRIBUTIONS

The research reported in this thesis seeks to make several contributions, both academic and practical. These are summarised below.

Research Path	Research Value
Initial set of implementation issues	Systems that are desirable and feasible
Design and development of systems	Systems that are achievable
Formulation of evaluation framework	Systems that can be broadly assessed
Case study evaluations	Systems that are practicable and effective
Retrospective analysis and assessment	Systems for the future

FIGURE 1.6 : *Progression of the stages of the research*

(a) *Bringing together fragmentary literature under well-defined headings.* Many efforts at using informatics in health and medicine in developing countries have been full of good intention, but low on focused effort. This thesis has made its stand on the need and aims of health informatics in developing countries clear. In addition, it has identified several interrelated issues that determine the degree to which an implementation can be judged successful.

(b) *Identifying appropriate technologies.* The research has identified robust computer technologies, simpler than the advanced state-of-the-art systems currently available, but which can operate under difficult field conditions. It has therefore included one often neglected resource - affordable and reliable hardware.

(c) *Developing two systems that work in the field.* The two systems developed can stand as models for other developers. While not perfect, they exploit the resources available, while accepting certain limitations on their scope.

(d) *Proposing an evaluation framework for summarising and communicating results.* A lot more work is required in the area of evaluation, but a comparative base will be lacking without a framework for describing experience. The framework developed here highlights the value of careful experimental design, but does not neglect 'softer' social factors. It provides a standardised method of reporting evaluation results.

(e) *Drafting guidelines for the design and implementation of future systems.* As a start towards the sharing of experience, the thesis suggests several guidelines, based on the results of the case studies, for future work in building health information systems.

1.7   THESIS OVERVIEW

The rest of this thesis takes up from here, introducing the two application areas, describing the two case studies, comparing the results, and concluding with suggestions for a more structured approach to health information system research.

Chapter 2 presents the two health information system application areas of interest: computer-assisted data collection systems and medical decision-aid systems. The need for such systems in developing countries is introduced, and some systems that have been developed described. For each application, some research directions, based on a review of literature in each area, are presented.

Chapter 3 summarises some approaches that have been suggested for evaluating computer-assisted data collection and medical decision-aid systems. It also presents the designs and results of some trials that have been carried out. It discusses the concept of evaluation and the need for evaluating health information systems from two perspectives - that of information systems specialists, and also from a health viewpoint. It then proposes a general framework for evaluating health information systems.

Chapter 4 describes the design and operation of the Field Data Collection System (FDCS). It presents a background to the case study, by giving a country and organisational context to the Medical Research Council Laboratories in The Gambia. The experimental trial design, methods and techniques used are described. The results are presented in different sections, describing the technical as well as the social aspects. The evaluation framework developed in chapter 3 is used to summarise the results obtained.

Chapter 5 describes the second system, ESTROPID - Expert System on TROPIcal Diseases, including details of how medical information was generated, manipulated and used to reach a diagnosis. A background to Kenya and the Kilifi District Hospital is given, as a contextual basis for the case study. The field trial evaluating the system is also described. The results discussed again include technical and social angles of the evaluation. Supportive results from a separate, independently carried out study is used to reinforce some of the case study

findings. The evaluation framework is used to structure and report the major findings.

A comparative analysis of the aims and the impact of both of these systems is made in chapter 6. The need for each system, and its expected role is discussed. The resource requirements in terms of hardware and user expertise are compared, as well as the complexity of the resulting software. The acceptability of the systems to the users, and the respondents or patients is presented. The effect of the system on the overall user productivity is assessed. The type and depth of evaluation required to assess the possible impact of these systems on health provision in developing countries is discussed. The chapter concludes with an assessment of whether the two case study implementations should be classified as successes or failures.

Chapter 7 first combines the main direction of each chapter, to present a cohesive overview of the entire thesis. It then discusses some of the practical and theoretical contributions of this research. It suggests some topics for further study, and concludes with a personal statement.

CHAPTER 2

TWO HEALTH INFORMATION SYSTEM APPLICATION AREAS

2.1   INTRODUCTION

This chapter introduces the two application areas which cover the case study systems implemented as part of this research. These are, first, a data collection system, to enable computerised error checking as data is entered in the field, increasing its quality, and second, a medical diagnostic-aid system to provide support for generalist health practitioners when treating out-patients. Before discussing the background to the research in these areas, it is important to consider, from the health angle, what makes health information systems appropriate for developing country contexts.

In 1978, the WHO proposed that, to achieve Health for All by the Year 2000, a Primary Health Care strategy must be adopted. They state that

> "Primary Health Care is essential health care made universally accessible to individuals and families in the community by means acceptable to them, through their full participation and at a cost that the community and country can afford",

intending that a basic yet inexpensive level of health care should be available, especially to rural communities. The gap between health care in industrialised countries and that in developing countries, as well as an urban/rural maldistribution within developing countries, has resulted in the rural poor and other vulnerable groups being put at a double disadvantage. A new strategy, directed towards their basic health needs, is "a burning necessity" [World Health Organisation, 1978].

For the purposes of this thesis, some of the components needed are restructured here under four headings: decentralisation, technology use, training, and evaluation.

(a)   *decentralisation* would mean re-orienting the national strategy towards the community, breaking down centralised control and management into a regionally organised approach, incorporating community-administered committees. This 3-tiered approach should allow for some organisation and decision-making by the recipients of health care, and enable them to report their needs to the regional centre. The regional centres would thus be better equipped to respond to the needs of communities, and would be responsible for reporting a region-wide profile to the centralised Ministry of Health.

(b)   *technology,* which should be used appropriately. Technology in their terms means "methods, techniques and equipment" which, to be appropriate, must be scientifically sound, acceptable to both the users and the beneficiaries, adaptable in varying contexts, simple and easily applied, as well as relevant to local conditions and resources. Not all sophisticated technology should necessarily be rejected as inappropriate on a basis of cost, rather, some deemed expensive may still prove valid and useful in other levels of the health system. Technologies for community health workers, however, must be proven, affordable and within their capabilities.

(c)   *training.* As much as possible, the workers must come from the communities they serve and should be trained close by, on courses adapted to their varying levels of literacy. This basic training should be concentrated in as short a period as possible, yet able to provide them with a minimum amount of useful and specific skills. Initial training should be supplemented by continuing training and by opportunities to share problems and ideas.

(d)   *evaluation* is the means by which the "success" of the program can be judged. Information flow, regular and relevant, is indispensable to making reasoned decisions. How relevant are the activities being pursued to the

needs and requirements of the communities? How much progress has been made in reaching the national goals? How has usage and coverage improved? Has the program resulted in a better health status? What has been its overall contribution to the community? For answers to these questions, certain indicators of change need to be continuously monitored at the community level and channels should exist for a two-way feedback between the various health levels.

To fit in with the needs specified by the WHO, appropriate information systems will be those that contribute towards meeting these ideals, i.e. those that

(a)   *encourage decentralisation*, breaking down the traditional centralised character of computer facilities in developing countries. Portable micro-computers, whether handheld or laptop, were used to take both these applications to the field - in a typical survey setting for data collection, and to an out-patients department for medical decision support.

(b)   *enable a sensible adoption of technology* that is affordable and can be sustained by the health system. Although the computers used are not the most technologically advanced, and are therefore available at a lower cost, they are still relatively expensive when considered alongside a typical rural health centre's budget. Their role has to be demonstrably validated, and suitable contexts more clearly defined.

(c)   *allow a training component*, transferring the skills necessary for national staff to perform the required tasks. People with little or no previous exposure to computers were trained to do their new computerised tasks within a short time.

(d)   *incorporate some means of monitoring and evaluation*, providing a means of judging a program's success in meeting national goals. The field trials, monitoring changes over a period of time, provide data that can be used to evaluate the potential of these systems in real-life.

With these considerations in mind, this chapter will turn to investigating further the role that computerised data collection systems and medical diagnostic aid systems can play in improving health services. The need for each system, and current research in both application areas will be presented. General research trends will also be identified and discussed.

2.2   COMPUTER-ASSISTED DATA COLLECTION SYSTEMS

2.2.1   *THE PROBLEM OF DATA COLLECTION*

A typical survey life cycle involves a number of activities which can be grouped into various stages. The progression is not strictly sequential and may involve some reiteration (Figure 2.1). Initially, in the design stage, the sample design has to be chosen and the questionnaire structure decided on. Preliminary ideas then need to be tested and probably refined before the start of the fieldwork. As data sets are completed, they are checked for completion and coding errors before entry onto a computer. Tabulations of the results are produced during the analysis stage, which will usually also include some statistical hypothesis testing.

The data collection case study focuses on the stages where errors may be introduced as data is collected - fieldwork, data cleaning, and data entry - ignoring the errors of bias that may be introduced by sampling design and nonresponse. Errors during data collection and editing may arise from various sources.

(a)   *the field workers interpretation of response* - what the respondent says and means may not always be what the interviewer understands and records.

(b)   *the coding and recording of the response* - if responses given do not directly correspond to the set of answers expected, the interviewer may either reinterpret them or pressurise the respondent to modify their responses to fit the structure. If the data is not clearly recorded, it may be interpreted incorrectly at the data cleaning and entry stages.

(c)   *data entry onto computer* - transcription errors may occur as data is being punched or, in cases of missing or ambiguous data, computer operators
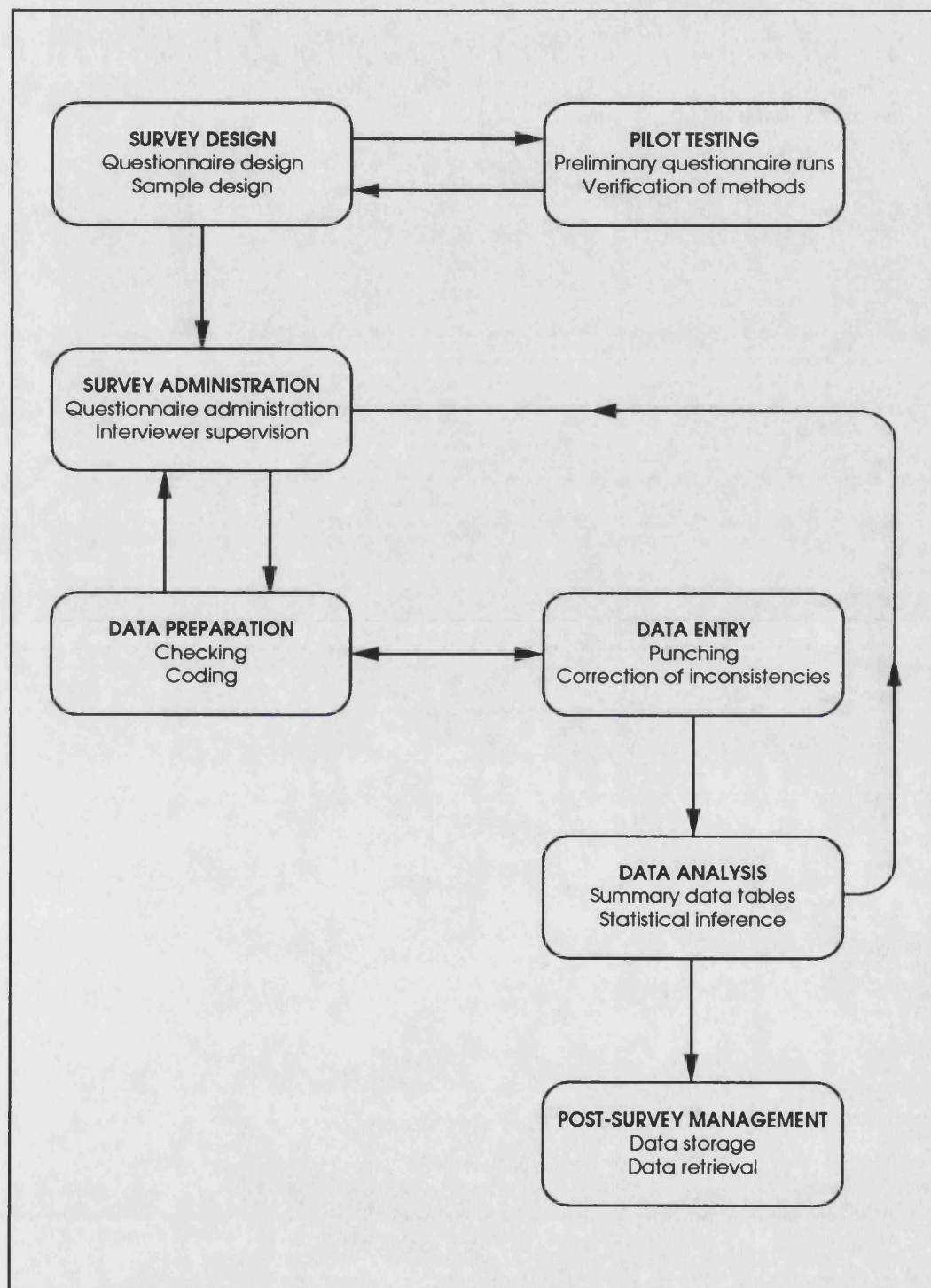
FIGURE 2.1 : Activities in the survey process

might attempt to substitute their own interpretation of the interviewer's meaning.

Common ways of detecting the last two types of errors are by double entry and plausibility checks on the data. Double entry requires the input of the same data set twice, preferably by two different operators, as a check on keyboard errors. The data entries are matched against each other in a verification routine and operators consult the original forms to correct mismatched entries. Plausibility checks are logical checks on the data to verify that they fall within acceptable and realistic boundaries and to check that there are no contradictions in the records. These are done on dates of birth, weights and other measurables and also to check that all entries belong to the set of expected answers for a question. When longitudinal surveys are being carried out, more complicated routines are possible, for example, to check for contradictions between the most recent entries and all previous ones. Not all errors, however, are necessarily detected at the end of this process. Often, a few will remain, though the residual error rate will usually be negligible.

Before the widespread availability of large storage devices, most data collected from studies carried out in developing countries that required large amounts of computer resources had to be sent abroad for data processing. This has become less likely and unnecessary because computer centres within developing countries can now be equipped with relatively low-cost equipment to meet their data processing needs. However, a problem may still remain. Where the data processing stage is divorced from data collection, access to the original records is often inadequate to correct all remaining errors. In such circumstances, the final percentage of invalid and unusable fields will be higher. Invalid fields within records may have to be discounted during analysis. The loss of such data may be significant in surveys with small sample sizes, reducing data quality and limiting the scope and applicability of the results.

Recent literature [Byass *et al.*, 1988a; Snow and Byass, 1988] has advocated a more localised approach to data collection in real-time so that punching, verification and processing is done within a short time of collection. The few residual errors that would pass undetected by the plausibility checks would be incorrect either as a

result of interviewer errors or input errors. A main advantage of rapid processing and checks is that incorrect or missing data excluded by range checks can be re-collected if necessary. A review of a survey carried out in The Gambia reports the following improvements by using this approach [Byass *et al.*, 1988].

(a)   the management of data collection and quality on a real-time basis,

(b)   feedback on project progress,

(c)   intermediate assessments and analyses of data are possible.

The authors attributed the high completion rates and low error rates to the easy interaction between computer, investigators and field staff.

The next step forward is more radical as it suggests the introduction of small hand-held computers in the field, in addition to a centralised office equipped with micro-computers [Ferry *et al.*, 1985; Reitmaier, 1985; Reitmaier *et al.*, 1987]. In the field, the computer can administer the questionnaire and validate the entries before storing them. It is suggested that this approach might significantly reduce errors and produce better quality data at the least cost in the shortest time. Figure 2.2 summarises how some of the sources of errors discussed above can be avoided by using this approach.

Good training for field staff and the provision of a very clear field manual outlining guidelines for questionnaire administration cannot be over-emphasised as the main way of reducing errors due to interpretation of response. CADAC systems may incorporate facilities for the inclusion of some of these guidelines as comments for online help. By using range checks and plausibility checks as the data is being entered in the field rather than during the verification stage at the central site, the number of interviewer errors and recalls back to the field would decrease. The use of at least one microcomputer for the co-ordination of a centralised database for all data collected and transferred from the portable computers used in the field is also an important aspect of this approach.

Sources of error during the data production process	Overcoming errors by using computers at the point of collection
1. Interpretation of response	Online guidelines for interviewer.
2. Recording of response	By using range and validity checks to verify entries fall within the acceptable range.
3. Completion of whole questionnaire	Forced entry to all questions and automatic skipping ensure that all appropriate questions are answered.
4. Data entry	Data already captured in a machine-readable form and can be downloaded easily.

FIGURE 2.2 : Overcoming errors in data production by using computers at the point of collection

2.2.2 AN OVERVIEW OF COMPUTER-ASSISTED DATA COLLECTION SYSTEMS

Governments, universities and commercial market research agencies in industrialised countries have been interested in using computer technology to improve data quality and speed up the production of statistics for almost two decades [Nicholls and Groves, 1986]. In the 70's, most of the attention was devoted to using computer-assisted telephone interviewing (CATI), and developing techniques for defining more complex questionnaires. Up until then, questionnaires with simple coded and short-answer open questions seemed to be the most appropriate type for computerisation. Research at this time also extended the capabilities of these systems to include survey design and call scheduling. For example, the United States Bureau of Census developed plans aimed at using CATI as an additional mode for data collection to be used in conjunction with mail questionnaires or personal interviewing [Groves, 1983].

Other aspects of computer-aided data collection are mostly referred to as Computer Assisted Personal Interviewing (CAPI), Computer Assisted Data Input (CADI), and Computerised Self Administered Questionnaires (CSAQ). These are all regarded as a family of Computer Assisted Data Collection (CADAC) systems [Nicholls and Groves, 1986; Lyberg, 1985]. Research has gradually become focused on producing integrated systems that can be used for the whole range of data collection methods.

CAPI has become feasible since widespread microcomputer adoption in the early 80's as an increasing number of small, portable yet powerful systems have been produced. A number of systems have been developed in various countries. In the Netherlands, several experiments have been carried out on NEC computers using a simple package, and subsequent research has produced BLAISE, an integrated system that runs on IBM-compatible machines [Denteneer et al., 1987]. The French SIC (Systeme Integre de Collecte) project has developed QUESTOR, which is now commercially available [Ferry and Cantrelle, 1988]. Several Swedish prototypes have been developed on various machines [Lyberg, 1985]. A few other systems are enhancements of some already available package to include some CADAC features, for example dBase IV, although these are generally limited to

CADI functions only. Most systems offer the following basic facilities: questionnaire specification to include question text, range and consistency checks, routing (detailing the sequence of questions to be asked depending on earlier responses); a guided interview complete with error messages; and some form of standardised output that can be used by statistical packages.

Several types of systems have been developed and tested by the Netherlands Central Bureau of Statistics. An early system called QUEST was developed for the NEC model with very limited memory and a 40 x 8 character screen [Bemelmans-Spork and Sikkel, 1985]. QUEST was written in BASIC and allowed questionnaire specification to be made in a syntactically structured text file. This permitted 3 types of questions, numeric, precoded or open, and also incorporated range and routing definitions. The syntactically correct questionnaire is then interpreted by the program. The data collected is stored in a compact form that is converted after data transfer to a larger computer.

The latest Dutch system, BLAISE, uses a special questionnaire specification language that is quite similar to Pascal in structure. A set of three different blocks defines a questionnaire - the QUEST block contains the question texts and the range of acceptable answers; the ROUTE block contains the routing; and the CHECK block contains consistency checks [Denteneer *et al.*, 1987]. Each question and each possible answer is given a label, so the specification has a very structured format. This then goes through a parser which checks the syntax and produces a Pascal program that is converted into executable code by the compiler. BLAISE also offers facilities for hierarchical questionnaires and can link up with other data sets for checking purposes. Ordinary Pascal procedures can be integrated into the code and very complex calculations can, in this way, be specified. It also allows tables to be set up, and a block of questions can be defined, displayed and handled in a spreadsheet fashion. The first version runs on IBM PC/XT compatibles and requires a Turbo Pascal environment.

A number of prototype systems have been developed in Sweden. Like many other statistical offices, the early systems were mostly tailored programs for a specific questionnaire. Newer systems have incorporated modules for questionnaire design, interviewing and a choice of sampling unit for the next interview

[Lyberg, 1985]. Tests have been done on various types of computers, including the Toshiba T1200 in 1988.

The French SIC system, also marketed as QUESTOR, operates on a range of MS-DOS and CP/M machines. The original system ran on the Husky Hunter, a hardy, and completely sealed, shock-proof machine the size of an A4 sheet of paper, battery powered, with a 8 x 40 character screen, and up to 496k of memory. The current commercialised version now runs on MS-DOS based machines. It has several different modules: a questionnaire generator, an interview module and an interface module [Ferry and Cantrelle, 1988].

The generator uses menus and windows to guide the survey manager through questionnaire specification interactively. The system allows hierarchical question-naire structures, where a subset of questions could be asked repeatedly for several individuals in a sampling unit; this is usual in demographic surveys, where basic information should be obtained for each adult member in a household. It could also be used in repeated surveys on the same sample, as data collected from previous waves can be used to check consistency in later waves.

The interview module can be operated by an inexperienced computer user to collect the data in a compact form that is formatted into a data file. This file can be used, together with another generated file containing a dictionary of all the variables, as raw data for common statistical packages. The French aimed to develop a system that could be used under "difficult" conditions, including tropical climates [Ferry *et al.*, 1985], and their system has now been tested in several West African countries [Ferry and Cantrelle, 1988; Crisan *et al.*, 1988; Balde *et al.*, 1988].

An example of a system that was designed specifically for developing countries is the child health map, but its application is restricted to producing certain kinds of health data [Reitmaier, 1985]. This runs on a battery-powered Sharp PC pocket computer used in conjunction with a cassette recorder for data storage and a colour plotter for printing. The system was developed to collect data on nutri-tional status, participation in immunisation programs and breast feeding rates of children. It has checking routines to prevent the entry of incorrect or incomplete

data, and immediately calculates the nutritional status while the child is being seen. The printer makes the system autonomous, enabling results to be printed out and, if required, compared on-site with data from other locations or years. Several other examples of single-purpose systems used in developing countries are cited in the literature [Ferry and Cantrelle, 1988].

2.2.3 RESEARCH DIRECTIONS

Some of the reasons suggested for using CADAC systems in developing countries are similar to the calls for systems in industrialised countries, for example as a means of improving data quality. Additional requirements in developing countries include independence from foreign experts and central scientific institutions after the implementation phase, low price and operating cost, independence from AC-power supply [Reitmaier, 1985], screen visibility in direct sunlight, machines able to withstand dust, humidity and high temperatures [Ferry *et al.*, 1985]. The educational level of interviewers is one of several other factors more directly concerned with the survey environment. These suggest some of the peculiar reasons why systems developed for industrialised countries may not be feasible for poorer countries.

In general, the approach to CADAC has moved away from single-application questionnaire development, which is costly in terms of programming time. The trend is towards generic development environments which can very flexibly be tailored to specific survey needs. The rapidly evolving technology has made a number of systems developed in the early 80's obsolete. Increasing sophistication is demonstrated in integrated systems incorporating many additional and often complex management functions. CADAC is now not only a feasible option, but achievable. The literature mentions that CATI is used widely by many market research agencies [Lyberg and Sundgren, 1987]; CAPI, however, is only catching up. In developing countries, however, implementing CADAC systems requires more stringent hardware specification and substantial benefits in view of the costs to be incurred.

Many systems for reporting health data in developing countries are flawed, and much of the data collected are of such a poor quality they cannot be validly used for policy making. When one-off or repeated surveys are conducted to redress this problem, data collation, analysis and interpretation usually proceeds at so slow a pace that planners cannot respond to emergency situations in time. This suggests why CADAC systems, able to provide high quality data within a short time, may make a striking impact.

The field data collection system reported in chapter 4 aims to provide CADAC facilities that could feasibly be used to conduct surveys in the field. Based on battery-powered handheld computers, it has facilities for supporting interviewers by guided interviews, thus exhibiting several of the desirable qualities of any such system for developing countries. While it could not be said to cover all aspects of a full CADAC system that aims to contribute to the whole survey process, it does include the essential elements needed to reduce errors during and immediately after data collection. The case study aims to evaluate its use in a field study, and assess its suitability for similar surveys in developing countries.

2.3   MEDICAL DECISION-AID SYSTEMS

2.3.1   THE ROLE OF COMPUTERISED MEDICAL DECISION-AIDS

Medical decision-aid systems have been advocated since the 1950's as a means of helping physicians choose between diagnostic and therapeutic options [Kleinmuntz, 1984]. The term *expert system* has been used to describe these systems, which were expected to demonstrate levels of expertise akin to that shown by human experts. Their skill lies in their ability to reason under uncertainty, using symbolic representations of medical knowledge, and techniques for manipulating the encoded knowledge. Medicine has been a subject for investigation since the early beginnings of research in *expert systems* because it was seen to exhibit qualities suitable for knowledge representation. It was thought that, as a field with domain specificity, boundaries of knowledge in various specialist areas could clearly be delineated and elicited. In reality, most serious clinical problems are

broad and complex, and do not lend themselves easily to accurate model-based solutions [Szolovits *et al.*, 1988].

Amongst the four earliest systems (INTERNIST-1 for general internal medicine, PIP for renal disease, MYCIN for therapy selection for patients with bacteraemia or meningitis; and CASNET for glaucoma) few, if any, are in routine clinical use outside the environments in which they were developed [Shortliffe, 1987]. Twenty five years on, medical expert systems have not fulfilled the early promises of enthusiasts, and many systems have remained prototypes under testing [Fox, 1984]. Sceptics suggest that the aims of providing accurate diagnosis in medicine is questionable, and dismiss the ideas behind computerised diagnosis as simplistic and currently unachievable [Lancet Editorial, 1989]. This vigorous debate has stimulated a change of expectations, reflected in new terminologies emphasising that their role should be seen as more of an aid for a clinician, rather than a replacement. The term *'medical decision-aid'* has come into use, and is employed in this thesis to include a range of newer systems with less ambitious, but more realistic, aims [Wyatt and Spiegelhalter, 1990a; Byass and Corrah, 1989a; Shortliffe, 1987].

Although medical decision-aid systems in industrialised countries are not in widespread use, it has, however, been proposed that they can be of use in developing countries [Goldberger and Schwenn, 1983]. Decision-aid systems, it is claimed, can be used to distribute medical knowledge to areas where it is most needed - rural areas in developing countries, to be used by paramedical staff with limited training, experience and facilities (Figure 2.3).

The argument is presented in terms of a chronic need for skilled expertise in medical decision-making. Many rural health workers are over-burdened with the numbers of patients they are required to see every day; Essex estimates up to 200 patients a day. Newly qualified staff may find it frustrating to be unable to deal with so many, often resulting in the symptomatic treatment of most illnesses and little differential diagnosis [Essex, 1975]. Essex claims that, in this context, diagnostic tools have a role to play in raising skill levels.

He goes on to point out the difficulties in the type of training medical staff receive. Hospital-based practice on inpatients using conventional methods of history taking and examination are unsuitable for out-patient practice for a number of reasons. Inpatients usually have diseases in their later stages and have already been selected into specialities. Also there is less of a time constraint because there is a possibility of daily follow-up. In out-patient clinics, on the other hand, patients present with many more minor complaints, and often have diseases in their early stages. Out-patients not properly investigated may not be seen again. Most students are therefore unprepared for real-life out-patient diagnosis after their training and need support they cannot obtain when alone in rural health clinics. Once again, decision-aids, computerised or otherwise, are the preferred solution.

Goldberger and Schwenn [1983] claim that it is difficult to persuade physicians and health workers to go to and stay where they are most needed. Also, that the training of health workers is superficial, insufficient and not supplemented by continuing training and therefore easily forgotten. Their explanation is that

> "the lack of economic and environmental support for rural health workers, the inadequacy of their scientific background, the relative isolation in which they must perform and the difficulty in providing competent teachers and regular supervision, all combine to ensure the failure of traditional attempts to provide them with adequate training".

A permanently resident source of knowledge for rural workers with a minimum of training could therefore significantly enhance their skills [BOSTID, 1986; Murray, 1985]. Proponents of this approach suggest that the reasons that no medical decision-aid system is in general clinical use in industrialised countries is because they were meant for physicians slightly less competent than the experts [Goldberger and Schwenn, 1983]. The applications in industrialised countries are generally within restricted domains and meant to be used in sophisticated medical settings. On the other hand, the nature of the diagnostic problem within developing countries aids the medical decision-aid approach because the number of diseases that can be diagnosed is limited by the clinical observations that can be made, the historical information, tests and treatments available. Thus, a natural knowledge boundary exists.
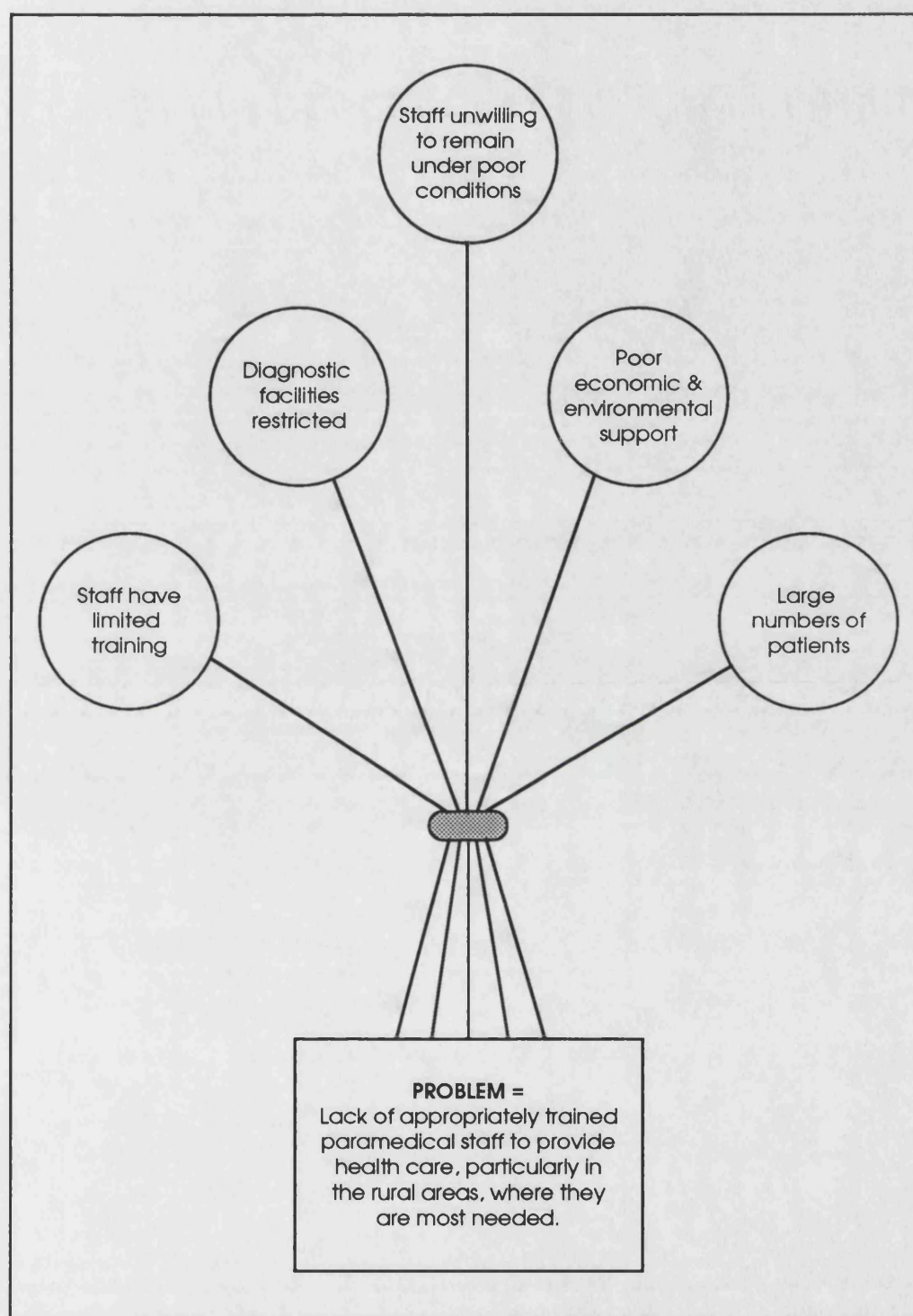
FIGURE 2.3   The suggested need for medical decision-aid systems
in developing countries

2.3.2 MEDICAL DECISION-AID SYSTEMS FOR DEVELOPING COUNTRIES

In the light of the arguments presented above, a number of research and development projects have been undertaken to explore the possibility of using medical decision-aid systems in clinical practice within developing countries. Four systems are summarised here. The first, a general medical decision-aid system for developing countries is TROPICAID, developed by Auvert and others at INSERM, France [Auvert *et al.*, 1986a; 1986b; 1986c; 1985a]. Subsequently other system prototypes have been developed by Byass and Corrah (The Gambia) [1989a], Porenta *et al.* (Austria) [1988], and Uplekar *et al.* (India) [1988]. The aim of each system, its knowledge base and representation, its method of inference, its current stage of development and the type of evaluation that has been performed will be described. This section highlights some similarities of the systems as well as ways in which they differ.

1. Identification of need

Porenta and Auvert both refer to the WHO Alma Ata conference, with respect to the need for improved health care in developing countries. Porenta states that the gap between health care between industrialised countries and developing countries seems to be increasing rather than decreasing. The problem, he continues, lies not with the lack of medical knowledge on how to treat the majority of cases, as these tend to be minor complaints that can be treated with simple remedies, but rather that current medical knowledge is maldistributed. Byass points out the shortage of skilled manpower. Auvert asserts that a further problem is to convince trained staff, physicians and paramedical staff to go to and remain in areas where they are most needed. Uplekar claims that, with simple training, a semi-literate village can provide primary health care, removing some of the health care burden from traditional providers.

2. The role of computer-based support

Porenta claims that a new role could be found for intelligent tutoring using knowledge-based systems to supplement traditional knowledge distribution in the form of books and personal tutors. Any such system could also offer on-the-job

training, as by continuous use, users could learn standardised consultation procedures. Auvert and Porenta see an additional potential if the systems can also act as a reference source. Byass states that computerised decision-support systems which are both clinically relevant and technologically feasible could be useful for diagnostic support at the primary care level. All the researchers propose computerised decision-support for rural health workers during consultations. Auvert claims that the use of such a system could increase the effectiveness of health workers if it offers immediate help without major changes in working conditions.

3.   Requirements of a computer-based support system

Auvert and Byass propose the use of hardware that is portable, lightweight, battery-powered, rechargeable by solar power, and virtually maintenance free. Auvert's software requirements include: knowledge on use of medications, treatment of a given disease, diagnostic decision, and collection of epidemiological data. Byass extends this even further: the system must cover all the problems that may be presented at a rural health centre (allowing for referral for difficult cases); must allow for multiple pathologies; must suggest a course of action rather than a diagnosis as an endpoint; must guide the user through the entire consultation; and must have a knowledge base that can be adjusted to meet local conditions. Porenta also mentions the need for health workers to recognise and treat only diseases within their capabilities, referring complicated cases to nurses or physicians.

4.   The construction of the knowledge base

All the proposals are built on a knowledge base containing encoded information representing the core of expert knowledge. Byass selected the indicators for his database from details entered on a questionnaire prepared for an out-patient screening clinic. His system contains general patient data on sex and age, the season of the year, 40 signs and symptoms, and 27 management strategies. Porenta's system combines several sources: textbook knowledge from two well-known medical authors, supplemented by the experience of a doctor who had worked for a short while in northern Ethiopia, and statistics collected on present-

ing symptoms of patients seen at an ambulatory care centre over 3 weeks. The project identified five problem areas: diarrhoea, worm infestation, eye diseases, skin diseases, and common infectious diseases, and fed the system with models of the presentation of these diseases.

Uplekar used case notes from patients seen by village health workers and at a general practitioner's dispensary to identify common symptoms, and compile a final list of 24 symptoms. The 25 drugs included were taken from the WHO Essential Drugs list for rural health workers, supplemented with additional information on dosage, side effects, and home remedies. They then prepared a program of algorithms for each symptom to extract relevant information about the illness sufficient to suggest a prescription, but do not give full details on how these were constructed.

The first version of TROPICAID based its knowledge on the Essex diagnostic flowcharts (see the example shown in Figure 5.1). The information contained in each node is a question, some of which have further explanatory messages. Details of 460 symptoms and 210 diagnoses are stored. Again, the list of drugs is taken from the Essential Drug list by the WHO, with information on names, indications, side effects and dosages added. The second version found the content of the medication and drug data bases to be adequate apart from slight modifications, but that of the diseases to be unsatisfactory. The new databases are larger, now with 500 diagnoses, 350 basic symptoms, 2000 clinical signs, and 160 essential drugs. Additional information includes therapy of each disease, a priority index for the treatment, a reference to a more generalised treatment, and a list of symptoms for which the treatment is effective.

5. *Knowledge representation*

Neither Byass nor Uplekar give details, in the papers available, on how knowledge is represented in their systems. In the first version of TROPICAID, each item of information was stored as a record in a packed thesaurus-like file, with a separate index file containing addresses to each. The fixed tree pathways were represented as nodes, each node containing the code number of the questions to ask, and for each possible answer, the code number of the next node, or the code

of a diagnosis. In the second version, the therapies are still stored in records, but the diagnostic database is now contained in frames. For each sign in the database, there is a corresponding question (now in one of several different categories classified to reflect the capabilities of the user). All the signs are linked into a network to preserve consistency.

Porenta chooses to represent the different disease groups in at least two different ways. An expanded decision network containing nodes, each with rules on how to proceed to the next node, is represented as frames. Each frame then has various slots with attributes flagging the presence or absence of a symptom. In addition, a rule-based approach is used. Each symptom has 2 certainty factors between 1 and -1, the first indicating evidence in strength of a disease and the second evidence against a disease. The decision network concept is used to represent the diarrhoea and worm infestation disease areas, and the rule-based approach for skin and eye diseases. How common infectious diseases are specified is not stated. Therapies are also represented as frames.

6.   *Inference mechanism*

Apart from stating that the system starts from symptoms and works its way through algorithms, Uplekar does not say more. Byass bases his on Bayes' theorem, using the unmodified probabilities of indicators in repeated calculations of advisabilities of the various strategies. He defines a certain threshold level above which a strategy should be pursued. The Porenta network contains nodes, all of which branch to other nodes based on whether the response to a node question is yes or no. At each node, an interpreter evaluates and weighs a set of symptoms included in a rule and determines which branch to pursue. With the rule-based strategy, a set of diseases that are potential candidates for diagnosis are chosen. These are ranked and worked on sequentially by the system, certainty scores being calculated for the disease under consideration by combining the factors using an algorithm. The diseases are reordered, the procedure stopping when a certainty threshold is exceeded.

TROPICAID-1 works on the same basis as the Essex flowcharts. A series of questions is asked and the branch to a subsequent question is dependent on

whether the answer is yes or no. Each flowchart chosen is examined exhaustively before it is abandoned. In TROPICAID-2, the system begins by selecting suspected diseases, based on the patient symptoms entered. The diagnoses selected are ordered and the most pertinent question is chosen. The answers update the findings base and allow some reordering, possibly the elimination of some diagnoses, the confirmation of some beyond the certainty threshold, or the addition of new ones. This question-reorder cycle is continued until a stop criterion is satisfied. A therapy is then chosen based on the degree of certainty of the diagnoses.

7.   *Interface*

The TROPICAID-1 interface was based on a series of menus with numbered choices, which trigger more menus or start the questioning process. The first menu allows the user to enter the main presenting symptom of the patient. Keyboard input is restricted to the use of three function keys and the numeric keys. TROPICAID-2 allows the user to enter all the patient's major symptoms using a series of selection menus. The Porenta system also uses a menu-based system for entry of patient symptoms, and then asks a series of questions to which the user may enter yes or no. The Byass system uses the basic entries from age, sex and season to determine the first question; subsequent questions are triggered by previous answers. The Uplekar system requires one major symptom to be entered at a time, the remaining symptoms to be added if they have not yet been mentioned.

8.   *Stage of development*

TROPICAID is the only system that has gone through a revision to a second version. All of the other systems were at the first prototype stage at the time of publication of the articles quoted here. Three have been evaluated to some degree, but require further modification before extensive evaluation can be carried out. The methods and results of these evaluations will be discussed in the next chapter.

2.3.3 RESEARCH DIRECTIONS

The systems differ in the type of knowledge that they aim to model. Byass'
system aims to model the actions of a paramedical worker, Porenta's knowledge
is based on what a rural health worker is expected to know, and the others on the
practice of physicians. Each of these has potential difficulties. If based on what
the paramedical workers do, there is an assumption that what they do is right. If
based on what the rural health worker is expected to know, practice may differ
significantly from training. If based on the practice of physicians, some of the
medical knowledge will be beyond their expertise.

The ideal could be based on what is taught during formal training, strongly
tempered with options in actual practice, for example, the ability to explain what
to do if the preferred drug is not available. The system's interface, providing or
requesting information, should have a firm basis in what they already know and
do. Unfortunately, this approach is rarely taken partly because few, if any, single
sources of such knowledge (for example, a book, a consensus group of experts, or
a code of accepted practice) that encapsulate all the required aspects exist. As a
result, many researchers are content with what information they can readily
obtain.

Several of these authors explain why they choose their particular methodology
against all the choices available. Byass rejects flowcharts as being too determin-
istic, and artificial intelligence (AI) methodologies as being too complicated.
Porenta *et al.* reject probability calculations as being too complex for the context,
but end up using three different representation methods for a relatively small
disease domain.

If these issues are to be resolved, it is clear that future research is needed. In
particular, factual investigations are essential to understand the real result of the
operation of any such system.

2.4   DISCUSSION

The current state-of-the-art in both these application areas is at different stages of achieving the aims of the systems. In addition, their potential within the developing world differs from that in industrialised countries. More accurately, the needs and resources of developing countries place constraints on the type of developments that can reasonably be expected to succeed in the short-term, and be sustainable in the longer term.

The case for computer-assisted data collection to increase data quality seems indisputable. These systems are well on the way to routine use in many statistical offices in industrialised countries. Computerised medical decision-aid systems, however, are in the centre of the contentious issue of machine intelligence. Furthermore, the problem of modelling human expertise is nowhere near solution. Computer-assisted data collection systems are quite acceptable because they can easily be shown to improve performance, without compromising professional responsibility. Medical decision-aid systems, on the other hand, have not been adequately shown to have dramatic improvements in performance, and have a higher risk of jeopardising the health and lives of patients. As a result, their introduction into routine practice is unlikely to occur in the near future.

CADAC research exists in response to a specific need with a fairly well-defined solution, more similar to conventional software programs. It is easy to define which of the survey tasks the systems will be used for. The state of research in computer-assisted systems is such that reputable systems are already available commercially, QUESTOR and BLAISE for example. Their immediate applicability in developing countries still seems hampered, but more by the existing conditions in the recipient environment, rather than the techniques required to computerise data collection.

On the other hand, there remain considerable problems of technique in the medical decision-aid domain. Rapidly evolving approaches to the modelling of medical knowledge and decision-making are being proposed. A view presented by Szolovits *et al.* [1988] of the historical progression of computer-aided diagnosis is that:

"By the early 1970s it became clear that conventional tools such as flowcharts, pattern matching, and Bayes theorem were unable to deal with most complex clinical problems".

These same "conventional tools" are those that are proposed by researchers in medical decision-aids for developing countries at the end of the 1980s. Why is this? One reason might be that the aims of decision-aid systems for such environments are simpler than those of the hospital-based specialist seeking computerised advice. This is not to say that the solution is simpler. It may well turn out to be harder to support semi-skilled generalists than specialist doctors in their decision-making.

Research in industrialised countries now seems focused on finding adequate AI techniques for modelling more complex representations of medical expertise. Research for developing countries, however, is directed towards using these now discarded techniques to investigate how suitable they might be for making comparatively simpler medical decisions.

In both areas, the approach is moving away from single-application programming to generalised software within industrialised countries. Commercialised packages for data collection are available, as well as so called 'expert system shells' (containing various reasoning mechanisms, but without the knowledge). Development time for applications can now be significantly reduced, concentrating instead on the interaction between machine and user, questionnaire and knowledge design.

The two case studies discussed in more detail in later chapters represent the best compromise of opportunity and design that could be achieved in investigating the potential of CADAC and medical decision-aid systems.

CHAPTER 3

THE EVALUATION OF HEALTH INFORMATION SYSTEMS

3.1 THE EVALUATION OF COMPUTER-ASSISTED DATA COLLECTION SYSTEMS

3.1.1 *AN EVALUATION APPROACH*

In their review of studies on computer-assisted interviewing, Nicholls and Groves claim that few comprehensive analyses of system feasibility exist [Nicholls and Groves, 1986]. In summarising the evaluatory literature, they describe most evaluations as speculative, merely discussing the potential advantages and disadvantages of computer-assisted methods with the support of weak evidence. They use the few studies that have been carried out under controlled experimental conditions using small samples to highlight deficiencies in research methodology. They suggest that the following factors should be evaluated - cost, timeliness, data quality, and added value (Figure 3.1).

Cost should be considered in terms of

(a)   installation and maintenance including the expected life of the hardware system, software development, maintenance, systems and programming support;

(b)   planning and set-up, including survey planning and general design which requires more detail for computer-assisted than paper for interviews;

(c)   survey interviewing, including interviewing productivity, initial training and supervisor time; and

(d)   post-interview processing, including costs of data entry and data cleaning.
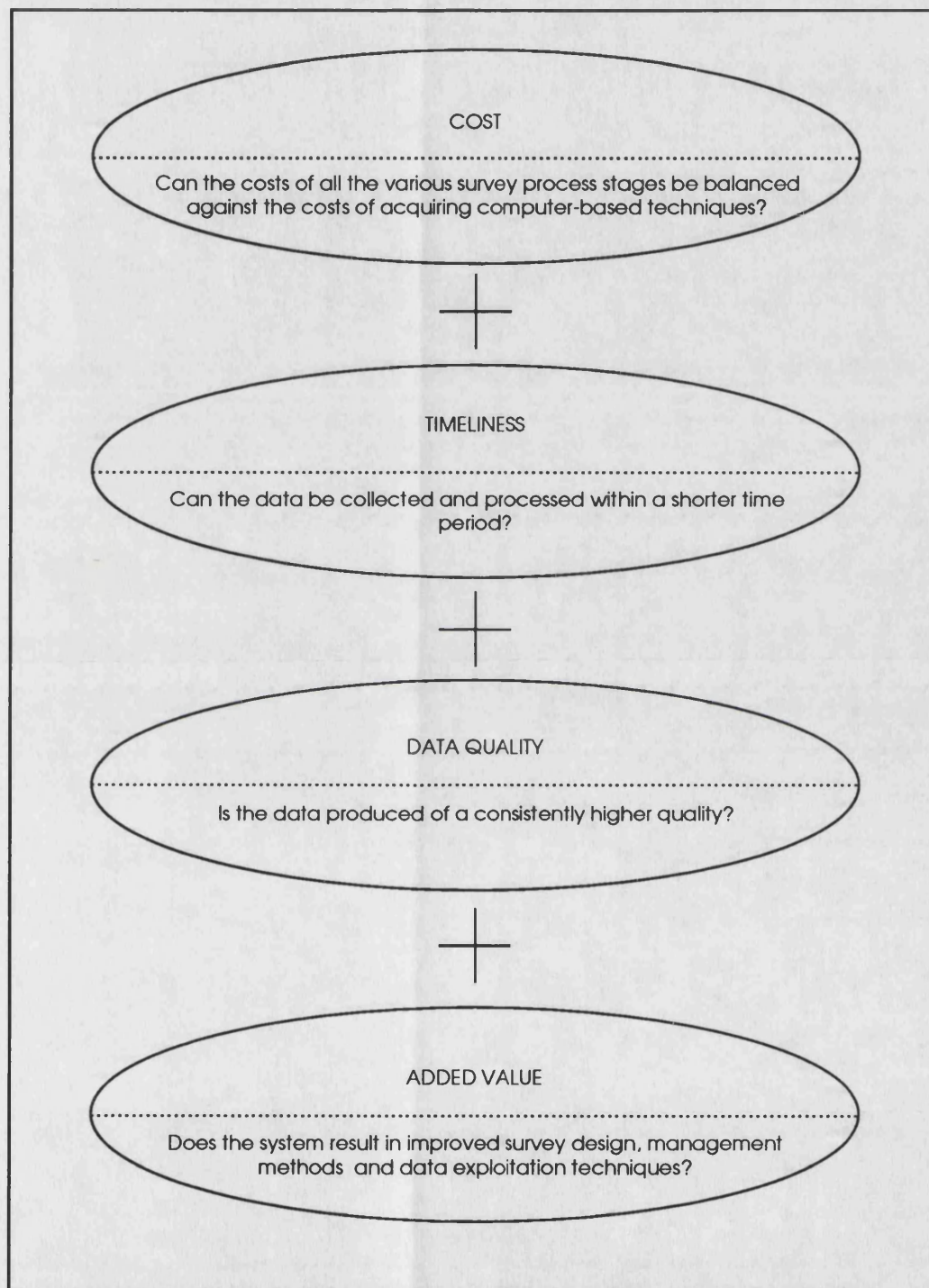
FIGURE 3.1   Evaluating computer-assisted data collection systems (based on Nicholls and Groves)

Timeliness is to be measured from the commencement of survey planning to the first tabulations of results.

Data quality is dependent on nonresponse (due to loss of sample cases and missing data items), and measurement error (caused by ambiguous question wording, incorrect recording by interviewers, or wrong answers from respondents). Nonresponse might be affected by:

(a)   respondent contact rate, if the use of the computer can increase productivity;

(b)   negatively by interviewers' and respondents' reactions to the use of the computer;

(c)   missing data items, if interviewers forget to ask certain questions.

Measurement error is affected by incorrectly recorded answers which could be introduced by the interviewer, data entry staff, editing staff or supervisors.

Added value would be the extra unsolicited benefits of using computer-assisted data collection methods. This may be in terms of improving survey management, or in providing extra information on the timing of interviews, for example.

But although these qualities would help to evaluate the cost-effectiveness of a computer-assisted data collection system, their evaluation approach ignores the social aspects, concentrating merely on the technical. Being more directed towards implementations in developed countries, they do not account for many factors that may prohibit success in developing countries as discussed in chapter one. Therefore, although the trials in European government statistical offices indicate some of the problems of implementing CADAC systems in general, they make a smaller contribution towards understanding how the same systems would work in developing countries. The trials of SIC/QUESTOR in some African countries discussed in the next section consider some of these problems in their assessments, but still fail to present a comprehensive picture of the conditions prevalent in a developing country setting.

3.1.2   SOME EVALUATORY TRIALS

The early Dutch trials aimed to test the reactions of interviewers and respondents, and evaluate whether suitable programs can be developed for flexible data collection [Bemelmans-Spork and Sikkel, 1985; van Bastelaer *et al.*, 1987]. Their experiments were on two different types of software - a customised program and a simple questionnaire specification program. The first experiment highlighted the problem of choosing what stages in the "data production" process should be computerised. Some of these tasks were identified as fairly unstructured, as in the information flow between the base and the interviewers, which could be more easily supported on paper. They also reported that the software had a slow response time and contained some software design problems. The hardware proved very limited, it had an inadequate memory, was too heavy, big and clumsy. None of the respondents objected to the computer, and most interviewers said that they would like to work with the computer if it could be improved.

For the second experiment data was collected on computer and paper, with no data cleaning so differences between the two methods could be observed. An analysis was made using the logging facility which recorded the instances when the interviewers returned to a question to correct an error [van Bastelaer *et al.*, 1987]. The results of this were used to assess the quality of the questionnaire. This analysis was based on data collected using the computerised method only, the results of the comparison between methods was not reported. Criticisms of the software were about its limited correction facility and delayed response. As in the first trial, few respondents objected to the use of the computer. The main difficulty was in finding a suitable hand-held computer with an adequate memory capacity that was shock, rain and temperature-proof. They concluded optimistically, that it was possible to implement simple surveys on inexpensive hand-held computers, although more complex surveys would require more powerful computers [Bemelmans-Spork and Sikkel, 1985]. They have since developed the more flexible BLAISE system [Denteneer *et al.*, 1987].

A British trial carried out to test the viability of using BLAISE on a complex repeated survey investigated respondent and user preferences. Experienced interviewers used the computer for either telephone or personal interviewing.

There was some difficulty in replicating the paper questionnaire on the computer, because the tabular facility was not fully developed in the software at that time. Also, questionnaire development was a lengthy process because the version of the BLAISE parser used only indicated one error at a time. Respondents were not averse to the system and some made positive comments. Interviewers were pleased to use the system after enough practice, and liked the continuous checking. During personal interviews, the interviewers sometimes had problems holding the machine as well as showing the answer cards to the respondent. Telephone interviewers thought computer interviews took longer, attributing this to inexperience in using the computer. However, personal interviewers thought them shorter [Foxon, 1988].

The Swedish statistical office has embarked on a CADAC programme and has conducted various tests on different systems [Lyberg, 1985; Lyberg and Dean, 1988]. In 1984, a test was conducted to assess interviewer reactions to using a computer for telephone interviewing. The software's response time was slow, although interviewers did not find computer interviews more time consuming. Interviewing speed decreased with computer use but this was thought desirable for improved data quality. This is an interesting result which suggests that the machine can act as a control mechanism, taking over the flow of the interview. The interviewers handled the technology fairly well, and no non-response due to the technology was detected.

In a similar test, carried out in 1985, the interviewer involved had a generally positive attitude to using the system, subject to some improvements. An ergonomic restriction of the system was discovered - the screen was sensitive to light, forcing the user to sit directly in front of it. As the options for computer-assisted interviewing seemed promising they made hardware specifications for their ideal system: a maximum weight of one kilogram, half the size of a desktop screen, battery powered, 512k of memory, a response time of less than two seconds, and a data communications capability. Unfortunately the manufacturers opted out on the grounds that the specifications were too demanding and that they could not produce, at that time, a machine light enough in weight with available technology and knowledge.

Consequently, the latest Swedish trials in 1988 used the Toshiba T1200 to test the data processing system and software. No major problems were encountered with the machine, although the weight of 5.2 kg was found to be still unacceptable and the batteries did not last as long as expected. The respondents in general were indifferent. Some interviewers were reluctant because they had reservations about the system's improvement of interviewing efficiency, the increase of physical stress, and the strict control of an interview; most, however, found working with the computer interesting.

These trials in European countries have varied in extent and scope, involving a varying number of interviewers, from one to thirty-five, and focusing on the use of complex survey designs. Trial results have generally tended to centre on qualitative assessments of respondent and user attitudes, as well as ergonomic issues. None of the trials reported here have aimed to quantitatively measure benefits, as suggested by Nicholls and Groves. Software is now available commercially for reasonably complex tasks, but still requires further refinement. The minimum hardware needed, at today's state of the art, is a laptop capable of desktop performance; most models currently available are still too heavy or too expensive.

The problem of suitable hardware was partly overcome by Reitmaier in using simpler hardware, trading performance for price. His system was used to investigate the effect of a computerised system for collecting data in improving the high error rate observed when collecting on paper. It also aimed to investigate whether the computer could produce final data sets and analysis faster than the manual system, eliminating a considerable time lag [Reitmaier, 1985]. An evaluation of this system in use was conducted in Cape Verde to collect information at maternal and child health clinics. It helped to detect data recording errors, and resulted in more standardised and accurate data [Reitmaier *et al.*, 1987]. Also, the authors claim that the nutrition maps produced led to faster and more controlled contextually adequate decisions. This can be related to Nicholls and Groves fourth component, the value added benefits of using this approach.

The aim of the French SIC project was to consider the effect of CADAC on an integrated approach to data collection, which would include survey planning and

management, questionnaire specification, training, data collection, communications, and analysis. As one of their specifications was for the use of the system in difficult environments, they have conducted various tests in francophone Africa. Experiments in Senegal [Ferry *et al.*, 1985] aimed to investigate the validity of the system at all of these levels. Three different experiments were performed and each of these was video-recorded and then investigated to identify problems in man-machine co-ordination and aspects for further refinement.

Ferry *et al.* summarised their results as encouraging, and noted no negative effects due to the machine during interviews. The hardware was durable, but consumed a lot more power than expected (apparently due to the high temperatures). It could also prove necessary to have some means of adjustable support for the computer to make it easier to use for longer periods without muscular fatigue. The keyboard was reprogrammed to make it possible to use only one hand for data entry. There were some difficulties with data transfer and communications protocols had to be changed. As the software was simple BASIC programming at this point, no facilities for returning to previous questions, question skips or error messages were incorporated. The interviewers, although inexperienced computer users, quickly became familiar with the system, and were generally pleased with the new method.

The subsequent commercialised system which emerged from this project, QUESTOR, has been tested in Guinea and Cote d'Ivoire as well as in France. A preliminary test in Guinea was to investigate whether it was viable to run QUESTOR on a Toshiba T1000 for collecting birth and death registration data [Crisan *et al.*, 1988]. In assessing the costs, their opinion was that it should be spread over 2 or 3 years, and also take into account other uses of the computer within the organisation. Preparation time and costs were reduced, it took a Guinean employee an hour to construct the computerised questionnaire. In addition, collection was faster on the computer than by writing and coding, and no time was needed for data cleaning. Their evaluation is not comprehensive and no other information is given on the working practices of interviewers, the reactions of respondents, hardware or software problems.

The Ivorian test involved collecting demographic information in both an urban centre and in a rural setting, also using Toshiba portables [Balde *et al.*, 1988]. They discovered from their evaluation that little training was needed, even for inexperienced users. The computer could last for only 3 hours per charge; though this was not a problem within the urban areas where they could use power sockets in respondents' houses. However, in rural areas, several supplementary batteries were required for a day's work. Their reliance on backup electricity in respondents' homes does not seem to be an acceptable approach in most circumstances; respondents may not have the right type of sockets, or may not be willing to offer free use of a costly resource. Several software problems remained, and the length of time for saving the questionnaire varied from 1.5 to 10 minutes depending on the type of computer model used. These saving times seem unjustifiably long, but no further explanations are given.

3.1.3   *ISSUES FOR FURTHER RESEARCH*

The trials described here show a clear progression in terms of technical capability of the systems and, as a direct result, the level of sophistication of the software. The main difficulty seems to be the weight of the portable computer hardware, but in developing countries, battery power is an additional problem. Time for basic training can be quite short, but most interviewers need more practice to boost their confidence for real survey work. Generally, interviewers liked to use the new technology, but because many systems were prototypes, their comments highlighted some software and hardware deficiencies. The majority of respondents are indifferent to the method used, some have a positive reaction, and very few negative ones are reported. Although it is clear that only certain structured tasks of the "data production" process can be computerised, the trials all seem to indicate that CADAC can now be considered a feasible option.

Nonetheless a number of questions are raised from this review of experience, and these would include such issues as ergonomics and organisational fit, which Nicholls and Groves do not mention in their review. It is important that the effect of keyboard design or screen readability on interviewer productivity, or other difficulties in handling the technology, are discussed as independent factors.

Most of the studies discussed above included qualitative observation but failed to justify the use of computer versus paper adequately in other measurable terms. These are a major shortcoming in the African studies, because if the advantages of introducing the new technology are not offset by the costs and effort involved, failure can be extremely damaging. Labour is inexpensive in developing countries, and interviewer and supervisor costs may be low enough to make better personnel management a more cost-effective means of improving data quality.

In evaluating the case study data collection system discussed in the next chapter, this thesis will examine both quantitative assessments of data quality and timeliness as well as more qualitative aspects such as working conditions and acceptability.

3.2   THE EVALUATION OF MEDICAL DECISION-AID SYSTEMS

While the type of system described above has a 'simple' rationale within the organisations promoting and developing them, medical decision-aid systems are vastly more complex. Hence, there are difficulties in the evaluation of decision-aid systems that arise in particular from the nature of the problem they set out to solve, namely to be able to reproduce the decision activity of an expert. While medical decision-aids are a larger area of academic study, there is no accepted method of how to evaluate them. The evaluations that have been reported have tended to be unstructured, subjective and of dubious value. Many suggestions for evaluation methods have been made, which include human aspects; ease of use and good interaction have often been stressed [Gaschnig *et al.*, 1983; Liebowitz, 1986; Hollnagel, 1989]. Putting these into practice is the major hurdle.

3.2.1   *ISSUES IN THE EVALUATION OF MEDICAL DECISION-AID SYSTEMS*

There are a number of problems in assessing the quality and performance of a decision-aid system [Chandrasekaran, 1983; Hollnagel, 1989; Wyatt and Spiegelhalter, 1990b; Gaschnig *et al.*, 1983]. Summarising the work of these authors, a number of issues for evaluation can be identified:

- a gold standard - choosing a standard against which the performance of the computer can be evaluated. It is usually unfair to expect the system to perform far better than the best human experts; thus some form of peer review would be preferable to comparison against absolute correctness.

- simple success versus failure evaluations based on the final answer may be insufficient as they do not allow for the possibility of acceptable levels of intermediate performance. It is sometimes the case that a correct answer may not be known because experts disagree.

- to ensure that test cases are correct and representative - must show that the cases cover a wide range of possibilities that will be met in practical use. It will be best to select a random number of cases as well as an additional set of difficult cases selected by a panel.

- controlling for various sources of experimental or evaluation bias. Some biases and other effects need to be controlled for. Biases may occur with unmatched control groups, as a result of a learning effect, developer bias, assessor bias, or circularity of test cases. Other effects that have to be controlled are: the Hawthorne effect, when the improvement in user performance is due to their awareness of being studied; checklist effect, due to the more complete and structured collection of data as a result of the trial; feedback effect, if users are told about their failures and successes.

- it is very easy to misinterpret the results of evaluation studies. Gaschnig *et al.* [1983] give an example to show how formal investigations, although statistically sound, can be misleading. Using a number of illustrative cases, they demonstrated that the success rate of the system depends on the inverse relationship between the number of symptoms accepted and the number of diseases diagnosed. Thus, the smaller the number of outcomes that the systems diagnoses, the higher its likelihood of getting the diagnoses right.

Both qualitative and quantitative techniques for analysing the results exist [O'Keefe *et al.*, 1987; Hollnagel, 1989; Liebowitz, 1986; Chandrasekaran, 1983]. Most of those mentioned can be related in some way to the basic Turing test. The basic principle of this test is that if an expert judge compares the output from a computer system and that of another human expert and cannot distinguish between them, the two can be regarded as functionally equivalent, and the

computer system can be said to exhibit "expert" behaviour. It has been used in a modified version to allow for standard criteria by several experts, making replicability possible. When performance is judged blindly on a numeric scale, statistical methods, such as paired t-tests, can be applied for analysis.

In discussing approaches to the design of full-scale evaluations, Wyatt and Spiegelhalter conclude that:

> "It is likely that, for many purposes, the most practical design for a field trial will be a multicentre before-after study, with staggered introduction and withdrawal of the decision-aid to correct for any significant trends".

Such an approach will avoid many experimental biases, having a firm initial baseline for comparison, and reducing the effect of a single, and usually unrepresentative, study.

Different aspects of a decision-aid system can be evaluated at various stages in its development. It is expected that evaluation should be a continuous process as a system evolves, and that it should be carried out on various prototypes as well as continuously on the final product. As it is easy to misinterpret the results of an evaluation, Gaschnig *et al.* [1983] have suggested that the purpose of the evaluation, the stage of development of the expert system, and what exactly is being evaluated is made clear initially.

3.2.2   TWO EVALUATION APPROACHES

Two approaches for evaluating medical decision aid systems are summarised below, both go further than mere technical measurement of performance, and include also the general impact of the system's use on the health structure (Figure 3.2).

Rossi-Mori and Ricci [1988], in discussing how medical decision-aid systems should be assessed, suggest four levels of evaluation.
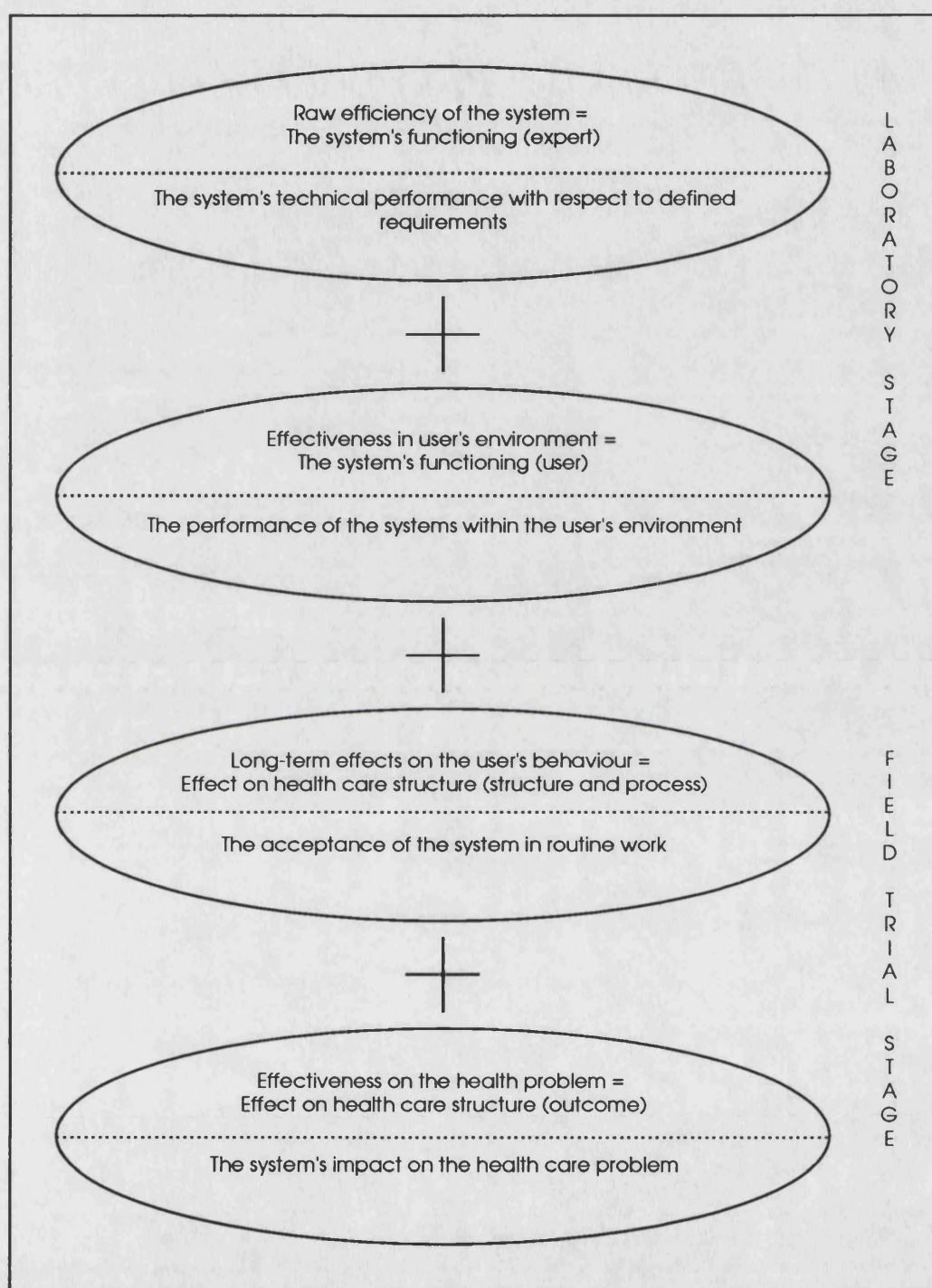
**FIGURE 3.2   Evaluating medical decision-aid systems
(based on Rossi-Mori & Ricci, and Wyatt & Spiegelhalter)**

(a)   *Raw efficiency of the system itself:* the technical performance of the system with respect to defined requirements and to a specific problem, apart from the user's general work context.

·   does the system meet its aim (for example, research, training, practice)?

·   does the system produce satisfactory performance?

·   is the knowledge base satisfactory?

(b)   *Effectiveness in the user's environment:* the performance of the system in the environment of a specific class of users, considering also the co-presence of other decision-aids.

·   are the system's suggestions helpful to the user?

·   is the reasoning process coherently presented to the user?

·   what is the routine behaviour of the user when operating the system?

(c)   *Long-term effects on the user's behaviour:* the acceptance or the rejection of the system in routine work and the features of the assimilation of the system's advice into decision-making.

·   does the medical decision-aid system continue to be used?

·   is it well integrated into the workplace?

·   has it educated the users?

·   has it produced passive agreement with the system's advice?

(d)   *Effectiveness on the health problem:* the impact on the specific health problem that the system is facing and, more in general, the global impact of knowledge based systems on the health field.

·   how much progress has been made towards the solution of the particular health problem which the system was meant to solve?

·   what does a cost-benefit analysis of the resources required (time, money, staff) with respect to the obtained benefits reveal?

·   are alternative improved means possible to achieve the same benefits?

·   are the health operators comfortable and satisfied?

·   are there any alterations to physician-patient relationships?

Wyatt and Spiegelhalter take a slightly different perspective, based on the work of Donabedian [1978]. They suggest designing appropriate questions around a

two stage evaluation process which utilises the concepts of structure, process and outcome.

(a) **Structure, process and outcome of the system's functioning** from the viewpoint of both the user and the expert. This is the first trial stage, which they view as the laboratory stage of the system's evaluation.

   User's viewpoint:

   - *Structure* - is it even worth considering?
   - *Process* - is it pleasant to use?
   - *Outcome* - does it give sensible results?

   Expert's viewpoint:

   - *Structure* - is the system of good quality?
   - *Process* - does it reason appropriately?
   - *Outcome* - are its conclusions safe and possibly valuable?

(b) **Structure, process and outcome on the health care structure.** This is the field trial stage of the evaluation.

   - *Structure* - does the system fit neatly into its intended environment and is it perceived by users as helpful?
   - *Process* - what beneficial or adverse effects does the system have on the processes of health care delivery?
   - *Outcome* - does use of the system have a measurable effect on health care outcome measures?

The major difficulty in achieving the widesweeping aims of these proposals is the lack of pertinent ways of assessing social and environmental factors. This is due not only to the lack of discussion on the matter, but primarily to an intrinsic difficulty in describing these factors in a comparative manner. In approaching the case study evaluation described in chapter 5, many of the long-term aspects of these approaches could not be tackled. Based on the investigation of some non-technical aspects, the discussion will be extended to considering cultural, social and organisational factors.

3.2.3 SOME EVALUATORY TRIALS

We will now refer back to the medical decision-aid systems for developing countries described in chapter 2, and summarise in each case, the levels of evaluation that have been carried out. The Porenta version has undergone some very informal evaluation at a prototype stage by its developers and some interested experts [Porenta *et al.*, 1988]. They list several items that they consider to be inadequate in the present version, some of which would have to be refined before the system can be tested in the field.

- There is an assumption of familiarity with medical terms, so that a smart dictionary might be required.
- There is a need for more investigation into whether a language-independent interface can be developed although current portable technology does not provide high resolution graphics.
- The method of question and answer is rigidly structured and requires more flexibility.
- Maintenance of the knowledge base is difficult, especially if more than one expert is involved.

The Uplekar system was tested in an urban dispensary on 900 cases, approximately half of which were seen by a medical assistant, the rest by a non-medical person, an engineer [Uplekar *et al.*, 1988]. All patients were seen first by a user on the computer and then seen by a general practitioner. The computer printouts and the doctors' notes were then compared. In discussing their results, Uplekar *et al.* listed a number of problems and possible revisions to be made before a larger field trial. No clinical examination or investigations are required by the system, which is based simply on presenting symptoms, producing a prescription only outcome. Although this defies medical principles, they defended their approach by maintaining that an intermediate step giving diagnoses would be of no use in this environment because therapy is what interests the patient. The knowledge base needs to be refined and the system has to eventually run on a portable machine. On the human side, patients use different expressions for the same symptom, and the system will need to be translated into local languages for use by paramedics not fluent in English.

The Byass system was tested in a laboratory-style evaluation using information collected from random consultations at an out-patient clinic [Byass and Corrah, 1989]. 500 cases were randomly selected from those used to build the database to check for internal consistency and another 500 cases were selected from the additional unused case histories. The cases were run through the system and the outcomes compared, and matched on whether the strategies proposed were in the same group. From the internal cases, the computer gave the equivalent management in 74% of cases. It could be said that the computer performed tolerably well in 88% of cases, as it incorrectly diagnosed 12% of potentially life-threatening cases. On further analysis of these life-threatening cases by another clinician, he agreed with the computer on 41% of cases, and with the out-patient clinic diagnoses on a third of cases; for the remaining cases, all three management strategies were different. From the external cases, the overall concordance level achieved, and the number of potentially life-threatening cases were slightly higher.

Byass' method of scoring "correct" and "incorrect" computer diagnoses illustrates the point made in the Gaschnig *et al.* paper that was discussed earlier. The system has 27 management strategies, but for scoring correct and incorrect diagnoses, they were grouped into 8 categories. If the computer's outcome lay in the same group as that recorded by the paramedical officer, it was regarded as a correct answer. The difficulty in this approach is that by reducing the categories from 27 to 8, a bias towards higher system accuracy is introduced. It remains to be shown how much better his system's methodology is than a simpler, randomly weighted method which allocates a patient to one of the eight groups depending on the number of management strategies present in that group. Thus the real benefit of the statistical model has to be shown to be significantly higher than a model operating merely on two simple rules of chance.

In summarising their difficulties, the Byass paper mentions that evaluating decision-aid systems is difficult because expert opinion differs and little comparative material is available from other implementations and evaluations. They claim the approach used in their system has been deliberately kept simple, because although more complicated AI procedures might improve performance, they will also undoubtedly increase complexity and the time needed for a consultation. In their analysis of the evaluation methodology, they mention that because this

evaluation was retrospective, other factors that may have influenced the diagnosis when the patient was seen were possibly unrecorded. Also, some of the discrepant cases were unusual, and probably included miscoded data. The system's specification of a threshold (a probability level at which a strategy can be accepted) has led to over-diagnosis of some strategies.

No formal criteria were used for evaluating the first version of the Auvert system, which was meant to gauge acceptability [Auvert *et al.*, 1986a]. The 50 doctors and paramedical workers who used the system in Chad found it easy to use, fast and beneficial. The hardware proved to be resistant and the batteries were powered by solar means with no difficulty. They concluded that a more complete evaluation was necessary to test performance, utility and integration into the health care system. They mentioned a number of remaining problems. The database containing the medical knowledge was satisfactory, but the diagnostic method had problems linked to the method of knowledge representation. They found the flowcharts and their representation method to be:

· too rigidly structured to take into account seasonal or geographic variability on disease frequency. Also, it could not avoid asking unhelpful questions;

· unsympathetic, giving diagnoses that could not be treated. This was because the diagnostic module was separate from treatments;

· simple, but does not correspond to the medical decision-making process - which is to formulate hypotheses based on a few signs, which are gradually reinforced or finally rejected;

· unable to offer an explanation because each disease point appears to be independent of the others, with no underlying disease model;

· not compact, knowledge duplication was common.

They concluded that making the system more powerful will require artificial intelligence (AI) techniques and increased complexity, leading to increased memory and time requirements. Currently, hardware remains comparatively costly for this type of environment, although price is falling while performance increases. In developing the second version of TROPICAID, these discoveries and considerations were taken into account.

3.2.4   RESEARCH ISSUES

No full-scale field study investigating in depth the problems of implementing medical decision-aid systems has been widely reported. Of the evaluations attempted of the four systems described here, Byass' is the most thorough. Uplekar *et al.* leave out many details of how the evaluation was carried out, and therefore offer no solid basis for comparison. Again, the 25 drugs that the system can prescribe are grouped into generic types, unveiling yet another source of misinterpretation. The Auvert and Porenta versions were largely informal. The conclusions are based on subjective assessments of the systems' knowledge base and inferencing capability. For TROPICAID, some unspecified, and probably unstandardised, method of gauging user acceptability is applied.

The above summary shows that little value has been placed on thorough evaluation to demonstrate the utility of systems. These evaluations have mainly dealt with the systems' performance, although Byass is in the process of conducting additional evaluation [personal communication]. The utility of the system should also include an assessment of the value attached to it by the users, and consider other implementation factors, for example, will the system be sustainable in the long-term?

This review of evaluatory approaches also reveals a lack of methods and techniques. Many techniques suggested in the evaluation literature may require complex and unnecessary procedures for the developing country environment. Where possible, the methods should be adapted, to propose simpler approaches that generate reliable data and provide a framework for accurate interpretation of the results. Byass' methodology is a start in the right direction.

In the general medical decision-aid literature, only the early stages of the simple evaluation frameworks suggested by Rossi-Mori and Ricci, and Wyatt and Spiegelhalter have been attempted. This indicates how far behind evaluatory research is in answering questions on the potential of medical decision-aid systems in clinical practice.

Perhaps the most important fact of all is that no system has yet been shown to significantly improve a clinician's diagnosis solely as a result of its diagnostic ability. Until this basic issue of making systems practicable is addressed, in no way can the current evaluation trials be considered adequate.

3.3 THE CASE FOR EVALUATION

In building systems, the developer not only has to build the right system, but also needs to know that it is right. This is vital in developing countries where there is both a greater risk of resource wastage if the system implementation is unsuccessful, and a greater opportunity cost. Evaluation is an important and useful stage of system development because it is a means of judging the system's worth or value. It can also serve as a learning process for future developments in similar environments.

In the attempt to assess whether performance is satisfactory, weaknesses in the system requiring further development may also be identified. Underlying most of the tools and methods of evaluation is the idea of measurement, and therefore the need to define what is to be measured [Hirschheim and Smithson, 1988]. Authors disagree on the role of evaluation, and some have used various alternative terms to distinguish between different types of evaluation. For example, the term *'verification'* has been used to mean "doing the right thing", against that of *'validation'*, to mean "doing it right" [O'Keefe *et al.*, 1987]. For most authors, evaluation has meant the exclusion of all but the technical aspects of a software system, but it has also at times been used as an umbrella term covering social, technical, organisational and other aspects that might affect the implementation of an information system.

Hirschheim and Smithson classify a continuum of approaches to the evaluation of information systems: ranging from what they call objective/rational to subjective/ political approaches. The zones in the continuum are as follows:

(a)   *efficiency* which they define as "performing a particular task well in relation to given criteria" includes hardware and software performance,

correctness of programs, and software reliability; all these are basic concepts in software engineering.

(b) *effectiveness* "deciding what tasks should be done", is more difficult to evaluate. One reason for this is the lack, in most cases, of an adequate initial definition of success or failure; another reason is that assessment at this level is often a subjective task, dependent on individual judgements of such things as the value of new information resources. Evaluations of this aspect have therefore tended to measure what is easy to measure rather than what is most important.

(c) *understanding* is an appreciation of the functions and nature of evaluation, and the limitations and problems in the process of evaluation. This arises because of a need to view evaluation as largely a social activity which evokes political and social issues.

Obviously, in terms of this continuum, the methods for conducting such assessments vary from informal (understanding) to formal (efficiency). The understanding aspect of the above classification is, as a result, more difficult to measure quantitatively. Effectiveness can be measured in a semi-formal fashion, and efficiency can be measured quite formally. These last two aspects are used and discussed often in the literature, but sometimes in such a way as to demonstrate a lack of appreciation of the true role of evaluation [Hirschheim and Smithson, 1988]. Although Hirschheim and Smithson seek to point to the need to redress the quantitative bias of the evaluation process, for the purposes of this thesis, it is more useful to consider understanding with relation to a *system*. Thus, reinterpreting Hirschheim and Smithson, evaluation should involve an appreciation of the functions and nature of a *system*, and the limitations and problems in the process of the *system*'s assimilation into an organisation. In addition, the political and social issues arising from the *system*'s use must also be investigated.

The WHO suggest that, from the health perspective, the purpose of evaluation is to improve services and guide planning, and suggest a process that includes the following components [World Health Organisation, 1981b].

(a) verify *relevance* or rationale for adopting the system within the general health policy;
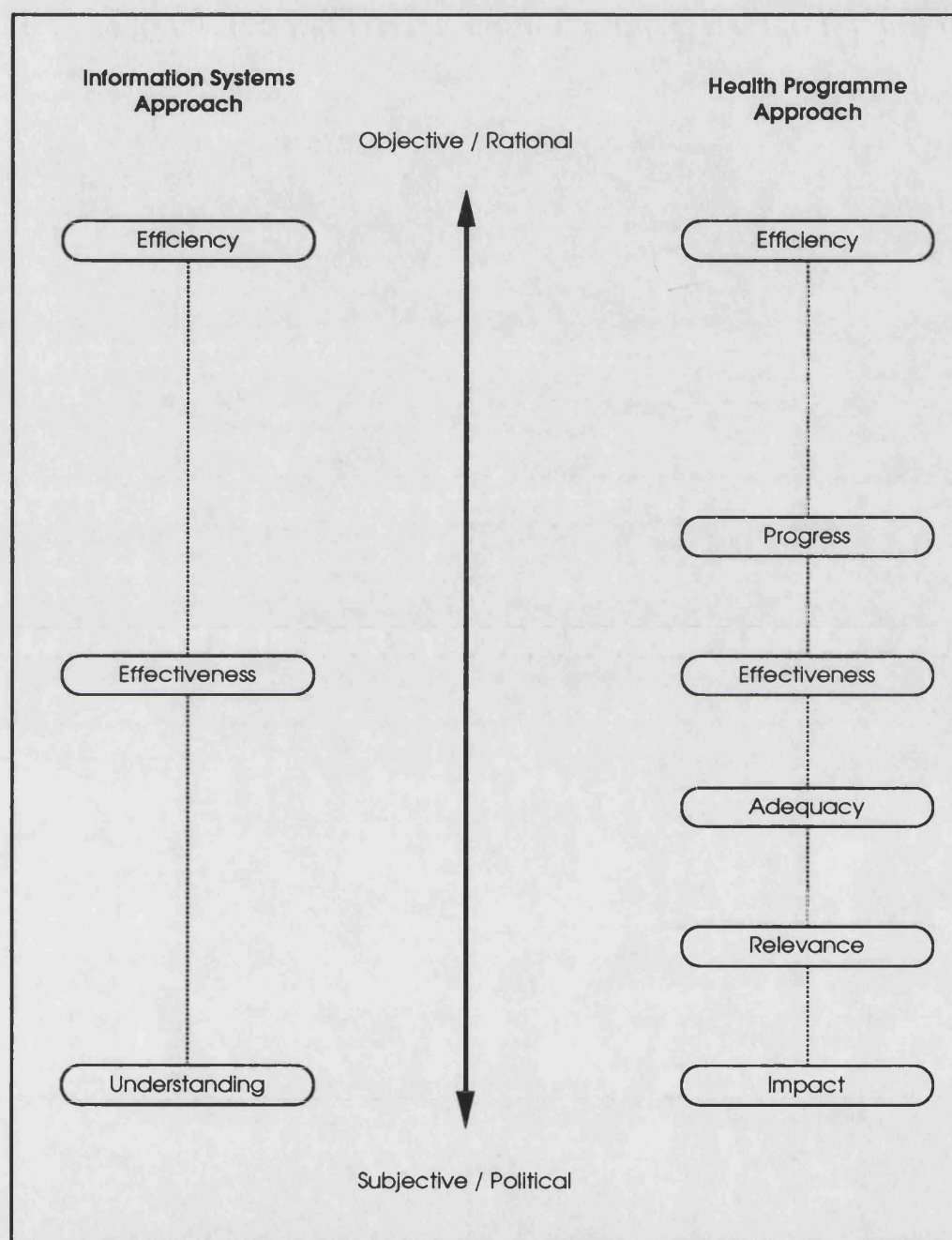
FIGURE 3.3 : Approaches to evaluation

(b)   assess *adequacy*, whether sufficient attention has been paid to all the details involved in implementation;

(c)   review whether *progress* is made by comparing actual activities with scheduled activities;

(d)   assess *efficiency*, whether the results obtained have been maximised in relation to resources used;

(e)   assess *effectiveness* at improving an unsatisfactory health situation, where feasible this should be quantified;

(f)   assess *impact*, the overall effect of the programme on health and related socio-economic development.

The WHO use some of the same terms as Hirschheim and Smithson, but order the components differently. In a sense, their approach embodies the understanding concept, emphasising the need to specify what is being evaluated, and then to assess the impact of a programme on the overall level of health. However, their main suggestion for aiding the evaluation process is in the use of indicators and criteria to measure change, thus implying a quantitative appraisal. A mapping of the WHO criteria to the Hirschheim and Smithson continuum might result in the spectrum shown in Figure 3.3.

Health information systems should perhaps therefore be evaluated to include both the information systems perspective and the health one. One general approach to the evaluation of a health information system has been proposed by Donabedian [1978] and was later used by Vallbona [1983]. The framework suggests that the impact of the information system on the health care system must be evaluated from three angles: the *structure*, the *process*, and the *outcome*. It must also be assessed from different stakeholder viewpoints: that of the *planners*, *providers* and *consumers*. The next section develops the Donabedian approach further, to be used as a framework for presenting the results of the two evaluation case studies in the next two chapters.

	System's functioning	Human perspectives	Health provision
Structure	Technical detail		
Process		Organisational setting and social interaction	
Outcome			Health status of target population
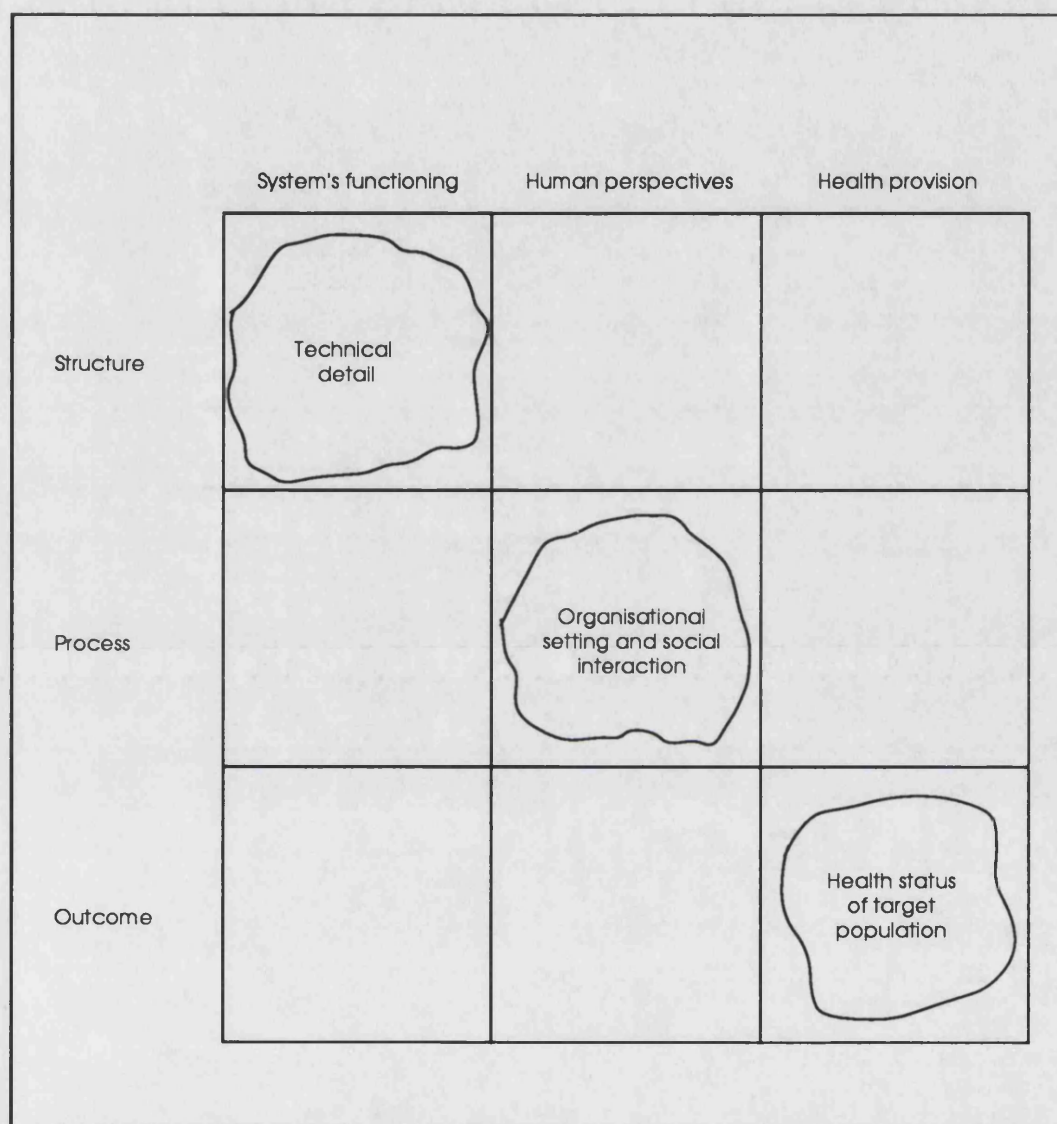
FIGURE 3.4 : Simple evaluation framework - the basic diagonal

3.4 A PROPOSED GENERALISED EVALUATION APPROACH

Based on the work described in this chapter and in particular that of Donabedian [1978], Vallbona [1983], and Wyatt and Spiegelhalter [1989], the following structure has been developed for evaluating the efficiency, utility and overall impact of a health information system. The justification for the approach described here can be made from two main angles. Firstly, it provides a procedure that assesses the technical and the social aspects of a system, as well as the long-term impact on improving health care. Secondly, it is a standardised procedure for reporting evaluation results that can be applied to both case studies. It incorporates all the points mentioned in Rossi-Mori and Ricci's outline for the evaluation of medical decision-aid systems. The Hirschheim and Smithson approach to information systems evaluation, and the WHO approach to health programme evaluation can also be accommodated within it.

The proposed evaluation approach is based initially on the three concepts of structure, process and outcome (SPO). These need careful definition:

(a)   The *structure* of a system is the manner in which it is constructed, or the whole of its essential parts. Put another way, the structure is 'the fixed and designed requirements'.

(b)   The *process* it performs is a series of operations by which a task is accomplished, and information transformed. This can be described more simply as 'the way things are done'.

(c)   The *outcome* of the system is the impact or visible effect. This can be rewritten as the 'general result of the system in operation'.

Conventional IT evaluation has concentrated on structure, whereas health evaluation has focused on outcome. The process is then the link between the two. To restate the flow between the three components, one wants to evaluate how the design of the structure has created or modified a process and achieved a set of outcomes.

This three-part analysis is then applied at three main levels: that of the system's functioning, human perspectives, and the overall impact on the health care system. One can observe a natural diagonal between the SPO components and these levels (Figure 3.4). System's functioning falls most directly under structure,
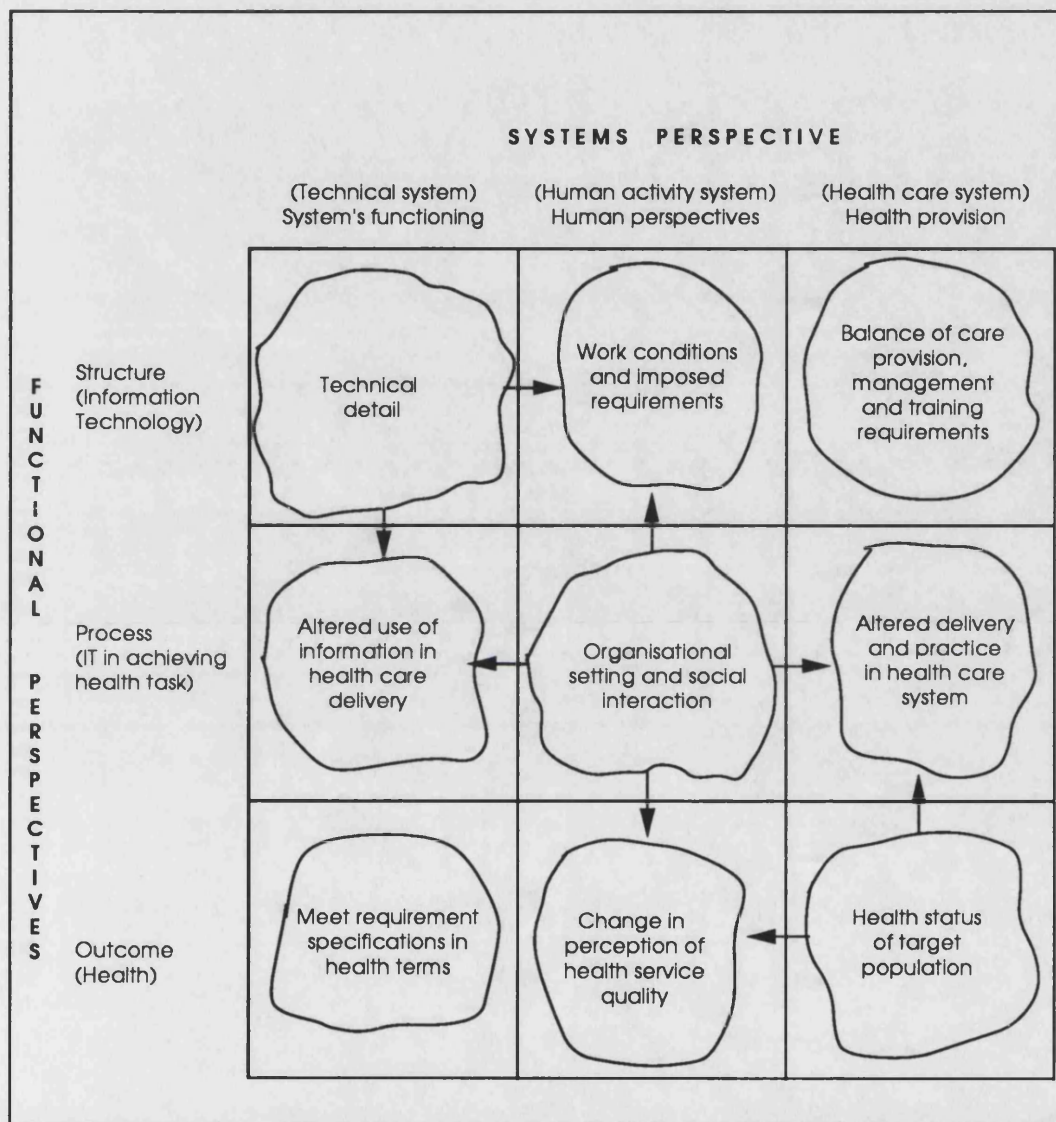
FIGURE 3.5   Full evaluation framework - the broader view

human perspectives under process, and the health care system under outcome. Developing this model more fully, and extending it into a full matrix, the approach used here sets out to capture a broader and more comprehensive view of the effects of a system (Figure 3.5). For example, when the structure of a system is considered from its overall impact on the health system, one can appreciate the opportunities and problems that arise in implementing IT in developing countries.

Considering each of these matrix items (levels and components) in turn:

1.   *The system's functioning.* This has been referred to as the raw efficiency of the system itself by Rossi-Mori and Ricci. Broken down into the evaluation procedure, the following aspects may be considered:

(a)   *Structure* - what are the hardware requirements and is the software structure understandable? Does the full set of system components work together in a technical sense?

(b)   *Process* - is the method by which the system transforms its data, the information processing, correct and valid?

(c)   *Outcome* - are the results relevant, applicable and reliable? Does it meet the requirement specifications?

2.   *Human perspectives.* This includes the acceptability of the system by the various stakeholders: the user, the recipient, and the administrator, and considers how the system's functions affect them. Rossi-Morri and Ricci include human perspectives in their outline, but only from that of the immediate system's user. This framework recognises three roles under human perspectives, which are sufficient for the case studies reported in the next two chapters, and help to preserve conformity. In other situations, there may be more roles worthy of consideration, in which case, the number of stakeholders has to be increased. Assessing human perspectives of information systems is not easy and these aspects are not easily measurable. Researchers must allow themselves both the freedom to identify sufficient stakeholders and the freedom of using qualitative judgements in their analyses when quantitative measures cannot be obtained.

The three stakeholders identified within this framework are explained here.

The *User* is the primary agent in the system implementation, who is indispensable for its proper functioning. Within the SPO dimension, this poses questions such as:

(a)   *Structure* - what are the changes to working conditions, in terms of the physical environment, skill requirements and so on?

(b)   *Process* - how is the user's mode of operation changed? Are these changes seen as desirable to the user as an individual, and in general to the user's organisational role?

(c)   *Outcome* - is the overall effectiveness of the user within the health care system enhanced?

The *Recipient* is the person who the system is expected to benefit, and who is often directly or indirectly affected by its implementation. This person will, in general, also be the source of the information processed. In a classic medical informatics application, this would strictly be the patient. Within a health informatics application, this may be more generally the population served.

(a)   *Structure* - are recipients required to modify their behaviour in any way?

(b)   *Process* - how is the recipient's experience of health care altered at the point of contact with the system?

(c)   *Outcome* - does the use of the system result in changes in the quality of service and better health for the recipient?

The *Administrator* is the planner responsible for the general management of the health unit. Note that, since this is under the human perspective heading, assessment at this level is focused on the management of the individual health unit rather than the whole health system. Thus, it is limited to the administrator's immediate sphere of influence.

(a)   *Structure* - is the system a reasonable, cost-effective and efficient alternative to existing structures?

(b)   *Process* - does the system imply change in the health care delivery activities for which the administrator is responsible? Does it change the character of the administrator's job?

(c)   *Outcome* - does the system improve specific health provision on a reasonable metric?

3.   ***Health care system.*** This involves a consideration of the impact of a system's use on the health care system as a whole, and on health itself. It concerns the national developmental level in its widest possible sense. In Rossi-Mori and Ricci's outline, this is the effectiveness at the "health problem" level of assessment.

(a)   *Structure* - does it change the balance between the functions of the different health care providers?

(b)   *Process* - does it affect practice and delivered quality of health provision?

(c)   *Outcome* - does it improve the health status and development potential of the population it serves?

Because of the constraints in terms of resources available to pursue the studies in this thesis, both of the evaluations reported in the following chapters had to be carried out within a short period of time. The field tests can contribute most in terms of the evaluation of the systems' functioning, the perspectives of users and recipients. The perspective of administrators, and the impact of the system's use on the health care system were not measured directly and the expected impact on these aspects are presented, based on discussion with relevant parties.

In summary, the evaluation framework proposed here permits a structured view of health informatics projects which recognises both the need to link an information technology grounded perspective with one that includes an understanding of the concept of health. It will be used as the vehicle for reporting a standardised summary of the two case studies, enabling a more structured comparison of the utility of the two systems.

CHAPTER 4

THE FIELD DATA COLLECTION SYSTEM:
DEVELOPMENT AND EVALUATION

4.1   PRELIMINARIES

All well-designed health information systems will require reliable data. This, however, remains a problem that is particularly acute in developing countries. The Field Data Collection System (FDCS) was designed to meet this need, providing a means of obtaining high quality data with reduced lead time to analysis.

This chapter describes the design of FDCS, which can be used to enter questionnaire data in a computer-readable form at the point of collection. In this way, the data is immediately checked for errors as it is punched, and is later downloaded to a micro-computer ready for analysis. The system allows for two levels of users: the field manager who designs the questionnaire and runs the survey, and the field worker who administers the questionnaire and collects the data. Interviewer access is limited to the outermost level by password control. The system can be used for a variety of purposes: ranging from continuous monitoring to quick, one-off surveys.

This chapter also gives some background to the case study, describing the country (The Gambia), and the organisational context (Medical Research Council Laboratories) within which the evaluation was undertaken. It then presents the results of the experimental trial. This includes a description of the system's functioning, the various users' reactions, and the respondents' (interviewees) attitudes. The results are further summarised under the structure, process, and outcome evaluatory framework introduced in the previous chapter.

4.2   THE DESIGN OF THE FIELD DATA COLLECTION SYSTEM

FDCS incorporates many basic CADAC facilities, the major limitations on its functionality being due to the hardware chosen. In choosing hardware, there are trade-offs between price, performance and ruggedness as well as between input/output devices, storage media and weight. The Psion Organiser II XP handheld computer was chosen because of its cost, portability, data storage and transfer capabilities. It is only slightly larger than a standard sized calculator and weighs 250g without battery. It has a full set of alphabetic and numeric keys, as well as cursor and function keys, protected by a hard sliding case, making it less susceptible to dust when not in use. It operates on alkaline or rechargeable batteries, and has an optional mains adapter. With a storage capacity of up to 128K per storage device, a data transfer capability and an inbuilt procedure-based BASIC-like programming language, it offers a range of possible applications and functions. A basic machine costs approximately £140, but storage devices are relatively more costly, starting from £35 (January, 1989 prices).

The Psion has certain limitations, however. Its screen is small with a limited display of 2 lines of 16 characters each; the keys are also quite small. Its hardware specific language may be regarded as another disadvantage; most data collection systems use more generic languages such as Pascal. Since the beginning of this project, additional products have been offered that greatly enhance the Psion's potential. Psion printers are now available that use thermal paper and are capable of printing 80 characters per line at a rate of one line per second. In April 1989, a new Psion series was released with a screen twice the original size and double the amount of internal memory. Prices of all models have fallen. On balance, the Psion Organiser offered a compact yet flexible basis for developing the FDCS software.

The FDCS software is menu-based and divided into three levels - an outer level provides the functions needed to conduct an interview, leading via a user-specified password to the inner menu which contains functions for questionnaire specification. Beyond this, there is a supplementary level providing additional and optional features [Appendix A]. This design allows two levels of password protection to be set by the survey manager, one at the outer level for interviewers

to use the data entry functions, and another at the inner level to be reserved for questionnaire definition by the manager.

The inner level menu can be used as an illustration of the menu-based approach. A series of options are displayed:

<div align="center">

ADD/EDIT/REMOVE

DIR/OPTIONS/QUIT

</div>

which allow the user to define or append to a questionnaire file (ADD), to change a file (EDIT), to delete a file (REMOVE), obtain a list of files (DIR), access the supplementary level of functions (OPTIONS) or leave this menu level (QUIT).

1.   *Questionnaire specification*

In FDCS a questionnaire is stored under a file name, up to 8 characters long, as specified by the survey manager. The length of a questionnaire is limited by the size of the storage device and the expected size of the data file. Questionnaire specification involves three basic components for each question: the question text, the answer type, and the explanation. These can be enhanced by specifying the routing (the next question depending on the answer), and consistency checks between questions.

- The question text is limited to a maximum of 16 characters so it can be comfortably displayed back on the screen without any scrolling.
- Answers can be of various types - date (with range limits that can be defined in years, for a particular month of a certain year, or by using the Psion's internal date clock), number (integer ranges, decimal ranges of up to 8 decimal places, or a specified number of digits), string (specifying a string length of up to 20 characters) or coded (allowing single characters to represent different options).
- The explanation facility allows a longer text string with a maximum of 100 characters that can be used as a help option, for example, to specify a longer version of the short question text or to explain one-character codes. The explanation is horizontally scrolled onto the screen and the display can be controlled by using the cursor keys to stop or start the scrolling.
- The supplementary level menu is accessed by choosing OPTIONS. Skip conditions for routing and consistency checks can be specified here, as well

as passwords, automatic stamping of records with password identification, date and time.

2. *Questionnaire administration*

To administer a questionnaire, the interviewer chooses the run option from the outer level menu. This level also contains other options to allow the interviewer to choose between different questionnaires, to view the contents of a data file, and to tally data entries within specified variable ranges in a data file.

·   When the FDCS option is chosen, the interviewers may need to identify themselves using one of the passwords, if this has been preset by the survey manager.

·   When the run option is chosen, the user is told how much space remains for storing data for the currently loaded questionnaire, in terms of complete records, before the interview starts.

·   Questions are displayed sequentially (unless a skip condition has been defined) on the first line of the screen and a range or list of possible answers as a prompt on the second line.

·   Input is checked and appropriate error messages generated if it is invalid. The use of an ESCAPE menu during input offers options of returning to previous questions, viewing the question explanation, or stopping the interview.

·   At the end of each interview, the user is given an option of viewing or editing the record contents before saving it. If answers to skip conditions are changed, the user is prompted for a confirmation; if this is given, all subsequent answers are deleted and the interview is resumed from that question.

3. *Summary*

As described above, the Psion is in a sense a limited machine. Some of its disadvantages are overcome by the software, for example, messages longer than the width of the screen can be scrolled. For program development, the use of a microcomputer and cross-compilation means that complex programming tasks can be undertaken in a better environment, with debugging facilities and support for

software transfer. Overall, its advantages far outweigh the disadvantages. In terms of robustness, compactness, weight and cost, it is ideal for arduous developing country environments. The software offers an ambitious and flexible option for survey management by allowing two levels of use and a comprehensive range of functions using a menu-driven approach.

4.3   BACKGROUND TO THE CASE STUDY

The FDCS system described above was developed in particular for use in health surveys in developing countries. It was evaluated on a child malaria morbidity study in The Gambia. It is therefore appropriate, before getting into the details of the trial, to describe briefly the country and organisational context. The statistical data presented is drawn from a number of sources [Graham, 1986; Hill, 1987; Gowers, 1987; The World Bank, 1981], but chiefly from the World Development Report [The World Bank, 1989b].

The Republic of The Gambia is located on the west coast of Africa, and forms a 11,000 sq.km. enclave into Senegal. The country is dominated by the River Gambia which runs throughout its length. The land adjacent to the river is marshy, changing to savannah type grassland as it slopes away from the river. The climate is semi-arid, being in the Sahel region immediately south of the Sahara Desert. Most of the annual rainfall of 100cm falls between June and October.

The Gambia gained its independence from the British in 1965, adopting a parliamentary system of government. It has been governed by one President, Sir Dawda Jawara, since independence. Apart from a coup d'etat attempt in 1981, the country has been politically stable, being a multi-party democracy. Administratively, the country is partitioned into 5 divisions, further split into districts covering some 1000 villages and urban councils.

The majority of the estimated population of 800,000 is rural (63%), but there is increasing migration towards the urban areas at a rate of 8.5% annually. The annual population growth rate is 3% and, characteristically of developing coun-

tries, most of the population (45%) is under 15 years of age. Three-quarters of all children are enroled in primary school, but this is heavily biased towards the urban areas, where enrolment is as high as 90% in some districts; in the rural areas, enrolment figures rarely exceed 20%. Only one in five children attend secondary school.

The economy is strongly agrarian, groundnuts and groundnut oil provided two-thirds of the country's earnings in 1987. An increasing amount of foreign exchange is now generated by a growing tourist industry, a sector actively promoted by the government. The country is still not self-sufficient in food production, and the value of food imports constitutes 40% of all merchandise imports. However, exports are currently increasing annually at a rate of 12.5%, and imports are slowly decreasing. GNP per capita has decreased in the past decade and was $220 in 1987, but is making a modest recovery, now increasing at under 1% a year. Total long-term external debt runs at $273 million, one and a half times the GNP. The amount of aid has been steadily increasing, at $103 million, the country receives one of the highest per capita levels of aid in Africa ($129).

In The Gambia, life expectancy at birth is 43 years, but this increases to over 50 years after the age of five. Infant mortality has dropped by a quarter over the 25 years since independence, but is still high at 145 per 1000 live births. Malaria, diarrhoea, measles and respiratory infections are the highest causes of infant morbidity and mortality. The country has come a long way to achieving its goal of 80% immunisation coverage, as a means of reducing the high childhood risk of contracting infectious diseases. There is one physician to 11,690 people; although the proportion per nursing person is higher, there are no recent figures. Of total government expenditure, 7% was on health in 1980.

In 1981, a phased programme for primary health care began, with aims of national coverage by 1986. The programme operates in a pyramidical structure. At the base are health posts staffed by village health workers and traditional birth attendants in each village with over 400 inhabitants. At the next level, a group of villages is supervised by a community health nurse, who resides in a key primary health care village, and runs a sub-dispensary. Dispensaries operated by

nurse/dispensers are at the next level up, providing basic care, maternal and child health services, and vaccinations. The main institution in rural areas is the health centre staffed by qualified medical personnel. Two general hospitals operate at the apex as major referral centres, also providing outpatient facilities. The government-run service is supplemented by several specialist units run by non-governmental organisations [Graham, 1986].

At the end of the Second World War, the field base of a British military hospital was converted into a field laboratory of the British Medical Research Council. The laboratories operate on a block grant from the British home organisation, supplemented by grants from international agencies, to conduct research into tropical diseases. Clinical, laboratory-based and epidemiological work is done on major tropical diseases such as malaria, schistosomiasis, hepatitis, and AIDS. Additional research on nutrition is conducted by an affiliated unit. The MRC operates several field stations, as well as the base in Fajara.

As a research oriented institution with non-governmental funding from abroad, it is not representative of the type of health organisations found within developing countries. The difference will generally be noticeable in terms of managerial style and resource availability. However, these differences were less significant for this case study because both the types of resources needed from the MRC - staff, and computing facilities - are approximately equivalent to what is available within the Gambian Ministry of Public Health. The MRC had already experimented with various ways of improving data quality, and was therefore a suitable environment within which to undertake the field trial, providing an opportunity to evaluate the utility of FDCS with a structured approach.

4.4   THE FDCS FIELD TRIAL DESIGN

The aim of the study was to evaluate the advantages of computerised field data collection against the standard paper and pencil method currently used. Data was collected using both methods and used to compare error rates, interview lengths and questionnaire completion rates. A logging facility, to trace the sequence of questions taken during interviews, was added onto FDCS during the trial. Also

investigated were the training requirements, the views of the interviewers on the software and the technology, the reactions of respondents to the technology, the impact of the technology on working conditions, and the environmental tolerance of the hardware.

1.   Project design

The trial was over a period of 7 weeks (in January to March 1989). This period was divided into one week of training and 6 weeks of data collection on a malaria morbidity study. Four of the 8 interviewers involved were working on the study before the field trial, and the others were employed for the duration of the project only.

All the interviewers were resident within the study area - Bakau - and the experienced interviewers were already allocated to one of four areas, each covering approximately 125 children. The inexperienced interviewers were paired with the experienced ones depending on the area they were most familiar with, and therefore more able to locate the residences of the study children. In each pair one interviewer was allocated a Psion and the other used paper for data collection each week, alternating between the methods each week (Table 4.1).

	Pair 1		Pair 2		Pair 3		Pair 4	
Interviewer	1 NEW	2 OLD	3 OLD	4 NEW	5 NEW	6 OLD	7 OLD	8 NEW
Week 1	C	P	C	P	C	P	C	P
Week 2	P	C	P	C	P	C	P	C
Week 3	C	P	C	P	C	P	C	P
Week 4	P	C	P	C	P	C	P	C
Week 5	C	P	C	P	C	P	C	P
Week 6	P	C	P	C	P	C	P	C

LEGEND			
Methods:	C = Computer P = Paper	Pairs:	OLD = Experienced interviewers NEW = Inexperienced interviewers

TABLE 4.1 : Survey Design

Each interviewer was given a list of children to see each week, updated to exclude the children who had left the study and to include those who had joined. Data collected on the Psion was transferred to a Toshiba T1200 portable microcomputer

whenever the interviewers returned to base. Data collected on paper was punched, and cleaned by the computer centre staff as is normally done. When complete, a copy of the paper data set was stored on the Toshiba. Both data sets were stored in a database and subjected to a rigorous checking program to locate residual errors. All interviewers were given lists of the errors in the data they returned, and asked to correct them if possible. To evaluate acceptability by both interviewers and respondents, the interviewers were observed at work, and asked to complete weekly questionnaires to monitor their opinions on both methods; selected respondents were questioned on their reactions.

2.   *Questionnaire design and specification*

The morbidity questionnaire around which the trial was based is regularly used to interview the mothers or guardians of the study children and record incidences of illness, so that malarial episodes can be isolated. The questions ask for the child's symptoms over the past week, and in particular, for any symptoms of fever and details of treatment given. If the child has a high temperature at the time of interview, a blood film is taken and later checked for malarial parasites.

The original questionnaire was amended to allow testing of a range of FDCS features. It was expected, based on prior experience, that most errors would be introduced in the first five questionnaire items (for child identification). The MRC has devised a system using sticky labels pre-printed with the child's details as form identifiers, to reduce transcription errors introduced when the details are copied from pre-printed lists in the field. To test whether the use of the computer in the field might have an effect on these identification errors, the printed list method was re-introduced for all interviewers. It would have been possible to use a database with FDCS, but this was not done, as it was felt better for the comparison to be based on the more common way of providing interviewers with a respondent list.

The amendments to the pre-existing questionnaire were as follows. Start and end times were added to the original forms. More numeric items were added as only one was included in the questionnaire. Two measurements - mid upper arm circumference and mid thigh circumference were introduced, to give some

BAKAU MALARIA STUDY

TIME OF STARTING INTERVIEW   *1* :__:__: hr :__:__: min

NAME   *2* :__:__:__:__:__: *3* :__:__:__:__:__:

SURVEY NUMBER   *4* :__:__:__:__:__:

COMPOUND NUMBER   *5* :__:__:__:__:__:

WEEK NUMBER   *6* :__:__:

Is the child: (1) Staying in the compound   (2) MAF
          (3) Dead   (4) Visiting - social
          (5) Visiting - medical   (6) Admitted   *7* :__:
          (7) Refusal
          If visiting, specify where ....   *8* :__:

Is the child well today? (Y/N)   *9* :__:

If unwell, which symptoms does the child have:
(record no. of days during past week if Y)   (Y/N)   DAYS

Fever / hot body.. *10* :__: *11* :__:
Headache.. *12* :__: *13* :__:
Vomiting... *14* :__: *15* :__:
Diarrhoea.. *16* :__: *17* :__:
Chest pain / difficulty breathing.. *18* :__: *19* :__:
Bad cough.. *20* :__: *21* :__:
Fresh cold... *22* :__: *23* :__:
Other.. *24* :__:
       25 If yes, specify ....

Has the child had a fever in the past week?   (Y/N) *26* :__:
       Was a blood film taken?   (Y/N) *27* :__:
       Result?   Neg N, Pos Y   *28* :__:

WERE YOU ABLE TO CONFIRM THE BLOOD FILM   (Y/N) *29* :__:
RESULT BY CHECKING THE HEALTH CARD?

Has the child received chloroquine in the past week?   (Y/N) *30* :__:
If YES, from whom? (1) MRC (2) BHC (3) RVH
(4) Other - (specify) *32* ....   *31* :__:
       WERE YOU ABLE TO CONFIRM THIS BY SEEING
       THE TABLETS OR CHECKING THE HEALTH CARD   (Y/N) *33* :__:

Did the child sleep under a bed-net last week?   (Y/N) *34* :__:

RECORD THE CHILD'S AXILLARY TEMPERATURE   *35* :__:__: . :__:
If temperature is above 37.5˚C, please take blood film.

       DID YOU TAKE A BLOOD FILM?   (Y/N) *36* :__:
       IF Y, BRADY NO   *37* :__:__:

Middle upper arm circumference (cms)   *38* :__:__: . :__:
Middle thigh circumference (cms)   *39* :__:__: . :__:
Health card number   *40* :__:__: / :__:__:__: / :__:__:
Birth weight (kgs)   *41* :__: . :__:
Date of birth (dd/mm/yy)   *42* :__:__: / :__:__: / :__:__:

             DATE   *43* :__:__:__:__:__:

             FW   *44* :__:

       TIME AT END OF INTERVIEW *45* :__:__: hr :__:__: min

FIGURE 4.1 : The adapted Bakau Morbidity questionnaire

indication of the errors and variability produced when measuring and recording. Health card details were also included to indicate errors and variability when recording constant information. It was important that no major changes in structure should be made so that the old interviewers would not be confused. The new questionnaire is shown in Figure 4.1, reduced to fit on one side of a page instead of two. It includes consecutive numbering for the data items, not present in the questionnaire used in the field trial. These are used here for easy identification of the items as they are referred to in the text.

3. Interviewer profiles and training

All the interviewers in the study were in their twenties ranging from 21 years to 28 years (average age is 23 years), 6 of them were male, 5 had been to Secondary Technical School and 3 to High School. In The Gambia, the secondary education system is divided into Secondary Technical Schools and High Schools. In Secondary Technical schools, students qualify after 4 years with a Secondary Four exam, with grades divided into distinction, credit, pass or fail. In High Schools, students take the General Certificate of Education (G.C.E.) at the Ordinary Level ('O' levels) after 5 years that is graded from 1 to 9, 1 being a distinction, 2 to 4 being credits, 5 to 8 passes and 9 a fail.

All of the interviewers had passes or credits in at least 3 subjects. The most highly qualified interviewer graduated from High School with five passes and 3 credits. Two of the experienced interviewers had been working with MRC in Bakau for 3 years, the other two had been employed for 8 months on the Bakau Malaria Morbidity study. Only one of the new interviewers had some field experience, having worked as a survey technician for 2 years. All of the experienced interviewers had seen computers before in the MRC Computer Centre. One of the inexperienced interviewers had seen and used one for about 30 minutes at school. None of the others had ever used a computer.

Training was over 5 days and included the following:
Day 1: Basics on surveys, the purpose of the Bakau Malaria (MALBAK) Study and the aim of the project; introduction to computers, and to the Psion; Practice exercises on a hypothetical anthropometric survey.

Day 2: Detailed instruction on other features of FDCS - editing entries, going back to correct answers, getting longer explanations of questions; Types of errors and error messages.

Day 3: Instruction on the full-length MALBAK questionnaire, and explanation of questions; Instruction on how to measure; Practice of MALBAK on the Psion.

Day 4: New interviewers practice individually in the field with paper question-naires; Discussion session with old interviewers on common errors, general approach and hints; Practice in pairs (experienced and inexperienced) in the field with paper questionnaires.

Day 5: Revision of PSION with a reference sheet summary of all functions; Practice in pairs in the field with Psion.

There were a few functions on FDCS that they did not learn how to use during training. Instead, the instruction sheets explained the additional ones (LOAD - to change the questionnaire being used, VIEW - to browse through the data file, SUM - to generate tallies of variables). These were not introduced during the training sessions as they were not essential to data collection, and it was felt interviewers could learn in their own time if they wished.

At the end of this period of training, it was clear that FDCS was a system inter-viewers could use. They had reservations about whether the system would result in less errors, but were willing to try to use it. Their responses to a detailed questionnaire are given later in this chapter.

4.5   RESULTS - THE SYSTEM'S FUNCTIONING

Following training, the trial itself was undertaken over a 6 week period with the alternating design described above. Data on the performance of the FDCS and the paper system was collected weekly. No major incidences upset the experimental design and the Psion achieved its intended effect of reduced error rates within less time. The results presented in this section concern the system's technical perform-ance, analysed in terms of errors in the data collected, average interview lengths, contact rates, data entry and verification, software and hardware assessment.

1.   *Errors in data collected*

A program was written to logically analyse questionnaire data and check that the correct skip sequences were followed. The program also checked the form identifier details against the list of cohort children to be seen that week for incorrectly entered details. As expected, most of the errors occurred in the first 5 questions, and were often transcription errors. All the errors were tallied and the number of field errors for each interviewer each week counted and listed alongside the number of completed questionnaires.

Table 4.2 shows the number of field errors per 100 questionnaire items for each interviewer for the three attempts at both methods. An implicit assumption in the design was that the differences between the interviewers who used the Psion in the first week immediately after training, and those who used paper in that week, was negligible.

Attempt	Method	Interviewer							
		1	2	3	4	5	6	7	8
1	Paper	0.80	1.16	0.65	0.55	1.26	1.11	0.89	1.34
	Psion	0.29	0.48	0.22	0.46	0.42	0.46	0.37	0.61
2	Paper	0.49	0.89	0.25	0.32	0.43	0.72	0.57	1.00
	Psion	0.27	0.30	0.29	0.15	0.41	0.36	0.38	0.17
3	Paper	0.18	0.58	0.49	0.25	0.48	0.65	0.41	0.44
	Psion	0.15	0.31	0.15	0.15	0.23	0.27	0.24	0.10

TABLE 4.2: Number of item errors per 100 questionnaire items

At the first attempt at both methods, the use of the Psion halves the paper error rates for seven of the interviewers. This effect is remarkably consistent for the experienced interviewers, with a 58% to 66% reduction. For the inexperienced interviewers, the reduction range is much wider, starting from 14% to 66%. At the second attempt, the Psion's effect on error reduction is less evident for the experienced interviewers, as their paper error rates decrease by half, but accompanied only by a slight change in the Psion's rates. Except for one interviewer, the Psion's error rates continue to be lower than that of paper. This effect is also evident at the third attempt.
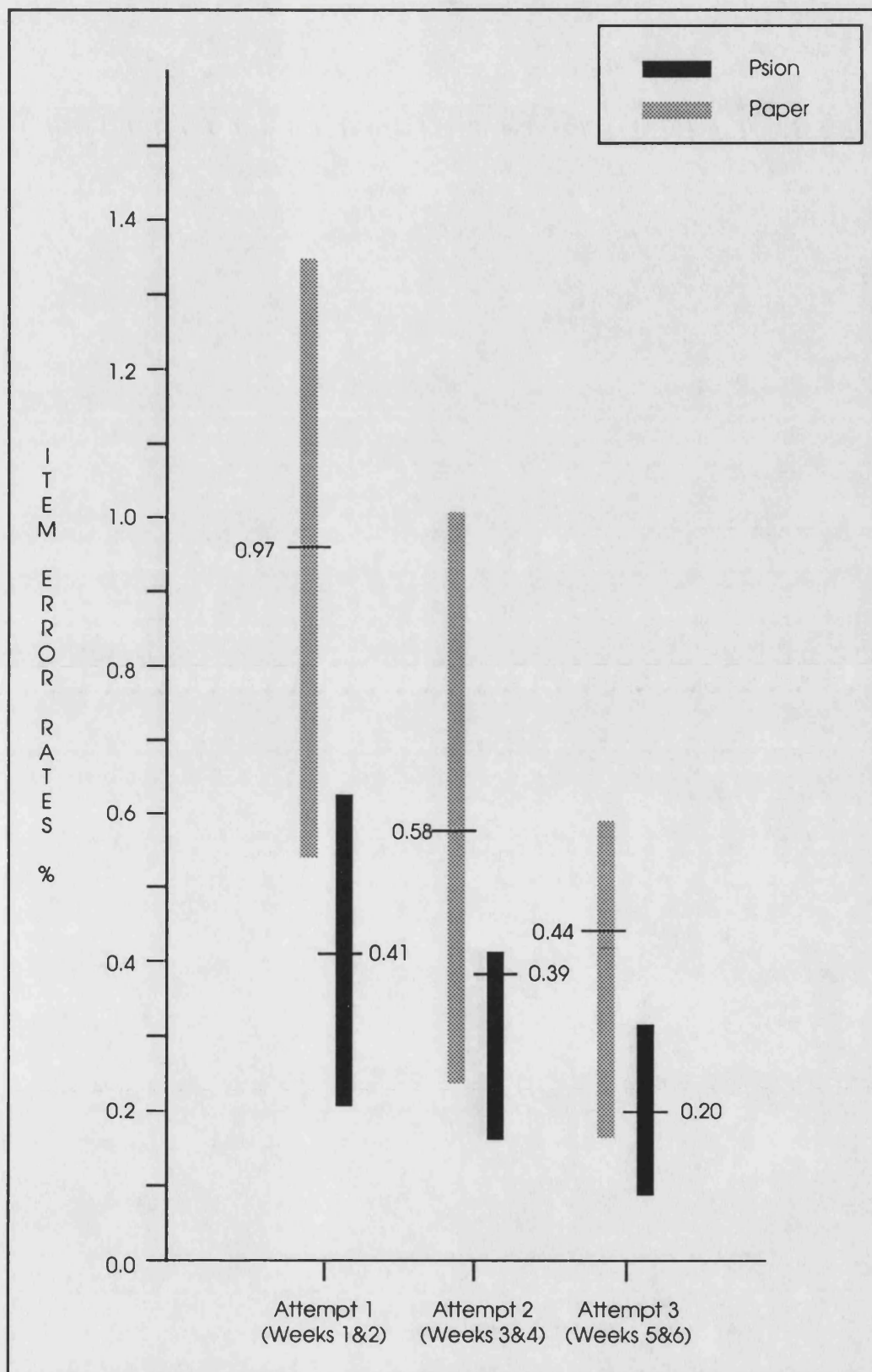
FIGURE 4.2 Means and ranges of error rates

Figure 4.2 shows the means and ranges of the error rates using both methods. This illustrates clearly the learning effect for both methods, and shows that the variability of errors on paper is greater than that on the Psion and that the error rates on the Psion are consistently lower than those on paper. FDCS results in a more standardised performance from interviewers of varying abilities.

An analysis of variance assuming a Poisson distribution confirms that there is a highly significant weekly trend accounting for 36% of the total observed variation (prob < 0.001). The differences between methods is also highly significant, accounting for about a quarter of the total observed variation. There is also some effect due to individual interviewers, the pairs, and their interactions with the method used. Put simply, it can be inferred that some interviewers and pairs were better at their work than others, and also that some felt more comfortable with one method than another.

Sources of errors for each method show that the computer makes a difference not only for branching errors, but also reduces questionnaire identification errors. This result seemed at first surprising since the error rates for identification errors were not expected to be significantly different for the two methods. There are perhaps two reasons for this result: firstly, the computer forces the interviewers to think about what keys to press, therefore increasing accuracy; secondly, it also eliminates script errors and misunderstandings due to unclearly written letters. Table 4.3 summarises the number of errors occurring in 4 sections of the questionnaire. Overall, the computer decreased the number of identification errors by 26% during the survey period.

Question groupings	Question numbers	Method	Wk1	Wk2	Wk3	Wk4	Wk5	Wk6
ID questions	Q2-Q6	Paper	101	94	46	68	35	38
		Psion	77	52	48	53	34	18
Symptoms	Q7-Q25	Paper	41	39	28	29	22	21
		Psion	8	4	0	7	0	2
Treatment and Measurement	Q26-Q42	Paper	45	40	10	19	10	20
		Psion	8	8	3	6	4	3
Date and Time	Q1,Q43-Q45	Paper	25	25	9	25	16	15
		Psion	0	0	0	0	0	0

TABLE 4.3 : Summary of sources of errors for sections of questionnaire

In summary, FDCS reduced the number of errors in the data sets and also standardised the data quality. Its effect is most marked at the start of the survey but is sustained throughout. The data is inadequate to describe what may happen in the long-term, but it is unlikely that the number of errors on paper will be reduced to equal that on the computer merely because the computer eliminates several potential sources of human error that persist on paper.

The observed reduction in error rates suggests that the advantages of the computer will be most pronounced for short-term studies that use structured questionnaires with many skips, and least for long-term surveys with no or very few question skips. This is because interviewers produce better quality data as they become accustomed to the forms, especially if there is continuous supervision and feedback of their errors and difficulties. In developing countries, the need for data often leads to short-term studies of one type or another, and hence FDCS may be of particular relevance.

2.   *Average interview lengths*

Average interview lengths were calculated for each interviewer on a weekly basis. These are calculated from times recorded by interviewers at the beginning and end of each paper interview, and the times recorded for interviews by the Psion.

Additional variability was introduced into the interview length figures from various sources. On paper, interview times were sometimes not recorded or incorrectly recorded, for example, in some cases the time at the end is earlier than the time at the beginning; these are ignored. On the Psion, the interview length was sometimes recorded as longer than it actually was if the next questionnaire was activated before the interviewer had located the respondent. This happened regularly in the first few weeks, the interviewers were later instructed to initiate the questionnaire only just before an interview. Thus, the length of Psion interviews at the start of the survey vary more than those at the end. The data is shown in Table 4.4.
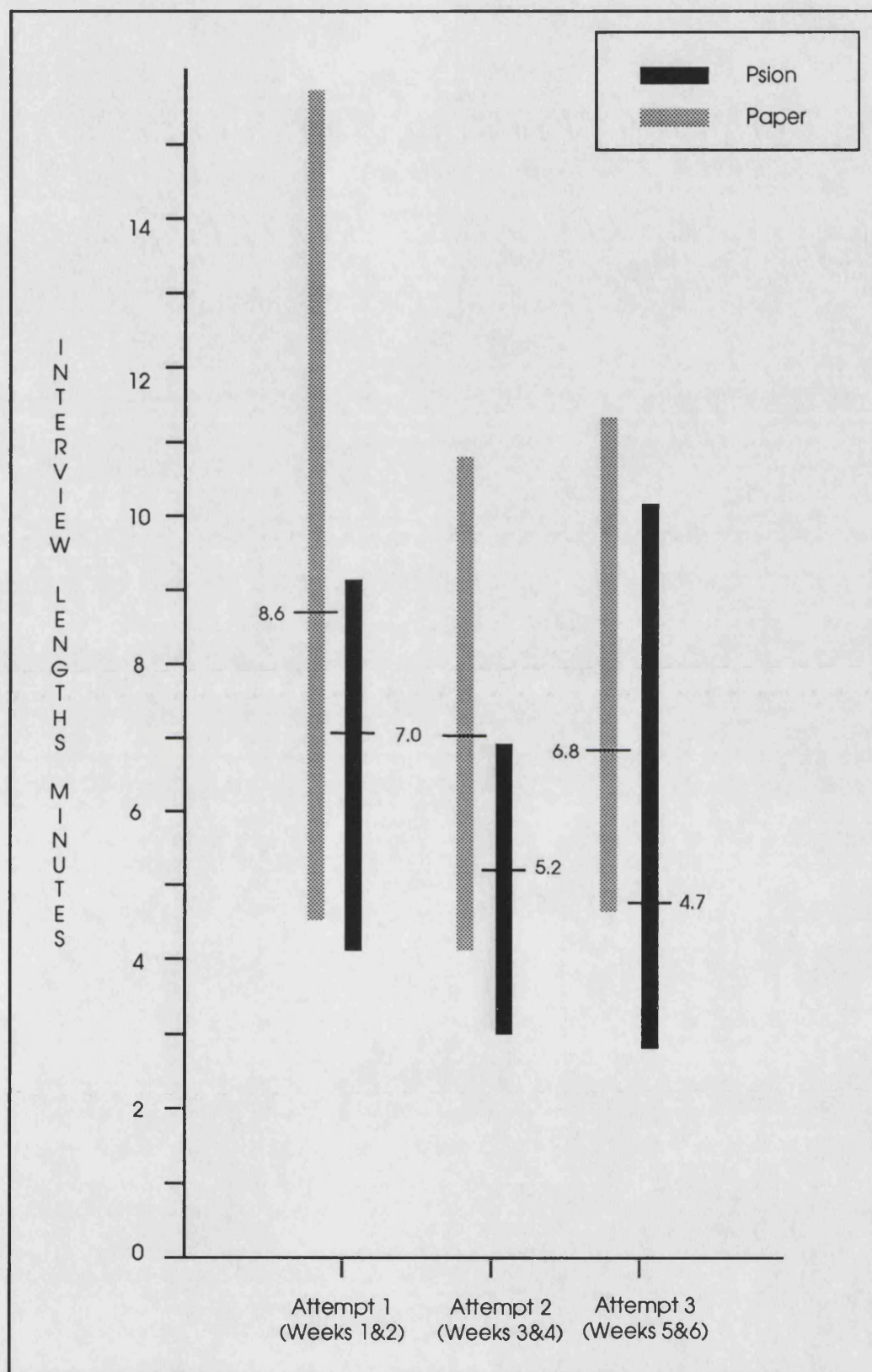
FIGURE 4.3 : Means and ranges of interview lengths

As a result, the impact of the computer on interview lengths is less clearly defined, but still evident (Figure 4.3). The computer again tends to reduce and standardise interview lengths.

Attempt	Method	Interviewer							
		1	2	3	4	5	6	7	8
1	Paper	6.7	4.5	6.9	15.7	5.6	8.2	7.3	13.8
	Psion	7.2	5.2	5.4	7.6	8.9	4.1	8.4	9.0
2	Paper	5.4	10.7	5.9	4.1	5.8	8.2	5.3	10.4
	Psion	4.2	3.0	4.7	5.9	4.5	6.0	6.4	6.8
3	Paper	6.2	5.1	8.0	5.0	4.6	7.9	6.0	11.5
	Psion	3.1	2.8	4.2	4.5	4.1	3.2	5.3	10.1

Table 4.4 : Average interview lengths

3.   Coverage Rates

The non-contact rates are based on the number of children allocated to the interviewer for that week, but not seen (Table 4.5). Figure 4.4 shows that the non-contact rates for the computer are higher than that for paper. The six highest non-contact rates all occur when the Psion is used, most of these occurring at the first attempt at using this method. The experienced interviewers account for five of these extremely high non-contact rates. On the other hand, the three 100% contact rates were achieved by the inexperienced interviewers. Within each pair, non-contact rates were expected to be roughly similar, as the same children were allocated to be seen each week by each member of that pair. In fact, the inexperienced interviewers have far lower non-contact rates. The obvious explanation is that they worked harder!

The occurrence of these high non-contact rates cannot be attributed to the computer requiring more interviewing time overall (Figure 4.3 shows that computer interviews take less time). It may be possible that the interviewers had less field time available because they sometimes needed to return to base to have the batteries replaced. Alternatively, they might have spent less time in the field overall, because they expected to finish earlier, and therefore made fewer recall visits. This unexpected finding cannot be fully explained because it was not
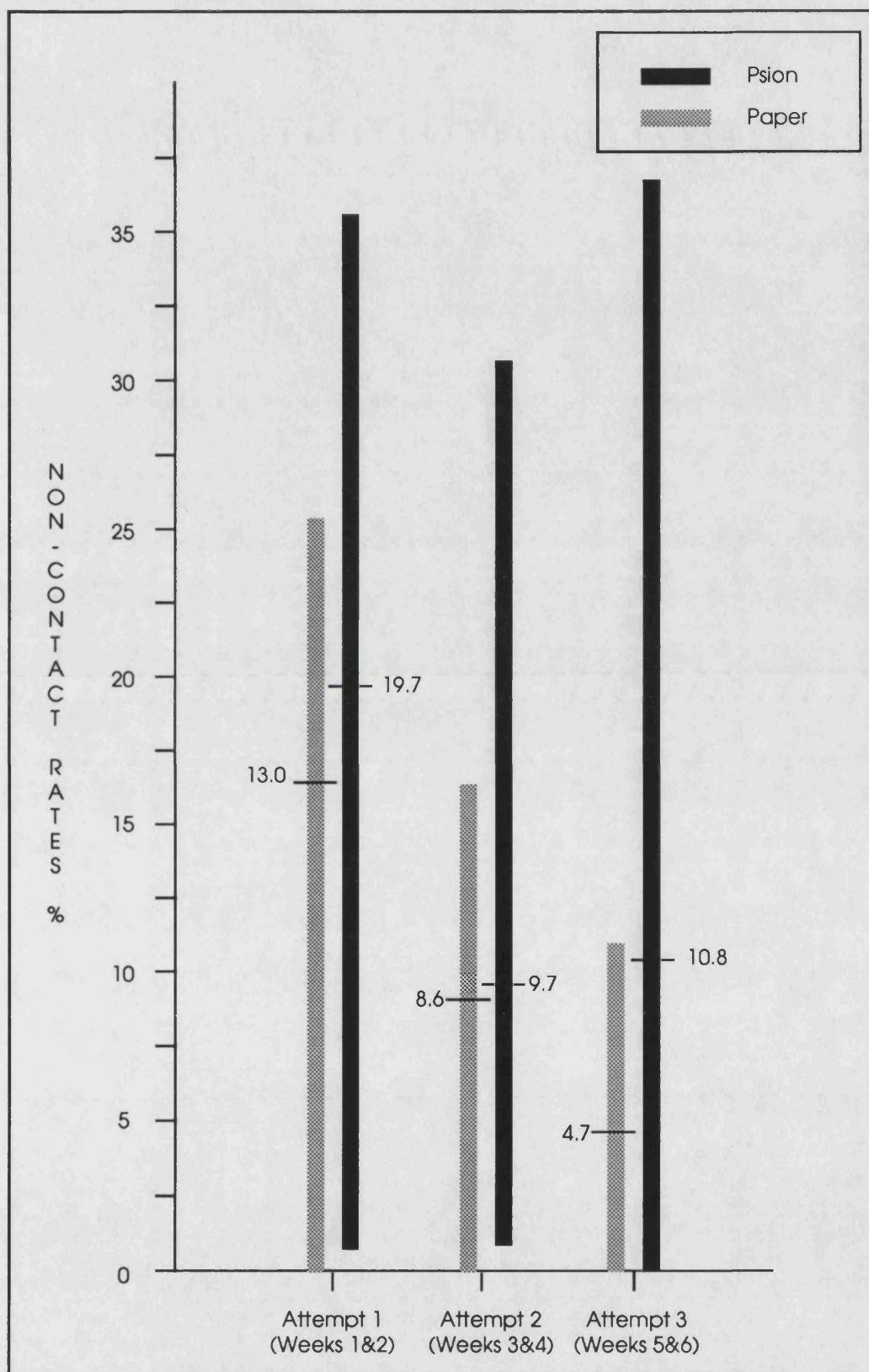
FIGURE 4.4 : Means and ranges of non-contact rates

evident during the survey period and no attempt was made to control it by informing the interviewers of their performance.

Attempt	Method	Interviewer							
		1	2	3	4	5	6	7	8
1	Paper	17.7	11.8	3.6	0.8	17.8	9.2	17.2	25.6
	Psion	1.7	31.9	35.4	12.2	1.7	34.8	11.6	28.5
2	Paper	0.0	15.0	1.6	4.8	10.9	10.3	9.3	16.7
	Psion	2.5	11.8	1.6	3.9	1.7	6.7	30.9	18.6
3	Paper	0.0	6.5	0.8	2.5	0.0	11.2	8.7	7.9
	Psion	2.4	5.8	4.1	6.5	0.9	36.8	11.1	19.1

Table 4.5: Non-contact rates (%)

4. Data Entry and Verification

An estimate of the time taken to punch in the data from the paper questionnaires was made in the last week of the study, when the computer centre manager took measurements (without the computer operators' knowledge) of the time taken for each stage of data entry - punching, verification and correction (Table 4.6).

Number of forms	Punching	Verification	Correction
123	145 mins	35 mins	140 mins
102	118 mins	25 mins	60 mins
190	220 mins	45 mins	140 mins
49	60 mins	15 mins	30 mins
464	543 mins	120 mins	370 mins

TABLE 4.6 : Estimated time taken for each stage of data entry

It took about 9 hours to punch in the data once, about 2 hours to run the verification once, and about one and a half hours to do the first corrections. These measurements would vary depending on how many users are logged into the computer network, and the user's time-slice allocation. Normally, one operator is responsible for the Bakau Malaria Study data, and punches them in twice, verifies and does corrections. This takes a whole week, and is his sole responsibility.

A number of data entry errors were discovered in the final paper data set. Most of these were found when verifying the field errors using the logic check program and are shown in Table 4.7. A few were found by accident, and were mostly insertions of values by data entry clerks where none were written by the interviewer. Errors due to unclear writing are attributed to the interviewer, not the data entry operator.

	Wk1	Wk2	Wk3	Wk4	Wk5	Wk6
Number of fields with punching errors	34	8	10	6	3	5

TABLE 4.7 : Residual errors trapped by consistency checks after data entry

The errors found after verification were very few. Obviously, the real rate is higher, but there is little evidence to suggest that it is considerably greater. It is perhaps also a reflection of good layout of the questionnaire, structured for easy data entry [Byass, 1986]. For this particular survey, the use of the Psion would make the double entry and verification procedure unnecessary and reclaim the time of one computer operator dedicated to this task. It would also eliminate the time lag, a week in this case, to obtain a clean set of data.

5. Hardware and Software Assessment

Four Psions were used continuously in the field for 7 weeks, and were left in the charge of the interviewers who took them home. They fell down several times, usually in sand but sometimes on a hard surface. They survived with a few scratches on the screen, and at the time of writing, are still in working condition. The print on the keys grows faint with use, particularly on the enter key. Towards the end of the trial some of the machines used in the field did not work when plugged into the communications link, probably due to dust accumulation on the contact pins of the communications slot. This may be put right by thorough vacuuming.

The batteries did not last as long as was expected. Duracells (alkaline batteries) lasted for 7 working days on average, and rechargeables a maximum of 2.5 working days. An additional week of data collection by the experienced inter-

viewers using the Psions was made at the end of the main survey period to verify this. All the logging facilities were disabled and the interviewers were given spare batteries and instructed on how to change them. All four interviewers managed to change the batteries on their own at least once during that week and reported back that this was a more satisfactory arrangement than an enforced return to base. The logging facility employed during the trial used up additional storage space and battery power, but two rechargeables per week seems to be the minimum requirement.

The central server (database) onto which the Psions downloaded their data was a Toshiba T1200. It performed well in spite of voltage fluctuations, and the battery backup proved adequate to continue working on the hard disk without mains power for about 15 minutes. It was also carried around a great deal, but suffered no ill-effects.

The software used in the trial contained various bugs, but most of these were found during the training sessions or in the early stages of the trial. One may have remained, however, as an interviewer complained of a STR TO NUM (string to number conversion) error message on the screen soon before the completion of the study, but this error could not be reproduced. During the first two weeks, a particular software error sometimes caused data loss before storage onto a storage device. Several interviewers complained of this in the first week of data collection, but always re-recorded the data if the Psion did not indicate that it was saved. All other errors reported were located and corrected.

In transferring data between machines, two fatal errors were made and data was lost irretrievably. The first incident was in deleting all the data files (the equivalent of *.* in MS-DOS) before transfer. The second was because of a power surge that occurs when the communications link is plugged in. Storage devices must be removed before attaching the cables as the surge is strong enough to wipe out the volatile memory. The Psion must be switched on and off before the packs are plugged in and transfer begins. For one set of data, fortunately a small one with only 2 questionnaires, the pack was plugged in too quickly, resulting in a loss of data. The data communications cable did not work once perhaps because it was left plugged in the socket after use.

The hardware proved robust in the dry and dusty environment, and could be safely entrusted to the interviewers to take home. However, in the long-term, they might need some cleaning of the more exposed components. A field study of QUESTOR in francophone Africa attributed the rapid decline of battery life to the high temperatures. This might be true as the batteries lasted for a shorter time than expected. A few early software faults were responsible for some data loss in the field, but these were corrected as the survey progressed. Another potential cause of data loss was in the susceptibility of storage devices to erasure if there was a surge of power. This danger might be eliminated by keeping to a strict stepwise procedure, or by improved hardware design.

4.6   RESULTS - USER PERSPECTIVES

The technical issues described above give one perspective of the trial, but the opinions and interests of the various people involved was also sought to give a fuller picture.

1.   *Training Questionnaire*

At the end of training, all the interviewers were interviewed individually with a detailed questionnaire, with questions on their opinions of the hardware, notes, question design, and their general expectations of the differences between using the computer and using paper. The questionnaire used is shown in Figure 4.5. The summary of their responses is shown in Table 4.8.

They were split in their opinions about whether they found it convenient to use the computer in the field, four giving positive responses and the others negative ones. More of the experienced interviewers (3) gave positive replies about the Psion than the inexperienced interviewers, 3 of whom gave negative replies. None of them thought the computer was heavy or difficult to hold. All of them said the screen was readable, and most also added that they found it easy to adjust, if necessary, when used in direct sunlight.

TRAINING ASSESSMENT QUESTIONNAIRE

Hardware

1. Do you think the Psion will be convenient to use in practice in the field?
2. Do you find it heavy or difficult to hold?
3. Is the screen easy to read?
4. Was getting used to the keyboard easy or difficult at first?
5. Some people find it difficult to enter a combination of letters and numbers..
    i.   Did you have this difficulty at first?
    ii. Do you still find it difficult to do this?
6. Do you now find it easy to locate the keys you want to press?

Notes

7. How easy was the manual to use at first / when learning?
8. What changes do you think must be made to the notes?
9. What sections of its are still hard to understand?
    i.   ...... paging back, returning to an earlier question
    ii. ...... using the explanation/help facility
    iii. ...... loading different questionnaires
    iv. ...... browsing/looking/going through records
    v.   ...... editing a record before saving
10. Do you now feel you can get help from it when you need it?

Question design

11. Are the questions clear from the screen?
12. Is there enough information on the screen to answer the questions?
13. How often do you refer to the explanation facility?
14. Do you use the paper version of the questionnaire?

Attitudes

15. Do you now feel confident that you could use the Psion Organiser to collect data on your own?
    a. No
    b. Only with some help
    c. Yes, reasonably confident/well
    d. Yes, completely confident
16. Do you think using the Psion will lead to refusals?
17. How do you think respondents will react to being interviewed with the Psion?   for example, suspicious / interested / neutral.
18. How do you think the Psion will affect your interviews?
    i.   length
    ii. quality
19. How do you think the Psion will affect the amount of corrections you have to make on the data you collect?
20. Has the training made you more interested in computers?
21. Do you like using the Psion?

FIGURE 4.5   : Training assessment questionnaire

Most of them found it difficult to get used to the keyboard at first, especially in entering a combination of numbers and letters (to do this, one has to hold down the SHIFT key and then press an alphabetic/numeric key). Finding the keys became easier as the survey progressed, though most had not mastered this fully at the end of the training session.

Some seemed to find the instruction notes handed out during the training sessions a bit difficult to understand to begin with, but easier as they re-read them in their own time while also practising on the computer. In general, they found the notes easy to use and helpful as a reference.

Question	Answer	All int.	Old int.	New int.
Convenient	positive	4	3	1
	negative	4		
Heavy	no	8	4	4
	yes	0		
Keyboard difficult at first	no	6	3	3
	yes	2		
USING FUNCTIONS				
going back	yes	7	3	4
	no	1		
explanation	yes	7	3	4
	no	1		
load qnaires	some idea	3	1	2
	no idea	5		
view answers	some idea	3	2	1
	no idea	5		
edit	yes	5	2	3
	no	3		
Confident?	completely	3	2	1
	reasonably	5		
Psion shorter	yes	2	2	0
	same	1	1	0
	no	5	1	4
Less mistakes made on Psion	yes	3	1	2
	same	2	1	1
	no	3	2	1
Less mistakes found on Psion	yes	3	2	1
	same	0	0	0
	no	5	2	3

TABLE 4.8: Summary of interviewer responses to training questionnaire

On being asked to demonstrate how to do the various functions on the PSION, none of them completely understood how to use all of the functions on the outer

level. The inexperienced interviewers learnt slightly more than the experienced ones (see the section on functions in Table 4.8).

In general, they did not find the screen questions difficult to understand as it was easy to recall the long questions from the abbreviated versions. Six of them did not use the explanation facility at all during their practice session. Four of them found they did not need to refer to the paper version of the questionnaire during their interviews. Yet when asked whether they were confident in the use of the computer without help, a minority said they were completely confident. The rest said they were reasonably confident.

Answers to the questions on expectation of respondent reactions all indicated that they did not expect any refusals because of the machine, because they meant to explain its use before starting interviews. They did expect some respondents to express surprise at the new method and ask questions. One reported that a mother was surprised to learn that all the writing on paper could be stored on such a small machine.

All of the inexperienced interviewers and two of the experienced ones thought that the Psion interviews took longer, and their explanations were that the computer controls the speed of data entry and that they had to wait until it had gone through each step, so they could not do several questions ahead. They also remarked that one had to be more careful to press the right keys on the computer and not cause errors. Of the remaining two, one thought paper took longer, the other that interviews were about the same length. Three thought they would make more mistakes on paper because they expected the computer to check their entries, one drawback being that going back to correct answers takes longer than on paper, although paper corrections were generally more messy. Another three thought they would make more mistakes on computer, and the rest that the number of errors would be roughly equal.

Asked about how many errors would eventually be spotted by the supervisor, most of them thought that the number of errors found on the computer would be higher because they were more familiar with paper. The rest thought there would be more mistakes on paper because corrections were more likely to be unclear.

All of them liked using computers and thought that using the Psion had given them some new experience. One commented that it was good to find out how computers now make things easier; another that he liked using the Psion because the information is recorded free of mistakes; and that changing mistakes on paper is more difficult after leaving the field, whereas on the machine mistakes were changed immediately.

The training of one week was sufficient to learn all the basics and become confident about using the computer without any help. Practice sessions in the field proved invaluable in consolidating their confidence. At the end of training they could use the keyboard and all the basic functions, as well as refer back to the notes for help when required. In general, the inexperienced interviewers, who were less enthusiastic about the new technology, understood and learnt more than the old interviewers, but were naturally less confident about their ability to work alone in the field.

2. *Working patterns and practices*

During the first two weeks, the interviewers were accompanied to the field to observe how they worked in practice using either method. The checklist in Figure 4.6 was developed to record interviewing conditions, work procedures and attitudes.

Interviews were mostly carried out in a house, with the interviewers sitting down and using their laps as a support for writing. Of the 80 interviews observed, 49 (61%) were under these conditions. Very few interviews were carried out in direct sunlight, only 5 (6%) of these were observed. The remaining interviews usually took place outside, but sitting or standing in the shade. Most interviews were in the early afternoon around lunchtime when mothers would normally be at home, and late afternoon when the older children would be expected to be home from school.

PSION OBSERVATION CHECKLIST

1. FW NAME: ....

2. DATE: :___:___: / :___:___: / :___:___:

3. Number of completed interviews: :___:___:___:

Interviewing conditions

4. Position
   ...... in house   ...... outside   ...... crouching
   ...... shade   ...... direct sunlight ...... in hand
   ...... standing   ...... sitting
   ...... using support   ...... on lap

5. Ask for health card   ...... before starting   ...... after starting
   Ask for blue card   ...... before starting   ...... after starting

6. Explanation of Psion
   ...... without being asked
   ...... after being asked
   ...... unnecessary, mother seen before

7. START TIME: :___:___: / :___:___:

8. Study number: :___:___:___:___:___:

9. How many symptoms asked?
   ...... all
   ...... most
   ...... a few mentioned
   ...... none

10. When is temperature taken?
    ...... at start of interview
    ...... before starting symptoms
    ...... when question reached
    ...... waited until read
    ...... took other measurements

11. Using the tape measure
    ...... fiddling   ...... slipping   ...... repeating   ...... OK
    ...... baby   ...... small child ...... reasonable size

12. Number of beeps on the Psion: :___:___:
    Why?

13. END TIME: :___:___: / :___:___:

14. What does mother do?
    ...... waits all the time
    ...... waits some of the time
    ...... wanders off, has to be called

    If next interview not immediately after,

15. Is Psion turned off?   ...... yes   ...... no

16. How is Psion carried?   ...... bag   ...... hand   ...... pocket

REMARKS

FIGURE 4.6 : Psion observation checklist

A number of differing interviewing styles and procedures were observed. Interviews varied on a number of factors:

(a)   The interviewer may have asked for the blue registration card (which states that a child is involved in the study) and health card before starting the interview, and therefore could continue without stopping to request these. If the interviewer asked for the health card only when that particular set of questions was reached, there was usually some delay while the mother searched for it.

(b)   The interviewer sometimes explained the use of the Psion, without being asked, before starting the interview. If no explanation was given, some respondents interrupted to ask questions about the machine during the interview.

(c)   The set of questions relating to morbidity were often skipped without enough probing. In quite a few instances, the respondents said the child was well, and only mentioned the occurrence of a symptom if it was specifically asked for. The interviewers were therefore instructed to ask all the morbidity questions, despite the mother's answer, and to record the child as well only if the answers to all symptoms were negative. This was not always done - the interviewer sometimes chose not to ask any symptomatic questions, or would ask for only those that were judged 'common' and more likely to occur.

(d)   The interviewers differed as to when they took the child's temperature. Sometimes, they preferred to do so at the beginning of the interview and ask the mother to hold small children steady while they asked the rest of the questions. At other times, they only took the temperature when they reached the question. This created a difference of about two minutes between the interview lengths of these two procedures. They sometimes took the other measurements at the same time if the child tended to keep still. Recording is not as easy on the Psion because each question is asked sequentially, and the interviewer has to memorise the measurements and enter them when the question was asked, perhaps introducing an error.

(e)   Measurements could take longer if the child was a baby, or difficult to hold still. The amount of clothes the child had on could also introduce a slight delay if there were several items that needed to be removed before measuring.

(f)   Some mothers would wait throughout the interview, and be available to help with the child if necessary. Others would wander off to do other household chores, or occasionally, attend to visitors.

The interview length is therefore a function of all of these variables and differs amongst interviewers. In an effort to standardise procedure, the interviewers were instructed to:

(a)   ask for the blue cards and health cards before starting.

(b)   ask all symptoms to confirm that a child that is said to be well has no minor symptoms.

(c)   to always take the temperature only when the question is reached, and to wait for the required two minutes, ensuring the thermometer is in the right position, and holding the child still if necessary.

Working patterns and styles varied for each interviewer as they developed individual ways of conducting their interviews.

In practice, a sequential series of questions was subjected to a variety of working techniques. Some interviewers took shortcuts and asked a selection of the morbidity questions for probing, and often did measurements in parallel. The use of the computer system forced a formalisation of procedure, but the need for standardised working practices was due to the nature of the field trial. Any computerisation process has to take into account the fact that people tend to adapt what they are expected to do if they find an easier way of working.

3.   *Weekly Progress Monitoring Questionnaires*

The questionnaire shown in Figure 4.7 was used at the end of weeks 2 through 6. Questions 2 and 3 were asked only in weeks 2 and 3, and were dropped for subsequent weeks as they did not reveal any interesting answers.

The number of interviewers who felt their workload was decreased with the use of the Psion rose from 3 in the first week to all 8 in the last week, although 6 of them thought the workload was the same in the second week. During the first week, two thought the workload had increased, but none did so subsequently. The majority of interviewers felt that the number of errors on the Psion had

---

WEEKLY PROGRESS MONITORING QUESTIONNAIRE

1. How did you find the Psion compared to last week?
   a. a lot easier
   b. a bit easier
   c. no difference
   d. a bit difficult
   e. a lot more difficult

2. What problems did you have?

3. How have most respondents reacted to the Psion this past week?
   inquisitively / suspiciously / irritably / neutral, indifferent

4. How do you feel the Psion has affected the amount of work you have to do?
   a. increased workload
   b. decreased workload
   c. no difference

5. Do you feel the Psion has affected the accuracy of the data you collect?
   a. no
   b. yes
   c. no difference

6. Which interviews take longer? Those with the
   a. Psion
   b. paper and pencil
   c. no difference

7. Do you find it easier to transport and return information on the Psion?
   a. yes
   b. no
   c. no difference

8. Which method do you now prefer?
   a. Psion
   b. paper and pencil
   c. no preference

FIGURE 4.7   :   Weekly progress monitoring questionnaire
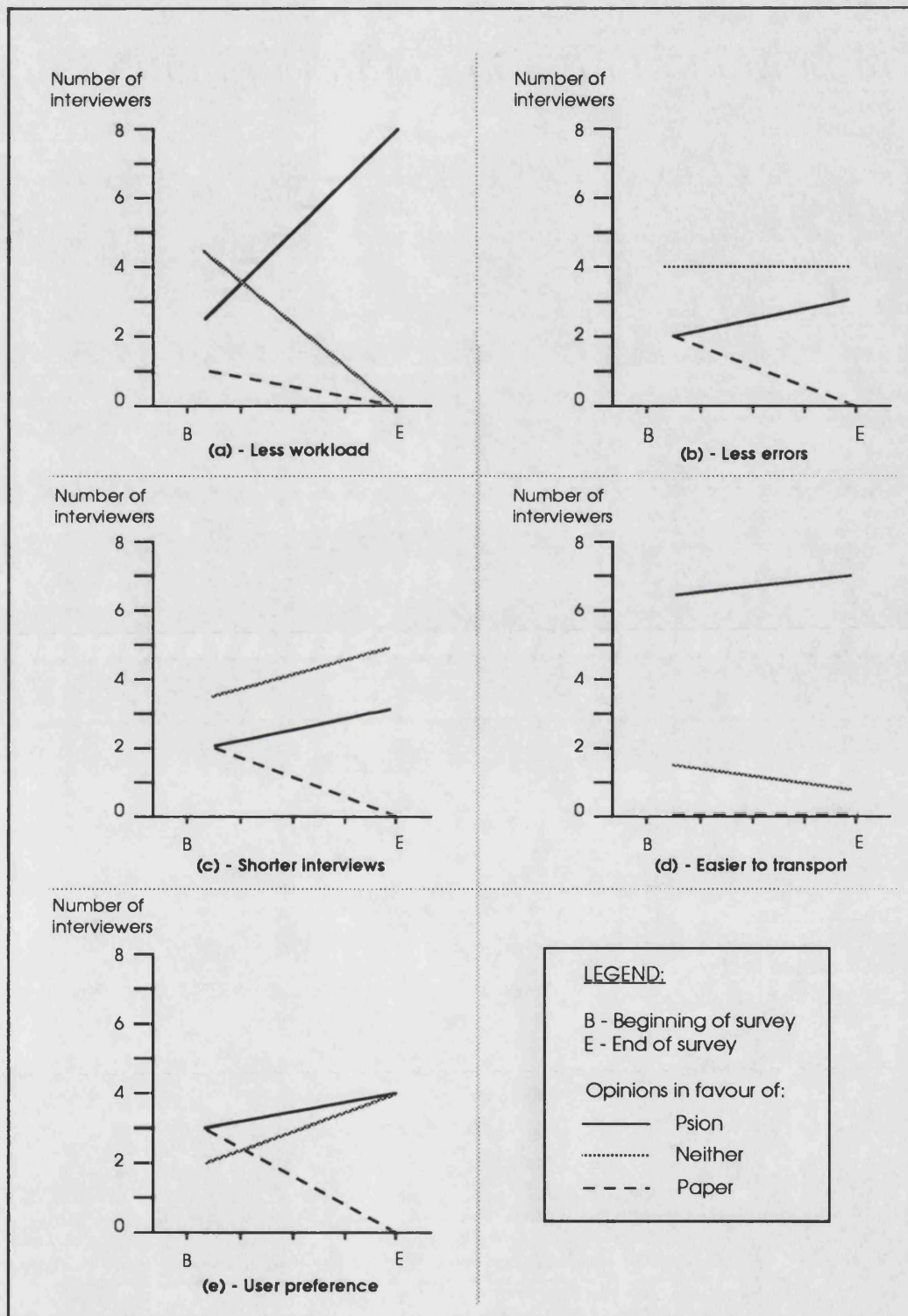
FIGURE 4.8 : Changes in interviewer attitudes over the survey period

increased each week. In the sixth week there was an even split between those who thought Psion decreased errors and those who thought it made no difference.

During much of the survey period, the majority of interviewers agreed that the interviews were about the same length. In the second week, most said that paper took longer, but the balance shifted the other way a week later. By the last week, no one said the Psion took longer. No interviewer thought paper easier to transport at any time during the study. The number who said the Psion was easier remained high, between 6 and 7, over the whole period. The number preferring Psion varied from week to week, but by the last week, none of the interviewers expressed a preference for the paper.

The summary of changes in Figure 4.8 shows rises in favour of the Psion for all five factors over the period of the study. The rise from neutral to strongly in favour was steepest for workload, see Figure 4.8(a). For ease of transport, the Psion was always thought highly of, with no votes for paper being easier to carry around, see (d). Most interviewers thought there was no difference in interview times, see (c). The number in favour of the Psion for decreasing errors and the number preferring Psion both rose to 4, see (b) and (e).

Attitudes in general reflect a clear learning effect. At the end of week 2, unfavourable answers begin to change to either a neutral or positive position, presumably as they got more familiar with the computer and began to see the results of its use, with the last week ending on the highest note of all.

4.   *Opinions at the end of the study period*

At the end of the field trial, each interviewer was again interviewed individually and asked the questions in Figure 4.9 which includes some open questions designed to allow the interviewers to freely express their opinions. The interviewers mentioned the following as aspects they liked most about the computer. A pile of paper questionnaires is difficult to carry around and has the tendency to scatter but the Psion is easy to carry, being one instrument "requiring only a finger". The software allows changes to be easily made during the interview, questions are skipped automatically and time and date recorded, mistakes are

corrected in the field and require no further checking before submission to base. Some also mentioned that the Psion had given them new experience with computers.

On what they disliked most about the computer, seven of them mentioned low batteries, and the need to return to base to have them changed. Also mentioned were the constraint on the number of questionnaires that can fit onto one storage device, time waiting for the next question to appear, and not knowing what to do if there were software errors.

On what they liked most about paper, some mentioned the speed with which errors can be corrected, others the ability to see many answers at once, and not having to wait for the next question. On what they disliked most about paper, they mentioned that it was easier to miss some questions; and pointed out some of the various inconveniences of paper: difficult to hold many sheets at a time, heavy to carry, easily scattered, dirtied, torn or blown away.

On the differences in the amount of time it took them to do interviews, most mentioned that they had found it easier to do interviews at the end of the survey than at the beginning; the experienced interviewers had to get used to the new format and extra fields, the inexperienced ones to the whole process of surveys. More of the inexperienced ones said the computer was always quicker than paper, even at the beginning; the experienced ones tended to find the computer slower at the beginning. Most said their mistakes lessened as the survey progressed, and that by the end they made more mistakes on paper than on the Psion.

During the discussion, each interviewer was shown their answers to the weekly progress questionnaires and their average interview lengths and errors for that week, and asked to give likely explanations for what happened. Various reasons were given for variations in interview lengths: mothers may wander off, more care had to be taken when inexperienced at the start of the study, there were problems with using the unfamiliar computer at the start, also the questions became more familiar as the study progressed.

END OF SURVEY INTERVIEWER QUESTIONNAIRE

1. What do you like most about using the computer to collect data?

2. What do you dislike most about using the computer to collect data?

3. What do you like most about using paper to collect data?

4. What do you dislike most about using paper to collect data?

5. Have you noticed any changes since the beginning of this trial in:
   a. the time it takes to do interviews?

   b. the number of mistakes in the data you collect?

6. A discussion during which they are shown their answers to the weekly questionnaire, their averages for interview lengths, their average error rates, and also how each of these relate to the whole group.
   They are then asked to give likely explanations for any noticeable trend. or for any contradictory answers.

FIGURE 4.9 : Final interviewer questionnaire

Some found that the computer reduced the amount of work they had to do, both during the interview in skipping questions, and recording time and date, as well as afterwards, with no further need to check on the forms. Most agreed that the computer rapidly reduced the number of errors found on their forms.

None of the final preferences were decidedly in favour of paper; some did not mind either method, a few still enthused about the computer. Of the ones who preferred paper at the beginning, they mentioned that their main difficulty was in getting familiar with the use of the computer, and being able to use it as quickly as writing, especially when correcting mistakes.

4.7   RESULTS - RESPONDENT ATTITUDES

The opinions of the respondents were also taken into account. Respondents were interviewed with the questionnaire in Figure 4.10 in two blocks, either at the beginning of the survey, or at the end. Mothers were selected indirectly, by first randomly selecting cohort children and then including their mothers if they had not previously been selected. Lists of 35 mothers were drawn up for each batch of interviews. All respondent interviews were done using an interviewer not involved in the trial as interpreter.

In the first batch, 11 of the randomly selected mothers were not interviewed because they had gone away with the child, had moved away from the area, had not yet been interviewed with the computer, or did not notice the interviewer using it. 24 mothers had seen and been interviewed with the Psion either once or twice. Occasionally the usual respondent was a relative other than the mother, and this person was interviewed instead.

22 respondents had no concerns about the use of the computer. The following comments were made: "a bit strange because I have never seen anything like it", "but was surprised to see something different", "not worried because [I could see that interviewer asked the same questions as before". 2 said they had objections and both gave as the reason that they had "doubts about it because have never

RESPONDENT QUESTIONNAIRE

COMP NUMBER: :__:__:__: DATE: :__:__: / :__:__: / :__:__:

1. Do you have any objections to the interviewers using this machine?
    a. Yes
    b. No
2. Have you seen anything similar before
    (for example, calculators or other computers?)
    a. Yes
    b. No

    If yes, what?

3. What do you think the fieldworker uses this machine for?

4. Did the fieldworker explain what it is used for?
    a. Yes
    b. No

5. Would you like to know more about what it does?
    a. Yes
    b. No

    If yes, then give a short explanation.

6. Which interviews do you think take longer?
    a. With this machine
    b. On paper and pencil
    c. About the same

7. Which do you prefer the fieldworkers to use?
    a. Paper and pencil
    b. Computer
    c. No preference

FIGURE 4.10 : *Respondent questionnaire*

seen anything like it". One of these had never seen a calculator before but the other had seen adding machines in a shop.

20 respondents had seen calculators before, mostly in the market or in shops and offices; sometimes their husbands or other relatives had calculators in the compound. 2 respondents said they had seen something like the computer; one said she had seen "something that looked like a draught board with numbers", the other that she had seen it "in a video film". 2 respondents had never seen calculators or computers.

Asked about what they thought it was for, some answers were:

"I know its something to do with the job - when being interviewed, the interviewer types in what I say",

"something interviewers use for writing",

"a small radio, I noticed that after typing, the interviewer pressed the last key, waited for a while and then asked another question",

"I think they use it to know if something wrong happens; I rang up telephone company once and abused them, they phoned back to say the number was noted on their computer and gave me a warning",

"I know everything is available - because tape recorders record voice, I assume this is similar",

"I thought at first that the interviewer was going to use it for bleeding",

"they use it for checking mistakes therefore this is better".

Eight respondents gave answers about it being something to do with the health of the child, "to find out whether the child is sick or not", and about "medicines for the children". Three thought that it was used to count or calculate something; "thought at first that it was a calculator, but when I saw the letters I knew it was different". Four respondents said they know it is part of the interviewer's job, but "don't know what its for". Two respondents said "don't know" after being probed for a guess.

20 respondents said the use of the computer had not been explained to them, 4 said it had and that "it was used for writing", "so that what is on the form is what is on the computer and the computer can help correct mistakes". 12 respondents

thought the interviews with paper were longer, 11 thought they were about the same, 1 could not say as she "wasn't paying attention".

12 preferred the computer, and gave as reasons "because of your explanation", "it is quicker than writing", "because if the machine gives better answers, it is better for all", "they always make mistakes on paper, especially with date of birth, but since they started with computers, there have been no mistakes", "because according to your explanation, there will be less mistakes", "does not skip any questions". 12 said they had no preference, "because my answer stays the same no matter what method is used", "it's their job, how can I have a preference?".

In the second batch of interviews, question 5 was asked at the end of the interview to examine the difference an explanation made to the answers to questions 6 and 7. Four respondents had travelled, moved away finally, or had changed address and could not be traced. 31 respondents were questioned. 2 respondents had never noticed the interviewer with the computer, although they had been visited and interviewed several times during the project. The other 29 respondents had noticed the interviewer using the computer on average 4 times.

28 of them said they did not have any objections. 1 said she "does not know what it's for, so can't tell" whether she has any objections or not. 25 of them had seen calculators before, 4 of them had not. Although the majority had not been given an explanation, about half of them thought the interviews were about the same length with either method, approximately half thought paper took longer, only one thought the computer took longer. The majority had no preference to which method was used, only 2 said they preferred the computer.

The results from the two batches of questionnaires correlate well (Table 4.9). Very few objected to the use of the computer, none of those who objected had been given explanations of the Psions' use. All of them said their doubts stemmed from not understanding its use. The large decrease from 50% to 7% of respondents who preferred the computer might be directly attributed to the explanations they were given before being asked for their preferences in the first batch; the procedure changed in the second. Many approved of any way of improving their children's health. There is a remarkably large exposure to modern micro-

electronic calculators as over 90% of respondents had seen them with shopkeepers, traders or members of their families. Very few had heard of or seen computers.

	First block	Second Block
Number of respondents interviewed	24	29
Number with no objections	22	28
Number who had seen calculators	20	25
Number who had seen computers	2	0
Number who were not given an explanation	20	24
Number who thought computer took longer	0	1
Number who preferred computer	12	2
Number who had no preference	12	27

TABLE 4.9 : Summary of respondents answers

About 50% thought paper took longer than the computer, and this figure does not change over the two blocks of interviews. It is also interesting to note that a few respondents did not notice the computer at all whilst being interviewed; although the mothers were occupied with household chores, they remained within earshot and sight for much of the interview, thus demonstrating the Psion's unobtrusiveness. In general, the respondent interviews suggest a large capacity to accept the use of new technology, perhaps influenced by its similarity to calculators, which a large proportion had seen before.

4.8   STRUCTURE, PROCESS AND OUTCOME ANALYSIS

To summarise the efficiency, and impact of FDCS within the structure, process, and outcome approach, some questions will be formulated within that framework and the answers produced by the field trial discussed.

1.   *The system's functioning.*

It needs to be asked whether the hardware performance was satisfactory, and whether the other resource requirements were met within this context.

(a)   *Structure.* Does the hardware and software used in this study demonstrate adequate capabilities for use in developing country survey applications?

The hardware, though less powerful than the commonly used Toshiba portable for similar work, proved to be durable in spite of a few falls and being left in the charge of the interviewers for the duration of the study. However, there are a few unexplained reasons for difficult data communications between some Psions and the Toshiba. The rechargeable batteries also lasted for far less time than was anticipated. The volatile storage devices were expected at the outset to use less power than they did in practice. The duration of the batteries directly influenced the amount of data that could be collected at a time, and this inadequacy needs to be further dealt with. Some software problems at the start were solved during the training sessions, a few residual problems remained which were discovered during the first two weeks of data collection.

(b)   *Process.* Does the use of FDCS alter the characteristics of data gathering for the better?

Questionnaire definition was straightforward. The physical restrictions of the Psion did not noticeably impede data entry in the field as the interviewers found it easy to conduct interviews with short prompts, and found the compactness of the machine a bonus for carrying around. Data transfer needed some care, but was generally done without much difficulty within half an hour. With few minor exceptions, the data was successfully transferred onto a microcomputer after being returned to the base.

(c)   *Outcome.* Did the system deliver solid benefits?

The expected benefits of the system have proved real in practice, in reducing the number of errors in the data collected as well as the time taken to conduct

interviews. The improved survey cycle time cut out the week's work required for data entry and checking. The system thus has much potential for short duration surveys.

2. *The user's perspective.*

For the user's perspective, the general issue is how well did the system suit the interviewers and did it increase their efficiency without adversely affecting their work?

(a) *Structure.* Were the users given adequate support for their work?

The training in the first week was adequate as an initial introduction and the Psion proved easy to learn. In a more established study, the interviewers should have further training after they had mastered the basics, for example, to learn to replace batteries and use other outer level menu functions. The menu-based software proved easy to learn and use for those with no previous computing experience. When using FDCS, the focus of work shifted completely to the field, no additional checking by the interviewers was necessary before the data was 'handed in'.

(b) *Process.* Did FDCS satisfy the interviewers in their day to day operations?

A change in the attitudes of the interviewers can be observed during the study. Although their initial expectations were highly individualistic ranging from conservatively pessimistic to optimistic, they all tended to express opinions more in favour of the Psion at the end of the study. Some of the things that were disliked by the interviewers were related more to trial management than intrinsic weaknesses in the software system, for example, a full storage device with no more storage space for questionnaires. The various things listed as positive by interviewers covered all the expected benefits - automatic question skipping, error checking, automatic recording of time and date, as well as unexpected by-products, the interviewer only needed to carry one instrument considerably lighter than paper questionnaires and easier to handle when interviewing. Some of them found the system initially difficult, and these were the most resistant to its use at

the start. At the end of the survey, all of them could mention how some aspect of the software could help their work, and none expressed an opinion against the use of the Psion.

(c)   *Outcome.* Did the use of FDCS improve the quality of the interviewers' work?

Overall, it would seem, from the point of view of the interviewers, that once they had become familiar and confident with the system, it resulted in few negative changes in their jobs and improved their efficiency. As their work output, interviewers could return data on a chip instead of a pile of heavy papers, which most mentioned as a bonus.

3.   *The respondent's perspective.*

From the respondent's perspective, very little of the structure changed. The use of the system made no difference to the normal interviewing pattern or procedure. As to the process, respondents did not object to the use of the computer and became noticeably more enthusiastic when its use was explained to them. The main outcome was that the length of the weekly interviews was reduced some-what.

4.   *Expected viewpoint of administrator.*

The field work did not include a detailed investigation of the opinions or attitudes of the survey administrators or planners. The researcher very largely occupied this role. However, it is possible to put forward the following general observations.

(a)   *Structure.* Has the system proved to be cost-effective, and a means of maximising resource use?

A system for 4 interviewers would need 5 fully equipped Psions as well as a mains adapter and a communications link, amounting to an initial expenditure of £1,200. Batteries could be seen either as a recurring expense if Duracells are

bought, at the cost of £1.50 each, £6.00 per survey week. Alternatively, two sets of rechargeable batteries could be bought for each machine for £7 per battery, and a recharger for £14, a total of £84. Balanced against the cost of one data entry clerk on a maximum salary of £50 per month, the initial costs can be paid for after a minimum of 2 years. Computer supervisor time is hard to measure and is taken here as negligible. While these figures may look bleak, clearly the value of improved data quality and the possibility of faster response needs to be added to the equation.

As a result of this trial, a number of ways of using the Psion to improve survey management have been discovered. In using such a system, the manager would tend to develop general computing experience, which is a vital part of the modern data processing cycle.

(b)   *Process.* Has the system been shown to be appropriate and reliable?

As a field trial, the work reported is intended to show flaws as well as successes. Some measures against data loss were identified and implemented: as a precaution, the user was not allowed to change the questionnaire data after it has been stored, although the contents of the file could still be browsed through. This was done to ensure that there was no tampering, intentional or unintentional, with the data after the interview. However, two easily identifiable areas for improvement remain: (i) software design - incorporating a step by step method of conducting data transfer. This stage can never be entirely fool proof if volatile storage devices are used, as the user must always refrain from plugging it in before connecting the communications link. There is a possibility of using improved technical facilities, and newer storage devices such as smart cards. Data security might be guaranteed if other types of storage devices are used but this might require more battery power. (ii) in management - implementing a battery backup system by giving each interviewer a spare replacement battery.

The analysis of error rates and interview lengths between the two methods has shown that FDCS reduces both of these. Although not all of the interviewers are convinced of this, many respondents agree with this. However, the coverage rate during the trial decreased when the computer was used, although this could

perhaps be linked to the duration of battery life, less field time being available for recall visits. The supervisor's role in survey management changes slightly with the use of the Psion, requiring less overall checking of the finished questionnaires, but also a new skill to be acquired in transferring data (although this could still be allocated to the computing centre).

Properly managed, FDCS could be a significant improvement in the process of data collection and management. It would encourage more sophisticated sample design, allowing more adaptability in questionnaire design, testing and modification.

(c)   *Outcome*. How well does CADAC meet the requirements for health information gathering?

CADAC systems produce better quality data in less lead time. The computer-readable data in ASCII format feeds naturally into analysis software. FDCS offers a quick start-to-end way of conducting surveys in rural areas of developing countries, helping to make the survey process more easily managed. It also has some implications for more general health monitoring activities in the long-term.

5.   *Expected impact on the health system.*

(a)   *Structure*. How does it affect the roles of different health personnel? The use of the system could reduce the data collection load on paramedical staff if it encouraged the demand and use of specific items of data for management purposes. It would enhance the role of data processing personnel.

(b)   *Process*. How could it affect the quality of health provision? The system offers a method for generating data of an acceptable quality for various types of surveys, including short, one-off surveys. By improving the accuracy and timeliness of data, problem-oriented policies could be based on the information produced. Data might be produced to assess the extent of a problem, but also to evaluate the effectiveness of any interventions proposed and implemented.

(c)   *Outcome*. How does it improve the health status of the population? Administered well, FDCS could ultimately improve the overall health status of a population if the cycle of demanding data, using that data to design and implement policies, which in turn produce more data, is maintained by decision-makers.

4.9   CONCLUSIONS

The malaria morbidity questionnaire used in this study is relatively straightforward, being short and without many complicated skips. That the system has proved to reduce errors in this type of survey is a major achievement. In other situations with more potential for error (for example, longer questionnaires with more skips, or those that are administered over a short survey period), its impact might be more dramatic.

The technology used in this study has worked quite successfully within a developing country context. Some of work reported here could have taken place anywhere in the world with similar results. The sense in which developing countries are special is that resource constraints and environmental factors make the implementation of informatics difficult. These then are the factors that constrain technology and software choice. At this level, the FDCS implementation on the Psion proved effective.

The use of the evaluation framework to summarise the results illustrates its power to highlight the main findings of the study. It also provides a firm basis for comparison with the results of the second study presented in the next chapter.

CHAPTER 5

EXPERT SYSTEM ON TROPICAL DISEASES:
DEVELOPMENT AND EVALUATION

5.1   PRELIMINARIES

ESTROPID, an Expert System for TROPIcal Diseases, was developed as a self-sufficient system that could be used by trained nurses or other similarly qualified health personnel during routine consultations. It is intended for a particular environment, the primary health care clinic in tropical Africa. The system is designed to deal with the spectrum of diseases that can be treated at the out-patient level, and that require a minimum amount of clinical skills to understand the medical procedures and investigations it requests. The aim of the development is to provide diagnostic support and management advice, to help improve service to patients, and encourage better resource usage and management at the primary health care level.

The primary focus of the evaluation given here is to assess the system's potential for clinical benefit. The scope of the evaluation, and consequently the interpretation of the results, is limited by several factors. The trial only included one user (one portable computer was available, and no other clinical officers were free to participate on the project on a full-time basis). Only two weeks could be allocated to the trial by both the clinical officer and the doctor. However, these constraints do not undermine the aims of the thesis, which are exploratory rather than definitive in nature. The trial, although small, raises interesting questions in this research area which are discussed in this chapter.

5.2   THE DESIGN OF ESTROPID

The system was developed in muLISP-86, to be used on any IBM-compatible microcomputer. It is best described in four segments: the knowledge base, the inference mechanism, the interface, and how all these three components are used in its operation.

1.   *The knowledge base*

The knowledge used in ESTROPID is solely extracted from the Essex flowcharts [Essex, 1980]. This provided a handy, concise and compact collection of medical knowledge, whose development was supported by the World Health Organisation. The set of 64 flowcharts each comprise of a chart heading, a series of decision boxes with signs or symptoms to ask for, with two arrows (one for positive answers, and the other for negative ones) leading to either another decision box or a diagnosis box (Figure 5.1).

This seemed an ideal source of knowledge for a computer-based system, and indeed ESTROPID is not the first medical decision-aid system to exploit Essex's work [Auvert *et al.*, 1986a]. The knowledge base stores the contents of Essex's flowcharts in a series of nested lists which comprise of the question texts for each decision box, and a unique identifier for each question made up of the chart number, box number, and additional code to show the ordering (Figure 5.2). An indexing structure was also created to enable rapid access to question texts on the basis of their keywords. In this way, the inference engine could traverse the tree and specific questions could be located more quickly.

2.   *The inference mechanism*

The knowledge base structure described above provides a representation of Essex's work. In order to operationalise this in a computer program, it is necessary to provide an inferencing mechanism to navigate through the question texts and to arrive at a conclusion (diagnosis).

56 Shock with severe abdominal pain

Refer to hospital ☐   Admit to the nearest health centre or hospital △   Treat in outpatient clinic ○

| Trauma AND shock | YES → | Ruptured kidney or spleen | ☐ | 56.1 |

NO ↓

| History of missed period | YES → | Fever AND bad smelling bloody vaginal discharge | YES → | Septic abortion | ☐ | 56.2 |

NO ↓

| Early pregnancy | YES → | Ectopic pregnancy | ☐ | 56.3 |

NO ↓

| Late pregnancy hard painful uterus | YES → | Antepartum haemorrhage | ☐ | 56.4 |

NO

| Female of childbearing age | YES → | Dark blood from vagina | YES → | Ectopic pregnancy | ☐ | 56.5 |

NO

| Urine can be tested for sugar | YES → | Sugar in urine | YES → | Diabetes | ☐ | 56.6 |

NO

| Sickle cell disease present in population | YES → | Jaundice without bile in urine | YES → | Sickle cell crisis | △ | 56.7 |

NO

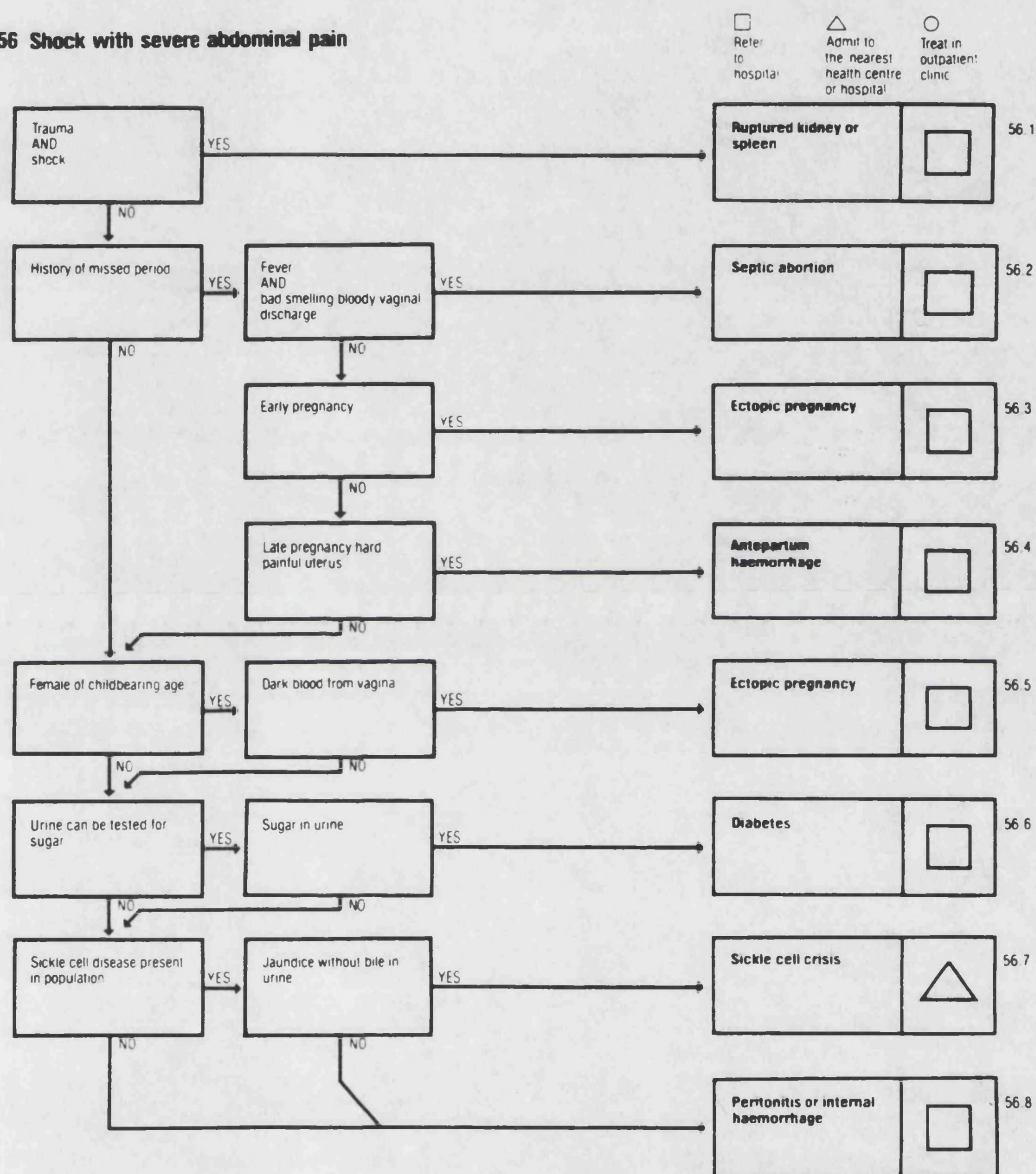| Peritonitis or internal haemorrhage | ☐ | 56.8 |

FIGURE 5.1 : A flowchart from Essex

```
(SETQ
...

C56 '(
(TRAUMA S5601A2)
(SHOCK S5601B2)
(HISTORY OF MISSED PERIODS S5602)
(FEVER S5603A2)
(BAD SMELLING DISCHARGE FROM VAGINA WITH BLOOD S5603B2)
(IN FIRST THREE MONTHS OF PREGNANCY S5604)
(IN LAST THREE MONTHS OF PREGNANCY WITH HARD PAINFUL UTERUS S5605)
(OF CHILDBEARING AGE **ZS=F** S5606)
(DARK BLOOD FROM VAGINA S5607)
(URINE CAN BE TESTED FOR SUGAR S5608)
(SUGAR IN URINE S5609)
(SICKLE CELL DISEASE IN POPULATION S5610)
(JAUNDICE WITHOUT BILE IN URINE S5611) )

...
)
```

Figure 5.2 : Static knowledge representation in nested list structure

```
(DEFUN FC56 ()
(IF (SAND S5601A2 S5601B2)
   (DIAGNOSIS RUPTURED KIDNEY OR SPLEEN 1)
(IF (SIS 5602)
   (IF (SAND S5603A2 S5603B2)
      (DIAGNOSIS SEPTIC ABORTION 1)
      (IF (SIS S5604)
         (DIAGNOSIS ECTOPIC PREGNANCY 1)
      (IF (SIS S5605)
         (DIAGNOSIS ANTEPARTUM HAEMORRHAGE 1) ))) ))
(IF (CONNECT)
(IF (SIS S5606)
   (IF (SIS S5607)
      (DIAGNOSIS ECTOPIC PREGNANCY 1) )))
(IF (CONNECT)
(IF (SIS S5608)
   (IF (SIS S5609)
      (DIAGNOSIS DIABETES 1) )))
(IF (CONNECT)
(IF (SIS S5610)
   (IF (SIS S5611)
      (DIAGNOSIS SICKLE CELL CRISIS 2)
   (DIAGNOSIS PERITONITIS OR INTERNAL HAEMORRHAGE 1) )
(DIAGNOSIS PERITONITIS OR INTERNAL HAEMORRHAGE 1) )))
```

Figure 5.3 : Flowchart dynamics as represented in a function

Each of Essex's charts also has a separate decision function specified in the knowledge base representing the dynamic flow structure through the decision boxes using IF..THEN..ELSE constructions (Figure 5.3). Given a set of symptoms to start with, the system evaluates the keywords and decides on an order of access to the knowledge base, using the frequencies with which a chart is mentioned. It ranks the chart referred to most frequently the highest, and this is then accessed first. It confirms the heading of the chart, uses the appropriate decision function, accesses any other functions indicated (but keeping track of those accessed so no circular path is followed) until it reaches a diagnosis. This is thus a 'forward chaining' mechanism, starting from symptoms, which are then used to confirm a particular diagnosis.

3.   *The interface*

The interface is based around the use of a keyboard, allowing text entry, numeric options from a menu list, and Y/N answers. It is a fairly basic input facility, illustrating none of the sophisticated pull-down menu facilities that are currently available in many modern programming environments. There are no help facilities, as it was felt the questions were to be fairly straightforward. No options for backtracking through questions are allowed, the system therefore does not offer opportunities for mistakes to be corrected during a session.

4.   *Operation*

The system commences operation by requesting basic information about a patient. Then a series of menus, based around the broad groupings used by Essex [1980], prompts for the presenting symptoms mentioned by the patient. In some cases the symptoms are not in the menu lists and an additional text entry facility can be used, although this is quite limited. In this mode, the user types out the symptoms in simple English; if the set of words used is not matched directly in the knowledge base, the system then proceeds in an interrogatory fashion, to ask all the questions in its knowledge base that contain the keywords of the user entry. If none of these are confirmed, the user is informed that the entry is not understood. The system operator chooses when to stop entering symptoms and requests a move to diagnosis. The user is then asked a series of questions

requiring Y/N answers. The outcome of the system is a list of diagnoses, as well as the advice given in Essex on whether the patient can be treated at the clinic, or requires hospital-based treatment. As such the system does not take diagnosis onto treatment.

5.3   BACKGROUND TO THE CASE STUDY

The Republic of Kenya is located on the east coast of Africa. Straddling the equator, and spread over 583,000 sq. km, the country experiences a variety of climatic conditions, ranging from desert in the north to temperate highlands in the central region. The country has two rainy seasons, the long rains are between March and May, and the short rains in November and December.

Kenya became independent from the British in 1963. The party involved in the independence movement has remained in power ever since, changing the presidency from Jomo Kenyatta to Daniel arap Moi in 1978. Politically, the country operates a one-party state that allows controlled electoral choice over parliamentary candidates. Administratively, the country is divided into 8 provinces, further divided into 41 districts composed of locations and sub-locations.

Three-quarters of Kenya's 22 million people live in rural areas, but the urban drift is gaining momentum, currently at almost 9% per year. The annual population growth rate is slowing down, but at 3.9%, the population is still set to double within twenty years. Again, the age structure is characteristically young, with approximately half of the population under 15 years of age. 94% of all children are enrolled in primary school, and one in five children attend secondary school. Almost two-thirds of adults are literate.

The economy is heavily dependent on agricultural exports, the major agricultural products are coffee, maize and tea. Today, it produces slightly less food per capita than it did at the turn of the decade, and the country receives over a 100 thousand metric tons of cereal food aid. The country's mineral wealth is poor and a fifth of its total imports is spent on fuel. It has a relatively strong industrial

base, contributing 30% of its gross domestic product (GDP). Kenya exports manufactured goods to a value of $700 million in 1987, most of which are to middle income and other low-income countries. A large proportion of foreign exchange is generated by a growing tourist industry. GNP per capita is $330, but has recently been decreasing, partly as a result of increased population pressure. Total long-term external debt runs at $4978 million, about two-thirds of its GNP. The amount of aid currently stands at $565 million, most of which (71%) is in the form of block grants from donor countries and agencies.

Life expectancy at birth is 58 years. Infant mortality has dropped dramatically by almost half since independence, and is the lowest among low-income African countries at 74 per 1000 live births. The country has achieved relatively high immunisation coverage rates ranging between 60% to 86% for children before their first birthday. Causes of morbidity and mortality in childhood are highly dependent on geographical areas, but malaria and respiratory infections are dominant. There is one physician to 10,100 people, and one nursing person to every 950 people in 1984. Of total government expenditure in 1980, almost 8% was on health.

The Kilifi District Hospital is one of two general hospitals serving the Kilifi District. It provides a range of services including out-patients, immunisation, maternal and child care, oral, orthopaedic, gynaecology, ophthalmology, physio-therapy and occupational therapy clinics. The hospital is staffed by several clinical officers (with three years of medical training, half of that required for formal medical doctors) and nurses of varying specialties.

The Kenya Medical Research Institute (KEMRI) Coastal Unit is part of an indepen-dently run, government-sponsored research organisation set up in collaboration with the Wellcome Trust. Currently research on malaria is being carried out to investigate clinical and socio-economic factors of the disease presentation and progression as well as the cultural aspects and understanding of the disease and steps taken for treatment. The unit is based on the grounds of the Kilifi Hospital, and works closely with the hospital's staff.

5.4   THE TRIAL DESIGN

The trial of ESTROPID was based at the Kilifi District Hospital and conducted in collaboration with doctors at the KEMRI Coastal Unit during November 1989. A clinical officer was trained to use ESTROPID on a Toshiba T1200 over a 3 day period. Training included a general introduction to the computer, the aim of the system, the basis of its knowledge, as well as some trial runs using hypothetical cases.

Normally, patients are seen by either a doctor, or more generally, a clinical officer (CO). For this study however, patients were seen by both. In this way, the diagnosis given by the CO using either method could be compared with that given by the doctor. In a pre-run pilot study, 10 patients were seen by both the clinical officer and a doctor (normal consultations), and 10 more by both, but with the clinical officer using ESTROPID (computer consultations). The following patient flow procedure was tested and then used for the main study (Figure 5.4):

1.   Patient arrives at the out-patient clinic and joins the queue. The patient at the top of the queue sees the next CO available, who may be the study CO, or one of the others on duty that day. Patients who go to any of the other clinical officers are out of the trial.

2.   Number allocation. Clerk takes down patient details, name, age and sex, and allocates the next serial number. The patient is given a paper slip with this number. If there are two or more patients to see, for example two siblings and a parent, all are allocated individual numbers.

3.   Consultation with clinical officer. The CO notes the time of arrival on the study form (Figure 5.5), fills in the rest of the form as the consultation proceeds, and enters the time the consultation ends. For normal consultations, the patient's complaints, examination notes, diagnoses and treatments are recorded; the complaints, and diagnoses only are recorded for computer consultations.

4.   Consultation with doctor. The doctor requests and keeps the patient's number slip, fills in the study form for the patient, prescribes, and also writes a consultation summary on the patient health card. If necessary, he also requests further tests.
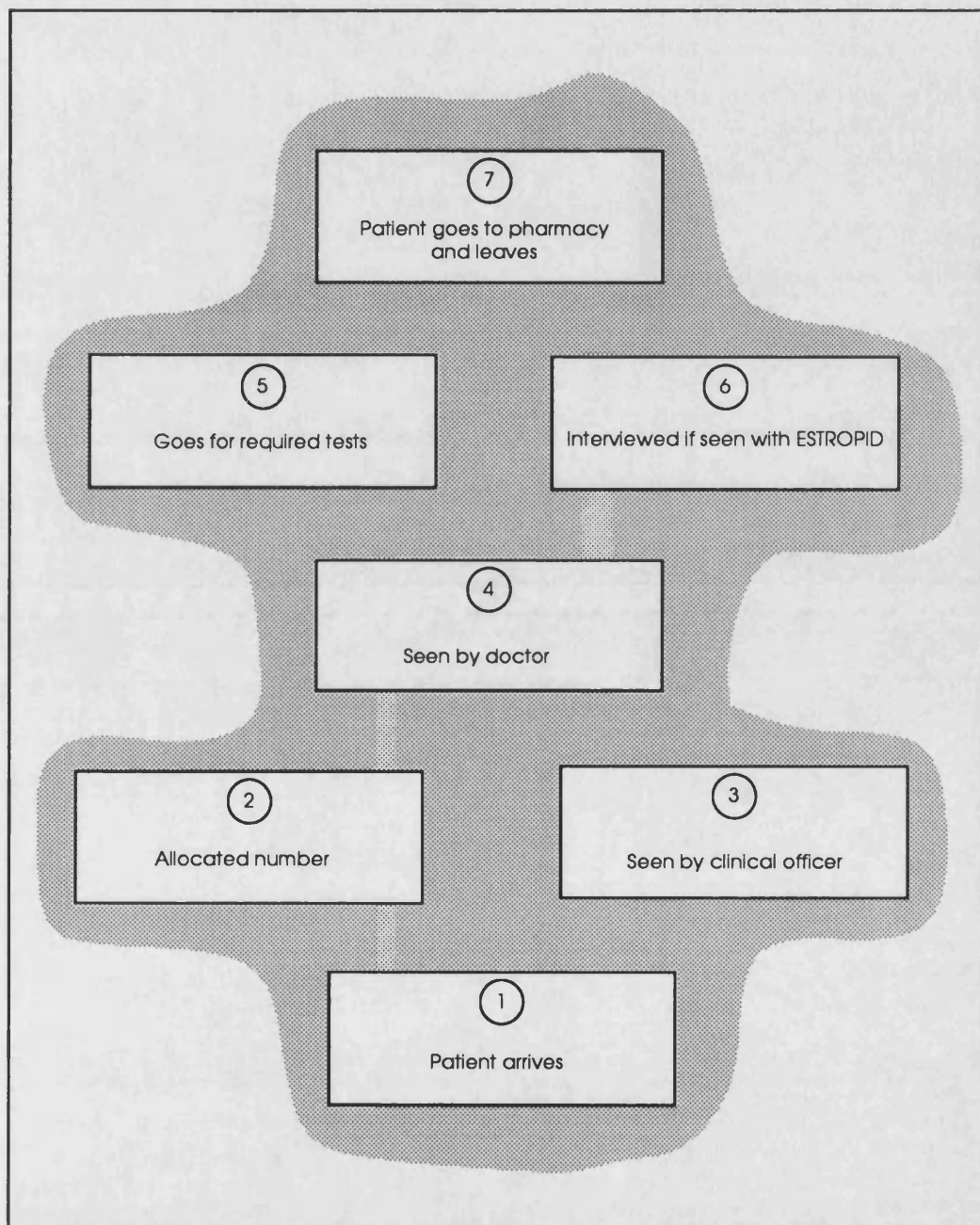
FIGURE 5.4 : Patient flow procedure

COMPUTER-AIDED DIAGNOSIS STUDY - PATIENT DETAILS FORM

1.   Date   :_:_: / :_:_: / :_:_:

2.   Time at start of consultation   :_:_: : :_:_:

3.   Patient study number   :_:_:_:

4.   Patient 's firstname   :_:_:_:_:_:_:

5.   Patient's lastname   :_:_:_:_:_:_:

6.   Sex   :_:

7.   Age   :_:_: yrs :_:_: mths

8.   Major complaints
     1.   ....
     2.   ....
     3.   ....
     4.   ....
     5.   ....

9.   Notes
     ....
     ....
     ....
     ....
     ....
     ....
     ....

10. Diagnosis
     1.   ....
     2.   ....
     3.   ....

11. Treatment
     1.   ....
     2.   ....
     3.   ....
     4.   ....

12. Time at end of consultation   :_:_: : :_:_:

13. Completed by
     (1 = Clinical Officer, 2 = Doctor)   :_:

FIGURE 5.5 : Form for filling in details of patient consultation

PATIENT QUESTIONNAIRE

1.   Have you been to an outpatient''s clinic before?

2.   Did you notice anything different this time?

      If yes,

3.   What did you notice?

4.   What do you think it is used for?

FIGURE 5.6   :   Questionnaire to interview patients seen with ESTROPID

5.   Further tests. The patient may be required to have a confirmatory blood test, urine, stool, haemoglobin, or any other rapid tests that could be provided by the laboratory. The patient returns to the doctor with the results.

6.   Patients consulted with ESTROPID are interviewed using the short questionnaire shown in Figure 5.6.

7.   The patient reports to the hospital pharmacy for the prescribed drugs and leaves.

The main study took 8 days, with the clinical officer seeing patients using ESTROPID on the first and every alternate day. On the other days, he saw them using normal methods. The clinical officer and doctor worked in different buildings and were blinded to each other's diagnoses and management recommendations.

Coding schemes were devised with the help of a paediatrician, familiar with the working of the clinic, for each section of the form (Figures 5.7 to 5.9, which include only the symptoms, diagnoses, and treatments used during the study). The coded information was then entered onto a database. Comparisons were made between diagnoses and treatments (where applicable) and the number of matching ones noted. The differences were assessed by the paediatrician, who judged the disagreements on a scale of 0 to 4. The clinical officer was interviewed at the end of the trial and a number of normal consultations reviewed and assessed by repeating the consultation on ESTROPID using the notes.

5.5   RESULTS - THE SYSTEM'S FUNCTIONING

At the start of the trial, it was estimated that 50 patients could be seen in a day. Therefore, the goal was for 400 patients, half of whom would be children and the other half adults. As a result, all adults who presented sick whilst bringing their children to the paediatric department were also seen. In practice, however, only 25 patients could be comfortably seen in a day, and in the time available, the initial goal could not be met. In the data presented here, the adult consultations will be excluded as they form too small a sub-group to demonstrate

Code	Symptom	Code	Symptom
01	fever/sweating	32	itching/scratching
02	loss of appetite	33	blisters
04	pica (eating soil)	34	scabies
05	general body pain	36	fits
06	oedema	37	headaches
07	swollen glands	38	dizziness
08	increased appetite	39	hypopigmentation
09	lethargy/body weakness	40	swelling/lumps/abscesses
10	cough	41	cuts
11	difficulty breathing	42	burns
12	sneezing/runny nose	43	sore mouth
13	sore throat/pain in throat	50	chest pain
14	ear ache/ear problems	51	dysuria
15	ear discharge	52	palpitations
16	eye problems/red eye	53	joint pains
17	eye discharge	54	toothache
18	pain in eye	55	backache
19	purulent sputum	56	neck pain/stiff neck
20	nausea	57	trauma
21	vomiting	60	vaginal discharge
22	diarrhoea	61	dark urine
24	mucus stools	62	red urine/blood in urine
25	abdominal distension	63	ulcers on external genitalia
26	abdominal pain/colic/ache	64	swelling on genitalia
27	constipation	65	jaundice
28	anal irritation	70	nose bleeds
29	worms in stool or vomit	80	ingestion
30	sores or lesions on body	90	miscellaneous
31	rash		

FIGURE 5.7: Coding sheet for symptoms

Code	Diagnoses	Code	Diagnoses
01	malaria	40	scabies
02	anaemia	43	tineasis
03	hepatosplenomegaly	45	impetigo
04	splenomegaly	46	measles
05	marasmus	47	other skin disease
06	kwashiorkor	48	septic cord
07	viral disease	49	eczema/dermatitis
08	scurvy	50	cellulitis
09	allergy	51	abscesses
10	URTI/flu	52	lymphadenopathy
11	rhinitis	53	burns
13	tonsillitis	54	laceration
14	LRTI/acute resp infection	55	mumps
15	(broncho)pneumonia	57	neonatal infection
16	broncholitis/bronchitis	58	herpes simplex
17	bronchospasm/asthma	59	acute nephritis
18	whooping cough	60	febrile convulsions
19	TB	61	meningitis
20	otitis media (acute)	62	epilepsy
21	otitis externa	63	congenital heart disease
22	eye infection/conjunctivitis	70	septicaemia
23	gingivostomatitis	71	hepatitis
24	oral thrush	72	glandular fever
25	chronic otitis media	75	fractures
26	ear problem	76	soft tissue injury
27	dental infection	80	ingestion
28	allergic conjunctivitis	81	blisters
29	mastoiditis	82	bacterial infection
30	gastroenteritis	83	schistosomiasis
31	diarrhoea	84	sickle cell disease
33	worms	90	miscellaneous
35	constipation	91	neonatal jaundice
36	peptic ulcer/dyspepsia	93	hepatomegaly
37	dehydration	92	malnutrition (unspecified)
38	renal colic	98	no diagnosis made
39	urinary tract infection		

FIGURE 5.8: Coding sheet for diagnoses

Code	Treatments	Code	Treatments
01	paracetamol/panadol	77	water and glucose mixture
02	aspirin/asa	78	mist expectorant
10	chloroquine	80	benzyl benzoate
12	fansidar	81	gentian violet
20	crystapen/pen g	83	whitfield ointment
21	pen v	86	oral rehydration solution
22	ampicillin	89	sennokot
23	amoxil/amoxycillin	93	hydrocortisone
26	procaine penicillin	96	adrenaline
27	triplopen	98	sofradex
31	metronidazole	A1	plasma protein fraction (PPF)
32	septrin	B1	epanutin/phenytoin
33	tetracycline	B2	phenobarbitone
34	nystatin	C1	valium/diazepam
35	chloramphenicol	D1	tepid sponging
36	allerdex	D2	dress wound
50	ketrax	D3	clean with spirit
52	mebendazole	D4	incise and drain abscess
55	metrifonate	D5	plaster of paris
60	other antacid mixtures	E1	tetanus toxoid
62	magnesium trisilicate	M1	admit
64	buscopan/hyoscine	M2	dietary advice
65	piriton	M3	hygiene advice
70	actifed	M4	blood transfusion
71	franol	M5	malnutrition clinic
72	ferrous sulphate/fersolate	M6	dental clinic
73	folate/folic acid	M7	plenty of fluids
75	glucose	M8	phototherapy
76	multivite	M9	verbal reassurance

FIGURE 5.9: Coding sheet for treatments

representativeness of adult out-patients. Overall, 180 children and 20 adults were seen. The age distribution of the patients is shown in Table 5.1.

Age	Frequency	Cumulative percentage
0	47	23.5
1	45	46.0
2	35	63.5
3	21	74.0
4	11	79.5
5	6	82.5
6	3	84.0
7	2	85.0
8	2	86.0
9	2	87.0
10	3	88.5
11	0	88.5
12	2	89.5
13	1	90.5
18+	20	100.0

TABLE 5.1 : Age distribution of children seen in trial

All of the clinical officer's consultations were longer than those of the doctor. Table 5.2 shows that the overall mean for the CO's computer consultations was 8.3 minutes compared with 5.0 minutes when the same set of patients was seen by the doctor.

		Attempt at using the method								Overall mean	
		1st		2nd		3rd		4th			
Computer	Clin Off	8.5	(1.9)	8.5	(2.4)	7.3	(1.5)	8.9	(3.1)	8.3	(2.2)
	Doctor	5.5	(2.4)	4.9	(1.5)	5.4	(2.3)	4.9	(1.3)	5.0	(2.0)
Normal	Clin Off	7.0	(1.8)	8.0	(2.1)	6.6	(2.0)	6.4	(2.0)	7.0	(2.0)
	Doctor	5.4	(1.4)	5.2	(1.6)	4.2	(1.0)	5.2	(0.7)	4.9	(1.6)

LEGEND
( ) - standard deviations

TABLE 5.2 : Mean consultation lengths for groups of patients seen on different days

For normal consultations, the corresponding means were 7.0 minutes and 4.9 minutes. Both of these differences between means were significant ($p < 0.05$) when a paired t-test was done. An unpaired t-test was used to compare the computer

and normal groups seen by the clinical officer. This was significant, confirming that the computer consultations took longer than the normal ones. As expected, the same test was not significant for the doctor's consultations, the length of all consultations were uniform on average regardless of whether the patient was previously seen with a computer or not, confirming that no bias existed. The already existing difference between the consultation lengths of the two clinicians was increased when ESTROPID was used. No obvious trend can be observed to demonstrate a learning effect as the trial progressed.

The difference between the mean number of symptoms recorded by the CO and that of the doctor when the computer was used was 0.5 (Table 5.3). This difference was significant, showing that more symptoms were recorded by the CO than the doctor. When normal methods were used, the difference was 0.2 and not significant. The unpaired t-test comparison between the CO's computer and normal consultations was significant and that of the doctor not significant. These results suggest a checklist effect on the CO's symptom recording when ESTROPID is used, i.e. by prompting for symptoms, ESTROPID elicits a higher number of symptoms than when normal methods are used.

		Attempt at using the method								Overall total		Overall mean
		1st		2nd		3rd		4th				
Computer	Clin Off	41	[27]	82	[61]	88	[57]	76	[48]	287	[193]	3.2 (1.0)
	Doctor	33	[27]	78	[61]	71	[57]	57	[48]	239	[193]	2.7 (0.9)
Normal	Clin Off	60	[49]	73	[56]	58	[45]	51	[36]	242	[186]	2.7 (0.9)
	Doctor	60	[49]	63	[56]	58	[45]	46	[36]	227	[186]	2.5 (0.9)

LEGEND
( ) - standard deviations   [ ] - number of matching symptoms

TABLE 5.3 : Total number of symptoms recorded and means for groups of patients seen on different days

The symptoms recorded by the CO and the doctor intersect for 73% of the total recorded symptoms when the computer was used, and 79% when normal methods were used (Figure 5.10). A t-test on the difference between these proportions was significant.
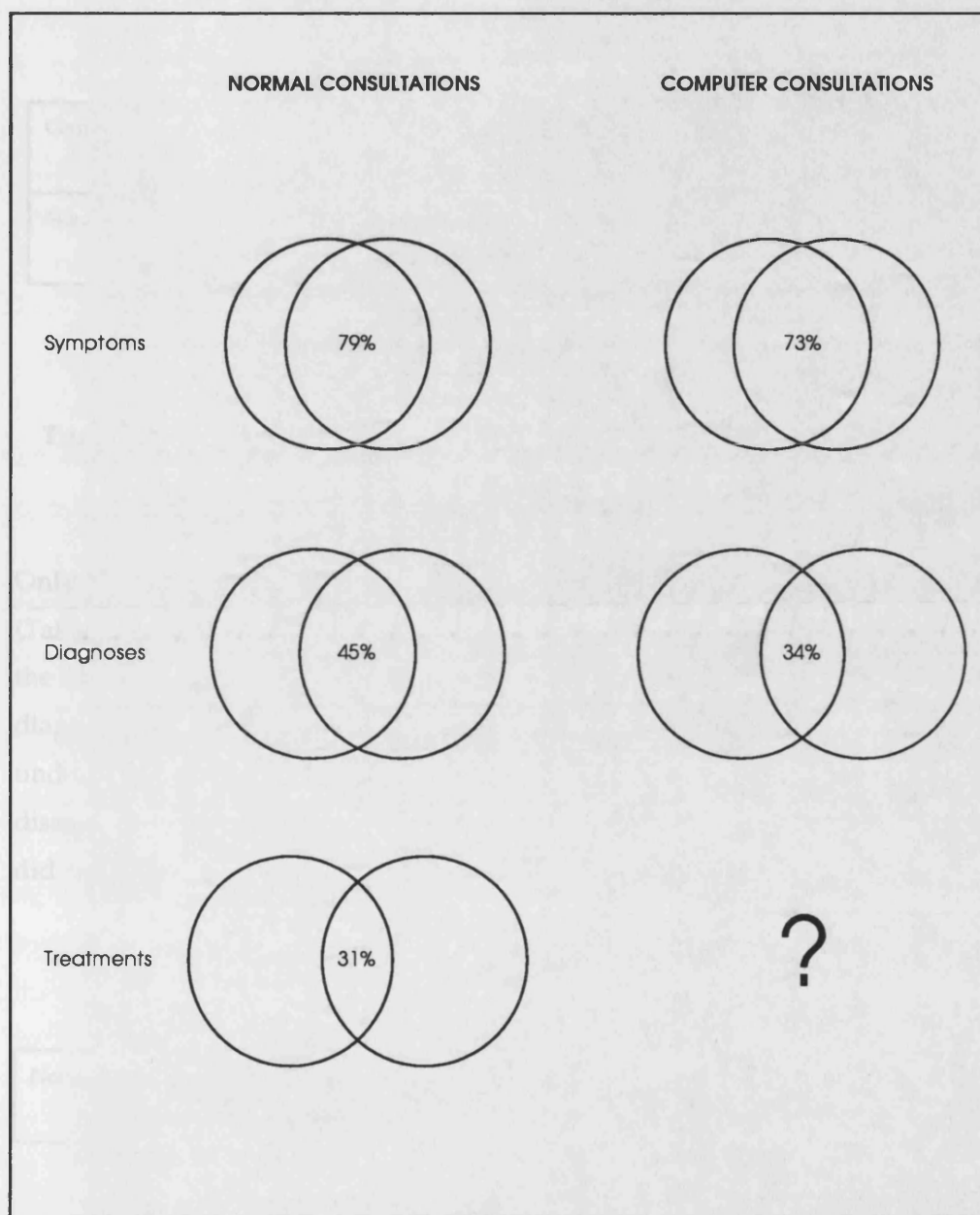
**FIGURE 5.10   :   Comparison between recorded items for both methods
- considering matching items**

The corresponding figures for the diagnoses recorded was 34% for computer and 45% for normal (Table 5.4). This difference was also significant.

		Attempt at using the method								Overall tot.	
		1st		2nd		3rd		4th			
Computer	Clin Off	31	[12]	52	[14]	52	[17]	56	[16]	191	[59]
	Doctor	21	[12]	48	[14]	40	[17]	43	[16]	152	[59]
Normal	Clin Off	43	[18]	54	[24]	50	[21]	36	[14]	183	[77]
	Doctor	39	[18]	44	[24]	43	[21]	34	[14]	160	[77]

LEGEND
[ ] - number of matching diagnoses

TABLE 5.4 : Total number of diagnoses recorded and matching for groups of patients seen on different days

Only 31% of all the treatments recorded for the normal consultations matched (Table 5.5). The summary data in Figure 5.10 shows that during normal practice, the 21% variation in symptom recording is followed by a 55% variation in disease diagnoses, and a 69% variation in treatment prescriptions. This run seems understandable, since a disagreement early in the flow must tend to lead to disagreements later, and thus the variation will build up. In this trial, ESTROPID did not improve these figures.

		Attempt at using the method								Overall mean	
		1st		2nd		3rd		4th			
Normal	Clin Off	60	[14]	70	[23]	62	[22]	43	[21]	235	[80]
	Doctor	70	[14]	74	[23]	69	[22]	61	[21]	274	[80]

LEGEND
[ ] - number of matching treatments

TABLE 5.5 : Total number of treatments recorded and matching for groups of patients seen on different days

An earlier decision to compare diagnoses assuming that of the doctor to be right was abandoned after discussion with several doctors at the KEMRI Unit. Even for out-patient diagnoses, the assumption that a doctor would be right for almost all

cases was false and unlikely to be useful. It was infeasible to use the expert panel method suggested by other authors to reach a consensus diagnosis because of the time commitment demanded of several clinicians. A practical option was for one clinician to score each pair of diagnoses and treatments on a severity of disagreement scale as described here.

After coding and entry onto the computer, each set of diagnoses and treatments (where applicable) was displayed blind to the clinician on a computer screen as Person 1 and Person 2. He then made two scoring judgements for each pair. He first considered how severe a misdiagnosis Person 2 would have made if Person 1 was right; and then considered Person 1's error if Person 2 was right. Each score was made on a scale of 0 to 4: 0 meant no difference, 1 a slight difference, 2 a moderate difference, 3 a strong difference, and 4 a very strong difference. The scores for Person 1 were recorded as negative and those for Person 2 as positive. The modulo for each pair of scores was added up to give a single severity of misdiagnosis or mismanagement score for each patient. These combined scores are used in the analysis below.

The figures in the rest of this section are presented as the scale of disagreement per patient. Both the clinical officer and the doctor recorded equivalent symptoms for 26% of the patients seen first using ESTROPID. A higher percentage (37%) agreement was reached for normal consultations (see Table 5.6). Equivalence for the symptoms noted, however, was not a strong indicator for equivalence in diagnoses since in contrast to the overall 31% for symptoms, only 19% of all patients seen had indistinguishable diagnoses (Table 5.6a).

The marginal distributions of concordance levels of diagnoses for computer and normal consultations (Table 5.6b and 5.6c) shows no significant difference on a chi-squared test ($p < 0.05$). So, although the differences between the number of matching diagnoses recorded (as discussed for Table 5.4 above) is significant, when compared for the severity of misdiagnoses, the two types of consultations are similar.

Agreement on diagnoses

Agreement on symptoms	C.A.	M.A.	S.A.	C.D.	Marginal total
Complete agreement	10	24	19	3	56 (31%)
Moderate agreement	22	40	43	4	109 (61%)
Slight or no agreement	3	7	3	2	15 (8%)
Marginal total	35 (19%)	71 (40%)	65 (36%)	9 (5%)	180 (100%)

(a) Symptom and diagnosis comparison for all patients seen

Agreement on diagnoses

Agreement on symptoms	C.A.	M.A.	S.A.	C.D.	Marginal total
Complete agreement	5	9	7	2	23 (26%)
Moderate agreement	15	20	20	2	57 (63%)
Slight or no agreement	2	4	3	1	10 (11%)
Marginal total	22 (24%)	33 (37%)	30 (33%)	5 (6%)	90 (100%)

(b) Symptom and diagnosis comparison for computer consultations

Agreement on diagnoses

Agreement on symptoms	C.A.	M.A.	S.A.	C.D.	Marginal total
Complete agreement	5	15	12	1	33 (37%)
Moderate agreement	7	20	23	2	52 (57%)
Slight or no agreement	1	3	0	1	5 (6%)
Marginal total	13 (14%)	38 (43%)	35 (39%)	4 (4%)	90 (100%)

(c) Symptom and diagnosis comparison for normal consultations

Agreement on diagnoses

Agreement on symptoms	C.A.	M.A.	S.A.	C.D.	Marginal total
Complete agreement	5	6	1	1	13 (14%)
Moderate agreement	18	5	13	2	38 (43%)
Slight or no agreement	18	3	14	0	35 (39%)
Complete disagreement	2	1	1	0	4 (4%)
Marginal total	43 (48%)	15 (17%)	29 (32%)	3 (3%)	90 (100%)

(d) Diagnosis and treatment comparison for normal consultations

LEGEND
C.A. - Complete agreement
M.A. - Moderate agreement
S.A. - Slight agreement
C.D. - Complete disagreement

TABLE 5.6 : Patient concordance groupings for symptoms, diagnoses, and treatments
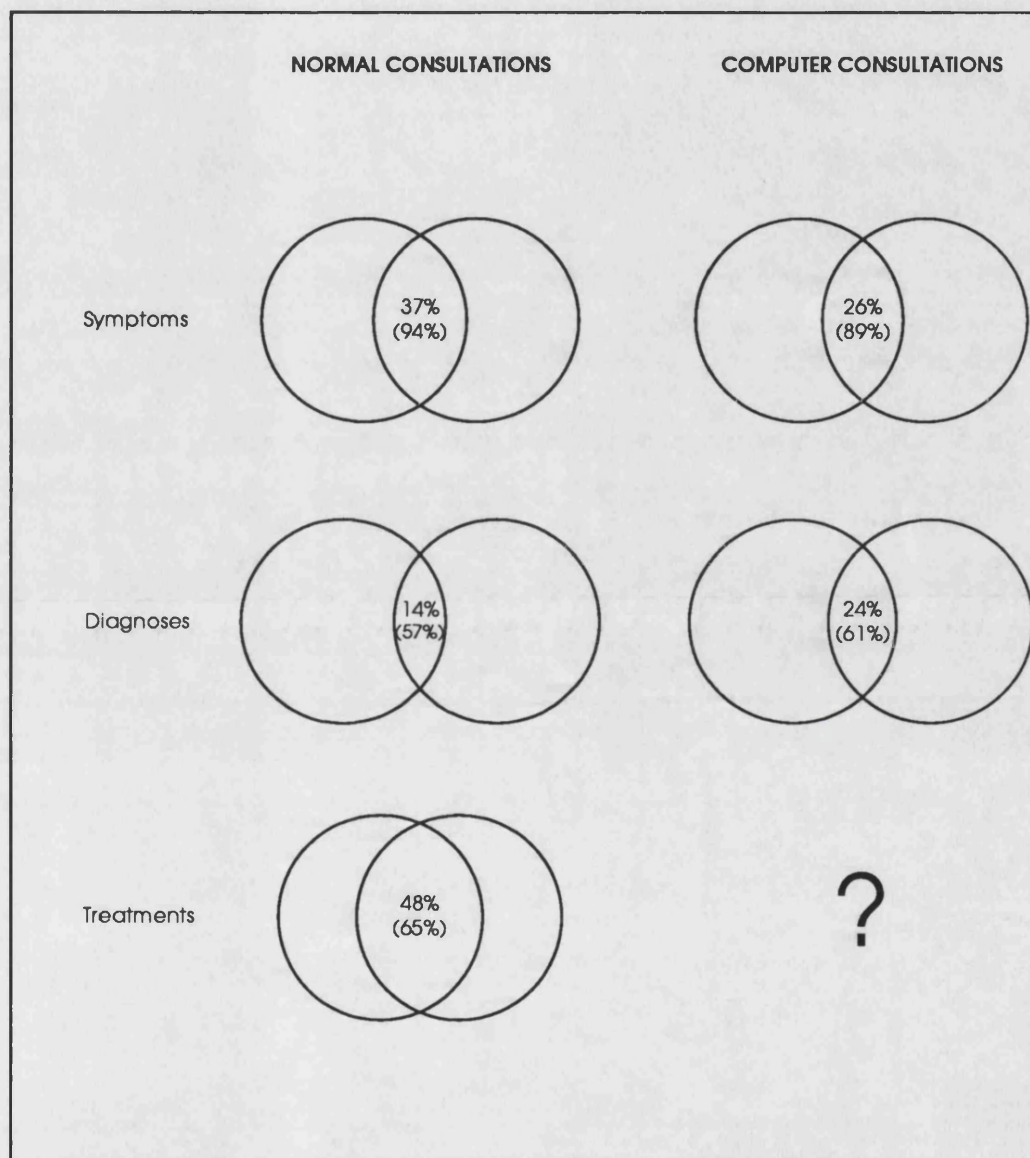
FIGURE 5.11 Comparison between recorded items for both methods - considering disagreement per patient

(Brackets contain combined `no disagreement' and `slight disagreement' figures)

Although only a sixth of the patients seen as normal consultations had no difference in diagnoses, almost half of all patients had indistinguishable treatments (Table 5.6d). In a very small proportion of patients (3%), the treatments suggested differed very strongly; differing strongly in an additional 32%. Figure 5.11 summarises the results obtained by basing comparisons on disagreements per patient.

There are wide discrepancies between the diagnoses and treatments recorded when a comparison is done on a strict equality basis - for normal consultations, a 45% agreement on diagnoses, followed by 31% agreement for treatments. When, however, a comparison is done matching equivalent diagnoses and treatments, a 14% agreement on diagnosis is followed by a 48% agreement on treatments. Adding on slight disagreement figures, 'acceptable' agreements rise to 57% for diagnoses, and 65% for treatments. A more accurate picture is drawn with the second analysis, showing that the final outcome differs only on approximately half the patients seen. The difference is strong and sometimes worryingly so for the remaining 35% of patients seen. It is towards reducing these differences in final treatment that diagnostic support aids should aim, if their impact is to be at all significant.

This section has described two ways of analysing the results between the diagnoses and treatments of CO and doctor. It illustrates a potential snare in presenting such comparative material, which can be used to produce seemingly different conclusions.

5.6   RESULTS - HUMAN PERSPECTIVES

The preceding section presented a technical analysis of the diagnostic problem, and the differences observed when ESTROPID was in use. This was not the only aim of the case study exercise, which was also intended to highlight the difficulties of implementing a medical decision-aid system in practice. In this section, the discussion is presented from the viewpoint of the clinical officer, and of the patients involved in the trial.

5.6.1   THE CLINICAL OFFICER

1.   Observed practice

Several of the consultations by the clinical officer and the doctor were observed
and transcribed if an interpreter was available. The consultation rooms were not
secluded or private and often had other people around, either fieldworkers or
nurses doing various other tasks in the same room. Thus it seems hardly likely
that the process of observation greatly changed the usual consultation practice.
Three consultations are mentioned here to illustrate their working practices.

As an example of a normal consultation, we have the following: (C=Clinical
Officer; M=Mother; *=some action)

C :   Who is sick - the child only?

M :   Yes.

C :   How old is he?

M :   1 year 3 months.

C :   What's wrong with the child?

M :   Coughing, diarrhoea, fever, and vomiting.

C :   When did it start?

M :   Yesterday.

*   Examination: chest auscultation, abdominal investigation.

C :   How many times did the child have diarrhoea yesterday, and how many
      times today?

M :   Yesterday, twice, today none.

C :   How many times did the child vomit today?

M :   None.

C :   Does the child have younger siblings?

M :   No.

C :   Where do you go for clinics?

M :   Vipengo.

*   The clinical officer consults the child's growth chart, a record of the child's
      weight since birth.

C :   Were you told about the graph and was the condition of the child
      explained to you?

M :   Yes, as the condition has not improved, we have to attend every Friday.

*   Fills out rest of form, and patient departs.

Here it can be noted that the clinical officer obtains the symptoms, performs a quick examination, probes for more information, peruses the child's growth chart and is ready to offer nutritional advice to the mother, all in the 6 minutes the consultation took. Any computer system has to match all of these services in an equivalent time span for similar, routine consultations.

As an example of a consultation with ESTROPID:

C :   How old is the child?

M :   Born last year.

C :   Are you the mother?

M :   Yes.

C :   What's wrong with the child?

M :   Crying at night, was burnt with hot porridge.

*   With this consultation, none of the symptoms could be entered into the menus. When entered as text, the information could not be used by the system, so it generated a series of questions to confirm any of the headings to the flowchart. The first time, the clinical officer skipped skin ulcer, and had to stop that consultation, and try again. On choosing the skin ulcer chart, not the most appropriate but the only reasonable sounding option, and answering the questions, the system pronounced "Bacterial infection".

This illustrates a situation in which the computer does not cover the symptoms, and was much more of a hindrance than a help. The child would obviously cry at night if it was unable to sleep or was in pain. The obvious management strategy is to treat the burn. The clinical officer could not find an easy way of entering this as a complaint and as a substitute, answered yes to a chart he thought would cover the symptom. Consequently, the computer pronounced an incorrect diagnosis.

These can be compared with one of the interviews observed for the doctor (D=Doctor):

D :   What is the child's name?

M :   ...

D :   What is the child's age?

M :   ...

D :   What's wrong with the child?

M :   Fever, cough.

D :   What else?

M :   Difficulty breathing, hot body during the night.

D :   What else?

M :   Swelling of the eyes, abdominal pain.

D :   For how long?

M :   3 days, she used to refuse to eat, but now she's eating well.

*   Examination, and prescription, then patient leaves.

The two example consultations of normal practice for both the CO and the doctor do not seem to differ significantly in their approaches for routine out-patient cases. The standard procedure is to elicit symptoms, probe for further details if necessary, examine the affected body system (i.e. respiratory system (RS), ear, nose and throat (ENT), cardiovascular system (CVS), central nervous system (CNS), and so on), and prescribe. One would reasonably assume that the approach differs with more difficult cases, and investigations will be related closely with the depth of the clinician's medical knowledge, understanding and experience. To investigate the consultation process further was beyond the scope of this study.

2.   *The post-trial interview*

The post-trial interview with the clinical officer reveals that he thought that ESTROPID was 'good' at diagnosing scabies, malaria, respiratory tract infection, broncho-asthmas, and measles. It was 'bad' at diagnosing pneumonia (when few of the signs of pneumonia had been asked for), bronchitis (only once was there a correct diagnosis), fluid loss (which it suggested for many patients with diarrhoea), and cholera (also generally overdiagnosed).

He thought a system like this could be of some use, but it needed a shorter diagnostic path. At present, many of the questions asked by the computer are of no use, for example, when a patient comes with diarrhoea and vomiting, the

computer may ask "is there any constipation?". In an out-patients department, the aim should be to reduce the time taken to attend to a patient, but the computer actually increased the time taken. Questions asked should be relevant to the particular complaint and the body system of interest, for example respiratory system or cardiovascular system. As it is, the computer is quite rigid and could hardly be used to see 90 to 100 patients per day. The system would be of even more use if it could clerk in patients automatically (kept the notes).

Asked about the system's usefulness for prescriptions as well as diagnoses, the CO gave the opinion that it would be better for the user to give the treatment because he knows the patient's condition best, but it would be good if the treatment suggested by the CO could also be printed out for the pharmacy.

The CO was asked to explain how the Vipengo clinic, where he was normally based, was different from the out-patient's clinic at Kilifi District Hospital, where the trial was held. The Vipengo clinic is several miles from Kilifi, and is more representative of the facilities to be found at many district clinics. He replied that the size of the hospital's out-patients department was much larger than that at his clinic. Vipengo's laboratory facilities are limited to routine investigations only. In clinics smaller than Vipengo, i.e. dispensaries, nurses, not clinical officers, often see patients. He did not think that computer-aided diagnosis could routinely be of any use to nurses at the clinic level, because they would rely too much on it, so that if it misdiagnosed, they would mistreat. In his opinion, these nurses have a superficial knowledge of medicine, and have been deployed to these centres to see patients because there is a shortage of manpower.

When asked to describe his routine consultation method, he explained it as follows. When seeing a patient, he uses the symptoms mentioned by the mother, forms an idea of what body system could be affected, and then does a general standard and quick examination (head to toe), first the general condition, then body systems of interest. If complaints suggest a particular system, that system would be examined more thoroughly. It could take 5 minutes to diagnose a patient with minor complaints; for more difficult cases about 10 minutes; to admit a patient to ward would take up to 15 minutes.

When asked if there were any specific conditions that are difficult to diagnose where a diagnostic aid such as a flowchart might be useful, he replied that medicine cannot be learnt from a book, 70% has to be learnt practically on patients, and that internal knowledge and experience is more important than textbook notes.

He did not think it was a good idea for recently qualified staff to use computerised decision-aids because they have more theoretical knowledge than practical experience and they need personal exposure to a range of conditions to gain that experience. For example, different areas have different disease prevalences; direct contact with all of these during supervised internships are the best basis for this. A computer, however, could help a fully qualified clinical officer; one can forget medical knowledge, and it could help if it asked just a few relevant and accurate questions.

Asked about the method of representing medical knowledge in flowcharts, he pointed out two specific failings. If a patient comes with vomiting, coughing and fever, and the computer asks "Vomit after bout of coughing" and Yes is entered, it immediately diagnoses whooping cough. On serious chest diseases (pneumonia and bronchitis) the computer missed out on a number of important signs, for example, rhonchi, and bronchial breathing. The computer requires too little data to make a serious diagnosis.

He thought the current method of representing the questions could be adequate, but good only if the questions are specific. One of the potential sources of difficulty in the phrasing of Essex's questions is that the answers may be unknown. When asked, the clinical officer could not remember any situations where he was unable to answer 'Yes' or 'No' to a question, i.e. when his reply should have been 'Don't know'.

His rough estimate of the proportion of times he felt the computer was right, and the times it was wrong, he put at right 60% of the time, and wrong 40%. He went on to say that the computer might perform better if it used signs according to clinical examination and findings instead of the symptoms reported by the mother. For example, if the informant says "breathlessness", the patient may not be

breathless at the time of examination. One is not led to diagnosis by complaints alone, the clinical signs and findings are more important. One may sometimes need to find out whether the disease has been chronic, in which case additional information about the duration, and the onset (whether acute or gradual) has to be sought. Normally, one tries to obtain as thorough a symptom history as possible before starting to examine the patient. The depth of history taking and the importance put on certain symptoms depends on the patient's condition.

3.   *Test cases*

The forms from the last set of patients seen as normal consultations were used to test ESTROPID and compare the system's results with the clinical officer's diagnosis. Surprisingly, when comparing the computer diagnoses with those pronounced by clinical officer and doctor, the number of cases with matching diagnoses was slightly higher for the doctor (Table 5.7), although of course this is not statistically significant ($p<0.5$), and therefore not a consistent and replicable result.

	Clinical Officer	Doctor
Matching diagnoses	10	12
Partially matching	4	5
Unmatched	5	2
Miscellaneous	2	2

TABLE 5.7 : Comparison of computer diagnoses with those of clinical officer and doctor

The 2 in the miscellaneous category had to be restarted about five times, and are put here into a different category because they presented particular difficulties for both the clinical officer and the system. At the final attempt for each case in this category, one matched the clinical officer's diagnosis and the other did not.

During each entry, the comments and opinions of the clinical officer were noted down. 21 consultations were assessed in this fashion. Some of them are discussed here because they highlight certain deficient aspects of ESTROPID.

CASE 1:

This case illustrates an intrinsic difficulty with ESTROPID because it bases its diagnostic strategy solely on patient symptoms. A patient presented with fever, refusal of feeds, coughing, body rash, and vomiting worms. Two of these symptoms could not be entered into ESTROPID - refusal of feeds and vomiting worms. One of the diagnoses was conjunctivitis, which was noted by the clinical officer, although not mentioned by the mother. This is frequently associated with measles, as was the case for this particular patient. The clinical officer's diagnoses of measles with bronchopneumonia and conjunctivitis, was matched with the computer's of measles with pneumonia.

CASE 2:

This case illustrates the inadequacy of ESTROPID's handling of text entries. The body rash symptom had to be entered in the text mode. It was noticed that, because the system had no entry with just 'body rash', it asked, instead, a lot more questions including the word rash than it would have done if an entry of 'skin rash' had been specified. This happens because there are more questions associated with the word 'body' than the word 'skin'. The questioning mechanism used to clarify text statements is inadequate. It has only crude means of sifting through the questions, and some improbable or previously confirmed ones still get asked.

CASE 3:

This illustrates the problem of terminology. The clinical officer had written down on the form 'altered bowel habits', by which he actually meant constipation, not constipation or diarrhoea. 'Altered bowel habits' could not be entered into ESTROPID, so constipation was entered under the text entry mode. This additional information was actually not used during the session because of another problem - an error in ESTROPID's logic which ignores text entries if symptoms have been entered under the menu mode.

CASE 4:

This case illustrates the inconvenience caused by the system not allowing the user to backtrack. In this case, a long list of questions was asked, to most of which the answer was 'No'. The clinical officer kept his finger over the 'N' key and

mistakenly chose 'No' instead of 'Y' in answer to the question for fever. ESTROPID has no facility to correct previous answers, and the session had to be aborted and restarted.

CASE 5:

This illustrates the difficulty of reported symptoms not being corrobated on examination. A mother reported that a child had fever, and this was entered in the menu mode. However, on examination, the child was found to be afebrile. The clinical officer entered 'Y' to the fever question, but this exposes the problem of the system being unable to eliminate a reported symptom based on stronger contradictory clinical examination evidence.

CASE 6:

This case illustrates a problem with the flowcharts. A patient presented with fever, breathlessness and vomiting. The system accessed the appropriate chart for fever. However, on replying to only two questions, the first on "creps" (the answer 'N'), the second "breathlessness without severe diarrhoea" (the answer 'Y'), ESTROPID immediately pronounced pneumonia, but without any history of coughing or creps. This was a gross omission as the breathlessness was not confirmed clinically, and the child actually had malaria.

CASE 7:

This case illustrates the incompleteness of the knowledge base. A woman presented with breathlessness, dizziness, rapid heart rate and buzzing of ears. Only the first symptom in this list could be entered. The computer's diagnosis of hookworm anaemia matched the clinical officer's of severe anaemia and hook-worm infestation. However, ESTROPID could have reached that diagnosis sooner if the other symptoms, all indications of anaemia, could be entered and used.

CASE 8:

This illustrates another inadequacy of the flowcharts, in not including common associations between symptoms. In this case, the flowchart translation into a computerised system deals with the reported set of symptoms in a rather clumsy manner. A patient presented with joint pains, headache and fever. During the first session, the clinical officer was instructed to restart, replying 'No' to joint

pains in the second session. The patient had malaria, which is often associated with joint pains and headache.

What can we note from these reruns of some normal cases? These problems can be attributed to one or a combination of four sources: how patients report symptoms, how medical personnel conduct consultations, the flowchart structure, and the design of ESTROPID.

(a)   Mothers do not report all the symptoms that the child has, an obvious one such as conjunctivitis may be neglected although it deserves treatment.

(b)   ESTROPID was rather rigid in the number of initial symptoms it could accept, and some important symptoms could not be entered.

(c)   The flowcharts have been unable to include all the common associations between symptoms; when translated to a computerised medium additional information that would quickly have guided the system to a diagnosis is missing.

(d)   After the initial symptom entry the system design could only accept additional findings during the consultation when it requested it.

(e)   The procedures for using text entries was inadequate, and ESTROPID could not cope with contradictory information (for example, as mentioned above, reported symptoms that were not confirmed on investigation).

(f)   One incorrectly entered response means the whole session had to be restarted; there are no facilities for changing responses during a session.

(g)   There is the additional problem of terminology - if the system is to accept text entries, the same condition could be reported to the clinician in different ways, or the same condition as reported to the clinician in another language may be translated into English in different ways, or similar translations may be recorded on the form in different ways.

(h)   In explaining the use of the flowcharts, Essex instructs the user to initially only make use of the symptoms reported by the mother. This is against clinical practice, as there are often obvious signs that could quickly be noted by the practised medical worker, and taken into account in reaching a diagnosis.

5.6.2 THE PATIENTS

The parents or guardians of 58 patients were interviewed. Only one of them said she had never been to the out-patients department before, so the rest can be assumed to be familiar with clinics. When asked whether they noticed anything different about the consultation that day, 17 of them said they noticed the clinical officer using a machine, 16 of them mentioned either a typewriter or a machine used for writing, 9 (16%) of them used the word computer, and 14 said they saw "something" but that they did not know what it was used for.

One patient guessed that the machine was a "generator used for oil". Some of the other comments were that the clinical officer was seen "touching numbers", "writing with a typewriter instead of a pen", "touching something strange never seen before".

Some of the answers to the questions showed their annoyance at the delay: "asked so many questions and then asked to come here", "too much time taken because of so many questions". Ideally, their permission should have been sought before being used in the study, but before the start of the trial it was felt that this would bias the selection of patients. As it did not initially seem that the patients would be subjected to much delay, this ethical element was not introduced into the design. To compensate in some way, an attempt was made to explain the purpose of the study to each patient after they were interviewed.

This patient interview exercise did not reveal any other significant insights into patient behaviour or expectations.

5.7   A CONFIRMATORY STUDY

Another research project undertaken at the same location provides interesting confirmatory evidence for the ESTROPID trial [Snow, 1989 (unpublished data)]. In this section, some of the characteristics of (paediatric) out-patient diagnosis in the tropics as gleaned from this out-patient malaria diagnosis study are presented.

OUTPATIENT STUDY

1. Number   :__:__:__:__:
4. Date of birth   :__:__:__:__:__:__:
5. Sex   :__:

What are the complaints of the child:
14. ....   :__:__:
15. ....   :__:__:
16. ....   :__:__:
17. ....   :__:__:

Does s / he have :
18. Fever ....   :__:__:
19. Cough...   :__:__:
20. Difficulty with breathing................................   :__:__:
21. Runny nose...   :__:__:
22. Sore throat..   :__:__:
23. Ear problems...   :__:__:
24. Poor appetite / feeding...............................   :__:__:
25. Vomiting..   :__:__:
26. Diarrhoea..   :__:__:
27. Bloody stools..   :__:__:
28. Convulsions...   :__:__:
30. Weakness..   :__:__:

Other symptoms:
31. ....   :__:__: :__:__:
32. ....   :__:__: :__:__:
33. ....   :__:__: :__:__:

 Clinical Diagnosis : Doctor
52. 1 ....   :__:__:
53. 2 ....   :__:__:
54. 3 ....

Clinical Diagnosis : Clinical Officer
55. 1 ....   :__:__:
56. 2 ....   :__:__:

FIGURE 5.12 : Excerpts from the study questionnaire

Several of the findings discussed here simply confirm the results from the ESTROPID trial; others contribute additional knowledge.

The data is from a small study to investigate the accuracy of malaria diagnosis for paediatric out-patients between doctor and clinical officer (CO) at the Kilifi District Hospital. 388 children under the age of ten were seen in two out-patient clinics over a 3 day period, first by a CO and then by one of the 4 doctors available for the study. Under-fives were routed first through the maternal and child unit before joining the main queue, so some of the children seen were well and had reported only to be routinely weighed and examined. The doctor recorded patient details, complaints, made an examination and recorded a diagnosis. Relevant sections of the original questionnaire are shown in Figure 5.12. Later, the CO's diagnoses were also added on the form, and the symptoms and diagnoses coded.

The majority of out-patient clinic attendance consists of paediatric patients, as shown in Table 5.8 and confirmed in the Coastal Province Annual Report [Kenyan Ministry of Health, 1988]. Four-fifths (79%) of the children seen were under five.

Age (yrs)	Frequency	Cumulative percentage
0	28	7
1	149	45
2	88	68
3	43	79
4	35	88
5	17	92
6	10	95
7	9	97
8	6	99
9	3	100

TABLE 5.8 : Age distribution of paediatric patients

The average number of complaints (including those patients who were well) was 2.2. The distribution of the number of complaints reported for each patient is shown in Figure 5.13(a). The 6 most common complaints are listed in Table 5.9. Fever and cough were reported in over half of the children. Fever was reported as a single symptom in 2% of the children, and cough alone in 3%. Both were reported as the only two complaints in 9% of the children seen. These two
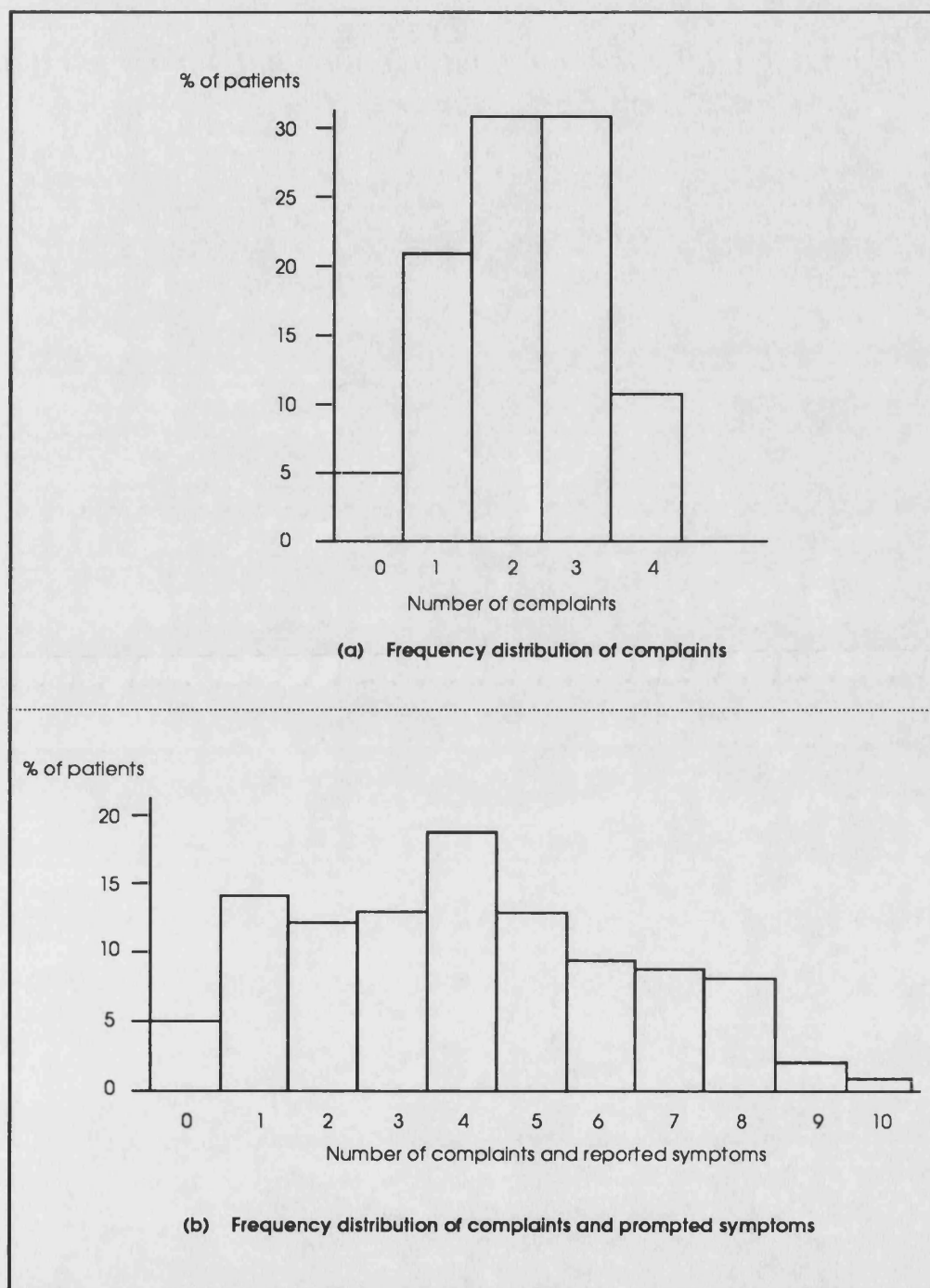
FIGURE 5.13 : *Number of symptoms reported per patient before and after prompting*

symptoms serve as a very poor discriminant for what a paediatric patient has, and a clinician would have to use clinical findings to support or suggest a diagnosis.

Complaint	% of patients reporting this complaint
Fever	56
Cough	50
Difficulty breathing	15
Diarrhoea	14
Abdominal pain	13
Vomiting	10

TABLE 5.9 : Frequency table of most common complaints

As to be expected, when prompted, the number of symptoms reported as well as the incidences of symptoms both rise (see Figure 5.13(b) and Table 5.10). The mean number of symptoms now rises to 4.0, almost double. An opportunity to record further symptoms after the list of prompted symptoms stimulated some additional reporting of previously unmentioned symptoms for some patients. This illustrates the increased likelihood for patients to reply in the affirmative when asked for the presence of very specific symptoms.

Symptom	% of patients reporting this symptom unprompted	% of patients reporting this symptom when prompted
Fever	56	62
Cough	50	63
Difficulty breathing	15	33
Runny nose	4	47
Sore throat	<1	4
Ear problems	6	18
Poor appetite	1	31
Vomiting	10	24
Diarrhoea	14	27
Bloody stools	2	9
Convulsions	<1	5

TABLE 5.10 : Increases in reported incidences when symptoms prompted

The data presented above also illustrates the role that a few common symptoms have to play in out-patient diagnoses. Many consultations have to start from an average of two symptoms only, with cough or fever or both included as one of the reported symptoms in fifty percent of patients. However, when prompted the average number of symptoms reported increases to four, with an over-representation of those symptoms that are included in the prompt list. Here, there is a danger of introducing a bias in favour of prompted symptoms, which are more

likely to be reported as present than those symptoms that are not explicitly asked for.

Not all of the symptoms reported will be important in reaching a diagnosis. Some reported symptoms are misleading, and may sometimes be related to some past illness, not the current one. On examination, other symptoms may be discounted because they play no contributory role to the diagnosis. The symptoms reported may also lead to another error, that of omission, when significant symptoms are not mentioned at all. Some of these may emerge from further questioning during the course of the consultation, but this depends on the accuracy and depth of history taking. In general, the role of reported symptoms must be seen to be that of guidance, many symptoms need also to be substantiated by physical examination or confirmed by further questioning. This is often not the case for medical decision-aid systems.

An investigation into the patterns observed in the number of diagnoses made by CO's and doctors in the outpatient malaria diagnosis study, reveals that their mean does not differ much (Table 5.11). Although the doctor is expected to record up to 3 diagnoses, on only 5 occasions is this done. The general tendency of clinicians to avoid over-diagnosis is upheld here.

Number of diagnoses recorded per patient	... by doctor		... by clin. off.	
0	6	(2%)	17	(5%)
1	291	(75%)	242	(62%)
2	86	(22%)	129	(33%)
3	5	(1%)	–	
Mean no. of diagnoses	1.23		1.29	

TABLE 5.11 : Number of diagnoses recorded per patient

An examination of the primary diagnoses made by doctor and CO reveals that 35% of those recorded match. When adjusted to include matching diagnoses placed on different ranks, the concordance level rises to 45%. This is the same figure as that obtained for normal consultations when the total number of diagnoses recorded was matched (Figure 5.10). Not enough information was

available to do a matched comparison for each patient, as was done for Figure 5.11.

Table 5.12 presents 10 of the diseases with the highest frequencies, showing the number of patients diagnosed by either doctor or CO with a primary diagnosis of the disease shown, as well as the number of times the two primary diagnoses match. The high percentage of diagnoses by both doctor and CO for URTI and malaria seen in Table 5.12 reflects the disease profile presented in the Coastal Province Annual Report [1988]. It is interesting that these two diagnoses are generally indicated by the two most common symptoms: cough and fever. Diagnostic methodologies that provide no opportunities for further examination will again bias the diagnostic outcome, if the possibility of additional contributory factors not obvious to the parent is not allowed for.

Disease	Diagnosed by both		Diagnosed by doctor		Diagnosed by clin. off.	
Upper Resp. Tract Inf.	41	(11%)	81	(21%)	96	(25%)
Malaria	25	(7%)	50	(13%)	68	(18%)
Well	12	(3%)	42	(11%)	26	(7%)
Otitis media	2	(1%)	9	(2%)	18	(5%)
Other skin disease	8	(2%)	10	(3%)	13	(3%)
Lower Resp. Tract Inf.	0	(0)	24	(6%)	2	(1%)
Pneumonia	3	(1%)	9	(2%)	12	(3%)
Bronchitis	1	(<1%)	8	(2%)	16	(4%)
Gastroenteritis	6	(2%)	18	(5%)	6	(2%)
Worms	2	(1%)	5	(1%)	7	(2%)

TABLE 5.12 : Concordance of primary diagnoses between doctor and clinical officer for some common diseases

In summary, it appears from this set of research data, that patient reporting of symptoms is useful as a guide for initial hypothesis formulation, but is dependent on what signs and symptoms are prompted for. In general, initial hypotheses have to be formed on a basis of two symptoms on average. A cursory physical examination should immediately narrow the range of hypotheses, as well as include obvious signs that would aid diagnosis. This step is often ignored in computer-aided diagnostic systems, as in ESTROPID, wrongly placing the initial emphasis solely on patient symptoms reported.

5.8 STRUCTURE, PROCESS AND OUTCOME ANALYSIS

Here again, the structure, process and outcome analysis framework can help to summarise and highlight the major discoveries of this case study.

1. *The system's functioning*

(a) *Structure.* The Toshiba portable worked well with a stabiliser. The interface was not flexible enough to handle the range of symptoms it was presented with. The flowcharts were overambitious in attempting to cover more diseases than could be treated at the out-patient level, sacrificing more useful detail for treatable diseases. The flowchart structure resulted in a knowledge representation method that was repetitive and cumbersome. They also did not include many important relationships between symptoms, that help to reinforce or reject diagnoses. Thus, the incompleteness of the knowledge base is entirely due to the incompleteness of the set of flowcharts.

(b) *Process.* ESTROPID was slow in all interaction, in accepting entries and displaying questions. The interface was very basic, and it could not deal with contradictory information, handle extra information during the consultation, or allow any backtracking. Its reasoning mechanism is very opaque, not demonstrating human-like intelligence in its line or method of questioning. The richness and complexity of the real diagnostic task left ESTROPID looking very amateurish indeed.

(c) *Outcome.* In spite of its obvious defects, ESTROPID's concordance with the doctor's diagnoses was not much worse than those of the CO for normal consultations. But the obvious conclusion from its performance is that a system has to incorporate many more real-life aspects to be at all useful.

2. *The user's perspective*

(a) *Structure.* What difference did ESTROPID make to the clinical officer's normal working conditions, and did it impose any new requirements?

The computer needs to be kept in a secure place as it cannot be left safely in an unlocked room. This meant taking the computer and stabiliser each day from the store room, to the consultation room. Under normal conditions in a rural health clinic, this might be an important consideration if the computer equipment cannot be transported easily. The layout of the room does not change much, as the computer is placed on the desk the clinical officer usually writes on.

(b)   *Process.* How does ESTROPID affect social interaction, and the clinical officer's behaviour towards patients?

The use of the computer does not change the initial detail gathering stage in a consultation. The patient is still asked what is wrong with them, in the usual manner. Probing, in taking the rest of the history, changes with the use of ESTROPID, as the clinical officer now uses the items on the menu to guide his interrogation. Routine physical examination is now relegated a less conspicuous role, and the clinical officer only examines the patient when ESTROPID requests it. As was discussed earlier, the system's structure is not a good model of actual practice. Patient symptoms need to be verified by physical examination as often as possible. The consultations take longer, and are now related to the clinical officer's typing and reading abilities, the speed with which the patient details can be input, and the questions displayed on the screen read.

(c)   *Outcome.* Did ESTROPID improve the quality of the clinical officer's diagnoses?

The ESTROPID diagnoses, although not acted upon, did not significantly improve the clinical officer's concordance with the doctor. The clinical officer felt he disagreed with ESTROPID's diagnoses 4 out of 10 times, indicating that if he were meant to implement ESTROPID's diagnoses, he might have chosen not to.

3.   *The patient's perspective*

(a)   *Structure.* It was the trial design itself that presented most inconvenience to the patient rather than the use of ESTROPID. They had to see two clinicians, instead of one, and were subjected to some delay.

(b)   *Process.* ESTROPID consultations are longer and are less personal since they exclude many of the additional functions a CO does very quickly, for example, in passing information to reassure patients, or advise on nutrition.

(c)   *Outcome.* Although the system has certain systematic errors, these do not result in significantly worse diagnoses overall. However, ESTROPID nowhere reaches the speed and competence with which the CO performs his job while retaining personal contact and scope for additional health advice where necessary.

4.   *Viewpoint of administrator*

No formal evaluation structure was set up, but many informal discussions took place, to suggest what factors are important when a manager of a health unit evaluates the potential of such medical decision-aid technology.

(a)   *Structure.* The administrator needs to be satisfied that ESTROPID is a cost-effective and efficient alternative. Implemented on an IBM-compatible computer and requiring a stabiliser, it does not at present offer such an alternative. If the system is dedicated to diagnostic support during working hours, it cannot be used for other purposes to justify the initial expense. The ESTROPID software itself is in addition inefficient, being unreliable in its diagnoses and slow in operation. An evaluation needs more base-line data as a more balanced basis for assessment.

(b)   *Process.* Any system in routine use should be protected from misuse, whether unintentional or otherwise. The questions and other output of the system must be clear and understandable and the action proposed as an outcome should be within the user's capabilities. ESTROPID will need to have a questioning strategy that can examined, to ensure that the user can assess the outcome, and possibly challenge the system in some cases.

(c)   *Outcome.* The system should be able to offer at least the same quality of care to patients, and also be used to see as many patients as a CO currently sees in a similar time span. The performance of ESTROPID in this trial has not been startlingly impressive. It needs to be considered whether effort is best directed in this area of research, or whether resources might be better expended in other

means of improving the quality of health delivery and support using information technology.

5.   *Expected impact on health system*

This technological application seems to be a non-starter in this context, for a number of reasons. Current technology remains too expensive for ordinary health centres. Software design still falls far short of the required level of expertise. In this study, employing a skilled clinical officer, there seems to have been more of a likelihood of challenging the system on disputable cases. In other situations, his experience may not be the overall norm, and may result in the transfer of accountability from a health worker to a machine that has not proved its worth.

5.9   DISCUSSION

Expert systems are notoriously difficult to evaluate and this trial has proved to be no exception. The trial design had several limitations as discussed in the introductory section of this chapter.   A major difficulty with ESTROPID is that observed clinical practice does not correspond well with its structural assumptions. The performance of ESTROPID was dependent on its structure, design and logic mechanism as well as the basis of its knowledge - but there are inherent problems with the design of the Essex flowcharts. Not surprisingly therefore, it did not result in a significant improvement in the user's performance.

Most medical diagnostic-aid systems for developing countries attempt to improve or support the diagnostic ability of nurses or other paramedical staff. None have met their stated aim of increasing the quality of routine decisions made at the primary health care delivery level. It is doubtful that such systems could have an impact if the health personnel were not trained well enough to question its decisions. Well-trained staff could well do without a computerised aid, and perform better, and faster, delivering care with a more personalised approach than any computer-assisted system could provide.

This analysis does not conclude that health personnel do not need any help. It argues instead that their need of computerised decision-aid technology is questionable. The list of proposed diagnostic-aids should be extended beyond computerised systems, for example, to investigate the construction of manual reminders or checklists which may be more useful.

It is essential that "normal" practice is investigated before considering how any decision-aid could assist decision-making in this context. It is important to establish what the normal concordance levels between diagnoses made by health personnel of different abilities are. This should be supported by additional methods of weighting the severity of disagreements between diagnoses. The problem of establishing a scale of discordance between diseases is not trivial and requires further investigation. It will also be useful to explain the sources of the discrepancies found.

After this initial investigatory step, the role of the diagnostic aid can be properly evaluated. The concordance levels when the diagnostic aid is used can therefore be compared to that of "normal" practice. A basis for the severity of misdiagnosis, having been established, provides another factor for comparison.

The implications of the above results for medical decision-aid system design are for both user interface design and the reasoning mechanism. A more realistically designed interface should ask not only for the symptoms reported by the patient, but also provide some general outline for patient examination. Based on this initial data, hypothesis can be generated in either the manner of statistical reasoning methods, or that of rule based strategies. The reasoning mechanism must therefore be able to allow for an initial set of data values to be passed to it for processing.

Indefatigable optimists are well-warned. Should they wish to proceed, there are a number of preliminary issues that need to be investigated. Accurate assessment of the impact of a diagnostic aid can be made only if it can be compared with data collected from preliminary base-line studies. They must also be prepared to find that their system, having ironed out all the technical difficulties, can be matched or bettered by a set of low-cost initiatives (such as manual aids, and fora for

training and discussion) which leave the decision-making firmly in the hands of the health professional.

Whereas these results and conclusions may not be startling to the seasoned medical practitioner in the tropics, these issues have generally been unrecognised or ignored in the literature on medical decision-aids for developing countries. It is time that this deficiency is redressed.

CHAPTER 6

COMPARATIVE ANALYSIS

6.1   INTRODUCTION

This chapter aims to judge the degree of success of each of the case study systems described in this thesis, and offer explanations for the observed outcomes. The two application areas are compared, and the discussion is extended to considering the generic health information system application.

A number of aspects that influence successful integration of a health information system in developing countries will be considered in the following sections. The aims of each system and how well they fitted their roles; the resources that are required and whether they form too strenuous a demand on the infrastructure to be sustainable; the design of the software and its efficiency; the social context, as success is critically dependent on the co-operation of the people affected by the systems; the effectiveness of the system in improving the work of health personnel; the adequacy of the evaluation process to assess potential; and finally the impact on the health care problem. In short, in the context of these factors, can the case study systems be regarded as successes or failures?

6.2   THE SYSTEMS' ROLE

A prerequisite to considering a health informatics application is to look at the environment in developing countries. The desirability of any system must be considered within the context of the issues and constraints discussed in chapter 1. It is important to specify the aims of the system, and to consider whether it is desirable, a priority, and achievable.

The context for FDCS is a lack of adequate information for development activities, the main problem being in producing data of a suitably high quality within a short period of time. The need for this has been stressed by several multi-lateral agencies. As with many ideas generated from outside the developing world, a possible objection to this is that, although international agencies may regard data as vital for decision-making, policy makers within developing countries may not. This mismatched sense of values, between the imported and the homegrown, can be a recipe for failure. This thesis relies on the interest shown in survey work in developing countries, and a growing acceptance of its value. In addition, research projects on primary health status provide an opportunity for investigating more thoroughly whether computerised data collection might have a beneficial role.

Interest in data collection systems in industrialised countries is generally in response to a well-defined and accepted need by organisations. The initial indications are that computer-assisted interviewing will be integrated into the survey process by many statistical offices and survey organisations in industrialised countries. The systems for these environments are ambitious, covering many stages of the survey process and allowing complex questionnaire structures, and survey management functions. Although a range of flexible packages are now commercially available, few systems have been designed specifically for the developing world. Those that have been introduced in Chapter 2, although more contextually adequate, require further refinement of the software, and improved hardware support.

The aims of FDCS do not differ much from other computer-assisted data collection systems, in some respects. The system is proposed as a replacement for paper and pencil, and a means for improving interviewer data quality. In addition, it provides a number of added benefits for a supervisor in terms of survey design and management. The need to structure a questionnaire adequately before CADAC specification may be a chore, but a very necessary one. It ensures that all skip and input options are well thought out beforehand, which is a common deficiency of survey design. In terms of survey management, a pre-written checking procedure can be applied to a data set as soon as it is downloaded from the CADAC system to a microcomputer. FDCS offers a challenging alternative to achieving some of these CADAC goals within developing countries.

ESTROPID, like many other decision-aid systems for developing countries, was designed to address the need for skilled medical decision-making in out-patient practice. Chapter 2 has discussed the issues raised by researchers and what they see as the problems of providing curative primary health care services within developing countries. These include the inadequacy of training given to paramedical workers, the unwillingness of staff to remain where they are most needed due to a general lack of incentives (financial, resource support, and value recognition). It is suggested by advocates of medical decision-aid systems that, under these conditions, computer-assisted diagnosis can be useful, to provide rural health workers with a tool for dealing with everyday health problems. This is a gloomy assessment, presenting a scenario in which technology comes to the aid of a poorly operating human system - not an ideal situation.

Medical decision-aid systems that have been suggested for developing countries differ in aims from those for industrialised countries. Systems for developing countries are meant for paramedical generalists with limited medical knowledge that has to be applied to a wide range of common ailments. On the other hand, systems for industrialised countries are usually within narrow domains, to be used as a second opinion for specialists, or alternatively, as consultants for doctors with other specialities. The systems within industrialised countries have rarely fulfilled their role with a demonstrably acceptable level of expertise. However, some researchers suggest that applying artificial intelligence techniques to the 'simpler' problem of diagnosis within developing country contexts will be a learning experience. Prototype systems for developing countries aim to provide advice on the conditions and possible treatments for patients, to act as a reference source, as well as collect medical records for further study.

ESTROPID aims to do less than other systems, it does not maintain comprehensive patients records or give a treatment strategy (although it suggests the level of skill required to treat the patient's condition by specifying where it should be treated). It was intended for use by paramedical staff based in out-patient clinics attended by large numbers of patients daily. It was meant to be suitable for all consultations, to elicit symptoms, guide questioning, and pronounce a diagnosis of a patient's condition.

Although ESTROPID and FDCS are designed primarily to meet an existing need, there is a difference in focus of their application contexts. The problems in survey and questionnaire administration are a well trodden path, with a clear potential for computerised benefit, whether in industrialised or developing countries. For medical diagnosis, there are problems and opportunities, but these are not well understood or defined, and it is unclear whether computerisation can be a significant benefit.

The analysis in this thesis suggests that, to identify a problem area and see a scope for computerisation, based on the experiences or state-of-the-art in industrialised countries, is an inadequate foundation for proposing a full-scale computerised solution in developing countries. These two application areas *are* clearly based on real problems, but the question is are the systems described here real solutions?

The answer partly depends on the ability of the recipient organisation to absorb the technology and the social structures in general, and not only on proof that the systems do improve user performance. These issues will be discussed in the remaining sections of this chapter.

6.3   RESOURCE REQUIREMENTS

System implementation and lifetime operation require a number of resources, which include money, hardware, power supply, computer peripherals and consumables, support and maintenance, and user skills. System feasibility and sustainability depend in part on whether high levels of resource use are required. For example, whether the costs of hardware and consumables are high, whether the system would need a lot of support and maintenance from skilled computing staff over its lifetime, or whether users will be required to exhibit familiarity with computers.

FDCS operates on relatively inexpensive hardware, with no moving mechanical parts, therefore requiring little maintenance. Being battery powered, it has an independent power source and does not require any cables during field

operations. However, batteries and data storage devices will be a recurring expense. The maintainability of the system depends to some extent on whether all potentially susceptible slots are exposed to a lot of dust. During the trial, the minimum protection procedure was to slide the case back when the Psion was not in use. This proved adequate to keep the machines in working order, but it is likely that dust settled on the more exposed communications slot hindered data transfer with the field machines. Prior computing knowledge is not necessary for operation, and users are not assumed to have any keyboard skill or computing knowledge as these are easily acquired after training and a few days of use.

Commercial CADAC systems are usually PC-based, requiring IBM compatible hardware. The mechanical moving parts in these machines increase the probability of hardware failure in tropical regions. The peripherals generally required, disks and battery packs (for portable computers), are not extravagantly expensive by developed country standards. The two systems for developing countries discussed earlier require a number of components, which may be difficult for interviewers to keep dust-free and maintain in everyday use. Systems are generally claimed easy to learn, but it is not always clear whether this is from the interviewer's viewpoint, or that of the survey manager. Support for data transfer needs to be provided either by the survey manager (who may have to acquire new skills) or by the computing centre (if one exists). Compared to the requirements for more common IBM-compatible computers for similar purposes, the Psion has several advantages over current systems in terms of resource use.

ESTROPID operates on an IBM compatible portable Toshiba microcomputer. The main limitation on hardware feasibility and maintainability are high purchase and maintenance costs. The portable computer requires many more resources, and is less independent than the Psion. The Toshiba operates off the mains, with a battery backup, and needs a stabiliser for protection against voltage fluctuations. The software did not assume the user to have any computing skills, but keyboard familiarity was implicit in the system's design.

Other medical decision-aid systems for developing countries are now veering also toward IBM compatible machines, although Byass' system uses a non-IBM compatible laptop. Modern PC portables may prove to be more susceptible to

damage if used regularly in unregulated environments, with much dust and high humidity levels.

CADAC systems can be implemented on quite small computers with little processing power and complexity. Medical decision-aid systems, however, will tend to require more than this minimum to manipulate information. The difficulty with using non-standard hardware, however, is that systems may quickly become outdated. Costs are being reduced as performance increases, so it is likely that cheaper and smaller, yet powerful computers may be available in the near future, and more complex applications software can be produced at low cost. Icon driven technology at the laptop level is now being produced, and touchscreens and simulated mice are now also available. The trend is towards simple interfaces, which will require even less skill from users. Newer storage devices with less possibility for mechanical failure are also being developed, although these are still relatively expensive compared to floppy disks. Hardware systems are now equipped with more independent and longer-lasting power supply devices.

Health information systems with a high potential for survival will be those that require a minimum of available resources during their lifetime. Thus they should be as cheap as possible, requiring hardware that consumes little power, and preferably with few peripherals. In terms of human expertise, they should require as little maintenance as possible, and the user should quickly be able to learn how to use and look after the system. Technological trends are leaning in this direction, with newer systems containing many more desirable qualities. Costs remain prohibitive, as always at the initial stages of introduction, but these generally decrease as the technology becomes more familiar.

Nonetheless the widespread use of such computing technology in an average primary health care unit in the sub-Saharan region is a long way off. CADAC has more immediate potential in health, to be used for effective survey management, as resource utilisation may be offset by the benefits. Medical decision-aid systems will generally require resources far beyond the average budget of a single health unit, and provide no immediate promise for added benefits.

6.4   DESIGN ISSUES

A traditional classification of good software design comes from Boehm (as quoted in [Hollnagel 1989]) and includes the following components: portability, reliability, efficiency, human engineering, testability, understandability, and modifiability. Not all of these components are useful in the comparative discussion in this section. Reliability and understandability are the most important factors, followed by efficiency, human engineering and testability. Some reference will also be made to the complexity of the programming task.

The reliability of FDCS could be almost completely assured towards the end of the trial, when all remaining reproducible software errors had been corrected. Structured code, in a modular form makes identification and subsequent modification of errors easier. The interface was mainly based on a series of menus, with additional prompts when necessary. The FDCS managed to be quite ambitious on such a small handheld device, allowing for a minimum subset of CADAC functions as well as additional functions such as database cross-reference and data summary. To achieve this was a straightforward task, a 'normal' or conventional programming task that could be tackled in the usual way.

While the design of a basic CADAC system is simple, the addition of complex editing or management functions can introduce further complexity. For example, changing the answer to a question which decides what branch of the questionnaire to follow may invalidate the answers to branched questions further down in the questionnaire. Achieving this satisfactorily can be quite tricky, and not all systems succeed in this. Otherwise, the production of efficient CADAC software is an achievable task, well within the bounds of conventional software engineering. The level of skill required to maintain or modify the systems is within the abilities of expertise, however infrequently available, in developing countries.

ESTROPID was written in LISP code, not the easiest to understand generally, although well designed function names can be used to increase readability. The interface was based on a fairly basic principle of option numbered lists from which the user has to choose. It has many faults in its design, mainly to do with

the lack of compactness, and much duplicability. Thus the route to the final decision is not the shortest, and the system consequently does not display much of the desired 'intelligence' or 'expertise'. The system's testability is intricately linked to the complexity of the programming task. The more complex or unstructured a problem is, the more difficult it is to prove that the design meets the specification. For medical decision-aid systems, there is little agreement among researchers on how they should best be tested for consistency and performance.

The design of medical decision-aid systems that demonstrate expertise is an exercise fraught with difficulty, because it is by its nature an ill-defined problem. The choice of domain boundaries, knowledge representation, inference mechanism, and system outcome are in themselves difficult. Many systems consequently suffer from design faults because the diagnostic solution is complex.

Modularity of design is easy to achieve at the outset, as is understandability by using traditional software engineering techniques, which are the 'least worst' techniques available. However, it becomes progressively difficult to maintain throughout a system's lifetime. Medical decision-aid systems are also notoriously difficult to modify, and the maintenance of medical decision-aid systems is yet another unresolved issue discussed in the literature. Simply being in a developing country context and providing technology for paramedical workers does not mean any of these problems of software complexity are removed.

The two application areas differ in the level of software complexity required for system development. Good quality basic CADAC software can be achieved easily, although more complex functions will require some ingenuity. The design of medical decision-aid systems, even at a 'basic' level as suggested by some for developing countries, is not trivial. Furthermore, while it is possible to test CADAC software adequately, medical decision-aid systems are difficult to prove demonstrably complete as diagnostic tools. Structured programming, good software engineering practice and user-oriented interface design can easily be applied to both areas, but these can be just palliatives. The inherent complexity in medical decision-aid systems generally leads to their rapid exhaustion.

In system development, there has to be a tradeoff between what can be aimed for, and what should be settled for, given hardware and resource constraints. For many applications oriented towards developing countries, the hardware choice will often be the strictest limitation, and thus software issues necessarily take a second place.

6.5   THE SOCIAL SYSTEM

The analysis used in this thesis includes three types of stakeholders, the user, the patient or respondent (referred to as recipient), and the planner or manager. In the case of the user, it is of interest to investigate their opinions, behaviour, and change in working patterns. The patient is directly affected by the medical decision-aid system because it might influence the decision made by the paramedical officer, but the respondent in the FDCS system is only indirectly affected. In each case the planner has to be concerned with the expected impact of the system, and whether it improves the quality of the users' work with acceptable side-effects or additional costs, and whether it can be expected to make a significant impact on the whole health system in the longer term.

With FDCS the users did not object strongly to its use; opinions which were decidedly against FDCS at the beginning of the trial changed to a neutral or positive position by the end of the study period. The respondents did not mind which method was used, and it is likely that in this periurban area, most of them likened the Psion to a calculator. This result was similar to that in other studies, showing that strongly negative reactions to the use of computer-assisted interviewing techniques was rare.

With ESTROPID, the user was not satisfied with the performance of the system, which was way below his diagnostic speed and accuracy. The patients were put at some inconvenience, having to queue up twice, and therefore were subject to delays. Patient interviews sometimes reflected this annoyance with the long procedure. The ethical problem of very sick patients being forced to wait longer than necessary was overcome by the paramedical worker bypassing the study procedure if such a patient was seen. A surprising proportion of patients in this

provincial administrative town had either heard of or seen computers before. Of the other system evaluations, only TROPICAID has been evaluated in this dimension. Their findings, in contrast to the result presented here, were that paramedics were keen to use the system, though no formal field trials were organised.

Users in the case study trials were willing to try the systems, but only willing to use it in everyday practice if they could prove for themselves that it improved their work. As for changes in working conditions, the computer displaces paper, and forces a stricter adherence to formal rules of operating. This may be a good or a bad thing, good if there is a need to standardise practice, bad if some of these new changes are unnecessary and force users into making unanticipated and undesirable shortcuts. It is important that the systems are carefully monitored to check whether any of these effects are superfluous.

Recipients (patients or respondents) are ultimately concerned with improvements in their health or their family's health. Sensitive explanations of a system's use will tend to win them over. For medical decision-aid systems, there is a further ethical question of whether the patients should be asked for consent before being included in the study. As was discussed in chapter 5, this may result in biases in the evaluation results. This issue will have to be considered more carefully in future evaluations.

Soundly designed systems should have little difficulty in being accepted by the people involved, if the system can be demonstrated to improve the user's performance, and is explained as a means of improving the recipient's health. In the long-term, the system needs to be *shown* to perform its role well to continue to be accepted.

6.6 IMPACT ON EFFECTIVENESS

What difference does the use of the system make on the quality of the work performed? Does it help to improve health service to the community in any way?

FDCS was a clear improvement on paper and pencil, and helped to produce data of a consistently higher quality amongst a group of interviewers of varying abilities, also reducing the length of their interviews. It lessened the workload of interviewers, their supervisors and the survey manager in the checking of forms, and improved the timeliness of data, making it instantly available rather than subject to a week's delay.

While FDCS can be portrayed as a success at this level, ESTROPID was not. It lengthened consultation times, with no appreciable benefit in improved decision-making. Of the systems discussed earlier, Byass' system was the only one that aimed to evaluate diagnostic efficiency. Here, the computer's performance was not startlingly efficient either.

Where medical decision-aid systems have been shown to improve productivity, it has been difficult to establish a strict cause-effect explanation, simply because there exist a number of other contributory factors. For example, the process of ticking off patient symptoms and signs of a computerised list has been shown to have a significant 'checklist' effect. In this sort of situation, the solution may be found not in computerising the decision-making process, but to introduce a paper-based method for helping health practitioners do their jobs better. A similar effect can be found by the use of FDCS, when users take their time to enter in identifier details because they are using a keyboard. However, the bulk of the observed improvement in data quality can be traced to the effect of the software itself.

In the context of developing countries, partial success at improving just productivity is not enough. The resource implication of the introduction of a new system dictates that it has to be proven to have significant advantages before full implementation can be suggested. The benefits of the working system must be shown to outweigh the purchase costs, which are in general still high. The right type of activity must be computerised within the range of the users' functions, to apply computerisation only to those that can have a direct contribution in the provision of improved health care.

6.7   THE BENEFITS OF STRUCTURED EVALUATION

The amount of proof required to justify the implementation of a system is proportional to the risk attached to the systems use. The more strictly formal the procedure to be computerised is, the less has to be done to show it performs its expected functions. This however, does not change the fact that technical perfection does not imply social acceptability, so the evaluation has to go beyond fit to some specification.

The amount of risk attached to the use of FDCS is minimal, as is the prospect of misuse. There is always a danger that interviewers may fabricate data, but this can hardly be attributed to the recording method. The occurrence of fake data will be avoided or reduced by improved survey management. Thus the evaluation required to demonstrate technical performance will also be minimal. Despite this, the evaluations of CADAC systems so far have not aimed at demonstrably proving technical ability. Quantitative assessments have rarely been striven for, and unsupported qualitative statements have generally been the norm.

The amount of risk attached to the use of ESTROPID is large, as is the possibility of misuse if the system is not completely understood. Users, particularly as suggested for developing countries, may blindly obey the decision-aid systems without thorough consideration of the outcome in relation to the patient's condition. There is a lack of methodology to encourage more thorough system evaluation. The informal evaluations reported so far, with the exception of Byass', are of little relevance in attempting to compare the performance of systems.

Few systems have been evaluated, but this is a necessary precondition to validate the expenses incurred in setting these systems in place. The structure, process, and outcome evaluation framework is suggested in this thesis as a means of reporting work to establish some basis of comparability. By incorporating technical, social and the wider health care issues, it allows for a number of important aspects that have been recognised by some authors, but rarely integrated together or implemented in field studies.

6.8   CONCLUSION: SUCCESS OR FAILURE?

How can the design, implementation and evaluation of these systems be considered - as successes or failures in achieving their goal of better health provision? If a success, this chapter argues a system needs a design that is sensitive to the context in which the system should be used, an implementation that is practical in its resource demands, and an evaluation that quantitatively, in addition to other more qualitative measures, demonstrates that the system works.

Under these terms, FDCS has been a success, and has validated the use of simple hardware in developing countries. The design is simple, the implementation practical, and the evaluation case study has demonstrated its effectiveness. In health terms, it could have an important role in generating timely data, on which managerial decisions can be based.

On the other hand, ESTROPID could be described as a technical failure, even if a partial one, because its 'expertise' was way below standard. It also fails on the resource implementation level. On these two counts, it therefore has no prospects of being integrated into the health process as it stands. Further refinement to make its decision-making more 'natural' may improve performance, and implementation on a cheaper machine may make it more economical in its resource requirements. However, it is still unlikely to demonstrate its importance at the primary health care level.

What does the future augur for other computer-assisted data collection and medical decision aid systems? CADAC systems which incorporate a few basic facilities, and are used to full effect in well thought out health surveys can be put into place in the near future. Medical decision-aid systems are not a proven technological application, and their premature introduction may be detrimental. There is some potential for decision-aid systems, but certainly not in the near future, and not until more basic research questions as to the nature of the decision problem are addressed. In the meantime, alternative ways of improving performance within primary health using informatics should perhaps be investigated.

Based on the discussion in this chapter, the requirements for sustainable health informatics applications can be more strongly stated. The system choice should be based on priority needs in health, and have a high potential impact. Low-cost hardware is one way of maximising the use of scarce financial resources. Human resources should be encouraged, in the development of skills of indigenous health personnel. Above all, there is a need for formal means of assessing a system to verify its full and successful integration into the recipient organisation.

CHAPTER 7

CONCLUSIONS

7.1 THESIS SUMMARY

The route this thesis has taken was sketched in Figure 1.6. The more substantive parts of this work have concentrated on four of the outlined items. Of these, the promised illustration of the types of systems that are practical and effective was only met in part, as just one of the two case studies presented could be described as such. This chapter aims to provide the fifth and last step in the research path - guidelines for the systems of the future. It will do this through a discussion of the main contributions of this thesis to the research area, presented in terms of practical recommendations and research contributions. Before these tasks are pursued, it is helpful to present here a structured overview of the contents of each preceding chapter, building up the contribution of each in the overall framework of this thesis.

The introductory chapter has laid the foundation for establishing a generic description of suitable health information systems for developing countries. The problems of implementing 'health informatics' in 'developing countries' was discussed after definitions of the two terms were given. The bias of this thesis was clearly defined as being in favour of the use of health informatics applications to improve primary health care delivery. An initial set of implementation issues for health informatics applications was identified. Against this background, the research approach taken in this thesis has been developed.

The basis for the development and design of FDCS and ESTROPID was detailed in chapter 2. The state-of-the-art in both the application areas of interest - computer-assisted data collection systems and medical decision-aid systems - was

presented, with particular attention being paid to some systems for developing countries. Suggested reasons for implementation in the developing world were summarised, and differences between designs highlighted. This literature overview suggested some current research directions and justifies, in part, some choices made during the development of the two case study systems.

A major deficiency with most research projects in this field is in demonstrating that the aims of the system have been adequately set out and subsequently met. Some evaluation approaches have been suggested in the literature for both types of systems. These, even though not entirely satisfactory, encourage more disciplined assessment, but have hardly been used. Instead, there is a proliferation of ad hoc justification of systems. There is a need for more thorough investigation, including both quantitative and qualitative measures, to examine technical, social and organisational aspects of implementation. The evaluation framework presented here aims to provide a structured checklist for reporting evaluation results, as well as a means of comparison among systems.

The first case study was centred around the assessment of FDCS (Field Data Collection System). The menu-based design of this system, and its aims of delivering data of a higher quality in less time were presented. The background to the case study was described, covering both the country characteristics (The Gambia, its poor economic climate but richly distributed health care system), and the organisational context (the Medical Research Council, a research institute for tropical diseases). The crossover trial design, and the methods used for gathering information were presented. Analysis of the results showed that the use of FDCS reduced errors and interview lengths, and that it was accepted by the users and the respondents. This trial demonstrated that a well-designed system using relatively cheap technology can achieve a significant improvement in health data collection.

A medical decision-aid system, ESTROPID (an Expert System for TROPIcal Diseases), was the focus of the second case study. The knowledge source, the knowledge representation, the knowledge manipulation mechanism and the operation of ESTROPID were explained. The study was conducted in Kenya (a relatively prosperous African country, but with high population growth rates and

increasing pressure on health delivery services) at the Kilifi District Hospital out-patient clinic (serving a wide catchment area of the coastal province). A staggered introduction and withdrawal of the system design was used, but staff availability constrained the length of the trial period to about two weeks. To enrich the technical comparisons between normal and computer consultations, a number of additional investigations of the system's performance were undertaken. The results were that the system was widely short of being an 'expert', but also indicated what types of changes to design might be necessary to bring the systems closer to a practical design. Though not conclusive, the results question whether technical excellence, even if attained, will lead to practical and sustainable implementations for the developing world in the medium term.

The field experience gained with these two case studies provided a more solid basis for discussing the types of systems that will be practical and effective, and why. Comparisons were made between the two types of systems, their application areas and their aims. Designing a computer-assisted data collection system is a much better defined task than the fuzzy outlines of medical decision-making. Thus, it was easier to make FDCS meet the specifications than it was for ESTROPID, and the trials could more easily demonstrate FDCS' effectiveness than ESTROPID's. In the short-term, it would be more efficient to concentrate effort on easily definable, well-established technological applications like FDCS rather than the more risky areas of machine intelligence. There is a need to realise that computerisation is not necessarily the right solution for developing countries, and more efficient manual ways may be found to increase productivity. Future system development will need to combine a thoughtful consideration of the need, a relevant design for the context, and a sound basis for a thorough assessment of their value.

7.2   RESEARCH SYNTHESIS

Various strands have emerged during the research presentation in this thesis - some have a practical implication, others have more theoretical contributions. These findings will be synthesised here into 4 parts dealing with the stages that have to be undergone in any research project with a similar emphasis - choosing

an application area, developing a system, putting the system into practice, and assessing its effectiveness.

1.   *Prioritising needs*

Researchers need to consider a number of aspects before proceeding with a project. Is the system directly connected to a real problem in a priority area? Is computerisation the best answer, or are there manual solutions which are adequate? Is the system institutionalisable, that is, can it be maintained within the organisation?

The research in this thesis was in response to the problem of making basic health services accessible to the majority of populations in developing countries. Having identified health as a key area, the next consideration is whether the use of IT is appropriate. The claim of this thesis is yes, it can be, but one needs to consider more how it is used, with the focus firmly placed on health rather than medicine. The basis for conducting research into CADAC and medical decision-aid systems was in fact an inherited premise - from suggestions by other researchers that the computerised solutions in these circumstances is good and viable. The goal of the research was then to investigate whether these assumptions would hold in practice. By comparing the manual and the computerised methods, the case studies provided a means of weighing up the two alternatives.

To avoid a wasteful use of resources, a thorough preliminary investigation should be done at this stage, to analyse whether the organisation is ready, and willing to integrate the system into routine work. This may involve a consideration of the current procedures, a clear definition of the expected role of the system, its data inputs and how these would be provided, the information produced as outputs and how they would be used. A look at the preliminary issues in this way provides more opportunities to periodically review the system's chances of survival.

2. *Practical designs*

In recent years, there has been much attention paid to the microcomputer, and in particular the IBM type PC. What is sometimes forgotten is that there are other microcomputer options that are just as viable, outside the PC world. A system must be matched to the need, and small may prove to be more sustainable. It is important that the right choice is made, as the hardware element tends to very much affect all other areas of design, implementation and sustainability.

A common problem with hand held computers is the limitations of the interface, generally restricting the options to keyboard input and text output. The potential of newer interface technologies such as touch screens, icon-based menu options and sound need to be explored further. It is important to remember that such interfaces have a high imported cultural content, and ways of adapting these symbols to reflect indigenous culture would need to be found. Working with low-cost hardware may limit the system's input and output functions to a basic minimum, and inventive interface design will be needed to stretch a system to its limits.

At a superficial level, issues of software design for computer applications in developing countries do not seem to differ substantially from those for industrialised countries. However, there remain some key distinctions that have to be made. Design is a cultural issue and systems must fit into the social organisation at the most basic level. Thus in the design of ESTROPID, the role the clinical officer plays in reassuring patients, giving nutritional advice, and so on, the whole procedure of delivering *care* has been interfered with. There is also a need to make software that can continue to be supported by local staff, to increase its chances of being successfully grafted into the recipient environment. In the FDCS design, operation of the system was divided into two levels, the first of these dealing with the minimum of options for the interviewers to do their jobs satisfactorily, unencumbered by a load of extraneous functions.

3.   Sustainable implementations

A major consideration for many project implementations is how to keep the systems operational after its developers have left. Can the system be run by and maintained by the people trained to do so? This is an important aspect of implementation, as the system could hardly be considered a success if it falls into disuse. Key elements to support the continued use (i.e. sustainability) of a system are the emphasis placed on the adequate training of computer staff and users, as well as the value placed on their involvement in the evolution of the system. Another element is that of the flow of resources to and from the system.

An adequate amount of training, preferably structured and organised, must be a pre-requisite. However, it must not stop here, continuing training will provide opportunities for ideas to be reinforced, and for the system to be stretched to its limits. Training can also provide a valuable feedback channel on user experience. User involvement in adapting the interface, and refining the evolution of the system, will be an invaluable means of identifying errors and constraints, as well as for suggesting extended functions for the system.

During implementation, the whole of system input and output process must be catered for, and its full span of interaction taken into account. This includes organisational aspects, in ensuring that the data is available at the right time; operational aspects in the uninterrupted provision of computer peripherals and recurring resources; and the objective aspect, in enabling the recipients of the resulting information to understand and value the system's output.

4.   Disciplined evaluation

Too often in the past, researchers have labelled ad hoc attempts as evaluation, undermining its value as a means of judging a system's worth. A structured approach to evaluation can uncover and challenge hidden assumptions in the expected role, or the design, of a system.

A formal evaluation process must be incorporated into the development life cycle and not simply regarded as a final, but disposable, stage. The use of a framework

such as the one suggested in this thesis provides a starting point for reporting results in a logical and thorough manner, encouraging the production and communication of valid and reliable conclusions.

If a preliminary trial is successful, ongoing monitoring may be a useful way of assessing the system's impact on overall health care system in the longer term. As computerised systems tend to formalise work procedures, these need to be monitored to ensure that changes are necessary, effective and beneficial in the longer term.

The studies reported here, in particular ESTROPID, emphasised a further dimension to evaluation. Ethical considerations must be taken into account, and a means of ensuring that no damage is done to the people involved in testing the system. The problems of accountability and data security, probably not generally major concerns for developing countries at the moment, will have to be considered. This is a familiar prerequisite for many medical researchers, but it is less so for IT researchers.

7.3   TOPICS FOR FURTHER STUDY

The research synthesis above leads naturally on to the identification of areas requiring further study.

1.   *Research methodology*

Methodologies need to be considered for thoroughly assessing the potential of systems in improving health care. Longitudinal designs should be built into system developments, to allow for a long-term view of systems success. Many systems have been brought 'into life' as it were, but few continue to be monitored during their lifetime. This should change, and projects using more formal and rigorous evaluation methods will provide a firmer basis for making decisions. Results of such activities can act as a pointer for other researchers, or at the least stimulate debate and discussion in peer-reviewed circles.

There is a need to adapt current techniques of integrating user involvement in the system development process, or propose new ways of doing so. Participative approaches should follow an understanding and appreciation of the skills and levels of training of the users. Indigenous knowledge and familiarity with the social and organisational context should be tapped to flavour the evaluation process.

There is a problem of the validity of case studies and the difficulty of generalisability. More in-depth case studies of IT applications in developing countries, both success stories and dismal failures, can serve as slices of a richer picture, contributing to a better understanding of how to build sustainable systems.

2.   *Areas of application*

Looking at the problems of health in developing countries in a wider context reveals a very broad spread of possible applications for informatics. Researchers coming to the field would do well to consider the full range. Examples of such application areas would include the use of information technology in training, support for administrative tasks, including the maintenance of patient records, as well as the traditional management of the supply chain. Mapping this vast terrain is a valid and useful research activity.

CADAC systems need to be studied further to investigate how their integration into the survey process can help to achieve health aims. This may be the setting up of computer-aided mechanisms for monitoring the health status of a high risk group. In areas where rapid assessment and reaction is necessary, for example refugee camps or post-disaster regions, CADAC may have a significant role to play. The studies undertaken here indicate that CADAC applications are more valid as a near future research area.

Medical decision-aid systems pose many interesting problems, but as argued in this thesis, they have a fundamental problem of uptake. It may be worthwhile for researchers to consider their use in more specific high-risk areas other than general treatment at the primary health care level. Computerised diagnostic-aids,

for example, may be useful for AIDS/HIV screening, allowing for rapid risk assessment, as well as a quick way of collecting medical data for epidemiological or other research purposes. Other ways of exploiting artificial intelligence, such as in intelligent tutoring aids or database interfaces, administration or logistic planning, could be further explored. The use of newer software technologies such as expert system shells and hypercard media should also be pursued.

TOWARDS THE FUTURE

The era that heralded concern about the effective development of the majority of the world's population began relatively recently. It follows a growing realisation that no attendant economic miracles or dramatic progress in social well-being has followed widespread decolonisation in newly independent states. The logic of purely imitative developmental models, which have tended to result in the blind acceptance of technologies from the North by the developing world, has been challenged. There have been some spectacular technology transfer failures in the past. Many information technology applications have unfortunately often continued in this trend. It is my contention that these past debacles need not be repeated, and that technology can be selectively applied, if more thought is given to how it can be fully integrated into the social and organisational context. It is my hope that this thesis will make a contribution by directing the efforts of future researchers to priority areas such as health, and to the implementation of sustainable applications.

BIBLIOGRAPHY

AGBALAJOBI, F., [1980], The unstructured approach to medical computing in Nigeria, in: Kaihara & Lindberg (eds.), "Medinfo 80", North-Holland, 656 - 659.

AGBALAJOBI, F., [1983], Indigenisation of health informatics through adaptation, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 173 - 180.

ANTILL, L., [1984], Research methods for information processing, in: Mumford, E. et al (eds.), "Research methods in information systems", IFIP Conference Proceedings, Manchester, 1984, North-Holland.

ASSOCIATION INTERNATIONALE FUTURIBLES, [1985], Quelle informatique pour quel developpement? Bilan d'une analyse documentaire sur l'informatique dans les pays en developpement, Rapport de l'Association Internationale Futuribles, Paris.

AUVERT, B., AEGERTER, P., VAN LOOK, F., DU, L.T.H., BOUTIN, P., MONIER, J.L., EMMANUELLI, X., GILBOS,V. & BENILLOUCHE, E., [1986a], A hand-held decision-aid system designed for rural health workers, Computers and Biomedical Research, 19, 80 - 89.

AUVERT, B., AEGERTER, P., GILBOS, V., BENILLOUCHE, E., BOUTIN, P., DESVE, G., LANDRE, M.F. & BOS, D., [1986b], TROPICAID: A portable expert system for medical decision-aid in developing countries, in: Salamon, R., Blum, B., & Jorgensen, M. (eds.), "Medinfo 86", North-Holland, 222 - 224.

AUVERT, B., AEGERTER, P., GILBOS, V., BENILLOUCHE, E., BOUTIN, P., DESVE, G., LANDRE, M.F. & BOS, D., [1986c], Has the time come for a medical expert system to go down in the bullring: The TROPICAID experiment, in: Trappl, R. (ed.), "Cybernetics and Systems 86", Reidel Publishing Company, 727 - 734.

AUVERT, B., GILBOS, V., AEGERTER, P. DU, L.T.H., BOUTIN, P., PHILIPPE, A.M., BENILLOUCHE, E., VAN LOOK, F. & EMMANUELLI, E., [1985a], An expert system as an aid to decision-making by para-medical personnel in developing countries, in: van Bemmel, J.H., Gremy, F. & Zvarova, J. (eds.), "Medical decision making: Diagnostic strategies and expert systems", Elsevier (North-Holland), 81 - 84.

AUVERT, B., BENILLOUCHE, E., AEGERTER, P., DIQUET, B., BOUTIN, P. & MONIER, J.L., [1985b], A drug information system for general practitioners using a hand-held computer, in: van Bemmel, J.H., Gremy, F. & Zvarova, J. (eds.), "Medical decision making: Diagnostic strategies and expert systems", Elsevier (North-Holland), 129 - 132.

BAARK, E., [1986], Global trends and implications for developing countries - Constraints and options for the development of new information technologies, Bulletin of the Advance Technology Alert System No 3, Centre for Science and Technology for Development, June 1986, 136 - 140.

BALACHANDRAN, C.S., FISHER, P.F., & STANLEY, M.A., [1989], An expert system approach to rural development: A prototype (TIHSO), The Journal of Developing Areas, No 23, 259 - 270.

BALDE, O., CRISAN, C., & CANTRELLE, P., [1988], Saisie controlee sur le terrain avec micro-ordinateur portable: Une experience sur les registres d'etat civil de Conakry, presented at Congres africain de population, Union Internationale pour l'Etude Scientifique de la Population, Dakar, November 1988.

BANKOWITZ, R.A., MCNEIL, M.A., CHALLINOR, S.M., PARKER, R.C., KAPOOR, W.N. & MILLER, R.A., [1989], A computer- assisted medical diagnostic

consultation service - Implementation and prospective evaluation of a prototype, Annals of Internal Medicine, 110, 824 - 832.

BASDEN, A., [1983], On the application of expert systems, International Journal of Man-Machine Studies, 19, 461 - 477.

BAUMER, J.M., [1980], Optimizing the appropriateness of technologies, in: de Schutter, J. & Berner, G. (eds.), "Fundamental aspects of appropriate technology", Delft University Press, 178 - 187.

BEHRENS, R., [1988], Discussion and proposals for a system of health monitoring in community based interventions, Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Unpublished document.

BEHRENS, R., [1988], Enhancing nutritional surveillance using a field-based data collection system, Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Unpublished document.

BELL, S. & WOOD-HARPER, T., [1990], Cultural aspects of information systems development, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 23 - 40.

BEMELMANS-SPORK, M.E.J. & SIKKEL, D., [1985], Data collection with hand-held computers, presented at International Statistical Institute meeting, Amsterdam, 1985.

BENNETT, J.M. & KALMAN, R.E. (eds.), [1981], Computers in developing nations, Proceedings of IFIP Conference, Melbourne 1980, North-Holland.

BERNARD, J., [1976], The medical decision, in: de Dombal, F.T. & Gremy, F. (eds.), "Decision Making and Medical Care", North-Holland, 3 - 9.

BERTRAND, W.E., [1985], Microcomputer applications in health population surveys: Experience and potential in developing countries, World Health Statistics Quarterly, 38, 91 - 99.

BESSANT, J., [1987], Information technology and the North- South divide, in: Finnegan, R., Salaman, G. & Thompson, K. (eds.), "Information technology: Social issues", Hodder and Stoughton, 163 - 180.

BHATNAGAR, S.C. & BJORN-ANDERSEN, N. (eds.), [1990], Information technology in developing countries, North-Holland.

BINET, J., [1983], Technologie moderne et societe africaine, Afrique Contemporaine, No 126, 14 - 23.

BOAFO, S.T.K., [1987], Les transferts des nouvelles technologies en afrique subsaharienne, Revue Tiers Monde, 28:3, 643 - 650.

BOARD OF SCIENCE AND TECHNOLOGY FOR INTERNATIONAL DEVELOPMENT (BOSTID), [1986], Microcomputers and their applications for developing countries, Westview Press.

BOBROW, D.G. & STEFIK, M.J., [1986], Perspectives on artificial intelligence programming, Science, 231, 951 - 957.

BOYE, T., FENICHEL, A. & MENNASEMAY, M., [1988], Skill transfer and African development: A conceptual research note, The Journal of Modern African Studies, 26:4, 685 - 690.

BRITISH MEDICAL INFORMATICS SOCIETY, [1989], The validation and clinical testing of decision-aids in medicine, Abstracts of papers presented at a workshop in February 1989.

BYASS, P. & CORRAH, P.T., [1989a], Assessment of a probabilistic decision support methodology for tropical health care, in: Manning, P,, Zhineng, O. & Barber, B. (eds.), "Medinfo 89", Elsevier (North-Holland).

BYASS, P., [1989b], Choosing and using a microcomputer for tropical epidemiology. 1: Preliminary considerations, Journal of Tropical Medicine and Hygiene, 92.

BYASS, P., [1989c], Choosing and using a microcomputer for tropical epidemiology. 2: Study implementation, Journal of Tropical Medicine and Hygiene, 92.

BYASS, P., HANLON, P.W., HANLON, L.C.S., MARSH, V.M. & GREENWOOD, B.M., [1988a], Microcomputer management of a vaccine trial, Computers in Biology and Medicine, 18:3, 179 - 193.

BYASS, P., [1988b], A probabilistic approach to health care delivery in Tropical Africa, in: Reinhoff, O., Piccolo, U. & Schneider, B. (eds.), "Expert systems and decision support in medicine", Lecture Notes in Medical Informatics, 36, Springer Verlag, 136 - 141.

BYASS, P., [1987], Computers in Africa: Appropriate technology?, Computer Bulletin, June 1987, 17.

BYASS, P., [1986], The design of computer forms for tropical medical research, Methods of Information in Medicine, 25, 229 - 232.

BYTE, [1989], Small wonders - A review of portable computers, Byte, August 1989, 14:8.

CARR, M., [1980], Appropriate technology: Theory, policy and practice, in: de Schutter, J. & Berner, J. (eds.), "Fundamental aspects of appropriate technology", Delft University Press, 145 - 153.

CASLEY, D.J. & LURY, D.A., [1987], Data collection in developing countries, 2e, Clarendon Press.

CHANDRASEKARAN, B., [1983], On evaluating AI systems for medical diagnosis, The AI Magazine, Summer 1983, 34 - 37.

CHANDRASEKARAN, B. & SMITH, J.W., [1985], Tenth annual workshop on artificial intelligence in medicine: An overview, The AI magazine, Summer 1985, 84 - 89.

CHECKLAND, P., [1984], Systems thinking, Systems practice, John Wiley.

CHERNS, A.B. & DAVIS, L.E., [1975], Status of the case as science, in: Davis, L.E. & Cherns, A.B. (eds.), "The quality of working life - Cases and commentary", Volume 2, Free Press, 20 - 25.

COLE, S., [1986], The global impact of information technology, World Development, 14:10/11, 1277 - 1292.

COLLINSON, P.O., JONES, R.G., HOWES, M., NICHOLLS, J., SHEEHY, N., BORAN, G.R. & CRAMP, D.G., [1988], Of mice and men - Data capture in the clinical environment.

COMMONWEALTH SCIENCE COUNCIL, [1986], Status of and computer applications (Africa) - Status country papers presented or contributed in regional workshop on fundamentals of microprocessor-based systems, Nairobi, Commonwealth Secretariat, CSC Technical Publications Series No 208.

CONYERS, D. & HILLS, P., [1982], An introduction to development planning in the Third World, John Wiley.

COONEY, S., [1986], Informatics in a developing world - Small is necessary: The role of information technology in development communication, in: 10th IFIP World Computer Congress, Dublin 1986.

CRISAN, C., KEITA, L. & CANTRELLE, P., [1988], Observation demographique suivie avec des micro-ordinateurs portables sur le terrain. Une experience en Guinee, in: Congres Africain de Population, Union Internationale pour l'Etude Scientifique de la Population, Dakar, November 1988.

CROFT, D.J., [1972], Is computerized diagnosis possible?, Computers and Biomedical Research, 5, 351 - 367.

DAVIS, R., [1986], Knowledge-based systems, Science, 231, 28 February 1986, 957 - 963.

DE DOMBAL, F.T., [1976], How objective is medical data?, in: de Dombal, F.T. & Gremy, F. (eds.), "Decision Making and Medical Care", North-Holland, 33 - 40.

DE MADRIGAL, L.M., [1983], Minimum program for medical informatics education in a health services system, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 291 - 294.

DELAPIERRE, M. & ZIMMERMAN, J.B., [1985], Utilisations et besoins, Terminal, December 1985, No 26, 10 - 14.

DENTENEER, D., BETHLEHEM, J.G., HUNDEPOOL, A.J. & KELLER, W.J., [1987], The BLAISE systems for computer-assisted processing, CBS Select No 4, Automation in survey processing, Netherlands Bureau of Statistics, 67 - 76.

DICKSON, G.W. & SIMMONS, J.K., [1970], The behavioral side of MIS: Some aspects of the 'people problem', Business Horizons, 13:4, 59 - 71.

DONABEDIAN, A., [1978], The quality of medical care - Methods for assessing and monitoring the quality of care for research and for quality assurance programs, Science, 200, 856 - 864.

DOUKIDIS, G. & FORSTER, D., [1990], The potential for computer-aided diagnosis of tropical diseases in developing countries: An expert system case study, European Journal of Operational Research, Special Issue on Developing Countries, 49:2.

DUNCAN, K., [1983], Educational prerequisites for medical informatics, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 283 - 290.

EL KHOLY, A. & MANDIL, S.H., [1983], Microcomputers and health improvement in developing countries, WHO Chronicle, 37:5, 163 - 165.

EL SAYED NOOR, A., [1984], A framework for the creation and management of national computing strategies in developing countries, The Computer Journal, 27:3, 193 - 200.

ELLIOTT, K., [1986], The concept of appropriate technology for health - Past and present, in: "PHC technologies at the family and community levels", Aga Khan Foundation.

ENGELBRECHT, R., [1988], Status of and research needs for expert systems in medicine, in: Lecture Notes in Medical Informatics, 36, 361 - 366.

ESSEX, B.J., [1975], Approach to rapid problem solving in clinical medicine, British Medical Journal, 5 July 1975, 34 - 36.

ESSEX, B.J., [1980], Diagnostic pathways in clinical medicine, 2e, Churchill Livingstone.

EVANS, J.R., HALL, K.L. & WARFORD, J., [1981], Shattuck lecture - Health care in the developing world: Problems of scarcity and choice, The New England Journal of Medicine, 305:19, 1117 - 1127.

FERNANDEZ PEREZ DE TALENS, A.F.P., MOLINO, E. & SHIRES, D. (eds.), [1983], Health informatics in developing countries: Experiences and viewpoints, North-Holland.

FERRY, B. & CANTRELLE, P., [1988], L'utilisation de micro-ordinateurs de terrain pour la collecte en demographie, in: Congres Africain de Population, Union Internationale pour l'Etude Scientifique de la Population, Dakar, Senegal, Novembre 1988, Session 1.2, 14 - 30.

FERRY, B., BERGES, J.C., COUILLET, F. & CANTRELLE, P., [1985], La saisie et le traitement d'informations statistiques en milieu severe a partir de micro-ordinateurs portatifs, Projet Systeme Integre de Collecte (S.I.C.), Document de travail No 11, ORSTOM.

FISCHBACHER, C.M., [1988], Review of PCEA Chogoria Hospital information management system, Unpublished document.

FORESTER, T. (ed.), [1985], The information technology revolution, Basil Blackwell.

FORSTER, D., [1989], FDCS: The manual, Department of Information Systems, London School of Economics and Political Science, Unpublished document.

FORSTER, D., [1987], DIAGNOSE: An exploratory approach to computer-aided diagnosis in developing countries, B.Sc. Computing project, London School of Economics and Political Science, Unpublished document.

FORSTER, D. & DOUKIDIS, G., [1989], ESTROPID: An Expert System for TROPIcal Diseases, in: Hunter, J., Cookson, J. & Wyatt, J. (eds.), "AIME 89", Proceedings of the Second European Conference on Artificial Intelligence in Medicine, Lecture Notes in Medical Informatics 38, Springer-Verlag, 88 - 92.

FOUNTAIN, D.E., [1982], Infirmier, Comment traiter votre malade, Simavi, Haarlem, Holland.

FOX, J., [1984], Expert systems: Towards a routine technology, in: Roger, R.H., Willems, J.L., O'Moore, R. & Barber, B. (eds.), "Medical informatics Europe 84", Lecture Notes in Medical Informatics, 24, Springer-Verlag, 19 - 23.

FOX, J., GLOWINSKI, A. & O'NEIL, M., [1987], The Oxford System of Medicine: A prototype information system for primary care, in: Fox, J., Fieschi, M. & Engelbrecht, R. (eds.), "Proceedings of the First Conference on artificial intelligence in medicine", Lecture Notes in Medical Informatics, 33, Springer Verlag, 213 - 226.

FOXON, J., [1988], Field trials in the Labour Force Survey lap-top computer project in November 1987, Office of Population and Census Surveys, Unpublished document.

FREEMAN, C., [1979], Social and economic impact of microelectronics, in: Proceedings of the 1979 workshop of technology assessment - Its role in national and corporate planning, AGPS, 13 - 29.

GASCHNIG, J., KLAHR, P., POPLE, H., SHORTLIFFE, E. & TERRY, A., [1983], Evaluation of expert systems: Issues and case studies, in: Hayes-Roth, F., Waterman, D.A. & Lenat, D.B. (eds.), "Building expert systems", Addison-Wesley, 241 - 280.

GEVARTER, W.B., [1987], The nature and evaluation of commercial expert system building tools, Computer, May 1987, 24 - 41.

GHOSH, P.K. (ed.), [1984], Technology policy and development, International Development Resource Books No. 3, Greenwood Press, Connecticut.

GINZBERG, M.J. & ZMUD, R.W., [1988], Evolving criteria for information systems assessment, in: Bjorn-Andersen, N. & Davis, G.B. (eds.), "Information systems assessment", Elsevier (North-Holland), 41 - 55.

GOLDBERGER, H. & SCHWENN, P., [1983], Man-machine symbiosis in the assistance and training of rural health workers: A proposal, in: Pages, J.C., Levy, A.H., Gremy, F. & Anderson, J. (eds.), "Meeting the challenge: Informatics and medical education", Elsevier (North-Holland), 295 - 306.

GOLDERGER, H., SCHWENN, P., EMANUELLI, X., CHESNAIS, A., DHERBECOURT, N. & DE FERAUDY, H., [1983], Computer aided medical care in developing countries, Report to the Scientific Council, October 1983, Unpublished document.

GOOS, G., [1973], Documentation, in: Bauer, F.L. (ed.), "Advanced course in software engineering", Lecture Notes in Economics and Mathematical Systems, No 81, 385 - 393.

GOWERS, P., [1987], Selection of CHWs, A case study from The Gambia, Footsteps to Health, March 1987, No 5, 4 - 5.

GRAHAM, E.M., [1981], Comment, in: Streeten, P. & Jolly, R. (eds.), "Recent issues in world development - A collection of survey articles", Pergamon Press, 111 - 114.

GRAHAM, W., [1986], Health status indicators in developing countries - A selective review 1986, Commonwealth Secretariat.

GREMY, F., [1983], Medical informatics in developing countries, Balance sheet and prospectives, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 17 - 25.

GROSS-NIKLAUS, M., [1987], Benchtest: Psion Organiser II, Personal Computer World, August 1987, 114 - 118.

GROVES, R.M., [1983], Implications of CATI: Costs, errors, and organization of telephone survey research, Sociological Methods and Research, 12:2, 199 - 215.

GROVES, R.M. & NICHOLLS, W.L., [1986], The status of computer-assisted telephone interviewing: Part II - Data quality issues, Journal of Official Statistics, 2:2, 117 - 134.

GUPTA, G.K., [1987], Role of computer technology in developing countries, Information Technology for Development, 2:1, 43 - 58.

GUY, K. & ARNOLD, A., [1987], Global trends in microelectronic components and computers, United Nations Industrial Development Organisation, Technology Trends Series No 3.

HARRISON, P., [1983], The Third World tomorrow, 2e, Pelican.

HARRISON, P., [1990], Too much life on earth?, New Scientist, 19 May 1990, 28 - 29.

HILL, A.G., [1987], Use of demographic data and health statistics to inform primary health care, Centre for Population Studies, London School of Hygiene and Tropical Medicine.

HIRSCHHEIM, R. & SMITHSON, S., [1988], A critical analysis of information systems evaluation, in: Bjorn-Andersen, N. & Davis, G.B. (eds.), "Information systems assessment", North-Holland, 17 - 37.

HOBDAY, M., [1987], The international telecommunications industry: The impact of microelectronics technology and implications for developing countries, United Nations Industrial Development Organisation, Technology Trends Series No. 4.

HOLLNAGEL, E., [1989], Evaluation of expert systems, in: Guida, G. & Tasso, C. (eds.), "Topics in expert system design", Elsevier Science Publishers, 377 - 416.

HUDSON, D.L. & ESTRIN, T., [1984], Derivation of rule-based knowledge from established medical outlines, Computers in Biology and Medicine, 14:1, 3 - 13.

HURTADO, M.E. & NEWMAN, A., [1986], March of the barefoot microchip, South, September 1986, 111 - 114.

IKONICOFF, M., [1987], Trois theses erronees sur l'industrialisation du tiers monde, Revue Tiers Monde, 28, No. 110, 247 - 261.

JAIN, A.K., [1987], Expert systems: Prospects for developing countries, United Nations Industrial Development Organisation.

KAIHARA, S., [1983], Chairman's note with some experiences in Japan, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 243 - 246.

KALMAN, R.E., [1981], Eight strategic issues for informatics, in: Bennett, J.M. & Kalman, R.E. (eds.), "Computers in developing nations", North-Holland, 23 - 37.

KASTNER, J.K., DAWSON, C.R., WEISS, S.M., KERN, K.B. & KULIKOWSKI, C.A., [1984], An expert consultation system for frontline health workers in Primary Eye Care, Journal of Medical Systems, 8:5, 389 - 397.

KAUL, M. & KWONG, H.C., [1988], The computerisation of government in Africa, Commonwealth Institute report.

KEEN, P.G.W., [1981], Information systems and organisational change, Communications of the ACM, 24:1, 24 - 33.

KELSEY, R., [1988], Expert systems as a preliminary design tool for drinking water supplies in developing countries, Science Technology and Development, 6:1, 13 - 20.

KENYAN MINISTRY OF HEALTH, [1988], Coast Province Annual Report 1987 - Health information system.

KLEINMUNTZ, B., [1984], Diagnostic problem solving by computer: A historical review and the current state of the science, Computers in Biology and Medicine, 14:3, 255 - 270.

KLUZER, S., [1990], Computer diffusion in Black Africa: A preliminary assessment, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 175 - 188.

KOHLI, F.C., [1986], Information policy issues in developing countries, in: Kugler, H.J. (ed.), "Information processing 86", North-Holland, 591 - 597.

KORPELA, M., [1990], The role of information systems ergonomics in developing countries, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 265 - 275.

KORPELA, M., [1988a], Ergonomics of developing countries, Occasional papers 3, University of Helsinki, Institute of Development Studies.

KORPELA, M., [1988b], How a hospital computer and a health center computer began to talk to each other, presented at Informatics 88, International Conference on Informatics for Health, Havana, Cuba, February 1988.

KORPELA, M., [1988c], Research plan: Nigerian experiences in health care information systems development, A macroergonomical analysis, University of Kuopio, Finland.

KORPELA, M., [1988d], Travel report, Exploratory visit to Obafemi Awolowo University (Ile-Ife, Nigeria), University of Kuopio, Finland.

KRYSTALL, A., [1986], Introduction of microcomputers for management needs - Sharing the experience of Chogoria Hospital, Thunder and Associates Inc, Nairobi, Kenya.

KULIKOWSKI, C.A., [1983], Progress in expert AI medical consultation systems: 1980-1983, in: van Bemmel, Ball & Wigertz (eds.), "Medinfo 83", IFIP-IMIA, North-Holland, 499 - 502.

KURASHINA, S., ISHIDA, A., SUZUKI, S. & KAMINUMA, T., [1980], Comparative study of computer-aided clinical decision making systems, in: Lindberg & Kaihara (eds.), "Medinfo 80", North-Holland, 825 - 829.

LANCET EDITORIAL, [1989], Computer-assisted diagnosis, The Lancet, December 9, 1371.

LANGE, H.J., [1976], Decision-making by doctors and aided by computers, in: de Dombal, F.T. & Gremy, F. (eds.), "Decision Making and Medical Care", North-Holland, 11 - 21.

LENT, J.A., [1987], Technologies de l'information: Mythes, questions et alternative, Revue Tiers Monde, 28, No. 111, 613 - 625.

LIEBOWITZ, J., [1986], Useful approach for evaluating expert systems, Expert Systems, 3:2, 86 - 96.

LINDBERG, D.A.B., [1979], Information systems for rural health: Problems and potential benefits, in: Health Care Technology Evaluation, Proceedings, Springer-Verlag, 61 - 65.

LONGE, O., [1981], Computer applications to basic needs of developing nations, in: Bennett, J.M. & Kalman, R.E. (eds.), "Computers in developing nations", North-Holland, 189 - 193.

LYBERG, L., [1985], Plans for computer-assisted data collection at Statistics Sweden, in: ISI Proceedings of the 45th Session, Invited Papers, Volume LI, Book 3, Section 18.2.

LYBERG, L. & DEAN, P., [1988], The use of technology at Statistics Sweden, presented at Symposium 88: The Impact of High Technology on Survey Taking, Ontario, Canada, October 1988.

LYBERG, L. & SUNDGREN, B., [1987], The impact of the development of EDP on statistical methodology and survey techniques, ECE, Geneva.

MADU, C.N., [1989], Transferring technology to developing countries - Critical factors for success, Long Range Planning, 22:4, 115 - 124.

MANDIL, S., [1983], Present trends in health informatics in developing countries, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North- Holland, 37 - 39.

MARSDEN, M., [1984], Progressive technologies for developing countries, in: Ghosh, P.K. (ed.), "Technology policy and development, A third world perspective", International Development Resource Books No 3, Greenwood Press, 292 - 315.

MEINDL, J.D., [1985], Micros in medicine, in: Forester, T. (ed.), "The information technology revolution", Basil Blackwell, 359 - 371.

MILLER, P.L, [1988], Evaluation of artificial intelligence systems in medicine, in: Miller, P.L. (ed.), "Selected topics in medical artificial intelligence", Springer Verlag, 202 - 211.

MILLER, P. (ed.), [1988], Selected topics in medical artificial intelligence, Springer-Verlag.

MOHAN, L., BELARDO, S. & BJORN-ANDERSEN, N., [1990], A contingency approach to managing information technology in developing nations: Benefitting from lessons learned in developed nations, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 15 - 22.

MOON, Y.S. & LO, T.F., [1988], ROPES - A rule-based outpatient expert system, in: Reinhoff, O., Piccolo, U. & Schneider, B. (eds.), "Expert systems and decision support in medicine", Lecture Notes in Medical Informatics, 36, Springer-Verlag, 38 - 43.

MUMFORD, E., [1984], Researching people problems: Some advice to a student, in: Mumford, E. (ed.), "Research methods in information systems", IFIP conference proceedings, Manchester 1984, North-Holland.

MURRAY, G.D., [1985], Computer-aided diagnosis and teaching, in: MacFarlane, P.W., "Computer techniques in clinical medicine", Butterworths, 214 - 221.

NDONGKO, W.A. & ANYANG, S.O., [1981], The concept of appropriate technology: An appraisal from the Third World, Monthly Review, February 1981, 35 - 43.

NICHOLLS, W.L., [1983], CATI research and development at the Census Bureau, Sociological Methods and Research, 12:2, 191 - 197.

NICHOLLS, W.L. & GROVES, R.M., [1986], The status of computer-assisted telephone interviewing: Part I - Introduction and impact on cost and timeliness of survey data, Journal of Official Statistics, 2:2, 93 - 115.

NILSEN, S.E., [1979], The use of computer technology in some developing countries, International Social Science Journal, 31:1, 513 - 528.

O'KEEFE, R.M., BALCI, O. & SMITH, E.P., [1987], Validating expert system performance, IEEE Expert, Winter 1987, 81 - 89.

O'SHEA, T., & EISENSTADT, M. (ed.), [1984], Artificial Intelligence - Tools, techniques and applications, Harper and Row.

OBENSON, P., [1983], Computer science research in Africa: Needs, priorities and problem - The IAI setting, presented at UNESCO Regional Workshop on Research and Application of Informatics in Africa, Lagos, Nigeria, December 1983.

OKOT-UMA, R.W'O., [1988], Information technology applications in developing countries: A synthesis perspective, presented at British Computer Society seminar on information technology applications in developing countries, London, March 1988.

PAKER, Y., [1981], Informatics and development: Unesco's approach for the 80's, in: Bennett, J.M. & Kalman, R.E. (eds.), "Computers in developing nations", North-Holland, 11 - 20.

PANKHURST, R.J., [1980], Medical diagnosis in developing countries, Computers in Biology and Medicine, 10, 69 - 82.

PARTRIDGE, D., [1989], To add AI, or not to add AI?, in: Kelly, B. & Rector, A. (eds.), "Research and Development in Expert Systems V", Proceedings of Expert Systems 88, BCS Workshop Series, Cambridge University Press.

PEARSON, L., [1970], The crisis of development, Praeger Publishers.

PORENTA, G., PFAHRINGER, B., HOBERSTORFER, M. & TRAPPL, R., [1988], A decision support system for village health workers in developing countries, Applied Artificial Intelligence, 2, 47 - 63.

RADHAKRISHNA, S. (ed.), [1980], Science, technology and global problems: Views from the developing world, Pergamon Press.

RATEAU, O., [1981], Informatics: Evolution and use in the 1980's, with particular reference to developing countries, in: Bennett, J.M. & Kalman, R.E. (eds.), "Computers in developing nations", North-Holland, 3 - 9.

REDDY, A.K.N., [1988], Appropriate technology: A reassessment, in: Sardar, Z. (ed.), "The revenge of Athena: Science, exploitation and the Third World", Mansell, 290 - 308.

REINHOFF, O., [1983], A curriculum for short-term training in medical informatics for developing countries, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 295 - 304.

REINHOFF, O., PICCOLO, U. & SCHNEIDER, B. (eds.), [1988], Expert systems and decision support in medicine, Lecture Notes in Medical Informatics, 36, Springer-Verlag.

REITMAIER, P., [1985], Child health map: A computer-based method for the registration and mapping of the nutritional status, the participation in immunisation programs and the breast feeding rate in child populations of underdeveloped countries, University of Heidelberg.

REITMAIER, P., DUPRET, A. & CUTTING, W.A.M., [1987], Better health data with a portable microcomputer at the periphery: An anthropometric survey in Cape Verde, Bulletin of the World Health Organisation, 65:5, 651 - 657.

ROBEY, D., GUPTA, S.K. & RODRIGUEZ-DIAZ, A., [1990], Implementing Information Systems in Developing Countries: Organisational and Cultural Considerations, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information Technology in Developing Countries", North-Holland.

ROSSI-MORI, A. & RICCI, F.L., [1988], On the assessment of medical expert systems, in: Reinhoff, O., Piccolo, U. & Schneider, B. (eds.), "Expert systems and decision support in medicine", Lecture Notes in Medical Informatics, 36, Springer-Verlag, 292 - 297.

RUSH, H.J., [1984], Sources of technology in development, in: Ghosh, P.K. (ed.), "Technology policy and development, A third world perspective", International Development Resource Books No 3, Greenwood Press.

SALAMON, R., [1983], Medical decision making and developing countries, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 247 - 250.

SARDAR, Z. (ed.), [1988], The revenge of Athena: Science, exploitation and the Third World, Mansell.

SAVANE, L., [1984], Problems of scientific and technological development in Black Africa, in: Ghosh, P.K. (ed.), "Technology policy and development, A Third World perspective", International Development Resource Books No 3, Greenwood Pres, 371 - 379.

SCHWARE, R., [1987], Software industry developing in the third world: Policy guidelines, institutional options, and constraints, World Development, 15:10/11, 1249 - 1267.

SCHWARE, R. & TREMBOUR, A., [1985], Rethinking microcomputer technology transfer to third world countries, Science and Public Policy, 12:1, 15 - 20.

SHANKS, J.M., [1989], Information technology and survey research: Where do we go from here?, Journal of Official Statistics, 5:1, 3 - 21.

SHAW, M., [1987], Using and programming the Psion Organiser II, Kuma Computers Ltd.

SHIRES, D.B., [1988], Developing primary care information systems for Latin American/Caribbean countries, Dalhousie University, Canada, Unpublished document.

SHIRES, D.B., [1983a], Health informatics for all by the year 2000, Methods of Information in Medicine, 22, 61 - 62.

SHIRES, D.B., [1983b], Health information systems applications in technologically lesser-developed countries: An overview, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 57 - 63.

SHIRES, D.B., [1982], Health informatics in a changing world, AGORA, July/December 1982, 24 - 25.

SHORTLIFFE, E.H., [1987], Computer programs to support clinical decision making, Journal of the American Medical Association, 258:1, 61 - 66.

SNOW, R.W., [1989], Malaria concordance study, Unpublished data, KEMRI Coastal Unit, Kilifi, Kenya.

SNOW, R.W. & BYASS, P., [1988], Computerised morbidity surveillance in the field, in: Proceedings of the African Population Conference, International Union for the Scientific Study of Population, Session 1.2, Dakar, Senegal, November 1988.

SOFT WAREHOUSE, [1986], muLISP-86 manual.

SPIEGELHALTER, D. & WYATT, J., [1990], The evaluation of decision technology. 2: Laboratory testing, (in press).

STEPHENS, J., ALONSO, P.L., BYASS, P. & SNOW, R.W., [1989], Tropical epidemiology: A system for continuous demographic monitoring of a study population, Methods of Information in Medicine.

STEWART, F., [1977], Technology and underdevelopment, Macmillan.

STEWART, F., [1981], International technology transfer: Issues and policy options, in: Streeten, P. & Jolly, R. (eds.), "Recent issues in world development", Pergamon Press, 67 - 110.

STREETEN, P., [1984], Technology gaps between rich and poor countries, in: Ghosh, P.K. (ed.), "Technology policy and development, A third world perspective", International Development Resource Books No 3, Greenwood Press, 7 - 25.

STREETEN, P., [1981], First things first: Meeting basic human needs in the developing countries, Oxford University Press.

SUTTON, G.C., [1989], How accurate is computer-aided diagnosis?, The Lancet, October 14, 905 - 908.

SYMONS, V. & WALSHAM, G., [1988], The evaluation of information systems: A critique, Journal of Applied Systems Analysis, 15, 119 - 132.

SZOLOVITS, P. & PAUKER, S.G., [1978], Categorical and probabilistic reasoning in medical diagnosis, Artificial Intelligence, 11, 115 - 144.

SZOLOVITS, P., PATIL, R.S. & SCHWARTZ, W.B., [1988], Artificial intelligence in medical diagnosis, Annals of Internal Medicine, 108:1, 80 - 87.

THE COURIER, [1989], Dossier - Informatics and development, No 113, January - February 1989, 50 - 82.

THE ECONOMIST, [1989], Poor man's burden: A survey of the Third World, September 23, 1989, The Economist.

THE WORLD BANK, [1989a], Sub-Saharan Africa: From crisis to sustainable growth, The World Bank.

THE WORLD BANK, [1989b], World Development Report 1989, Oxford University Press, New York.

THE WORLD BANK, [1981], The Gambia: Basic needs in The Gambia, A World Bank Country Study.

TODARO, M. P., [1982], Economics for a developing world, 2e, Longman.

TORO, V. & CHAMBERLAIN, K., [1989], Recent work with microcomputers for census processing in developing countries, Journal of Official Statistics, 5:1, 69 - 81.

TOURETZKY, D.S., [1984], Lisp - A gentle introduction to symbolic computation, Harper and Row.

UEMURA, K., [1983], Statistical components and requirements for health information systems, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 199 - 203.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANISATION, [1986], The impact of expert systems, Regional and Country Studies Branch, Studies and Research Division.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANISATION, [1984], Technological self-reliance of the developing countries: Toward operational strategies, in: Ghosh, P.K. (ed.), "Technology policy and development, A third world perspective", International Development Resource Books No 3, Greenwood Press, 97 - 174.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANISATION, [1983], Microelectronics and developing countries: Towards an action-oriented approach.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANISATION, [1981], Implications of micro-electronics for developing countries: A preliminary overview of issues.

UPLEKAR, M.W., ANTIA, N.H. & DHUMALE, P.S., [1988], Sympmed I: Computer program for primary health care, British Medical Journal, 297, 841 - 843.

VALLBONA, C., [1983], Information systems for health care in developing countries, in: Fernandez Perez de Talens, A.F.P., Molino, E. & Shires, D. (eds.), "Health informatics in developing countries: Experiences and viewpoints", North-Holland, 87 - 100.

VAN BASTELAER, A., KERSSEMAKERS, F. & SIKKEL, D., [1987], A test of the Netherlands Continuous Labour Force Survey with hand-held computers: Interviewer behaviour and data quality, CBS Select No. 4, "Automation in survey processing", Netherlands Bureau of Statistics, 37 - 54.

VAN BASTELAER, A., KERSSEMAKERS, F. & SIKKEL, D., [1988], Data collection with hand-held computers: Contributions to questionnaire design, Journal of Official Statistics, 4:2, 141 - 154.

VAN BEMMEL, J.H., GREMY, F. & ZVAROVA, J. (eds.), [1985], Medical decision-making: Diagnostic strategies and expert systems, Elsevier (North-Holland).

VASQUEZ, S. & ZIMMERMAN, J.B., [1987], Information technology policies in the developing countries, in: "Report from Namur, Landscapes for an information society", IFIP Working Group 2.

VOGEL, L.C., MULLER, A.S., ODINGO, R.S., ONYANGO, Z. & DE GEUS, A. (eds.), [1974], Health and disease in Kenya, East African Literature Bureau, Nairobi.

VUISTER, M., [1988], Do's and don'ts for microcomputer hardware, Unpublished Communication.

WALSHAM, G., SYMONS, V. & WAEMA, T., [1988], Information systems as social systems: Implications for developing countries, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 51 - 62.

WHITE, K.L., [1985], Health surveys: Who, why and what?, World Health Statistics Quarterly, 38, 2 - 14.

WILSON, R.G., ECHOLS, B.E., BRYANT, J.H. & ABRANTES, A. (eds.), [1988], Management information systems and microcomputers in primary health care, Aga Khan Foundation.

WOFSEY, M.M. & DICKIE, P.M., [1971], Computers in less developed economies, Finance and Development, March 1971, 8, 2 - 6.

WORLD HEALTH ORGANISATION, [1987], How should information on health care be generated and used?, World Health Forum, 8:4, 409 - 438.

WORLD HEALTH ORGANISATION, [1981a], Health programme evaluation: Guiding principles for its application in the managerial process for national health development, Health for All series, No 6.

WORLD HEALTH ORGANISATION, [1981b], Managerial process for national health development - Guiding principles for use in support of strategies for Health for All by Year 2000, Health for All Series, No 5.

WORLD HEALTH ORGANISATION, [1981c], Development of indicators for monitoring progress towards Health for All by the Year 2000, Health for All Series, No. 4.

WORLD HEALTH ORGANISATION, [1980], Selection of essential drugs, Technical Report No. 641.

WORLD HEALTH ORGANISATION, [1978], Primary Health Care. International Conference on Primary Health Care, Alma-Ata, USSR, September 1978.

WYATT, J. & SPIEGELHALER, D., [1990a], The evaluation of decision technology. 1: The scope of decision technology and the need for evaluation, (in press).

WYATT, J. & SPIEGELHALTER, D., [1990b], The evaluation of decision technology. 3: Field testing, (in press).

WYATT, J., [1989], Lessons learned from the field trial of acorn, An expert system to advise on chest pain, in: Manning, P., Zhineng, O. & Barber, B. (eds.), "Medinfo 89", Beijing, Elsevier North Holland.

WYATT, J., [1987], The evaluation of clinical decision support systems: A discussion of the methodology used in the Acorn Project, in: Fox, J., Fieschi, J. & Engelbrecht, R. (eds.), "Proceedings of the European Conference of AI in Medicine", Heidelberg, Springer-Verlag, 15 - 24.

ZIMMERMAN, J.B., [1984], Politiques africaines de l'informatique, Politique Africaine, March 1984, No 13, 79 - 90.

ZIMMERMAN, J.B., [1990], Information technology policies for developing countries facing a worldwide IT industry, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North-Holland, 79 - 83.

ZWANGOBANI, E., [1990], Information technology in Africa: An overview, in: Bhatnagar, S.C. & Bjorn-Andersen, N. (eds.), "Information technology in developing countries", North- Holland, 167 - 173.

APPENDIX A

THE FDCS MANUAL

The Field Data Collection System (FDCS) is a package which provides a simple, adaptable and effective method of collecting data. It runs on a Psion Organiser, a small hand-held computer that is compact and easy to operate in the field yet powerful enough to store and manage a reasonable amount of data with space for up to 131,072 (128k) characters per storage device.

It strength lies in its versatility and error checking capability. Several questionnaires can be stored and used in rotation, as required in the field. Using the editing facilities available, questions and questionnaire formats can be changed at any stage during data collection. Formats for the responses expected are specified as well as the question texts; these formats are used by error checking routines during data entry to ensure that only valid input is stored.

The system includes a procedure that allows questions to be skipped if a pre-specified condition holds true. Simple edit checks can also be included to ensure no two logically impossible values can be entered. Other features include a varied level of password access, automatic date and time stamping of records, translation of date entries into ages, and a routine for calculating weight as a percentage of the NCHS standard.

At any point during data entry, a "paging back" facility allows the editing of any previous entries by using a simple sequence of keys. Help messages can also be scrolled onto the screen. A full record review and edit facility is possible before saving each completed questionnaire. Simple data summaries provide variable frequencies.

FDCS eliminates the need for data punching, as the information entered is already in a computer-readable form and can be readily transferred to a micro-computer. It reduces the effort required for pre-analysis data cleaning.

HOW TO USE THIS MANUAL

This manual explains in detail the facilities FDCS offers by giving clear demonstration examples as well as guidelines for correcting errors and explanations of the accompanying messages. It is basically divided into 4 parts.

The first part is introductory. Section 1 introduces the Psion, explains the functions of its components and how to operate the keyboard. Section 2 defines various terms that are used throughout the Manual, explains how to load FDCS and describes errors that may occur during installation of the package.

The second part explains the essentials to questionnaire specification and use. Section 3 explains how to construct simple questionnaire files; Section 4 how to edit questionnaire files; Section 5 how to input and edit a data entry; Section 6 how to view the contents of a data file; and Section 7 how to use the simple statistics facilities.

The third part explains the enhanced optional features of FDCS. Section 8 on adding passwords and how to automatically stamp records with passwords, the date or time; Section 9 on how to specify branching conditions; Section 10 on specifying editing checks; Section 11 on the indexing facility, and Section 12 on the NCHS library routines; and Section 13 explains how the use of these options affect data entry.

The final part concentrates on data transfer, and Section 14 explains the procedure for downloading data files collected and stored on the Psion to an IBM-compatible PC.

All the sections are annotated with diagrams showing how menus, questions, prompts and error messages will be displayed on the Organiser's screen.

SECTION 1
INTRODUCTION TO THE PSION

VARIOUS COMPONENTS

The Psion is enclosed in a hard protective case which pulls back to reveal the keyboard. Each of the 36 keys can be pressed singly or in combination with the SHIFT key.

Datapacks store various types of information (programs, data files, etc.) and fit into the 2 slots on the back of the Psion. Just above these is a dial for adjusting the contrast on the screen. With this, the brightness of the screen can be altered to make reading easier when the Psion is held at different angles.

The slot at the top of the machine slides back to reveal an opening through which the Psion can be connected with another computer via a cable and data transferred. The protective case can be removed by tugging it a bit harder. At the bottom is the battery compartment.

THE KEYBOARD

Most of the keys have double functions and produce the alternative character indicated in yellow above them when pressed simultaneously with **SHIFT**.

The *alphabetic* keys are printed in white. They produce capital or small letters depending on what mode the keyboard is set to. *Numeric* keys are on a blue background in the middle of the keyboard. The numbers are printed above the keys in black. *Punctuation and special* character keys are the alternative keys of the remainder of the alphabetic keys and are printed in white.

Function keys are printed in yellow. The most important of these (**ON/CLEAR, MODE, EXE**) are on yellow backgrounds.

ON	(with CLEAR written above it), turns the Psion on or clears whatever is at present displayed on the screen.
MODE	is an alternative action key. It's uses vary for the Psion's internal functions and will be explained where required.
EXE	indicates that you have finished giving the Psion some instructions and now want it to EXEcute these.
^ v < >	the keys with arrows allow you to move one row up, one row down, one character left, one character right respectively. Their effect varies slightly when different types of information is being entered. These will be explained in the relevant function descriptions.
DEL	usually DELetes one character to the left. When pressed in combination with SHIFT, deletes the character underneath the cursor.
SPACE	is a character in its own right and is printed as a blank space.
SHIFT	produces the alternative function when simultaneously pressed with a key that has another label above it. At any one time the Organiser's keyboard is set to one of 3 different states: capitals, small letters or numeric. When set to capitals, all alphabetic keys pressed will produce A, ... Z, This is the default state when the Psion is first switched on. It can be changed to small letters by pressing **SHIFT** and ^ (with

CAP written above it). Repeat this to reverse the effect. Pressing the shift key with one of the white keys will always print the alternative numeric key. The keyboard can be set to the numeric state by pressing **SHIFT** and **v** (with NUM written above it).

CHANGING THE BATTERY

Never remove the battery with the datapacks in place; always remove the datapacks first. Open the battery compartment at the bottom of the Psion and, if changing an old battery, remove it and slide in the new one as quickly as possible. If you take longer than 30 seconds, the internal storage of the Psion will be erased. The date and time will have to be reset as these are also cleared.

FITTING AND REMOVING DATAPACKS

Fit a pack by sliding it into a slot and clicking it in place. Remove it by gripping firmly on the roughened surface and sliding it out. Never remove packs when a program is running unless instructed to do so as this may corrupt data files.

CARING FOR THE PSION

Never leave the top slot open and always keep the Psion in its protective case when not in use. Protect it as much as possible from exposure to excessive dust, dirt and rain.

SECTION 2
HOW TO START

DEFINITIONS

First, some preliminary definitions of several words that are used in the rest of the manual.

The *cursor* is a mark that shows where the next piece of information will become visible on the screen. As it does so, the cursor moves on to the next position. On a menu, a flashing rectangle covers the first letter of an item and the cursor can be moved on to other positions by using the arrow or cursor keys ∧v<>.

A *character* is any number, letter or shape that can be made by the Psion. An *alphanumeric* character is one that is either alphabetic (A...Z) or numeric (0...9). A character can also be currency signs £, $ punctuation marks ; " or arithmetic signs % +. A *digit* is a numeric character, e.g. 4385 is composed of 4 digits. *Integers* are whole numbers, e.g. 4, 897, 0, and *decimals* are numbers with fractional parts, e.g. 6.9, 1.5.

A *record* is a collection of data items relating to an individual entity. e.g. a complete record of a patient admitted to a hospital may include date of admission, name of patient, sex, age, symptoms, diagnosis and treatment. Each of these headings for data items is called a *field*. A collection of such records is a data *file*.

A *prompt* is a message from the computer telling the user that some action is needed before it can continue. It might consist of words or a flashing cursor on a screen. A *menu* is a display that lists several alternatives for the user to choose from. On the Psion, it is a list of words any of which may be chosen to activate some function.

If a display is too long to be shown on the Psion's small screen at the same time, parts of it can be viewed in sections instead. The display will *scroll* or move along vertically, horizontally, continuously or a single line at a time.

GETTING STARTED

Press the **ON/CLEAR** button to turn the Psion on. If it has been pre-set to do so, the option FDCS should appear as the first item in the main menu (called "main" because it controls access to other functions)

```
FDCS FIND SAVE
DIARY CALC PROG
```
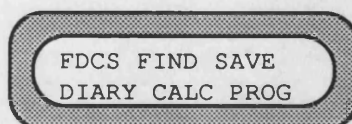
The flashing rectangle over the F is the cursor. To demonstrate how other options in the menu can be chosen, use the cursor keys ∧v<> to move the cursor to other menu items. You will discover that the menu list is longer than shown above. The screen scrolls up or down to reveal more options. To choose an option from the menu, you may move to the option using the cursor keys and then press **EXE**.

Alternatively, press the first letter of the option you require and it will be immediately executed. If there is more than one choice starting with the same letter, the cursor skips to the next option starting with that letter. In this case, you have to press **EXE** to execute an option.

If all goes well, the following submenu will appear

```
RUN VIEW SUMMARY
LOAD FILES QUIT
```

The following sections explain the functions of these options.

WHAT MAY GO WRONG?

If there is no flashing rectangle, the keyboard is set to receive numeric inputs and a cursor line will be displayed under an option instead. Either choose an option by pressing **SHIFT** and a key to trigger the alphabetic function or set the keyboard to alphabetic by pressing **SHIFT** and **NUM** simultaneously.
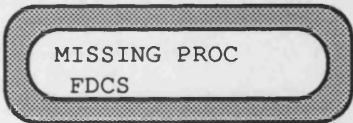
If the option FDCS does not appear in your list, move the cursor to the beginning of the menu and press the **MODE** key. The prompt to insert an item asks you to enter the name of the new menu item, carefully type in FDCS

```
INSERT ITEM
FDCS
```
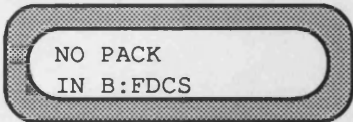
and press **EXE**. A menu will appear as described above. To choose FDCS, press **F** or **EXE** if the cursor is flashing on F.

If you have not inserted the correct datapacks and the Psion cannot find the instructions needed to run the package, the message

```
MISSING PROC
FDCS
```

will be displayed. If one of the packs is missing, the message

```
NO PACK
IN B:FDCS
```

will be displayed. In both these cases, press **ON/CLEAR**, turn the Psion off, insert the datapacks as required and repeat the above steps.

SWITCHING OFF

To do this, go back to the main menu by pressing the **ON/CLEAR** key or choosing the QUIT options in FDCS as many times as necessary. Press **O** to choose the OFF option. Alternatively, if the Psion is left off for 5 minutes without being instructed to do anything, it automatically switches itself off to conserve the batteries. When the **ON/CLEAR** key is next pressed, it carries on exactly where it left off.

SECTION 3
CONSTRUCTING A SIMPLE
QUESTIONNAIRE

Using the FILES option, you can define sets of questions which are stored under different filenames.   Each file can be edited, appended to or deleted.   On choosing this option the following submenu will appear

```
ADD EDIT REMOVE
DIR OPTIONS QUIT
```

CREATING A QUESTIONNAIRE

Use the ADD option to append questions to a file.   You are asked to
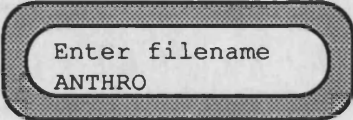
```
Enter filename
```

A valid filename is up to 8 characters long. To escape from this module, press ESC to clear any characters you may have entered from the screen, and press EXE, after which you will be returned to the FILES submenu. The first character must be alphabetic and no punctuation characters should be included. If your entry is invalid, the Psion will beep and you will be asked to re-enter filename. Choose names that will help you remember what is kept in the file.   Questions about feeding practices can, for example, be stored in a file called FEEDS.

Let us work through an example. If we wish to carry out an anthropometric survey and have designed the following questionnaire:
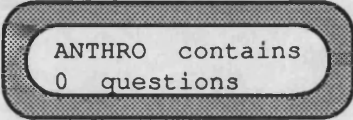
```
-----------------------------------------------
ANTHROPOMETRIC     SURVEY     EXAMPLE
1.   Child's name          |_|_|_|_|_|_|_|_|
2.   Age (in months)                |_|_|_|
3.   Sex (M/F)                        |_|_|
4.   Health card number    |_|_|_|_|_|_|_|_|
5.   Weight (in kg)            |_|_|_|•|_|_|
6.   Height (in metres)      |_|_|•|_|_|_|_|
7.   Field worker initials          |_|_|_|
-----------------------------------------------
```

If we wish to call the file ANTHRO, our answers to the first prompt would be

```
Enter filename
ANTHRO
```
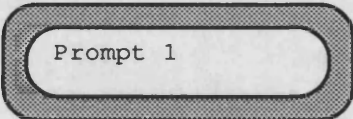
after which the display changes and informs you that

```
ANTHRO  contains
0   questions
```

There are two basic steps to question definition:
a. the question or prompt text
b. specifying the type of response expected
c. entering any explanatory text.
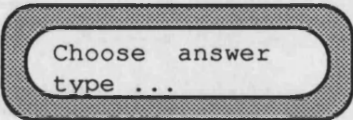
```
Prompt 1
```
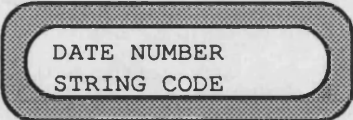
will be displayed when the Psion is waiting for you to enter a question. The number printed after 'Prompt' will change each time you are asked for a new question. Simply type in the text you wish to use. A maximum number of 16 characters is allowed so that the question can be comfortably printed back onto the screen. Long questions have to be abbreviated to shorter meaningful statements.

The message

```
Choose   answer
type ...
```

then appears. The three dots printed after the message indicate that the Psion is waiting for you to press any key when you are ready to continue. There are a number of format types you may choose from. The following submenu appears

```
DATE NUMBER
STRING CODE
```

1. DATE

Choose DATE if you expect a day, month, year entry. The data entry routine will do checks to ensure that the day lies between 1 and 31(or 28,29,30 depending on the month and the year), and that the month lies between 1 and 12. It then prints a prompt

```
┌─────────────────────┐
│ Translate  to age   │
│ (Y/N)?              │
└─────────────────────┘
```

Reply 'Y' if you want all date entries to be displayed as ages relative to the current date i.e. if today's date is the 6th May 1988, a date entry of 11th November 1983 will be displayed back as 4 years 6 mths. Reply 'N' if you want all date entries to be displayed as they were entered.

You will then be asked to

```
┌─────────────────────┐
│ Choose lower &      │
│ upper limits        │
└─────────────────────┘
```

After pressing a key, the following submenu appears

```
┌─────────────────────┐
│ NONE ONE-ONLY       │
│ BOTH                │
└─────────────────────┘
```

Choose the NONE option if any date entry will be considered valid. Choose ONE-ONLY if you wish to specify either a minimum or a maximum limit only; and BOTH if you wish to specify two limits.

The choice of ONE-ONLY leads to a further menu

```
┌─────────────────────┐
│ LOWER UPPER         │
└─────────────────────┘
```

Choose LOWER if you wish to specify a minimum limit on date for any of
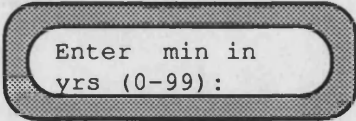
```
┌─────────────────────┐
│ YEARS MMYY          │
│ DATE-INPUT          │
└─────────────────────┘
```

Choose YEARS if all date entries have to be at least a specified number of years, and you will be asked to

```
┌─────────────────────┐
│ Enter  min in       │
│ yrs (0-99):         │
└─────────────────────┘
```

Use this for example, if you wish only persons over 18 to be considered as adults in a particular survey.
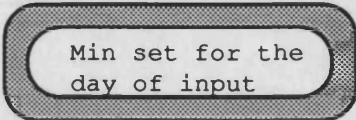
Choose MMYY if all date entries have to be after a certain month of a certain year.

```
Enter  min in
yrs (0-99):
```

The first two digits have to be a valid month, i.e., between 01 and 12, the last two digits can be any year from '00 to '99, assuming all dates to be in this century. To ensure that all children included should be born after January 1981, for example, specify a limit of 0181.
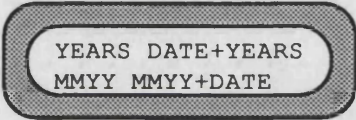
On choosing DATE-INPUT, you will be informed that the

```
Min set for the
day of input
```

A possible use is to enter the date of next appointment for a patient.

Similar menus will appear if you choose the UPPER option. As well as being able to specify an age or month and year range when the BOTH limits option is chosen, you can also choose two further options in

```
YEARS DATE+YEARS
MMYY MMYY+DATE
```

The DATE+YEARS option automatically uses the date of entry as a minimum and allows you to specify a maximum number of YEARS. The MMYY+DATE entry allows you to choose a minimum month and year with DATE-INPUT stored as maximum.

In our example, the questionnaire deals with children under the age of 5, a Date of birth field can be set to a maximum limit of 5. All input will therefore be checked against the current date (at time of input) to ensure that no entries for children more than 5 years old are stored.

2. NUMBER
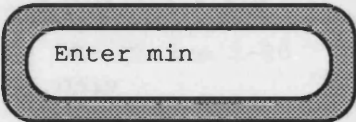
Choose NUMBER if you require numeric characters to be entered. The following submenu will appear:
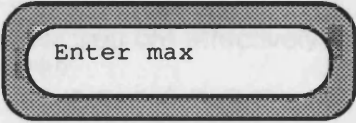
```
I-RANGE D-RANGE
MAX-DIGITS
```

I-RANGE allows you to set an integer range that would restrict all input for that field to values in between. Choosing this will lead to 2 more prompts:
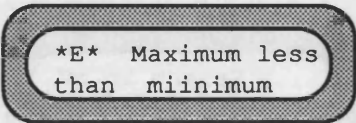
```
┌─────────────────────────┐
│  Enter min              │
└─────────────────────────┘
```

```
┌─────────────────────────┐
│  Enter max              │
└─────────────────────────┘
```
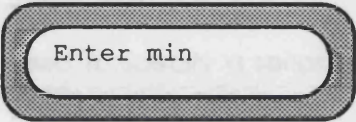
In both of these you are expected to enter a valid positive integer (i.e. between 0 and 32767 (the maximum integer allowed by the Psion). If the maximum value you enter is less than the minimum, the following error message will be displayed
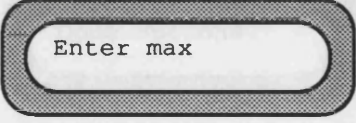
```
┌─────────────────────────┐
│  *E*  Maximum less      │
│  than   miinimum        │
└─────────────────────────┘
```

and you will be asked to re-enter the maximum value.

D-RANGE allows you to set a similar range which allows the input of decimal numbers. You will be allowed to
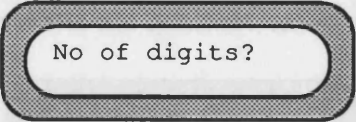
```
┌─────────────────────────┐
│  Enter min              │
└─────────────────────────┘
```

```
┌─────────────────────────┐
│  Enter max              │
└─────────────────────────┘
```
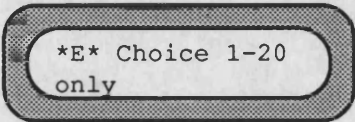
as before. Both the minimum and maximum values are to be specified with a maximum of 10 characters (including the decimal point). In spite of this restriction, you should still be able to specify a wide variety of ranges. During data entry, the Psion will indicate the number of decimal places allowed by printing trailing zeros after the decimal point to the required precision. The default value of decimal places is 1.

MAX-DIGITS allows you to specify a fixed number of numeric characters to be entered. Use this to enter health card numbers for example. The prompt

```
┌─────────────────────────┐
│  No of digits?          │
└─────────────────────────┘
```

will appear. You are expected to choose a number between 1 and 20. You will rarely need a large number of digits. Note that if more than 16 characters are required, the entry will be too long to be shown on one line on the screen and will be scrolled off. If your entry lies outside this range, the following error message will be displayed.
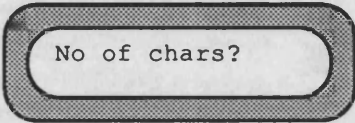
```
╭──────────────────────────╮
│  *E* Choice 1-20         │
│  only                    │
╰──────────────────────────╯
```

If you expect a 6 figure number, you are effectively limiting the entry to a number between 000000 and 999999.

3. STRING

Choose STRING if you expect a mixture of alphanumeric characters. Use this to enter addresses, drug descriptions, etc. You will be asked to enter
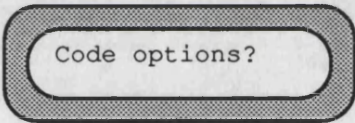
```
╭──────────────────────────╮
│   No of chars?           │
╰──────────────────────────╯
```

Your answer is again expected to lie between 1 and 20.

4. CODE

The CODE option can be used to specify a series of one character codes for entry. In response to the prompt
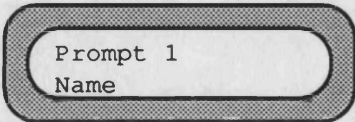
```
╭──────────────────────────╮
│   Code options?          │
╰──────────────────────────╯
```

you are required to type in up to 7 characters that have a coded meaning. For example, marital status may be single, married, divorced or widowed; we may enter SMDW as our code options in this case.

Step c in question definition allows you to enter any text up to 100 characters to be used as a help facility during data entry. Use this to explain shortened question texts or to explain codes. For the marital status example above, we can enter "s=single, m=married, d=divorced, w=widowed" as our help string. This is an optional step and you may press EXE without entering anything if you feel this unnecessary.
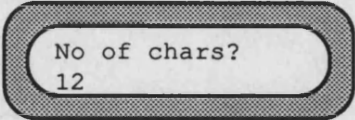
Back to our example, to enter the first question, we respond as follows

```
╭──────────────────────────╮
│   Prompt 1               │
│   Name                   │
╰──────────────────────────╯
```

We would expect a string to be entered as answer to this question. We would probably need about 12 characters, so we choose the STRING option from the menu and specify

```
No of chars?
12
```

As an explanation, we may enter "Child's name - 6 characters for first name and 6 characters for surname".
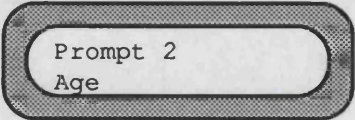
After you have entered a question, an answer type and an explanation, you then specify whether you wish to

```
ADD-MORE STOP
```
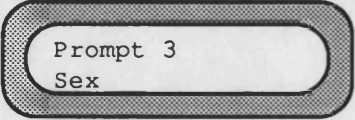
Choosing ADD-MORE will repeat steps a b and c, with the display 'Prompt ~ ' updated. In our example, we wish to continue entering some more questions so we choose this option.

The screen display will change and we enter

```
Prompt 2
Age
```

We choose as type of answer expected, a CODE with options MF and an explanation "m=male, f=female". Again we wish to ADD-MORE

```
Prompt 3
Sex
```

and choose this time, NUMBER. We expect the children in our survey to be between the ages of 0 and 60 months and choose I-RANGE to enter this.

Continue the above procedure for the rest of the example. Choose NUMBER and MAX-DIGITS for "Health card no"; NUMBER and D-RANGE).49 to 30 for "Weight"; NUMBER and D-RANGE 0.50 to 1.59 for "Height"; STRING 2 characters for "Field worker initials". Then choose QUIT.
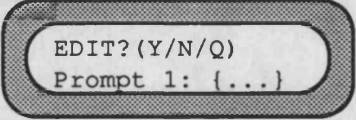
SECTION 4
CHANGING QUESTIONNAIRES
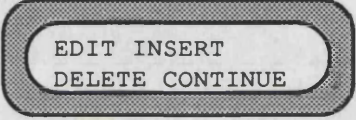
EDITING A QUESTION FILE

EDIT allows you to change the questions in a file. The question is scrolled on the bottom line with an EDIT?(Y/N/Q) prompt above it
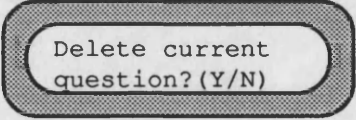
```
EDIT?(Y/N/Q)
Prompt 1: {...}
```

You can control the scrolling by doing the following. If scrolling to the left, you can pause the display by pressing the right cursor key >. If you press > again, the display will scroll to the right. Press the left key < to reverse the effect. At this point, you may choose the Q option to QUIT editing, the N option to go on to the next question, or Y (for Yes) if you wish to edit which leads to a further menu

```
EDIT INSERT
DELETE CONTINUE
```
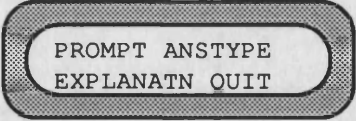
Choose the CONTINUE option if you decide not to edit. Choose the DELETE option if you wish to delete the current question from the questionnaire. You will be asked to confirm your choice by replying Yes to

```
Delete current
question?(Y/N)
```

after which the question will be deleted. If you reply No, you will be returned to the EDIT menu for another choice.

Choose the INSERT option to insert a question just before the current question. Enter a new question as in steps a, b and c explained in the previous section. On choosing the EDIT option, the
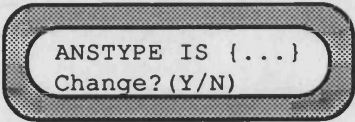
```
PROMPT ANSTYPE
EXPLANATN QUIT
```

submenu appears. You will be able to edit the prompt or explanation by using the cursor keys ^v<> and **DEL** to move around the field. Alternatively, press the ON/CLEAR key to clear the previous and then type in the new one. Press **EXE** when you have finished editing.

If you choose ANSTYPE, the current type is displayed on the first line and you will be asked whether you wish to change the answer type or not.
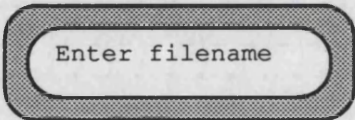
```
ANSTYPE IS {...}
Change?(Y/N)
```

On replying Y, the menu as for specifying answer types in ADD is offered for a new choice.

As an example, let us go back to our ANTHRO file and change the "Age(months)" entry to a "Date of birth" one.   We reply N to the EDIT?(Y/N/Q) prompt for Name. We reply Y to the EDIT?(Y/N/Q) prompt and choose EDIT from the submenu. We can now choose the PROMPT option. Press the ON/CLEAR once key to delete the previous entry and then type in "Date of birth" and press EXE. Then choose the ANSTYPE option, reply Y to Change?(Y/N) and choose the DATE, specifying Y to the translate option, BOTH limits in YEARS, with a minimum of 0 and a maximum of 10. Choose the Q option to Quit editing and reply Q to EDIT?(Y/N/Q) for Sex.
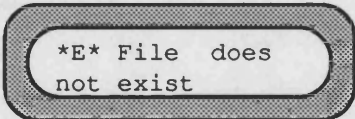
DELETING FILES

Use the REMOVE option. Enter a filename in response to the
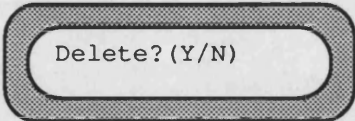
```
Enter filename
```

prompt. If the file does not exist, an error message

```
*E* File  does
not exist
```

will appear momentarily on the screen and the prompt will be repeated. Reply 'Y' to the

```
Delete?(Y/N)
```

to confirm your choice if you wish to do so. The deletion process may take some time if the questionnaire has been run and data saved. The data from the files are all archived and cannot be used within FDCS. They can be transferred using the TRANSFER command (see Section 14) to a micro-computer and deleted when transfer is complete.

LISTING FILES

Use the DIR, short for directory, option. After the first filename is displayed, subsequent filenames are shown if you press the v (down) arrow. You can go backwards and forwards in the list using this and the ^ (up) arrow. Press ON/CLEAR
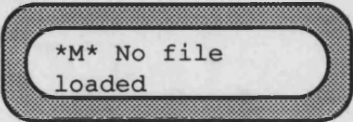
to go back to the main menu. If the ANTHRO file is the only file currently stored, it will be the only file listed and the arrow keys will have no effect.

SECTION 5
RUNNING A QUESTIONNAIRE

Data is entered using the RUN option on the menu. Questions in the currently loaded questionnaire file are usually displayed sequentially. If no questionnaire has been loaded, the message:
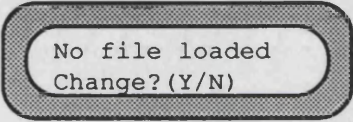
```
*M* No file
loaded
```

will be displayed momentarily on the screen.

LOADING FILES

Choose the LOAD option in the main FDCS menu. The currently loaded questionnaire will be displayed. If no file is current, you are informed that
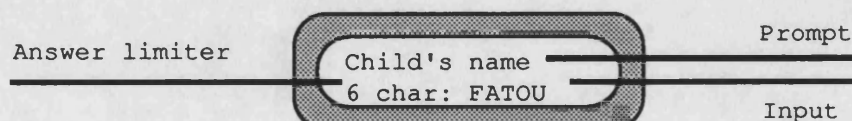
```
No file loaded
Change?(Y/N)
```

If you reply 'Y', a menu list of all questionnaire files available will be displayed. If the list is longer than can be shown on the screen, the cursor keys ^v<> can be used to scroll through line by line or item by item. Press the first letter of the file you want currently loaded or press EXE to choose the item the cursor is currently flashing on.

DATA ENTRY

As the program runs, all the questions are displayed in turn and prompts printed after each to indicate the type of input required. All input is checked logically against the predefined range of valid values or characters. If you enter invalid input, the Psion beeps and the prompt is repeated.

The screen layout is generally as follows, the question text is displayed on the top line and the type of answer expected indicated on part of the bottom line:

```
                                                     Prompt
Answer limiter      Child's name
                    6 char: FATOU
                                                     Input
```

The prompts will vary according to the type of answer expected and will be one of the following:

1. DATE

You are expected to enter exactly 6 digits, the first 2 should specify a day (DD), the next 2 a month (MM) and the last 2 a year (YY). The 3rd of March 1986 should be entered as 030386. For example,

```
DATE OF BIRTH
DDMMYY: 030386
```

Logical checks are done on each component of your 6 digit entry, DD should lie between 1 and 31 (depending on the month, the maximum value is 30 for April, June, September and November and 29 for February), MM should lie between 1 and 12 and YY can be any year up to the current year. If limits have been specified on the date field, further range checks are also done. If limits of 0 and 10 years have been defined and the current year is 1988, only year values between 78 and 88 are accepted as valid.

If an invalid date has been entered, an error message is displayed before the prompt is repeated and the entry displayed for editing. The entries listed on the list below on the left all generate the error messages on the right

```
DATE OF BIRTH
DDMMYY: 083188
```

```
*E* Invalid month
entered
```

```
DATE OF BIRTH
DDMMYY: 290283
```

```
*E* day > no of
days in feb
```

will be displayed because 1983 was not a leap year and February in that year had only 28 days.

```
DATE OF BIRTH
DDMMYY: 121212
```

```
*E* age must be
<=   10 years
```
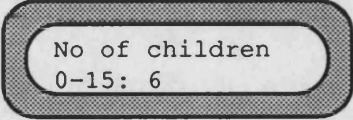
will be displayed because a date earlier than 10 years ago has been entered. This kind of error will happen only if a limits have been defined on the date when the questionnaire was constructed.

2. NUMBER

The prompt varies depending on whether you are expected to enter a certain number of numeric characters, or an integer or decimal value within a certain range.

If the answer type chosen was an integer range, the valid range is expressed as the lower limit - upper limit. An example of what might be displayed is:
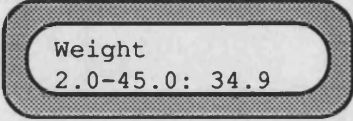
```
No of children
0-15: 6
```

You are expected to enter any integer greater than or equal to 0 and less than or equal to 15. If you enter 16 or -5 the Psion will beep twice, your entry cleared and the prompt repeated. It will do so until it has obtained a satisfactory answer.
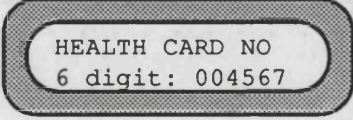
A decimal range answer type requires a numeric entry with or without a decimal point. The display shows the number of decimal places allowed by printing floating zeros to the precision previously defined.

```
Weight
2.0-45.0: 34.9
```

Any decimal value within this range is valid. Entries specified to a greater degree of accuracy are rounded up, say, for example, you enter 34.893 and only one decimal place is required, the stored entry will be 34.9.

Digit prompts require the entry of a pre-specified number of digits; as an example take
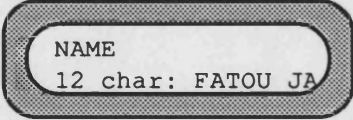
```
HEALTH CARD NO
6 digit: 004567
```

Here, you are required to enter a maximum of 6 digits.

3. STRING

The number of characters to be accepted as input is displayed, as in
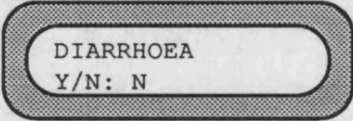
```
NAME
12 char: FATOU JA
```

Up to 6 alphanumeric characters can be entered in. You are allowed to press EXE only after entering at least one character.

4. CODED CHOICE

The 1 character options that are allowed are displayed and one and only one of these is accepted as valid input. For example,
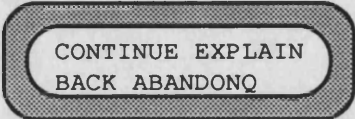
```
DIARRHOEA
Y/N: N
```

ENTERING UNKNOWN VALUES

If you do not know the answer to a question, press the ON/CLEAR key and then press EXE. A string of asterisks is stored as a value, e.g. "*****".
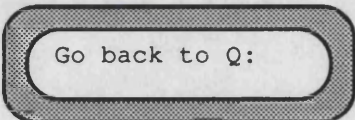
ESCAPE MODE

By pressing ON/CLEAR and then MODE just after the prompt, a submenu
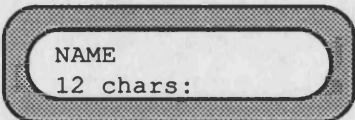
```
CONTINUE EXPLAIN
BACK ABANDONQ
```

offers you the option of going back to the prompt with CONTINUE, displaying a 100 character help string that usually offers explanatory text with EXPLAIN, or to stop entering data on that particular questionnaire by choosing ABANDONQ. The BACK option allows you to go back to a previous question and edit the entry. You are asked

```
Go back to Q:
```

The question entered is then displayed. You can use the ON/CLEAR key to clear the previous entry and enter a new one or the cursor keys and the DEL key. On pressing EXE, you will be returned to the current question.
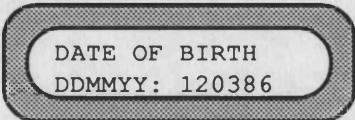
Try running your questionnaire as follows:

```
NAME
12 chars:
```

To find out more about what you are intended to enter, press ESC and then MODE to activate the escape menu. "Child's name - 6 characters for first name & 6 characters for surname" will be displayed scrolling to the left. Press the > arrow key to make the display stationary. Press the < arrow key to continue scrolling in the same direction or the right one to change the direction of scrolling to the right. Press ON/CLEAR to end and then type in FATOU JANNEH. Use the DEL key to erase mistakes and press EXE to go to the next question.

Go through the rest of the questionnaire, entering

```
DATE OF BIRTH
DDMMYY: 120386
```

```
SEX
M/F: F
```

```
HEALTH CARD NO
6 digits:
```

For this, press ON/CLEAR and then EXE to indicate that you do not know the answer.

```
WEIGHT
0.49-39: 20.7
```

```
HEIGHT
0.50-1.59: 1.02
```

```
FW INITIALS
2 chars: DF
```

After entering data for a complete questionnaire, the following menu will be displayed to ask whether you wish to edit your entry, save it (i.e. store it on the datapak).

```
SAVE-QNAIRE
EDIT-QNAIRE
```

If you choose EDIT, you will be allowed to peruse your entries and change any you wish. The first question and entry is displayed, after which the use of > and <, the right and left cursor keys will display one field to the right or one field to the left. As you pass the end of the questionnaire, the number of the current record is displayed. Press EXE if you wish to re-enter the entry that is currently displayed on the screen, change the field and press EXE to end. Press ON/CLEAR when you have finished editing the record. The EDIT/SAVE menu will reappear. You can edit as many times as you wish before choosing the SAVE option when the entries will be stored as a new record on the datapak.

You will then be asked whether you wish to continue entering in some more data into a new questionnaire or QUIT the RUN facility when the menu

```
NEW-QNAIRE STOP
```

is displayed.

SECTION 6
LOOKING THROUGH

VIEW allows you to flick through the entries that are in a data file. The cursor keys < > ^ v move left one field, right one field, up one record, down one record respectively.

On entry to this procedure the screen display will show the question text and the entry for the first field in the first record. Using the ^ or v keys will make the previous or the next record current. When the end of the data file is reached at v is pressed, the first record in the file is made current, and when ^ is pressed at the beginning of the file (i.e. record 1), the last record is made current and the current position in the file.

When a v or ^ key press in the middle of a record, the record number is displayed instead of the same field in the next record to indicate the change.

< and > change the field displays to the previous or the next in the current record. The prompt is displayed on the first line and your entry on the second. When the beginning or the end of the file is reached, the current record number is displayed.

Pressing the ON/CLEAR key will quit the option.

If as an example the following additional data is stored for our ANTHRO questionnaire:

LUCY, 090286, F, 346521 , 18.2, 1.04, DF
WILLIA, 280586, M, 003042, 24, 1.23, DF
CHRYSA, 011187, F, 002436, 21.8, 1.01, DF
DAVID, 130981, M, 00419, 22.1, 0.73, BS
ROWNEY, 230288, M, 021325, 4, 0.4, BS
RACHEL, 311184, F, 005211, 30, 1.18, CZ

We could visualise both the questionnaire and data file as follows:

	field 1		field 2	field3	field4	field 5	field 6	field 7
record 1		FATOU	JANNEH	120386	F	041239		
	20.7	1.02	DF					
record 2		LUCY	FATTY	280586	F	346521		
	18.2	1.04	DF					
record 3		WILLIASISOHO		011187	M	003042		
	24.0	1.23	DF					
record 1		CHRYSAZAINAB		130981	F	002436		
	21.8	1.01	DF					
record 5		DAVID	BARRY	230283	M	000419		
	22.1	0.73	BS					
record 6		ROWNEYCRUBAL		090888	M	021325		
4.0	0.40	BS						
record 7		RACHEL		311284	F	005211		
	30.0	1.18	CZ					

The following key presses will display the field on the right:

KEY PRESSED	SCREEN DISPLAY

V to enter VIEW

NAME
FATOU JANNEH

→

DATE OF BIRTH
2 years 11 mths

↓

record 2

←

NAME
LUCY FATTY

←

record 2

←

FW INITIALS
DF

←

HEIGHT
1.04

↓

record 3

↓

record 4

→

FW INITIALS
DF

↓

record 5

→

record 5

→

NAME
DAVID BARRY

→

DATE OF BIRTH
5 years 11 mths

ON/CLEAR

... leave VIEW

SECTION 7
SUMMARIES

We use the SUMMARY facility to find out the total number of occurrences of a certain string, value (or range)

Each prompt in the questionnaire is printed on the screen followed by a question on whether you wish to count for that prompt or not

```
no:question text
Counts?(Y/N/Q)
```

A 'Q' (QUIT) option stops the process and returns to the main FDCS menu. A 'N' skips to the next question. A 'Y' reply will ask

```
LESS MORE EQUAL
RANGE TOTAL QUIT
```

You may enter any sequence of valid characters as the prompt on the next line indicates (i.e. numeric, alphanumeric, within a specified range, etc.). Choosing any one of the first three options will lead to

```
Enter string
```
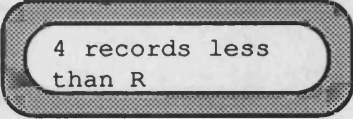
Tallies are made for all entries less than, greater than or equal to whatever value you enter. After the initial choice of string, choosing any other option will return the count for that condition on the same value. For example, for the data set used in the previous section 6,

```
1: Name
Counts?(Y/N/Q)
```
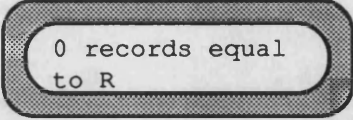
Press Y and L. Enter R in response to the string prompt

```
Enter string
16 chars: R
```

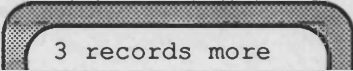to find the number of records less than R. The Psion displays

```
  4 records less
  than R
```

and returns to the menu. Press E to find out how many records are to R, and the following is displayed
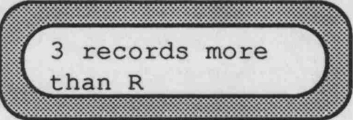
```
  0 records equal
  to R
```
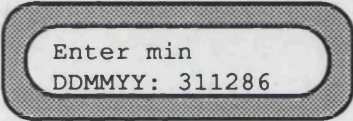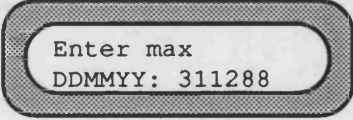
Choosing M will lead to the display

```
  3 records more
```

```
  3 records more
  than R
```

You can choose any opti~~~~~~~~~~~~~~~~~~~~~~~e. Choose Q to go on to the next prompt and reply Y to Counts?(Y/N/Q) for "Date of birth". You can choose the RANGE option to specify 2 different limits - a minimum and a maximum and a count is done for field entries within this range.
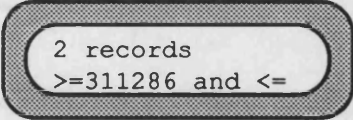
```
  Enter min
  DDMMYY: 311286
```

```
  Enter max
  DDMMYY: 311288
```

and the Psion will specify that

```
  2 records
  >=311286 and <=
```

Press Q to end and move to the next question, and Q again to stop summarising.

SECTION 8
PASSWORDS AND AUTOMATIC
STAMPING

—

PASSWORD PROTECTION

It is possible to install a procedure to allow the use of FDCS only if the user's password is in a pre-specified list. An additional level of protection can be set to restrict access to questionnaire files within FDCS.

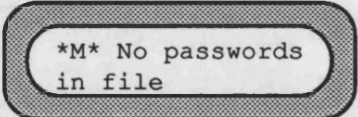Choose the FILES option then the PASSWORD option, and the menu

```
LIST ADD DELETE
STAMP QUIT
```

will be displayed. The LIST command leads to a further sub-menu

```
GEN-LIST
AUTH-LIST
```

where GEN-LIST is short for GENERAL LISTING and gives the details of all passwords that allow access to FDCS. The AUTH-LIST option (AUTHORISATION LISTING) returns the passwords that control access to the FILES option (which allows the user to change questionnaire structures). If there are no passwords in the list, you are informed

```
*M* No passwords
in file
```

ADD appends a new password to either of the access levels. Choose from

```
GEN-LIST
AUTH-LIST
```

to add a new password either to the general access level or to the files access level. You are asked to enter

```
New id:
```

and any alphanumeric combination up to 4 characters can be entered.
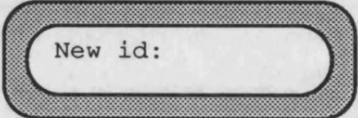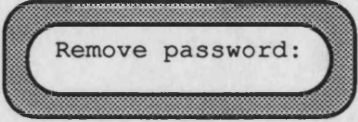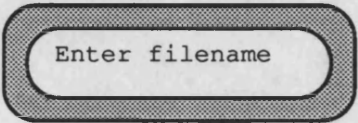
Use DELETE to remove passwords from both lists. After responding to both

> Remove password:

the specified password will be deleted from both lists.

STAMP allows you to choose the option of automatically adding an identification password to each record saved. Specify which questionnaire file when asked

> Enter filename
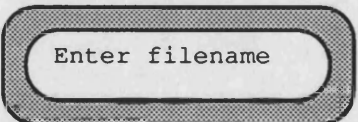
For example, in ANTHRO, we could delete the 7th question, and ask FDCS to automatically stamp the record instead.

Dates and times can also be automatically appended to records if requested within the DATE option in the OPTIONS menu. The menu

> DATE STAMPD TIME
> STAMPT QUIT

is offered. Choose DATE to see what the current date is and TIME to check the current time. For the STAMPD and STAMPT options, you must

> Enter filename

to indicate what questionnaire file is to be stamped. As an example, enter ANTHRO as filename, and choose STAMPD to add the date to all new entries.

SECTION 9
SPECIFYING BRANCHING
CONDITIONS

Often, a group of several questions cannot apply for a particular individual. Within FDCS, certain conditions can be specified for skipped such questions automatically without user intervention. Each condition is attached to a question. After that particular question is asked, FDCS checks whether the condition holds true. If it does, it jumps to the question specified.

On choice of BRANCH within the OPTIONS submenu, you are asked to

```
Enter filename
```

After doing so, you can choose from

```
LIST ADD DELETE
QUIT
```

to find out what conditions have already been specified, use LIST. If no branches are present you will be informed that there are

```
*M* No branches
in file
```

Use ADD to specify a new branch condition. If a previous condition has already been specified for that question, the old condition is overwritten. You will be asked to

```
Enter branch
question no:
```

and then to specify which condition must hold in

```
Enter branch
condition(<>=):
```

This can be any of the signs in the brackets or a combination of two of them, ie.e <> is not equal to , <= is less than or equal to, >= is greater than or equal to.

The question number and condition are specified next , with a prompt allowing you to specify a field entry

```
IF Q?
{prompt}
```

After which you must

```
Enter goto
question no:
```

Choose DELETE to remove a branch condition, specify which after the prompt

```
Enter question
no:
```

If you specify 0 as a question number for any of the above options, the option you are using is immediately aborted and you are returned to the BRANCH submenu.

Let us return to our ANTHRO questionnaire. Suppose we wish to allow for children who were not issued with a health card at birth and therefore whose age has to be estimated. Our question 2 is at present defined as "Date of birth". We now wish to insert new questions as follows:

Q2 : DOB known, code with options Y / N
Q3 : Age, integer range from 0 to 10

"Date of birth" now becomes Q4. We could leave the questionnaire as it is and instruct fieldworkers to enter unknown where appropriate. Alternatively, we can ask FDCS to skip to Q4 if Q2=Y and to skip over Q4 if Q3>0. It is best to write out the conditions before specifying them:

Q2:   DOB known
    If Q2=Y, then skip to question 4
Q3:   Age
    IF Q3>0 then skip to question 5
Q4:   Date of birth
Q5:   Sex

Use the EDIT facility. Reply ANTHRO in response to "Enter filename", N to "Edit prompt 1", and Y to "Edit prompt 2". Choose the INSERT option. Enter "DOB known" for the prompt, choose CODE for answer with YN. "Date of birth" now becomes prompt 3. Reply Y to "Edit prompt 3", choose the INSERT option again, enter "Age" for the prompt, choose NUMBER, I-RANGE with a minimum of 0 and a maximum of 10. "Date of birth" now becomes prompt 4. Reply Q to stop editing.

Now choose OPTIONS, then choose BRANCH, and type in ANTHRO as usual, as filename. Choose ADD and then respond as follows:

```
┌─────────────────────┐
│  Enter branch       │
│  question no: 2     │
└─────────────────────┘
```

```
┌─────────────────────┐
│  Enter branch       │
│  condition(<>=):=   │
└─────────────────────┘
```

```
┌─────────────────────┐
│  If Q2=             │
│  Y/N: Y             │
└─────────────────────┘
```

```
┌─────────────────────┐
│  Enter goto         │
│  question no: 4     │
└─────────────────────┘
```

Choose LIST to see what is stored and the following will be displayed

IF Q 2=Y'GOTQ 4

Choose ADD again and specify the next branch

```
┌─────────────────────┐
│  Enter branch       │
│  question no: 3     │
└─────────────────────┘
```

```
┌─────────────────────┐
│  Enter branch       │
│  condition(<>=):>   │
└─────────────────────┘
```

```
┌─────────────────────┐
│  If Q3>             │
│  0-10: 0            │
└─────────────────────┘
```

```
┌─────────────────────┐
│  Enter goto         │
│  question no: 5     │
└─────────────────────┘
```

Then choose QUIT.

SECTION 10
SPECIFYING EDIT CHECKS

It may be necessary also to ensure that no two fields contradict each other, and contain logically impossible entries. For example, in a survey on fertility and family planning, we wish to include these questions:

1. How many times have you been married?
2. What is your current status?
   m=married, s=single, d=divorced, w=widowed.

We might need to make sure that if someone responds 0 to question 1, they can only reply married or single for the second.

As usual, you will be asked to specify what file after the prompt

```
Enter filename
```

A submenu

```
LIST ADD DELETE
QUIT
```

will be displayed. LIST will list all checks, ADD will allow a new check to be specified, DELETE will remove a check. Before a check is specified, it will be best to write it out first, for example,

IF Q1=0 then Q2 must be (M or S)
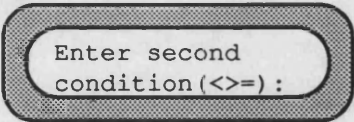
You will first be asked to

```
Enter first
question no:
```

```
Enter first
condition(<>=):
```

and to specify a field entry for comparison before you are asked to

```
Enter second
question no:
```

```
Enter second
condition(<>=):
```
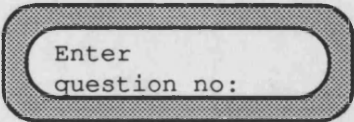
and then specify a second field entry.

Let us attempt to fit our fertility example to the above steps. Only 1 check can be specified per question, therefore we cannot say the following: if question 1=0, then question 2 cannot be equal to M or S. We need to modify our codes slightly to fit within this format. We may choose to use v=divorced instead of d, so that we can say that if question1 is 0, then question 2 must be less than or equal to S. Only M and S satisfy this condition.

If we list this check, the following should be displayed:

Q 1=0'Q 2<=S

If you choose DELETE, you will be asked to

```
Enter
question no:
```

of the particular check you wish to remove. The check condition is stored alongside the second part of the check and to delete it, you must specify the second question number.
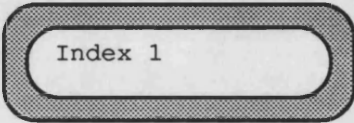
SECTION 11
INDEXING
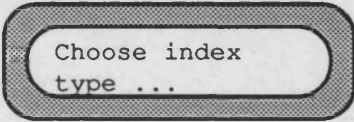
When interviewing the same group of people repeatedly over a long period of time, a register of respondents is always kept. It may be convenient to download this to be used as an index within FDCS.

To use an index, you have first to define its elements. these can be up to 9 fields, similar to those used in add and edit. You will be asked to enter

```
Index 1
```

then to

```
Choose index
type ...
```

and then enter an

```
Explanation 1
```

An indexed file must contain at least 2 fields, the combination of which must uniquely identify an individual entry. Only the field length expected is stored in the questionnaire file in the usual way. The answer types are stored separately in an index file to which indexed records can be appended by using the INPUT option.
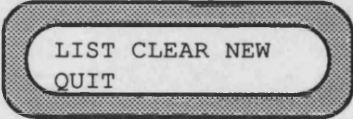
SECTION 12
USING THE NCHS ROUTINE
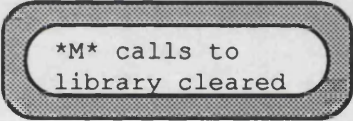
It is possible to incorporate an NCHS routine, so that during data entry, anthropometric measures can be calculated and displayed to the user, but not stored. This is particularly useful if the fieldworker must take urgent action for severely malnourished children.

To choose the weight for age option, you must choose LIBRARY within the OPTIONS menu, specify what file you are working with and then choose NCHS. The following submenu will be displayed
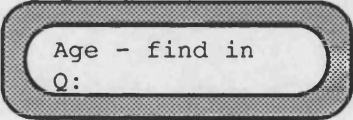
```
LIST CLEAR NEW
QUIT
```

LIST should indicate whether the NCHS routine has been used already, and if so what questions it uses. CLEAR allows you to delete the previous nchs choice. NEW allows you to specify a new one. You will be informed that all previous
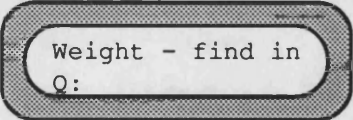
```
*M* calls to
library cleared
```

and then you will be asked to specify in which question the age should be obtained from

```
Age - find in
Q:
```

and then similarly for the weight

```
Weight - find in
Q:
```
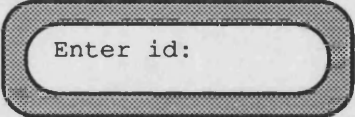
SECTION 13
RUNNING ENHANCED
QUESTIONNAIRES

This section will explain the effects of various optional choices on what happens during data entry, editing, viewing or summary. Any number of the following combinations are possible, in which case as many as specified may occur within the same questionnaire.
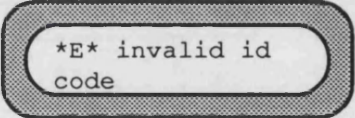
Password protection

If there are any passwords specified for the machine you are using, on choosing FDCS, you will have to
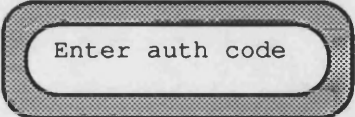
```
Enter id:
```

If this is not in the password list, you will be told that you have entered an

```
*E* invalid id
code
```

If you are also within FDCS and attempt to enter the option after it has been protected, you will be asked to

```
Enter auth code
```

If this is not present in the authorisation list, you will be informed that
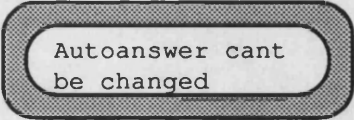
```
*E* You have been
denied access
```

Automatic stamping

This has no noticeable effect during normal data entry. If you attempt to page back , you will be informed that

```
┌─────────────────────────┐
│  Autoanswer cant        │
│  be changed             │
└─────────────────────────┘
```

You can view the contents of the field but, during editing of a record, you will not be allowed to change them and the above message will be printed. During summary, all of these questions will be skipped.

Skipping questions
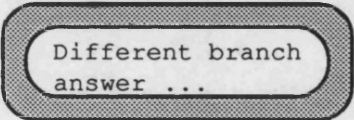
During normal data entry, you will notice that certain questions will not be asked if certain answers are entered. For all questions that are skipped in this fashion, strings of question marks "????" are substituted as entries. You will see these when you view records.

Paging back becomes slightly more complicated with question skipping. If for our ANTHRO example, we have entered the following

Q1:	FATOU SANNEH
Q2:	N
Q3:	5
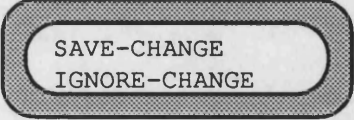
and we are currently at question 5. Question 4 is skipped automatically when an entry is > 0. We may discover at this point that the Date of Birth information can be obtained from the health card, even though the mother does not know it. We therefore wish to go back to question 2, change the answer to Y and enter the date of birth.

We press the ON/CLEAR key and then the MODE key to activate the escape menu, choose BACK and enter 2. We can clear the previous entry by pressing the ON/CLEAR key once only and then enter Y and press EXE. You will be told that you have entered a

```
┌─────────────────────────┐
│  Different branch        │
│  answer ...              │
└─────────────────────────┘
```

Press any key to continue and you will be offered the choice to

```
┌─────────────────────────┐
│  SAVE-CHANGE             │
│  IGNORE-CHANGE           │
└─────────────────────────┘
```

IGNORE-CHANGE retains the previous values and you are returned to question 5. Choose SAVE-CHANGE if you are sure that you want to replace the previous value with the new one. All entries after question 2 will be erased, a new route chosen as question 3 will now be skipped and you will be asked to enter the date of birth instead.
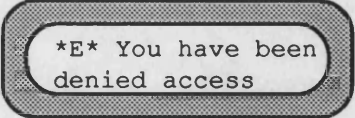
Editing checks

If , for example, all health card numbers start with different numbers to indicate whether a child is male or female. All males have numbers beginning with 1,2,3,4

or 5, and females have the rest. We would have therefore specified the edit check as follows

IF sex is female, then health card number must be greater than 599999

IF Q5=F, then Q6>599999

which is stored as Q5=F'Q6>599999. If we enter F and then contravene this condition by entering 345678, then the message "Check does not hold" will appear on the screen above the stored version of the check which will scroll continuously on the bottom line. Press any key to go back to question 6 when you will have the option of checking your entries, and editing whichever one is incorrect. If the wrong first digit has been entered, the '3' could be changed to a '7', for example. If you have entered the wrong sex, you will need to go back using the BACK option in the escape menu, change the sex to M, and then continue with question 6.

Indexing

The first question is asked normally. A search is made for the data entered and if found, the rest of the index entry is appended automatically to the current record and data entry will proceed from the first question after the last indexed field. If the data entered is not found, the questions will be asked as normal and not changed. Values entered for indexed fields cannot be changed or summarised, although they can be viewed.

NCHS routines

For the ANTHRO example, nothing noticeable will happen until you enter values for both the Date of birth and the weight. The data from any record with either field missing will not be acted upon. After the second value has been entered, the screen display will change to one specifying the percentage weight for age. If the age is outside the range covered by the NCHS data, you will be informed instead that the child is too old. Press any key to continue with data input in either case.

SECTION 14
TRANSFERRING DATA FILES

PREPARING FOR DATA TRANSFER

You need a copy of the Psion Comms Link (CL) package. This should be installed on your microcomputer. A manual is provided in addition to a RS232 cable. The two machines must be linked up, as described more fully in the manual, with the end of the cable fitted into the comms slot on top of the Psion and the other into the microcomputer's serial communications slot.

You must check that the PROTOCOL is set to PSION using the SETUP command within the COMMS submenu if you are using an IBM compatible computer. The CL program must also be running on the microcomputer before transfer.

APPENDIX B

ESTROPID TECHNICAL SUMMARY

B.1   THE ESSEX FLOWCHARTS

1.   *Development*

The Essex 64 flowchart set, which can diagnose about about 190 diseases and conditions, were constructed in the following steps [Essex, 1980]:

(a)   *Identification of the presenting symptoms.*

A survey was made of 2,960 out-patients in a Tanzanian hospital to investigate the principal symptoms patients complained of. Characters that could be precisely defined for quick and reliable use were chosen and grouped into: general, pains and irritations, discharges, chest symptoms and swellings.

(b)   *Identification of the diseases that cause symptoms.*

The diseases which cause the symptoms tabulated were identified and selectively included in the diagnostic scheme. The conventional method of teaching clinical diagnosis has been to detail diseases and their characteristic symptoms. The flowcharts aim to start from the presenting symptoms.

(c)   *Selection of the most useful symptoms and signs.*

Symptoms useful for differentiating diseases and making correct diagnoses are used. Tests take both time and money and as few of these as possible have been included, making the flowcharts largely dependent on symptom elicitation. Some symptoms are common to a lot of diseases and do not contribute to differentiating them. The most useful are those present in a few diseases and easy to identify.

(d) *Identification of diagnostic pathway.*

The diagnostic flowchart is essentially a decision tree in computing terminology. Diseases that are less common but more serious have to be eliminated first by using the most reliable disease characteristics first. The safest order for quick diagnosis is followed.

(e) *Allowing for diagnosis of a disease with many symptoms.*

Some diseases may be identified by several combinations of symptoms. Linkages or cross-references in the flowcharts must be allowed to correctly diagnose these from various starting points.

2. *Testing*

Two field trials were conducted: the first in Tanzania and the second in Guinea Bissau. In the first trial, the aim was to test the accuracy of the charts and the time taken for diagnosis. One medical student with no clinical practice examined and diagnosed about 1250 patients admitted to general wards using flowcharts. Each patient was also seen by an experienced doctor. 94% agreement was achieved. Of the wrong diagnoses about 3% would have had the same treatment or referral. The student took 1.9 minutes on average while the doctors took 13.7 minutes. A further test of repeatability was done with 20 medical students. Divided into 2 groups of 10, each group was allocated to charts each alternate day, using normal methods the rest of the time. Each student examined a newly admitted patient and made a diagnosis. 98% diagnosed correctly with charts; 70% conventionally. In addition, there was close agreement between the patients who used the charts and the accuracy of those in the non-chart group increased when they used the charts [Essex, 1975].

The Guinea Bissau evaluation was conducted in a hospital in the capital, where a doctor examined about 1600 patients over a period of 3 months using the flow-charts. Each patient was also seen by a specialist who often carried out further tests. 87% agreement was achieved; accounting for a systematic error in one chart, this led to an overall accuracy of 92.5% which was judged accurate enough. 5% of the patients were misdiagnosed because of errors in observing characters. Half of these would have had the correct treatment anyway. 2.5% of cases had symptoms not included in the charts.

3. *Use*

Essex stresses that the flowcharts must be used only after basic clinical skills have been taught. They aim to only identify present illnesses and not past ones. Using a patient's recollection of past symptoms and signs will result in wrong or totally useless diagnoses by the flowcharts. It is best to use them before any treatment is given, because this may change the physical signs present.

He also gave instructions on how they are to be used:

(a) There are 2 types of charts: those for symptoms and those for physical signs. The presenting symptoms must always be used first even if a patient does not complain of an evident sign. The symptom chart will identify the physical signs to look for and will also refer to other charts when necessary.

(b) Use the most important presenting symptoms. All of the characters are important when a diagnosis is being made. However, in the flowchart system, the emphasis is on using the patient complaints first as these can be assumed as most important. Most patients will usually be able to describe their problems accurately and concisely. If there are many symptoms, the patient should be asked to identify the most important.

(c) If in doubt about the presence or absence of a symptom, always answer 'no'. It is usually true that a symptom assumed false is absent if the correct method of examination has been used. An exception is when neck stiffness is suspected. The reply must always be 'yes' because this symptom is important in identifying meningitis in its early stages.

(d) How to answer the boxes.

OR - answer yes if any of the symptoms are true.

AND - answer yes only if all the symptoms are true.

COMBINED - split up into parts and answer according to the rules above.

B.2   LISP

LISP is an acronym for a LISt Processing language, widely hailed as an artificial intelligence (AI) development environment. As an interpretative language, each command is processed at runtime rather than precompiled into machine- readable code. This characteristic may impede speed of execution, but also allows for the freedom of building programs with an ability to construct self-executable code. Procedures and data are uniformly represented as list structures. An alternative interpretation of the LISP acronym - Lots of Infuriatingly Silly Parentheses - reveals another peculiarity. It is an embedded language, and distinguishes the elements of a list of lists with various levels of brackets.

Several dialects of LISP are currently available, and muLISP- 86 is one of these [Soft Warehouse, 1986]. Starting with a set of basic procedures, programs can be written within a debugging environment that allows the tracking of commands and variables.

B.3   KNOWLEDGE BASE

1.   *Data preparation*

The data was prepared so that the:

(a)   word subset, i.e. the set of all words 'understood' by the system, be kept as small, and therefore as manageable as possible.

(b)   All symptom descriptions were to be non-ambiguous, strictly defined and kept as close to the original definition as possible.

(c)   symptom descriptions were to be standardised by keeping to the noun form of a phrase as much as possible, instead of verbal of adjectival ones. For example, 'abdominal pain' would be translated as 'pain in abdomen'.

(d)   All symptom descriptions should be meaningful if asked independently of any other.

(e)   There must be a means of including items that are evaluated at runtime, after they are assigned values. Age is a good example. These are the **z????** items present in the knowledge base.

2.   *Data representation*

The flowcharts diagrams have been separated into 2 parts: the symptom description lists and the chart mechanism functions.

The chart headings are stored under the variable headings and are uniquely named as H01-H64, with some having an extra 'A' or 'B' tagged on to indicate an alternative.

The symptom descriptions are in the variables C01-C64. Each description has a unique identifier name. To keep the original chart structure intact, the boxes were numbered from left to right and top to bottom. When there is more than one description, i.e. in OR, AND or combined boxes, each box identifier has a combination of letters and numbers added on. In an OR box, 'A', 'B',... is added on, in an AND box with n elements, 'An', 'Bn', 'Cn',... is added on. A combined box is split up into its component parts and the same rules applied, in box 1 of chart 34, split up 'dyspnoea and fever OR dyspnoea and creps' into S3401A for the dyspnoea and fever bit and S3401B for the rest. The S3401A has A2 added on to indicate the first part of an entry requiring 2 elements, and that for dyspnoea and fever is S3410AB2. Similarly for the rest.

The flowchart mechanisms are expressed as functions containing IF functions with a combination of *sis* (which evaluates whether one condition is true), *sor* (which evaluates whether any condition in the list holds), *sand* (to evaluate whether all the conditions in the list hold) and *connect* (to verify whether a diagnosis has been reached in the current flowchart) commands.

3.   *Dictionaries*

From the lists of symptom descriptions, an index of all the words used in the system was compiled. A series of data files for each letter of the alphabet was

generated which contain occurrences of any symptom description with a word starting with that letter. This was the means by which duplicate symptom descriptions from different charts were traced. Also, they provided a way of supporting a basic free text entry facility.

B.4   OPERATION

The system operates by first asking for the patient's personal details. Based on these entries, some exclusion criteria, based on age and sex, are constructed. After the entry of the main symptoms via the menus, the user is asked whether there are any additional symptoms (not included in the menu), to be entered. These can then be typed in. The dictionaries are scanned for subsets of the words understood in the user entry. If no complete representation of the user entry is found, a series of questions are generated to confirm which of the symptom description lists most closely correspond to the user's intended meaning.

All the symptom entries are translated into the symptom identifier equivalents as each is entered. The set of identifiers is used in the next step to decide on which chart to access. This is done by a heuristic means, counting up the chart which occurs with the highest frequency within the symptom list. On deciding this, the flowchart functions are accessed. These access other charts as required, until a diagnosis is reached. The charts for any remaining major symptom entries are also accessed. All the diagnoses reached are pronounced for all the charts accessed.

B.5   A SAMPLE RUN

Patient details

Enter name of patient
Malamin Sonko

Enter sex of patient
1.   Female
2.   Male
->   2

Enter age group of patient
1.   Newborn   (under one month old)
2.   Infant   (1 month to 2 years old)
3.   Child   (3 years to 12 years)
4.   Teenager   (13 years to 18 years)
5.   Adult   (over 18 years)
->   3

Patient details

PATIENT DETAILS SUMMARY....

Patient's name   ->   (MALAMIN SONKO)

Sex   ->   2 (= male)

Age   ->   4 (= child : 3 years to 12 years)

Do you wish to change any of the patient's details? (Y/N) N

Symptom entry - menus

Does the patient complain of any of these
general symptoms of disease

1:     (FEVER)

2:     (GENERAL BODY WEAKNESS)

3:     (WEAKNESS OF LEGS)

4:     (DIFFICULT TO WALK)

5:     (LOSS OF WEIGHT)

Enter symptom(s):   1

Symptom entry - menus

Does the patient complain of any discharges (Y/N) Y

Symptom entry - menus

Does the patient complain of any of these discharges

1:    (DIARRHOEA)

2:    (DISCHARGE FROM PENIS)

3:    (BLOOD IN URINE)

4:    (INCONTINENCE OF URINE)

Enter symptom(s):   1

Symptom entry - menus

Does the patient complain of any of these discharges

1:    (BLEEDING OF RECTUM)

2:    (VOMITING)

Enter symptom(s):

Symptom entry - menus

Does the patient complain of any chest symptoms (Y/N)   Y

Symptom entry - menus

Does the patient complain of any of these
chest symptoms

1:     (COUGH)

2:     (BREATHLESSNESS)

Enter symptom(s):   12

Symptom entry - menus

You have stated the following as the patient's major symptoms...

1:      (COUGH)
2:      (DIARRHOEA)
3:      (FEVER)

Do you wish to change any of these entries? (Y/N)

Symptom entry - menus

Are there any other important symptoms you wish to mention? (Y/N) N

```
┌─────────────────────────────────────────────────────────────┐
│ Questionning                                                │
├─────────────────────────────────────────────────────────────┤
│ COUGH? (Y/N) Y                                              │
│                                                             │
│ BREATHLESSNESS? (Y/N) Y                                     │
│                                                             │
│ CREPS? (Y/N) Y                                              │
│                                                             │
│ DIARRHOEA? (Y/N) Y                                          │
│                                                             │
│ You have already entered this symptom                      │
│ RASH OF MEASLES? (Y/N) Y                                    │
│                                                             │
│ FEVER? (Y/N) Y                                              │
│                                                             │
│ You have already entered this symptom                      │
│ RASH OF MEASLES? (Y/N)                                      │
│                                                             │
├─────────────────────────────────────────────────────────────┤
│                                                             │
└─────────────────────────────────────────────────────────────┘
```

```
┌─────────────────────────────────────────────────────────────┐
│ Questionning                                                │
├─────────────────────────────────────────────────────────────┤
│                                                             │
│ You have already entered this symptom                      │
│ RASH OF MEASLES? (Y/N) Y                                    │
│                                                             │
│                                                             │
│ THE SYMPTOMS ENTERED INDICATE THESE DISEASES:              │
│ PNEUMONIA                                                   │
│ *** TREAT IN OUTPATIENT CLINIC                              │
│                                                             │
│ MEASLES DIARRHOEA                                           │
│ *** TREAT IN OUTPATIENT CLINIC                              │
│                                                             │
│ MEASLES WITH PNEUMONIA                                      │
│ *** ADMIT TO NEAREST HEALTH CLINIC OR HOSPITAL             │
│                                                             │
│ Quit program?  (Y/N)                                        │
├─────────────────────────────────────────────────────────────┤
│                                                             │
└─────────────────────────────────────────────────────────────┘
```

APPENDIX C

LESSONS LEARNT FROM THE SOFTWARE DESIGN AND DEVELOPMENT PROCESS

The work reported in this thesis is essentially based around seeking answers to questions about how IT based systems may serve primary health care in developing countries. As part of this, two computer-based systems were developed and tested in the field. The intention of this appendix is to document the stages of development of these systems, and to draw out the main lessons learnt.

C.1   FDCS: OVERVIEW OF THE SOFTWARE DESIGN AND DEVELOPMENT PROCESS

1.   *Preliminary discussion of problem*

Initial contact was established with a doctor (Dr. Ron Behrens of the London School of Hygiene and Tropical Medicine) who had worked in The Gambia on nutritional surveillance. He stated what he saw as a problem, the need to assess the status of children being seen in a nutritional survey, so that appropriate action could be taken immediately. Children would be weighed, their height measured, their age estimated, and this data could be compared against a set of international standards of the expected height:weight:age proportions. This would not necessarily be the only data that could be collected; other indicators of health status could be added.

2.   *Choice of hardware platform*

From this initial statement of requirements, it was clear that portable computers would be needed to implement any solution. Several types of hardware were considered, the most serious contender being the Husky Hunter, a durable shock-proof and weatherproof machine made for the armed forces. The price however, was prohibitive (£3,000+). This, and other laptop computers from a similar price range could not realistically be used as a basis for a system of this types in a poor developing country environment. The review of hardware thus turned to cheaper technologies. The first version of the Psion, the CP, had recently been replaced by the XP, a machine with the appropriate characteristics: a small handheld machine, its price (approx. £150), and its programmability represented the best all-round option.

The Organiser Programming Language (OPL) of the Psion is very similar to the BASIC (Beginners' All Symbolic Instruction Code) that was very popular in the early days of home computing. However, it has several other characteristics that are closer to more structured and disciplined languages such as Pascal. For example, OPL incorporates procedures, i.e. program segments, able to access and pass data to other segments as they are needed.

OPL has embedded functions that make the most use of the limited screen capability of the Psion. Menu handling facilities, and the controlled scrolling of messages allow for more flexibility in programming. Database handling functions are provided, for the storage of data into files, and the retrieval of items of data based on search strings. Data communications functions allow data transfer rates and error checking to be specified for copying standard ASCII character data sets from the Psion to any other computer. OPL was fairly easy to learn to program in, and the major source of irritation was its low level of command structure, unlike Pascal. The machine had 32Kb of RAM, and in practice there were no difficulties with the size of the working memory during the project.

3.   *Initial system specification*

As designer, the restriction of a data collection package to nutritional data surveillance seemed unnecessary. In addition, the integration of fixed fields for question texts and answer types, with predefined names and length, would produce a very inflexible system. The system specification therefore grew to developing software for the Psion that would allow question texts, types, and lengths to be defined by the user, and then used to generate prompts as the data was input. A facility to skip questions, and perform error-checking and cross-checking was also added. With this as a basis, system development commenced.

4.   *System development*

The Psion was deemed as an appropriate target hardware, but it alone presented a formidably difficult programming environment for its OPL language. The purchase of a microcomputer-based OPL emulator enabled programming to proceed on a standalone PC, and at a much faster pace. The emulator provided an environment for the writing, debugging and editing of programs, as well as a facility for compiling a set of procedures from OPL into machine code. A facility for creating an imaginary datapack containing data and programs on the PC, and duplicating all its contents directly onto a real Psion datapack was extremely useful.

The key issues of system development were mainly to do with the interface and security. The small screen required ingenuous use of characters for displaying the maximum amount of information possible to the user. In addition, security, protecting users from any false moves they might make, was important. Thus a password control was introduced, and the ability to find one's way back to the main menu using the ESCape character.

5.   *System review*

Initial versions of the system were demonstrated to several members of the Department of Nutrition at the London School of Hygiene and Tropical Medicine (LSHTM), used at an anthropometric workshop with students at the LSHTM,

shown to a statistician at the LSE for comments, as well as used by Dr Behrens to check for remaining programming and design errors.

6.   *System update*

An updated version was produced, incorporating suggestions from the various people who had seen it in operation. The most significant revision was the addition of a capability to go back to previous questions and to change answers. At this time also, the addition of password control, to separate the questionnaire specification levels from the data collection modules, and to allow the system to be used by interviewers without corrupting or unintentionally changing the questionnaire was introduced. A facility for data transfer was also proposed, but proved difficult to make foolproof. The systems as tested in the field made use of the Psion supplied data transfer routines which were not integrated into the developed package.

7.   *Finetuning the system*

Before the field study started, all the fieldworkers were instructed in the use of the system, and went on test interviews to try it out in the field. This exercise revealed several things. The interface had to be consistent and the names of menu items should clearly indicate their function. They had difficulty, for example, in using the option QUIT to exit some functions, but using the option OFF to exit from FDCS and turn the Psion off. All options were therefore changed to OFF. All of these changes were done in The Gambia using the OPL emulator, and the required procedures were then downloaded to the Psions.

8.   *Overall project perspective*

Initial design took about 2 months, system development about 6 months, testing another 2 months, refinement about 6 months, and two weeks of finetuning the system before the start of the full trial. During the total elapsed time of 16 months to the field work, other research work was pursued in addition to work on this project. The software architecture diagram on the next page shows the overall structure of the FDCS system.
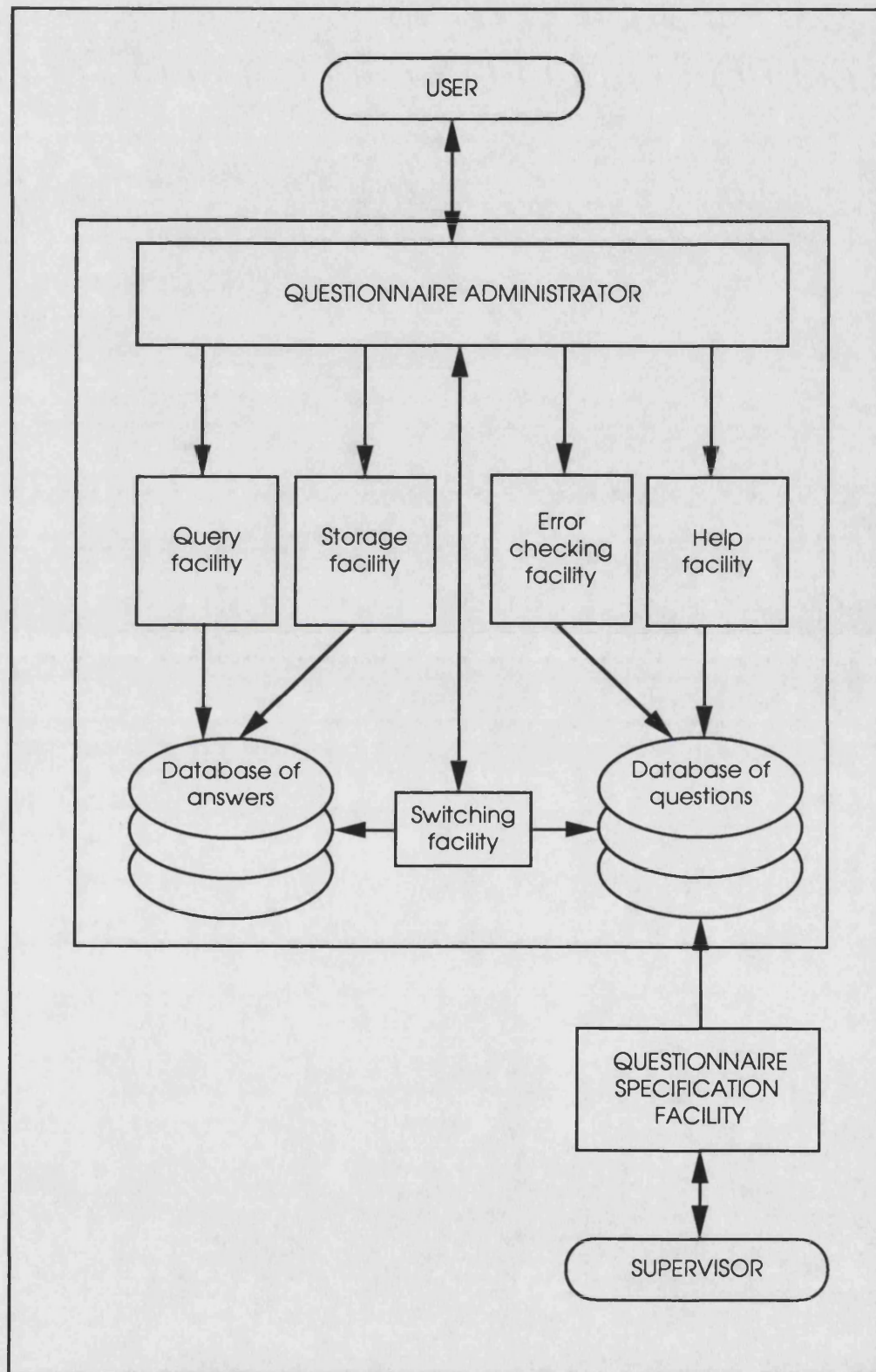
Software architecture of the FDCS system

KEY SYSTEM STATISTICS	
Software:	FDCS
Programming language:	OPL
Number of code lines:	2,873
Number of procedures:	92
Size of files:	Data = 0.5Kb
	Code = 39 Kb
Hardware:	Psion
Model:	Organiser XP
Screen size:	2 x 16 characters
Storage capacity:	up to 128Kb available on one datapack (at that time)
Weight:	250g

C.2 ESTROPID: OVERVIEW OF THE SOFTWARE DESIGN AND DEVELOPMENT PROCESS

1. *Preliminary discussion of problem*

The development of this software started from a B.Sc. Computing project. It grew out of an interest in development, and the use of computers in health. At that time, the area of medical expert systems seemed a new and innovative field that would be of some benefit. Unlike FDCS, the proposal did not come from a seasoned worker in the area. It stemmed instead from the interest of a systems developer in the beginning of a burgeoning and promising technology.

2. *Choice of hardware platform*

Choice of hardware platform for this system was almost inevitable. The IBM-PC was chosen as a generic and widely used technology, with many variants around the basic theme. Thus it seemed ideal for an environment where the use of computers would be encouraged, and other uses of the computer could be found outside diagnostic-aid functions.

3.   *Initial system specification*

The initial system specification was basically to develop a computer-based equivalent of the Essex flowcharts, with a view to further expansion to include more expert system features (e.g. explanation, help facilities, and so on). At the beginning, a simple natural language interface seemed a good choice to allow key words to be used to specify the patient symptoms.

4.   *System development*

The system was developed on an IBM XT machine, using the muLISP programming environment, with facilities for tracing and debugging functions. An initial difficulty was deciding how to put the flowcharts into the knowledge base. Knowledge representation, in terms of taking knowledge out of one form and expressing it in another, is non-trivial, but with no standard methods available. Thus, working out a consistent strategy for doing so demanded a fair amount of effort.

The first version of the system included a simple free text entry capability that could generate questions to verify the users intended meaning, and a facility for filing a record of all the symptoms entered (either positive or negative) that could be accessed during the consultation. The input/output functions in muLISP-86, an old version of muLISP, are particularly cumbersome, and much programming time was wasted constructing basic procedures that do functions provided by many recent packages. As such, the project reinvented a few wheels.

5.   *System review*

Aspects of the system's operation had to be changed to make it more usable. In reviewing the system with a colleague at the Medical Research Council Laboratories in The Gambia, several problems with the system were identified. The free format text entry mode could be cumbersome if a straightforward and common symptom was being entered. Problems of spelling, differences in the meaning of phrases could easily nullify its good intentions. ESTROPID could also not handle multiple pathologies, i.e. the presence of more than one disease in a

patient. Several limitations of the flowcharts were also noted: their handling of the time factor in symptom presentation, and inability to deal with seasonality of some disease presentations.

6.   *System update*

Two major revisions were done at the start of the Ph.D.. The first was to change the format of the user interface, to allow for faster data entry for simple cases. The interface was therefore split into two sections, the first a series of menu screens and the second remaining as the free text entry mode to cater for symptoms that could not be covered in the menu mode. The second change was with the inferencing mechanism. Instead of stopping after one diagnostic endpoint had been reached, the system was amended to use all the patient's presenting symptoms, and to access all the flowcharts relating to these symptoms. Thus, the outcome may be one diagnosis, two or more overlapping diagnoses describing the same basic condition, or diagnoses for two or more different diseases.

Learning from the Gambian experience of not over-burdening the user with extraneous functions, the facility to view earlier entries during a consultation was disabled. The original intention had been to explore the possibility of explanation in as natural a manner as possible. However, much innovative programming effort would have been required in an area which still has not achieved widely acknowledged and accepted techniques for doing so. It was decided that this was outside the scope of the research project.

7.   *Finetuning the system*

This was carried out with the aid of a doctor at the KEMRI unit, but the testing was geared more towards the diagnostic capability of the system, not its interface. Summary sheets of some of the patients seen at the hospital were used. The patients' presenting symptoms and signs, as recorded on paper, were used as the basis for test sessions, and the system's output was compared against the recorded diagnoses. In retrospect, the interface needed a lot more work to make it more natural for the clinical officer to use.
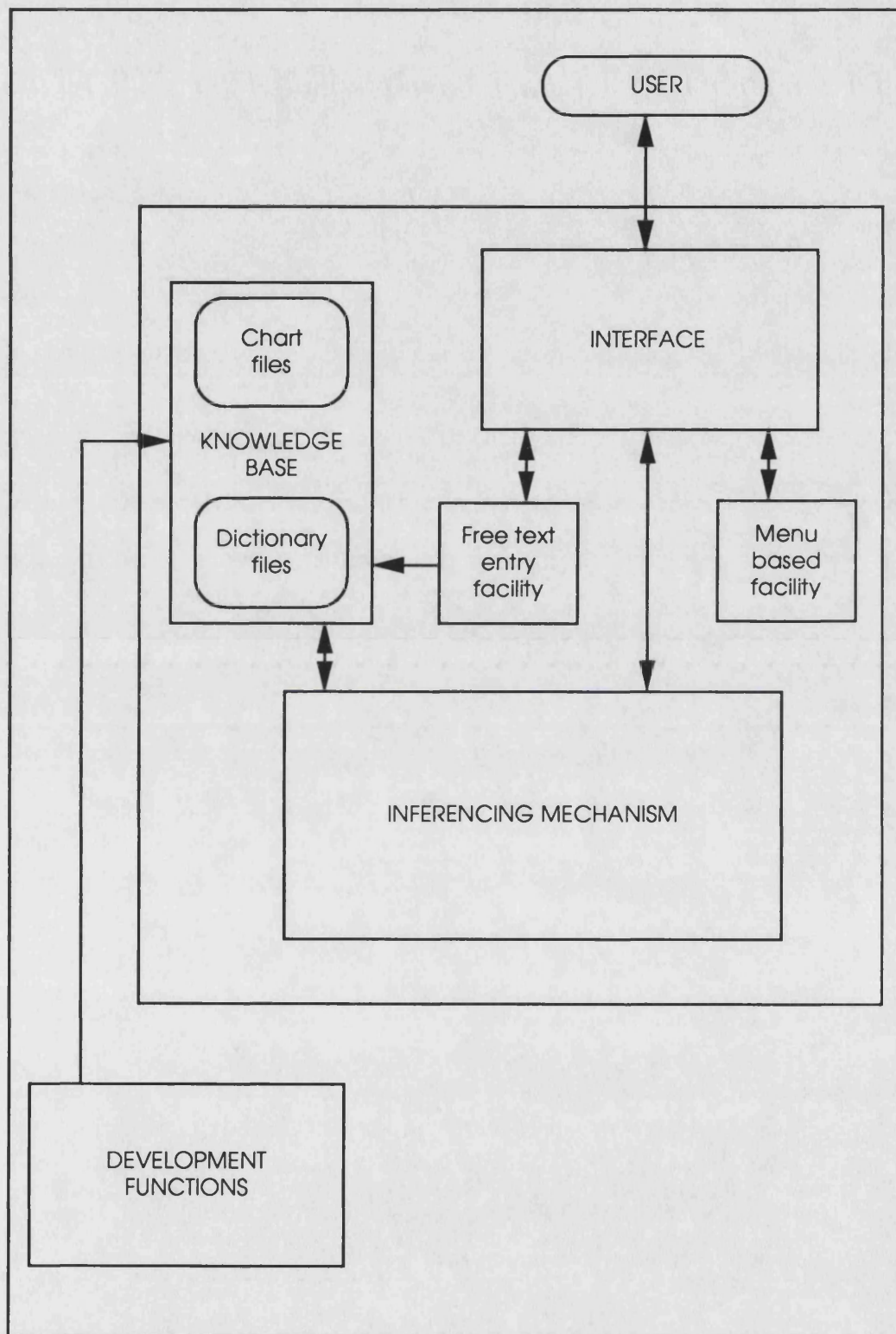
Software architecture of the ESTROPID system

8.   *Overall project perspective*

Estimating the time taken on this project is not easy, since the work was widely spread over four years. The initial pass through the first four stages were done during the B.Sc. course work over the period of nine months. System review at the commencement of the Ph.D. project took two weeks, and the system update required a further six months of program and development effort. Finetuning was done for two weeks before the field trial. The software architecture diagram of the ESTROPID system shown on the previous page presents its overall structure.

KEY SYSTEM STATISTICS	
Software:	ESTROPID
Programming language:	muLISP-86
Number of code lines:	1,089 (code only, including comments)
Number of procedures:	74
Size of files:	Data = 72 Kb (dictionary files)
	= 65 Kb (chart information)
	Code = 33 Kb
Hardware:	Toshiba
Model:	T1200
Screen size:	20 x 80 characters
Storage capacity:	20Mb hard disk
	740Kb floppy disk drive
Weight:	6.5kg (?)

C.3   LESSONS LEARNT

1.   *The importance of the interface*

In developing both systems for naive users, especially when using the Psion and its small screen, the importance of a consistent and simple interface has been highlighted. Simple concepts such as that menu items have to be meaningful, and that the same word should be used for the same purpose regardless of the level at which the software is operating, were learnt the hard way. Also, the point that ease of use for the end-user is much more important than the programming time required to achieve the desired effect came sharply into focus. Thus, in some

instances, code had to be changed, and often lengthened to allow for clearer communication with the user. Before the field trial, the procedure for going back to earlier questions for example, came up with a blank screen for several seconds before displaying the required question. It was much more useful to show the user that the system was doing something and the code was altered to display the question numbers as the system retraced its steps.

Similarly every option has to have an easy way out. The method for leaving an option should again be consistent at all levels, so that any function unintentionally activated can be exited gracefully. The first trial also revealed that access to too many functions at the beginning may impede a learner's progress. It is preferable to have a simple core, and relegate more complicated activities to a distinct 'advanced' level. A user is thus faced with a starting point that can easily be grasped, and a system that can be used to profit as soon as the basics are understood.

2.   *The importance of the programming environment*

A comfortable programming environment that encapsulates the essential tools for software development is a significant productivity improvement. By creating a medium for writing, editing and tracing the execution of a program, an environment reduces the time spent on locating errors. Automatic indentation of code eases the reading of programs, and makes familiarisation at a later date much easier. Both the muLISP-86 environment, and the OPL emulator included such basic functions. The developer of systems such as these is well advised to put effort into seeking out and learning to use a high quality programming environment.

3.   *The importance of a full development lifecycle*

In theory at least, good specifications should lead to consistent code and usable computer-based solutions. In practice, it is impossible to write the perfect specification, and consequently all code is flawed at the initial stages. One needs to have a specific testing strategy to pick out these flaws, and debug the code to a satisfactory standard. This error-baiting process needs to include the end-user

at some point, so that their difficulties with the system can be explored and corrected. Beyond this, the user needs to be tapped for further insight into the problem, and to aid refinement of objectives. All this takes time. Alternatives have to be thought through, further inadequacies identified, and the resulting amendments tested yet again.

This approach is similar in some ways to the classical prototyping approach advocated in software engineering literature. If these projects were to be done all over again, an explicit prototype approach would be adopted, to encourage the notion of progressively detailed sample systems that are to be improved.

With the decision-aid system in particular, the difficulties of working with an incomplete specification rapidly became clear. Converting a paper version of the Essex flowcharts to a computerised version was not as straightforward a task as presupposed. The inability to test ESTROPID thoroughly with potential end-users before the trial was also a considerable disadvantage.