AN ANALYSIS OF EVALUATIVE RESEARCH: THE CASE OF PRIMARY HEALTH CARE

by

JAMES BRODIE CHAUVIN

B.A., Queen's University at Kingston, 1973
M.A., Queen's University at Kingston, 1979

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE in THE FACULTY OF GRADUATE STUDIES (DEPARTMENT OF HEALTH CARE AND EPIDEMIOLOGY HEALTH SERVICES PLANNING AND ADMINISTRATION PROGRAM)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1985

© James Brodie Chauvin, 1985
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of HEALTH CARE AND EPIDEMIOLOGY

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date MARCH 3, 1985
The primary health care (PHC) model is being actively promoted as an effective and lower-cost alternative to conventional health care delivery systems in many developing countries. Despite the fact that over 300 PHC projects of varying scale have been implemented and reported on throughout the Third World over the past two decades, there appears to be little evidence available to support the popular hypothesis that the availability and utilization of primary health care services necessarily results in significant improvements in health. The objective of this thesis is to identify alternative strategies for evaluating PHC projects which will establish credible and useful results. The thesis reviews the evolution of both the PHC model and evaluative research methodologies, and then presents a critical analysis of a set of PHC project evaluations. The aim of this exercise is to identify some of the major factors which have limited the validity, utility and significance of the evaluation results. The thesis suggests that less rigorous evaluative research designs and evaluative techniques which use a combination of quantitative and qualitative data be used to enhance the credibility and utility of evaluation results.
# TABLE OF CONTENTS

**ABSTRACT** ........................................... ii

**LIST OF TABLES** ........................................ v

**ACKNOWLEDGEMENT** ........................................ vi

**INTRODUCTION** ........................................ 1

## CHAPTER ONE

**THE PRIMARY HEALTH CARE MODEL**

Introduction ............................................. 4
Primary Health Care Defined ............................. 4
The Evolution of the
  Primary Health Care Model ............................. 9
The Present Status of the PHC Model .................. 19
Summary ................................................... 24
References ............................................... 26

## CHAPTER TWO

**EVALUATIVE RESEARCH IN PRIMARY HEALTH CARE**

Introduction ............................................. 31
Evaluative Research Defined ............................. 32
The Rationale for Evaluative Research .................. 34
Evaluative Research in Primary Health Care .......... 36
General Analysis of Evaluative Research
  of PHC Programs ......................................... 37
Detailed Analysis of Five Primary
Health Care Evaluative Studies ......................... 47
  i. General Description of the
    Five Case Studies .................................... 50
  ii. Kavar Village Health Worker
    Project (Iran) .......................................... 56
  iii. Rural Health Development
    Program (Colombia) ................................... 62
  iv. Rural Health and Family Planning
    Project (Haiti) ....................................... 65
  v. Volunteer Health Workers (Thailand) .............. 68
  vi. Rural Teachers Health
    Program (Paraguay) ................................... 71
Analysis of the Evaluation Reports .................... 75
Summary ................................................... 95
References ............................................... 97
CHAPTER THREE

ALTERNATIVE APPROACHES FOR
CONDUCTING FIELD EVALUATIONS

Introduction .................................................. 103
The Level of Certainty Required in
Evaluative Research ...................................... 104
Alternative Approaches in Evaluative
Research
   i. Evaluative Research Designs .................. 110
   ii. Evaluative Methods and Techniques ..... 119
Summary ..................................................... 136
References ................................................... 138

CHAPTER FOUR

CONCLUSION AND RECOMMENDATIONS .............. 141

SELECTED BIBLIOGRAPHY ................................. 151
LIST OF TABLES

1. PHC CASE STUDY CHARACTERISTICS .... 52
2. CHARACTERISTICS OF EVALUATIONS .... 54
ACKNOWLEDGEMENT

I would like to take this opportunity to express my appreciation and thanks to my thesis advisors, in particular Dr. Nancy Waxler Morrison, for their assistance, constructive criticism, and encouragement in completing this task. This thesis is dedicated to my children, Caroline and David.

James B. Chauvin
April 1985
INTRODUCTION

Over the past fifteen years, there has been a shift in emphasis in the nature of health care delivery systems implemented in the majority of developing countries. Third World governments, many at the urging of foreign aid donor nations, are investing hundreds of millions of dollars into the planning, development, and implementation of primary health care (PHC) strategies. The primary health care model was approved as an alternative health care delivery system to conventional approaches, particularly in the developing countries, by the General Assembly of the World Health Organization in September 1978. The aim of this approach is to increase both the availability and accessibility of health care resources and services which are appropriate to the needs of, and which have a significant impact on the health of, the target population. Its success is predicated on the active participation of the target community in the definition of their health problems, and in the design and implementation of practical solutions. Additionally, the health paraprofessional is expected to assume many of the activities traditionally carried out by the health professional. (1)

To date, over 300 primary health care projects have been reported on. Often these case studies report only the most basic information, such as the number of health
paraprofessionals trained and deployed, and the number of people who utilize the services offered. Rarely is the effectiveness of this approach discussed in terms of its impact on health, much less its relative cost effectiveness in comparison to alternative and conventional health care delivery systems. Where impact evaluations have been conducted, their utility tends to be limited due to weaknesses in the evaluative research design and in the process involved in conducting the evaluation.

The purpose of this thesis is to investigate various strategies for conducting evaluations of primary health care programs in field settings. This will be accomplished through an analysis of the evaluation reports of several developing country-based primary health care programs. Based on the findings of this analysis, strategies designed to increase the credibility and utility of an evaluation will be proposed. While the focus of this research is on evaluative research conducted in the Third World, the findings and recommendations are intended to be applicable globally.

Chapter One presents an overview of the development of the primary health care strategy and the justification for its adoption as an alternative and/or adjunct to existing health care delivery systems in the developing countries. Chapter Two reviews the role of evaluative research in primary health care, and conducts an analysis of several primary health care program evaluations. The objective of this analysis is to identify
factors which tend to limit the validity and utility of the results of these project evaluations. Based on this analysis, Chapter Three presents alternative approaches for conducting field-based evaluations. These approaches are designed to minimize the effect of the factors identified in the previous chapter, while at the same time ensuring that the evaluations are credible. The final chapter provides a summary overview of the arguments and findings of the thesis. A plan of action for further study on this topic is also discussed.

REFERENCES

CHAPTER ONE

THE PRIMARY HEALTH CARE MODEL

Introduction

Over the past several years there has been a significant shift in the nature of health care service delivery systems being implemented in many developing countries. This is in response to the realization that conventional health care programs have not succeeded in improving the health status of the vast majority of their populations. The primary health care model is perceived as a more practical and less expensive approach. The purpose of this chapter is to examine this concept. The chapter begins by defining primary health care, traces its development as a health care delivery strategy, discusses its current status, and the means proposed to measure its success.

Primary Health Care Defined

Primary health care (also known by the acronym PHC), as defined in the Alma-Ata Conference Joint Report for the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), is

essentially health care based on practical, scientifically sound and socially acceptable methods
and technology made universally accessible to individuals and families in the community through their full participation and at a cost that the community and country can afford, to maintain at every stage of their development in the spirit of self-reliance and self-determination. It forms an integral part both of the country's health system of which it is the central function and main focus of the overall social and economic development of the community. It is the first level of contact of individuals, the family and the community with the national health care system, bringing health care as close as possible to where people live and work and constitutes the first element of a continuing health care process. (1)

The aim of this approach is to provide the entire population with essential health care. This is achieved by increasing both the availability and accessibility of health care services and resources which are appropriate to the needs and expectations of the intended recipient population within the bounds of available resources. From an operational viewpoint, this approach consists of the following basic elements:

1. activities pertaining to the primary care level of the formal or institutionalized health system;
2. community actions pertaining to the traditional or 'informal' health system which are directed at solving local health problems;
3. the concept of universal coverage and accessibility to appropriate health technologies; and,
4. mechanisms which must be created and supported for linking these three sets of activities and incorporating them into an integrated approach to socioeconomic development. (2)

The nature of the services provided and their mode of delivery are not restricted to any one strategy. This approach departs from the conventional curative, hospital-based and physician-focused model. In theory, the PHC approach consists of multiple services, encompassing such activities as health education, promotion, rehabilitation and preventive services, and family health care. These services include immunization campaigns, control of endemic vector-borne diseases, and mother-child health programs. Primary health care is intended to be part of the broader development process. The strategies for improving the health of the community are not restricted to health activities. This intersectoral approach is based on the tenet that improvements in health are not simply a consequence of medical or health care interventions. They are also influenced by factors such as nutrition, education, sanitation, water, and housing. These interventions are thought to have a synergistic effect on health, when such programs are undertaken simultaneously to health care programs. In essence, the approach relies on the coordination of activities within a variety of sectors, which ultimately influence the health status of the community.
The utilization of indigenous health paraprofessionals, who are expected to function in a village setting, is another element of this approach which sets it apart from conventional approaches. If possible, they are drawn from the village or rural community targeted to receive primary health care services. As such, those responsible for delivering these services have a personal stake in ensuring their correct delivery. Because of the many sided nature of the tasks involved, the paraprofessional's duties are not generally confined to health work. Rather, they will be related as much as possible to the many aspects of social and community life that affects an individual's well-being. (3)

The success of this approach also relies on the active participation of community members in the identification of local needs and the development and implementation of the means which respond to them. The inclusion of this element is based on the findings of the joint UNICEF/WHO study of alternative approaches to meeting basic health care needs. It noted that successful, simplified health care programs were characterized by extensive community participation. (4) As the Report to the International Conference on Primary Health Care states:

Community participation is the process by which individuals and families assume responsibility for their own health and welfare and for those of the community, and develop the capacity to contribute to their and the community's development. This enables them to become agents of their own development instead of passive beneficiaries of development aid. (5)
Control of the planning and operation of community health care services are therefore transferred from the medical profession and politicians to community members. This assumes that individuals are cognizant of the health problems which affect them, as well as the means to resolve them. It also assumes that the community will have the resources at its disposal to design and implement them.

No one PHC strategy is advocated for all situations. Local health conditions and ecological factors will define the most appropriate response. Each case may require different levels and types of curative and preventive measures, as well as varying modes of service delivery. Consequently, the nature and combination of elements within a primary health care program tend to differ from place to place. This flexibility is the strength of the approach.

Besides well-defined operational strategies, the primary health care approach also stresses the need for continuous evaluation of program performance. The objective of any evaluative exercise should be to ensure that the system is functioning 'correctly'. That is, that the elements which constitute a particular PHC strategy are effective in improving the health of the recipient community. According to the primary health care concept, as defined at Alma-Ata, evaluation is a continuous process, measuring the capacity of the system to
perform as intended. It also serves to identify weaknesses and strengths in the operation and maintenance of the system which affect its effectiveness. Such feedback is essential for proper management and policy decisions. As defined, the primary health care strategy is meant to serve as a comprehensive and well-developed means for delivering appropriate, low-cost health care services, particularly to underserviced populations.

The Evolution of the Primary Health Care Model

Certain elements of the primary health care approach are not recent developments. Community-based and supported services were operating in most developing countries long before the establishment of the European health care systems under the colonial administrations. The present popular label for such health care delivery systems is 'traditional medicine'. But indigenous medical paraprofessionals were also trained and deployed to serve within the colonial systems. For example, during the 17th Century in Jamaica, physician's attendants who had served apprenticeships and had passed examinations were allowed to sell herbal preparations and drugs, and could work as healers in the rural areas. Following emancipation, formal training for indigenous medical dispensers was established.(6) This system was copied with variations in several African
colonies. The training and deployment of medical assistants was undertaken in Uganda and other British colonies by the Church Mission Society.(7) In Senegal, the French colonial administration founded the Ecole Africaine de Medecine in 1918, where auxiliaries were trained for the 'Services de grandes endemies'.(8) Similar programs were developed and implemented in Madagascar, Nigeria, Zanzibar and Fiji.

These cadres of paraprofessionals were an effective and low-cost means of delivering health care services to the rural populations for two reasons. First, the indigenous paraprofessionals were effective since they cared for their own people in conditions these people were accustomed to.(9) Secondly, most European doctors, with the exception of a few dedicated individuals, seldom left the cities and towns in the colonies. The indigenous paraprofessionals therefore acted as substitutes for the European doctors in the rural areas. However, such programs usually existed only in the hinterland surrounding the major colonial towns.

Additionally, these programs tended to be sporadic, and received little support from government. Strategies designed to promote the legitimization and integration of the indigenous paraprofessionals into the 'European' health care system were actively resisted. Consequently, the indigenous practitioners tended to be adjuncts to the European doctors. As a result, the urban rich for whom the system was reasonably well suited
continued to receive adequate treatment, whereas most of the rural and low-income peripheral urban populations had little or no access to medical services. (10)

Up to the end of the Second World War, the development of European-styled health services in the rural areas of the 'colonies' rested with the missionaries and private business interests. However, these tended to be quite dispersed, and were limited in the range of services offered. The major emphasis was curative medicine. There was also little if any coordination among the medical missions, or between the mission-related services and government-supported health care services. (11) As a result, formal health care services were often duplicated, and there was a lack of integration into the regional or national health care delivery system.

While the industrial and agricultural sector-related health care services tended to include some public health activities in addition to their curative services, their objective was quite straightforward: to ensure a healthy and productive labour force. The rationale for the development of factory and plantation-based health services is stated in the remarks of a former vice-president of the United Fruit Company:

The work that has been done (in providing health services) was done for a very practical hard-headed reason - that of self-interest . . . sick people cannot work . . . It may have been an enlightened self-interest, but it was largely done because
(American business interests) could not get out the ore, or raise the bananas . . . unless these fundamentals were taken care of. (12)

Although these health programs served the need of providing a healthy and productive workforce, their effect was minimal on the health and well-being of community members in general. The range of services was usually limited to immediate health needs of the workforce, in particular emergency first aid. Health and medical services for women, children and the elderly were generally nonexistent. They were considered to be nonproductive members of the general population, and as such the foreign companies were not responsible for their general welfare. Only when it was perceived that local morbidity and mortality rates were extremely high and threatening the capacity of women to bear children (and thus replenish the labour force) were mother-child programs established. (13) The solution was technical and biological rather than social, political or economic. Success of such programs was not measured in terms of changes in the quality of life, but rather by quantitative increases in productivity. (14)

With the establishment of formal foreign aid programs following the Second World War, health was identified as one of the principal sectors to be developed. Ill-health was perceived as a highly visible problem, one which could cause social unrest. There was also a growing vocal demand for national independence in the Third World. The fear of 'Communism' and of
its potential influence on the nationalistic tendencies of some developing nations lent itself to the rhetoric of the period. The opening statement by the Dean of the Harvard School of Public Health at the Conference on Health Problems of Industries Operating in Tropical Countries held at Harvard in 1950 reflects this sentiment:

Powerful Communist forces are at work . . . throughout the world, taking advantage of sick and impoverished people, exploiting their discontent and hopelessness to undermine their political beliefs. Health is one safeguard against this propaganda. Health is not charity, it is not missionary work, it is not merely good business - it is sheer self-preservation for us and for the way of life which we regard as decent. Through health we can . . . prove . . . to the world the wholesomeness and rightness of Democracy. Through health we can defeat the evil threat of Communism. (15)

The politicization of health as a developmental instrument in the Third World was to have a major effect on the nature of health care services which developed thereafter. The strategy which evolved for developing the health care services infrastructure in developing countries was predicated on two assumptions:

1. that the best and most advanced western preventive and curative medical practices, and the institutional framework that provides these services, are universals that work equally well in all socio-cultural and economic settings; and,
2. that the people in developing countries will immediately perceive the advantages accruing to them if they give up old medical practices and adopt new ones. (16)

The ethnocentric perception that western civilization was superior in all ways to all other societies and that given the opportunity, people in the 'less fortunate' countries would clamour to adopt western models of health care and other social services delivery had a major effect on the direction of development of the social service infrastructure in most of the Third World. In the health care field for instance, western foreign assistance programs focussed on the construction of major medical centres, the implementation of massive public health programs, and a drive to orient medical education in the developing countries to adopt and conform to western standards. (17)

Emphasis on the establishment of centres of excellence for training and administration was legitimized by western medical associations, private foundations, and by the recently formed World Health Organization. These groups promulgated 'international' standards for curricula and for staffing in the health field. The standards were expected to ensure that medical and health care services established in the developing countries would meet the health needs of each respective
nation, while accommodating the economic and technical constraints on health care that had led earlier to the use of dressers, dispensers and medical assistants.(18)

The aid programs also served to reinforce the role of curative medicine, effectively restricting the scope and range of services offered. In many post-independence developing countries, the indigenous western-trained doctors who inherited the colonial health care systems chose to maintain them primarily for political and social reasons. Not only did this serve to ensure their dominance as a new elite, but also provided them the excuse to remain in the major urban areas, where the supporting facilities were located, "including good schools for their children and a lifestyle suited to the 'new colonials'".(19)

Accessibility to formal health services for the rural populations was further limited in some countries by the dismantling of the indigenous health care services. A number of newly independent nations not only sought to abolish what were perceived to be second-class services established under the colonial regimes, but some governments also outlawed the practice of traditional medicine.(20) Consequently, the operation and maintenance of a rural health care system ceased in many countries. What services continued to exist were provided either clandestinely by traditional practitioners, or were provided by missionary and private non-governmental
organizations and foundations.

During the 1960s and 1970s, billions of dollars of foreign assistance were directed towards the development of health care delivery systems in the developing countries. Despite the scale of these programs, the standard of living, and in particular the health status of the rural and peri-urban populations of these countries did not show improvement. (21) The expectation that these medical centres of excellence would produce doctors to serve farmers, production workers, and their families was not becoming a reality. (22) In 1974, the World Health Organization and the United Nations Children's Fund (UNICEF) issued a joint report stating that "the strategy so far adopted by many developing countries has failed". (23) This disappointing performance was attributed to a lack of commitment on the part of Third World governments to improve the standard of living and the provision of social services to the rural areas as well as to the internal limitations of existing rural health care services. (24) In general, the formal rural health care delivery system was considered ineffective and costly because it was poorly suited to the expectations of the population, and often inappropriate to their needs. Though these health care services did provide a broad range of curative services, activities such as immunization, health education, and community improvement
programs addressing the causes of illness and disease were lacking.

The financial burden associated with the maintenance of a hospital/clinic-based health care delivery system was enormous and limited the efficiency of the health care delivery system. Few Third World governments possessed either the financial resources required to maintain a system which conformed to accepted international standards, or the administrative capacity to oversee its efficient operation. While foreign assistance programs covered the capital costs of constructing the showcase facilities and the purchase of the attendant medical hardware, recurrent operating costs were the responsibility of the recipient government. In some cases, recurrent operating costs annually accounted for up to 25% of the original facility construction costs. One study on the financial burden of maintaining conventional medical facilities and services in developing countries reports that the operating costs of a single teaching hospital can absorb one-third of a nation's total health budget. In many developing countries, over one-half of the national health budgets was spent on maintaining urban-based curative health care services, even though the urban population accounted for less than one-fifth of the total national population. By the beginning of the 1970s, it was estimated that over 40% of the developing world's population was without access to western
biomedical health services, and that those living in localities of less than 20,000 inhabitants had access to only minimal services.\(^{(28)}\)

Besides being relatively inaccessible to the majority of inhabitants in the developing countries, the services offered by the conventional medical and health care system tended to be inappropriate to their ill-health conditions. For instance, the most common diseases in the developing countries during this period were intestinal, parasitic and infectious diarrheal disorders. In Egypt, Iran and Venezuela, for example, during the 1960s the recorded monthly incidence of diarrhea among preschool-aged children was estimated to be between 40\% and 50\%.\(^{(29)}\) The disease category 'bacillary dysentery and amebiasis, enteritis and other diarrheal diseases' was the leading identified cause of death in Paraguay, Guatemala and El Salvador in the early 1970s. In Pakistan and Bangladesh for 1972, the category 'all forms of dysentery' was the most frequently noted communicable disease. Diarrheal diseases also accounted for the majority of reported infant and childhood deaths in the Africa region for this period.\(^{(30)}\) It has been estimated that diarrheal diseases have been the primary cause of some 4,600,000 childhood deaths per year, resulting from about 600 million episodes of diarrhea.\(^{(31)}\) The second and third major disease groups in the Third World up to 1972 were air-borne diseases (respiratory ailments) and malnutrition
respectively. (32)

Together, these three disease categories accounted for the majority of reported deaths in the rural and peri-urban populations in the Third World, and particularly among children below the age of five years. What becomes evident upon even a cursory examination of these global statistics is that conventional approaches for the delivery of health services is inappropriate. The incidence of the most prevalent diseases and health status in general are contingent upon factors largely unrelated to the accessibility of curative medical services, such as inadequate nutrition, contaminated water supplies and substandard housing. Strategies which focus on identifying socio-cultural, environmental, and economic determinants of ill-health, and designing health care programs which address these issues is a more suitable approach. It was becoming quite evident as the 1970s progressed that conventional medical models were not producing the expected impacts on health. (33)

The Present Status of the PHC Model

Over the past fifteen years, governments have come to recognize that conventional health care delivery systems are inadequate in responding to the needs of the majority of the population in the Third World. Frustration with the 'trickledown' approach to development which was fashionable in
development theory and policy throughout the late 1950s and 1960s appears to have infiltrated the health field in the early 1970s. Recognition of the inequality in access to health care for significant proportions of many developing country populations, and the failure and high cost of existing health care systems, prompted international donors and Third World governments to search for a more practical strategy for delivering health care.

In 1971 the World Health Organization and UNICEF established the Joint Committee on Alternative Approaches to Meeting Basic Health Needs of Populations in Developing Countries. It identified several factors as contributing to the ineffectiveness of existing health care delivery systems. Among these were: the lack of clear national health policies and poor linkage of health systems with other sectors of national development; the lack of clear health priorities based on factual data; the resistance to change social aspects of health policy; the lack of active community involvement in planning and providing health care; and, the inappropriate training of health personnel. The Committee also cited a lack of realistic and proper health planning as responsible for the inefficiency of existing health care delivery systems. Among the elements identified as contributing to ineffective health care delivery were deficiencies of communication and transportation networks, lack of basic sanitation, insufficient use of health education,
and lack of adequate health information. (34)

The Report of the Committee was presented to the UNICEF/WHO Joint Committee on Health Policy in February 1975, and submitted to the World Health Assembly in May of the same year. Based on an analysis of several community-based health care delivery strategies, the Committee concluded that:

Despite the immense problems and the daunting economic situation, it is possible, using the resources available, to meet certain basic health needs of the populations in developing countries, achieve better health care coverage, and improve levels of health. (35)

This could be achieved, according to the findings of the Committee, by adopting a more holistic approach. Such a strategy requires a shift in emphasis to a curative-preventive approach on a national scale, and an integration of health services with activities in other sectors.

This study was used to promote discussion among health experts and to develop background materials for the International Conference on Primary Health Care, held in September, 1978, in Alma-Ata, USSR. The central focus of this Conference was the concern for the status of health care and the standard of living in the Third World. Representatives of 134 governments and several UN agencies and international and national non-governmental organizations attended. The Conference served to strengthen general support for a strategy
to develop a more suitable and realistic approach to the delivery of health care services worldwide, particularly in the Third World. The outcome of the Conference was a unanimous agreement, labelled the Alma-Ata Declaration, which asserts that health is a universal human right and that governments should pursue policies to provide accessible, affordable, and socially relevant health care to all, with a target date of the year 2000.(36)

Bennett (1979) has divided the developing world into five camps with respect to the response and acceptance of the PHC approach as the basis for national health policy: those countries where primary health care is a normal part of their program of social development; those where it is a part of the national health plan and has been recently added on an ad hoc basis; those where primary health care is still at the pilot project stage; those where it is said to be contemplated in the next few years; and lastly, those who have as yet little interest in the concept.(37)

To date, thirty-nine countries have signed regional charters pledging themselves to strive to achieve the goal 'Health for All by the Year 2000'.(38) Most efforts to promote the primary health care model have concentrated on providing basic services at the village level, using a local lay community member trained as a paraprofessional (the 'village' or 'community' health worker). In general, the village health
workers are recruited locally, trained at a central institute, and return to their home village. The training focuses on certain curative and preventive activities appropriate to the needs of the community he or she will serve. For example, 13,000 community health workers have been trained and deployed in the field in Burma within five years; in India, over 180,000 community health care workers have been trained since 1977. Similar programs have been implemented in Indonesia (150,000), Thailand (85,000), Peru (10,000) and Swaziland (1,500).(39) A recent study commissioned by the United States Agency for International Development (USAID) identifies twenty PHC projects funded by USAID which are national in scope.(40)

Many developing countries are experimenting with the primary health care model at the pilot project stage. These tend to be small-scale and experimental in nature.(41) Most are supervised or assisted by private non-governmental and volunteer organizations, religious groups, private foundations, local physicians and other individuals. These have little relationship to national health policy, and often serve only as local demonstration projects to show the impact of this approach on the standard of living of community members. Working with limited resources, they rely heavily on local leadership, voluntarism, and indigenous resources.(42) Only in a few countries, such as China, Tanzania, and Cuba, has there been a concerted effort to transform the prevailing philosophy
of health care delivery. These countries place a high priority on achieving a broad distribution of rural health services through the deployment of auxiliary health personnel and by decreasing dependence on urban-based medical facilities and personnel.(43)

Investments in the development and implementation of primary health care projects has increased significantly over the past five years.(44) For example, the USAID study reports that the Agency allocated $84.6 million to primary health care projects in fiscal year 1980, compared with $27.5 million four years previous.(45) For Canada, it is estimated that approximately 4% of that nation's overseas development assistance budget is allocated to supporting the operation of primary health care projects in the Third World.(46) This does not include funds invested by Canadian non-governmental organizations. While there is no firm estimate of the value of investment in PHC projects in the Third World, there is no doubt that it constitutes a significant proportion of health-related activity program funds.

Summary

The primary health care approach developed largely in response to a recognition that conventional approaches to
delivering health care were not benefiting large segments of the populations in developing countries. This strategy is increasingly being adopted as a means of delivering health care services to underserviced populations in many developing countries. Significant levels of resources are being invested in the development and operation of PHC projects. This strategy is actively promoted as a panacea to the limitations of the conventional health care delivery systems which characterize many of these countries. But as the following chapter will demonstrate, there is little evidence to support the claims that this particular strategy is any more effective than conventional strategies in improving health status.
REFERENCES


9. Ibid.


24. Ibid.


39. Ibid., pp. 8,9.


CHAPTER TWO
EVALUATIVE RESEARCH IN PRIMARY HEALTH CARE

Introduction

Decisions concerning the future direction of health care delivery systems in developing countries require accurate, valid and relevant information. Health services research, including the evaluation of existing primary health care programs, is the primary means of obtaining this information. In order to ensure that evaluation results meet these information criteria, care must be exercised that the evaluative research designs and methods employed are appropriate to the needs and objective of the exercise, taking into consideration the realities of the field setting in which the program exists.

The purpose of this chapter is to identify the principle weaknesses inherent in evaluative research on developing country-based primary health care programs conducted to date. The chapter begins by defining evaluative research, its rationale, and the need for evaluative research in primary health care. The chapter then provides a general analysis of evaluative research on primary health care projects in the developing countries, principally funded and executed by expatriate organizations and individuals. In order to provide a comparison, five primary health care program evaluation studies
conducted by indigenous researchers are then presented and analyzed. The chapter will demonstrate that the results of evaluative research on primary health care projects in developing countries tend to be limited in their validity and utility due to several factors, related primarily to evaluative research design faults, and weaknesses in evaluative methods and techniques.

Evaluative Research Defined

The definition of evaluation (and so, too, evaluative research) differs among experts in the field. For some, evaluation is the measure of the degree of success or the effectiveness of a program in achieving its predetermined objectives. For example, evaluation has been defined as an attempt "to answer statistically what proportion of the goal has been reached and how much of this can be credited to the program". (1) Likewise, it has been defined as "the process of determining the . . . amount of success in achieving a predetermined objective". (2) Evaluation has also been defined in broader terms, to include an assessment of program efficiency, adequacy, effort, performance, and impact. (3)

Program efficiency is the cost in resources of attaining a desired objective in the specified time period with this particular program. Adequacy refers to the capacity of a
program to meet total needs. Program effort relates to the
time, energy and resources expended to carry out program
activities. The performance of a program is an examination of
its constituent components to determine whether these
activities occurred as planned. An identification of the nature
and effect of the consequences (both the positive and negative
side-effects) of the program constitutes an evaluation of
impact. Perhaps Weiss (1972) provides the most succinct yet
comprehensive definition of evaluation:

An elastic word that stretches to cover judgements of
many kinds ... (it) investigates the consequences of
dynamic programmes that attempt to alter key variables
in peoples' lives. (4)

The aim of the exercise is to provide program policy makers,
planners and managers with valid and reliable data to improve,
extend, or in some way modify (including the possibility of
doing away with) the program. (5)

Evaluative research, in turn, is simply the set of
systematic procedures required for collecting and analysing
data which increase the possibility for 'proving' rather than
'asserting' the worth of some social activity. (6) In order to
conduct evaluative research, a set of tools and methods have
evolved, which in themselves have become a subject of study and
debate.
The Rationale for Evaluative Research

The use of evaluative research to assess the performance of social programs is a fairly recent phenomenon. Much of the interest in the field and most of the current evaluation techniques and models were developed in the United States. Social welfare programs borne of President Roosevelt's New Deal and post-World War II U.S. government administration policies spurred the development of models and techniques for evaluating their performance. (7)

A marked growth in both case studies of social welfare programs and the development and review of evaluation methodologies occurred throughout the 1950s and early 1960s. (8) The growth of the economy, and the number and complexity of social welfare programs implemented in the United States during the 1960s, particularly during the Kennedy administration, was unprecedented. Billions of dollars were invested in economic development programs such as industrial incentive programs, space exploration (particularly the development of military technologies), and in social welfare programs such as educational programs for the underprivileged, urban renewal, foreign assistance aid and environmental programs. The underlying assumption of these programs was that with enough dollars and the right efforts, sustained
economic growth would occur and the social ills of the country would be eliminated. But such a rapid and unprecedented scale of investment of public funds, especially into social welfare programs, for which the return on investment was uncertain at best, could not occur indefinitely without some control and accountability.

The take-off in evaluative research occurred in response to several factors:

1. a demand for greater accountability for funds invested in publicly-funded programs;
2. an increasing scarcity of resources, both financial and human, to maintain such programs;
3. the development of new technologies and approaches to service delivery of social programs which were challenging traditional approaches;
4. increasing interest among social scientists in the social relevance of conducting research in social settings as opposed to simulated experiments;
5. a shift in employment opportunities for academics as a result of funding constraints, forcing social scientists to become evaluators in search of research funds; and,
6. a recognition that social conditions and economic well-being were not necessarily improving for the majority. (9)
As the demand for program evaluations increased, so too did the need to legitimize the field of evaluative research. This perspective resulted in the promotion of scientifically-based paradigms. According to theory, such models should provide valid and conclusive evidence. To this end, the experimental paradigms came to dominate evaluative research.

Evaluative Research in Primary Health Care

Significant improvements in the health status of the intended beneficiaries can only occur after primary health care programs succeed in convincing the target population of the benefits accruing from the utilization of these services, and from their active participation in community-based and focused activities, such as educational and self-help programs, which seek to modify existing knowledge, attitudes and practices with respect to health-related behaviour. The burden is therefore on the primary health care program to demonstrate its capacity to overcome constraints which limit the ability of community members to participate actively in all aspects of program development and implementation.

According to the Report of the International Conference on Primary Health Care, evaluation is essential in order to ensure that PHC programs function correctly. Participant government representatives were encouraged to use evaluation results to
improve program effectiveness and efficiency. The Conference report also suggested that PHC programs be evaluated with respect to five criteria: relevance, progress, efficiency, effectiveness and impact. (10) However, while paying lip service to the importance of evaluation, as an essential tool for policy formulation and program management, the Report provided no guidance as to the 'how to' of designing or carrying out a valid and useful evaluation exercise.

General Analysis of Evaluative Research of PHC Programs

To date, over 300 developing country-based primary health care projects have been reported on in a variety of publications. (11) These range in scale from small, regional efforts focusing on specific health-related tasks to national programs covering a wide range of health-related activities. However, while most provide some statistics of progress and effort, few provide details of project impact. To date, only two studies have been published which provide an overview of the results of developing country-based PHC projects. (12) The review which follows is based on the findings of these reports.

The purpose of this overview is to identify some of the major factors which limit the validity and utility of evaluative studies on primary health care projects in developing countries. While by no means a definitive list, this
analysis attempts to identify some of the important weaknesses in evaluative research design and techniques which limit their validity and utility.

One factor identified as adversely affecting the usefulness and validity of PHC evaluations is related to who defines the goals, objectives, and the indicators of progress for these exercises. There is evidence to suggest that some evaluations are commissioned in order to account for funds expended on international assistance projects. The goals of the exercise and the criteria selected to measure attainment of the goals tend to be influenced by the needs of the funding agencies and organizations with little regard for local needs and expectations. In essence, the evaluation studies become ends in themselves. Instead of determining how project activities improve the health of the recipient population, they tend to report the number of facilities constructed, and personnel trained and deployed. It has been reported that some host country officials discount the importance of the evaluation studies because they believe evaluation is primarily the interest of the outside funding agency, and that the exercise is excessive if not irrelevant to health care services delivery. Consequently, the utility of the evaluative studies to the host government or agency responsible for the administration and management of the primary health care program is extremely limited.
Another limiting factor is the lack of defined and explicit objectives and goals for the program. It is often observed that little is known from the evaluations of many health care programs because the right questions are frequently not asked in the first place. (16) In the case of new programs being initiated, general, nonspecific objectives tend to be stated. This problem is illustrated in a review of evaluation studies of 180 primary health care projects in developing countries commissioned by the American Public Health Association (1977):

[A] large proportion of projects appear to be engaged in a variety of objectives which are largely implicit and unarticulated. Consequently, achievement tends to be measured, not in terms of movement in an intentional direction towards a stated end, but rather in terms of effort and energy expended. (17)

The review also found that the extent of change expected in indicators and the time period over which such changes are expected to occur are rarely specified.

Another factor which has been found to limit the value of evaluation studies is the lack of indicators which measure the effect of the project activities on the health and well-being of the intended target population. The United States Agency for International Development (USAID) study of 52 USAID-supported primary health care projects found that the evaluation documents reviewed contained few data that indicate the impact of the project activities on health status, and only limited information on other measures of progress. Such data, it was
found, are expensive, difficult, and time-consuming to produce. And as the time elapsed between the initiation of project activities and the conclusion of the evaluation was relatively short, it was difficult to devise short-term measures sensitive enough to detect changes in the health status of the population. It also found that because of factors not accounted for in the study, measured changes cannot always be definitely attributed to project interventions. As a result, changes in service use often serve as indirect or intermediate and sometimes surrogate measures of impact. As Milio (1977) states:

The most common and traditionally prevalent criteria are the familiar input-output measures, simply quantities of personnel, patients, visits, admissions, immunizations, contraceptives, etc. These gross quantities, as well as process measures, cannot of course indicate health outcomes.

In short, primary health care projects do not tend to be evaluated for their impact on the health and well-being of the target populations. At best, their impact is measured under the premise that the availability of health personnel and services, and the delivery of these services will necessarily result in improved health. While such data may be easier to collect, they are not necessarily suitable or reliable surrogates of impact.

Another limitation of these evaluation studies is that many have focused on a single area of primary health care such as immunization activities, nutrition, family planning or
But the primary health care model is an integrated approach. As such, evaluations should review all activities carried out under the program, and not assess success or failure on fragmented components. In the same vein, it is equally important to know how each of the various PHC components has affected the health status of the target population. Because these PHC activities often take place simultaneously, and in some programs at the same time as certain vertical programs, it is rarely possible to identify the effects of introducing peripheral, community-based services. Few project evaluations have been designed to provide feedback useful in evaluating each PHC component separately. Additionally, the influence of changes in other sectors on the health and well-being of the study population is rarely accounted for.

Evaluations have also been found to be inadequate due to the lack of personnel trained in evaluation design and techniques, the lack of reliable statistics and essential records, and insufficient interest on the part of health staff to conduct evaluations. Very often the evaluations were conducted over and above the normal workload of the individuals concerned. For example, many studies utilize those providing the services to collect data. Unfortunately, service providers are often ill-trained in data collection techniques, and tend to have little time or interest in this task. The lack of
adequately trained personnel is also responsible for a significant amount of unreliable data being collected. For example, the USAID study notes that several projects had to abandon plans to conduct follow-up field surveys because the baseline data were of poor quality, or the data collection techniques were too complex. (24)

The lack of personnel trained in data collection techniques also accounts for the restricted range of data collected. Standard surveys used in primary health care evaluations frequently tend to collect superficial data, but rarely probe the beliefs and attitudes which account for practices. (25) While projects collect inventory data of structures and personnel, the number of patient visits, or the utilization of staff, facilities or equipment, they fail to identify and explain changes in behaviour of the study population as a result of the increased availability and accessibility of health care services. The collection of such information requires a sensitive and culturally appropriate technique. Given the short time allowances for most evaluations, restricted budgets, and lack of trained manpower, such information is rarely collected.

Many projects experienced difficulties in establishing functional data collection and analysis systems. Many of these systems were too complicated for the capabilities of those responsible for the data collection and analysis tasks. Often,
more data than was necessary was collected, overwhelming those responsible for data analysis. The USAID-sponsored study cites a project in Guatemala where it was reported that the data system required the health promoter to complete more than 2,200 file report pages per month. The project supervisor was then expected to summarize 20 reports during the same period.(26) The APHA-commissioned study found that while large numbers of projects report collecting a variety of data, relatively few used the data as feedback to improve the operation of the program. Such data were often used instead to keep funding officials informed of project activities.(27)

These evaluations also tended to collect only quantitative data, often overlooking potentially significant qualitative information. 'Objective' survey data based on counting may not be appropriate to identify and explain identified behaviour patterns. As well, some activities, such as eating and defecation habits, are not always observable or quantifiable.(28) The scale of the project and the people involved in its management and administration are also critical elements in determining the success or failure of a project. Small-scale projects tend to be more manageable, and often are promoted by highly motivated local activists who are able to extract extraordinary efforts from project staff. However, such factors are often not accounted for.(29)

The time frame over which an evaluation is conducted, and
the point at which the evaluator becomes involved, can also affect the validity and utility of the results of the analysis. The first few years of a primary health care program are usually spent in developing operational cohesion, building community cooperation and trust, or coping with the aftermath of an unexpected situation. (30) Evaluators parachuted into the project often are not conscious of these processes. They may conclude that the project has failed because it did not achieve some predetermined objective or target by a predefined point in time. As Hyman and Wright (1971) suggest, conducting an evaluation during what they call the 'first cycle' of a project can result in a misinterpretation of events, as the project is not yet functioning at its maximum efficiency and effectiveness, and staff may be overly enthusiastic in implementing the project. Consequently, the results of evaluative research conducted during this period are not necessarily valid or meaningful for the entire program. (31) As Gwatkin et al. (1980) point out, this is an important factor which is often overlooked when evaluations are conducted:

The experience of these projects suggests that developing country villages more often than not differ from one another in ways that are potentially important for infant and child mortality, but which are often not obvious to even the most astute observers until they've actually been on the scene for a year, two years, or more. (32)
Two other factors identified as limiting the validity and usefulness of many primary health care evaluation efforts have been the use of small sample sizes and inadequate control areas for comparison with the experimental areas. (33) As one study reports:

Some of these differences (between study villages) can be partly controlled for by statistical means. But the application of such controls requires considerably larger samples than has hitherto been the norm, and . . . even then their use only rarely leads to fully meaningful persuasive results. (34)

However, the use of complex experiment-based research designs are expensive. Additionally, the manpower required to design and implement them are often not available in developing countries due to the academic qualifications required. Such approaches as often not politically feasible, especially when there are control areas where services are not to be introduced concurrently with the experimental communities. But even where quasi-experimental designs were used, other factors tended to affect their validity and utility. In some cases, certain evaluation activities were delayed or cancelled. Modifications to the design of many evaluations also occurred. These factors often had serious implications for the validity of the exercise. (35)

The fact that the majority of these PHC projects were designed, implemented, evaluated, and financed by external
funding agencies can also affect their inherent validity and usefulness. As the USAID study points out, conditions imposed by external funding agencies and the time taken to design and approve a proposal can seriously affect both the PHC program and its evaluation:

[The] bureaucratic procedures are so numerous and time-consuming that they, and not the substance of the program, become the foci . . . because of them . . . certain host-country realities (are) glossed over. (36)

By the time the project is approved, objectives and targets in the original planning document may be irrelevant and impractical.

Linked to this is the timetable imposed by funding agencies. For both USAID and CIDA, multiyear project plans are limited to a maximum of three years. (37) However, as the USAID report points out, the experiences of the PHC projects examined would suggest that this is an insufficient time horizon. Demonstrating health impact may in reality require five or ten years. Additionally, most projects are characterized by delays in project activities, due to a variety of reasons. The underestimation of a realistic timeframe for the project (and its evaluation) can significantly distort the interpretation of the results of the evaluation exercise.

The foregoing analysis has identified several factors which limit the utility and validity of developing country-based
evaluative research on primary health care. But the majority of these PHC projects were designed, carried out and evaluated by expatriates. Some of the problems discussed are directly related to this situation. Therefore, it would be of interest to examine a selection of PHC evaluations developed, implemented and evaluated by indigenous researchers to see whether they share the same problems.

Detailed Analysis of Five Primary Health Care Evaluative Studies

One of the factors which has limited the validity and utility of evaluative research studies on primary health care conducted in the developing countries is the lack of indigenous personnel adequately trained in evaluation design and techniques. The development of a capacity to design and implement realistic and pragmatic development programs requires the concomitant development of a capacity to design and implement program evaluations. Such programs must be evaluated in terms of their potential contribution to the development process. This is a prerequisite for a country to formulate coherent programs related to their development objectives.

The aim of developing cadres of developing country researchers and field staff trained in evaluative research methods and techniques also supports one of the elements of the rationale for foreign assistance programs: to make the
developing countries self-reliant and to a large extent self-sufficient, with respect to their capability to utilize indigenous personnel to design, implement and evaluate national social service programs. Such a strategy could serve to reduce the dependence on foreign experts and consultants to design and implement sophisticated, expensive and time-consuming program evaluations which may have limited utility to national decision-makers. Additionally, locally relevant evaluative procedures could be established, more sensitive to local conditions than imported techniques.

The purpose of the following analysis is to identify the factors which limit the utility and validity of evaluative research on primary health care conducted by indigenous researchers. This analysis does not purport to prove that evaluative research conducted by developing country researchers is better or worse than that conducted by expatriate researchers. The aim is to probe, to identify what elements of evaluative research are inadequate.

The five case studies that follow were not selected at random. They represent cases which are adequate for the present study, in terms of the nature, quantity, and quality of the information reported. The selection criteria for the evaluation studies reviewed here were that the primary health care project included a well detailed evaluative research component; that the evaluation had been conducted by developing country
researchers; and, that the evaluation report was available in either English, French, or Spanish. These criteria severely limited the number of project reports applicable to this study.

First, project reports which provide detailed descriptions of project-related evaluative activities are few. Many studies report general findings, but fail to provide information on the design and techniques used in carrying out the evaluation. Second, the number of evaluations conducted by developing country researchers, for which reports are available, is limited. Following a careful review of a dozen potentially applicable primary health care evaluation reports, five were chosen.

The projects selected for this study do not represent the totality of primary health care projects implemented in developing countries, only a subset. However, the author believes, based on his personal experience, that the shortcomings within the cases to be reviewed are representative of the majority of PHC program evaluations which are conducted in developing countries. Nor should this study be taken as a critique of the efforts of the research teams involved, nor the rationale for or activities of the projects. The objective of this exercise is simply to review and analyse the project evaluations. As such, this review confines itself to presenting data revealed in the project evaluation reports.
i. General Description of the Five Case Studies

The five primary health care projects to be reviewed took place in rural settings in a variety of Third World countries: three in Latin America; two in the Caribbean; and two in Asia. No primary health care evaluation reports from Africa satisfied the selection criteria. The projects took place between 1973 and 1980. The longest project period evaluated spanned five years, and the shortest two years. All the principal investigators for the projects and their evaluations were indigenous academics, connected with local Schools of Public and/or Community Health in their respective countries.

The projects did not vary significantly with respect to the nature and range of primary health care services offered. In general, the five projects concentrate on delivering preventive services, and all provided basic curative services. Most offered mother-child health activities, monitored nutritional status of children, provided immunizations, and promoted health education activities. Three of the projects also promoted family planning, and four of them included water supply and sanitation improvement activities.

Four of the evaluations sought to assess the impact of primary health care activities on the health status and well-being of the study communities. The other three were primarily concerned with conducting process and output
evaluations. Two of these reviewed the attitudes of community health workers (CHWs) in carrying out their assigned tasks, and the attitudes of community members towards the CHWs. The remaining study analyzed the output and activities carried out by primary health care dispensaries in the study area.

Most of the evaluations were concerned with service delivery rather than with outcomes. The indicators most commonly used were data on changes in health care facility and personnel utilization, availability, and accessibility, and in the knowledge, attitudes and practices of the study populations with respect to health-related behaviour. None of the evaluations used a true experimental design, although two originally intended to do so. The most frequently used design was a one group pretest-posttest design which compared conditions in the study villages before and after the introduction of the primary health care activities.

The projects and the results of their evaluation reports are presented separately. Each review consists of a brief background description of the project, its objectives, and a detailed analysis of the evaluative research design and techniques. A summary of project characteristics is provided in Table 1, and a summary of the characteristics of their respective evaluations in Table 2.
**TABLE 1**

PHC CASE STUDY CHARACTERISTICS

<table>
<thead>
<tr>
<th>Project</th>
<th>Duration</th>
<th>Services Offered</th>
<th>Total Study Population</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Kavar Village Health Worker Project (Iran)</td>
<td>1973-1974</td>
<td>- basic curative</td>
<td>10,000</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(24 mos.)</td>
<td>- referral to clinic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(phase 1)</td>
<td>- family planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1974-1978</td>
<td>- environmental health education</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(38 mos.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(phase 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Rural Health Development Program (Colombia)</td>
<td>1974-1979</td>
<td>- basic curative</td>
<td>10,600</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>(60 mos.)</td>
<td>- referral to clinic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- family planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- nutrition education</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- child immunization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- environmental health promotion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Rural Health and Family Planning Project (Haiti)</td>
<td>1975-1978</td>
<td>- basic curative</td>
<td>34,000</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(36 mos.)</td>
<td>- preventive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- family planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- child immunization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- nutrition education/rehabilitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- maternal/child health</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 1 (continued)**

<table>
<thead>
<tr>
<th>4. Volunteer Health Workers (Thailand)</th>
<th>1976-1978 (30 mos.)</th>
<th>- basic curative</th>
<th>- malaria treatment</th>
<th>- home births</th>
<th>- nutrition education</th>
<th>- prevention of communicable diseases</th>
<th>564 households (4/9 study villages. 50% sample for demographic survey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Rural Teachers Health Program (Paraguay)</td>
<td>1978-1980 (32 mos.)</td>
<td>- basic curative</td>
<td>- child immunization</td>
<td>- referral to clinic</td>
<td>- preventive health promotion</td>
<td>- environmental health promotion</td>
<td>23,958 100%</td>
</tr>
<tr>
<td>Project</td>
<td>Research Design</td>
<td>Outcome Measures</td>
<td>Data Collection Techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Kavar Village Health Worker Project (Iran)</td>
<td>- phase 1: one group pretest-posttest</td>
<td>- % change IMR</td>
<td>- direct surveys (researchers and VHWs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- phase 2: nonequivalent group posttest</td>
<td>- % change maternal mortality</td>
<td>- record systems (VHWs &amp; district health clinic)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- changes in KAP for personal hygiene &amp; health maintenance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- % change in birth rate and family size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Rural Health Development Program (Colombia)</td>
<td>one group pretest-posttest</td>
<td>- acceptance of VHws by community</td>
<td>- direct surveys (VHWs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- % of population covered</td>
<td>- record systems (VHWs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- frequency of VHW activities for same client</td>
<td>- monthly supervisory reports (research team)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- # activities conducted by VHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Rural Health and Family Planning Project (Haiti)</td>
<td>single time-series analysis</td>
<td>- % change IMR</td>
<td>- direct surveys (VHWs &amp; research team)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- % change infant morbidity</td>
<td>- record systems (VHWs &amp; district health clinic)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- % change adult morbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- infant growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- % change birth rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Study Type</td>
<td>Case Study/Program</td>
<td>Methodologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------</td>
<td>---------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Volunteer Health Workers</td>
<td>one-shot case study</td>
<td>- utilization of VHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Thailand)</td>
<td></td>
<td>- attitude of villagers towards VHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- use of alternative medical services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- # of activities conducted by VHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Rural Teachers Health Program</td>
<td>one group pretest-posttest</td>
<td>- % change child mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Paraguay)</td>
<td></td>
<td>- % change child morbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- utilization of teachers as VHW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- acceptance of teacher as VHW by community</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- change in environmental conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ii. Kavar Village Health Worker Project (Iran)

Of the five project evaluations to be reviewed, this is the most extensively reported. The project extended over a five year period (1973-1978) and was designed and implemented by personnel from the Department of Community Health, Pahlavi University, located in Shiraz, Iran. The project site was situated on the Kavar agricultural plain in southern Iran (Fars Province). The village health workers (VHWs) were expected to provide basic curative services (primarily first aid), referrals to the regional health station, family planning information, and environmental improvement and hygiene education. The ultimate objective of the project evaluation was to measure the impact of locally selected and trained village health workers.

The project was divided into two phases. The objective of the first phase, which extended over a twenty four month period, was to "study the feasibility of training villagers who possess basic literacy to provide primary medical care and preventive health care in rural areas". During the first six months of the project, study villages were selected, the training curricula and materials developed, VHW trainees selected, and the baseline survey completed. The trained VHWs were trained and deployed to the field during the second six
month period. At the end of the twenty-four months, the follow-up surveys were carried out, data analysed, and an evaluation report produced.

The researchers intended originally to use a pretest-posttest control group evaluation research design. In fact, the actual design used in the first phase evaluation was a one-group pretest-posttest. No explanation was provided as to why this change occurred. Of the fifty-five villages originally scheduled to be included in the study, fifteen were designated as experimental villages, and six as controls. No criteria for selecting these particular villages were provided.

Baseline surveys were carried out to collect demographic data, as well as information on prevailing environmental conditions (wells, latrines, etc.), and the identification of users of family planning methods in all study villages. A baseline knowledge, attitudes and practice (KAP) survey on health-related behaviour was conducted for 200 households in the study area (a 10% sample, randomly selected). No such study was conducted for the control villages. The researchers reported that "due to cultural reasons, females may be underreported in actual baseline population survey". However, no indication was given as to either the magnitude of the underreporting, nor the effect it had on the analysis. Follow-up surveys were scheduled to be conducted in the eighteenth month of the project period, and repeated at the end
of the project.

The indicators chosen by the research team to measure the success of the program were the utilization of services offered by the VHW, changes in sanitation and hygiene behaviour, and an increase in the acceptance of family planning. Accessibility was not measured since the VHW clinic was just a short walk from the houses in the villages. The researchers assumed that the utilization of services would serve as an adequate surrogate measure for the acceptance of the VHW by the experimental population. The VHWs maintained records on the number of monthly patient visits to the village clinic they were assigned to, as well as the age, sex and diagnosis of each case. They also recorded the type and frequency of environmental improvement, community education, home and school visits and family planning activities. They were also expected to prepare bimonthly reports on vital statistics for their assigned villages.

The project evaluation reported that 11,467 patients had consulted the VHWs during the study period, an average of 1.4 visits/person/year. It was also reported that persons aged 20-44 years tended to use the VHW services less than those offered at the higher level Health Center. The VHWs also tended to be consulted more for lower-order medical conditions. The project evaluation report noted that females used the village health workers' services in significantly greater proportion
(p>.01) than males, with this differential being most pronounced where female VHWs offered treatment services. The Phase 1 evaluation study also found that the number of patient visits to the district health centre decreased by 43.5% for the study villages over the first six months of the study, but that no such decrease occurred in utilization patterns in the control villages.

During the course of the first phase of the project, female users of contraceptives reportedly increased by 230%. Before the arrival of VHWs, 7% of women 15-44 years of age stated that they used some form of contraceptive; at the end of the Phase 1 study period, 24% of the same cohort stated that they used some form of contraception. The termination rate in the study villages was calculated as 25.6/1000 women 15-44 years of age using contraceptives, and 30.2/1000 for the same cohort in the control villages. According to the study, this point pattern demonstrated that the increase in pill use in VHW villages has not been among women who simply accepted a packet of pills because they were urged to by the VHW, but illustrated the effectiveness of the family planning education component of the project.

The researchers conducted a survey at the end of the project period to determine villager satisfaction with the VHWs. The vast majority of the respondents expressed satisfaction, and that most had had verbal contact with the VHW
sometime during the project period. Unfortunately, data on sanitation and preventive services were collected too infrequently to enable the researchers to evaluate the relationship between VHW activities and sanitary and hygiene behaviour.

Based on the results of the evaluative analysis, the researchers concluded that the "VHWs are highly productive and well accepted in the villages they serve" and that "the VHW concept is not only feasible, but would provide a strong foundation for Iran's developing health delivery system".

The second phase of the project, which extended over thirty eight months, sought to determine "if there was some indication that the services conducted by the VHWs were effecting a positive change on the health status of constituent populations". The research team suggested that a positive difference in vital statistics such as infant mortality rates between the experimental and control villages would indicate that the VHWs were effective. This phase was intended to be an extension of the first phase evaluative exercise. Baseline data collected in Phase 1 was to be used for comparison with data collected at the end of Phase 2. However, this did not occur. In fact, the evaluative research design used to measure impact of the second phase of the project was a nonequivalent group posttest.

The second phase evaluation report noted that the villagers
constructed many latrines during the project period. However, whether or not this was due to environmental control promotional activities of the VHW was not established. Nor were the acceptance and utilization of family planning methods evaluated in this phase. The evaluation also reported that the infant mortality rate for the experimental villages (84/1000 infants less than 12 months old) was less than that documented in the control villages (138/1000 infants less than 12 months old) at the end of the project. However, it is hard to understand how the researchers could conclude that the village health workers were responsible for positively affecting the health status of the study population. First, the evaluative research design used in the study did not compare pretest and posttest data for the study and comparative populations. Basing research results on posttest data between nonequivalent groups is a very weak design. Second, the data which was collected six months after the start of Phase Two would seem to contradict the conclusion drawn by the research team. These data indicated that the infant mortality rates were 65/1000 and 128/1000 infants respectively. In both cases, the infant mortality rates actually increased over the study period. Yet the researchers failed to conduct a comparative analysis of these data. No explanation was provided for this oversight.
iii. Rural Health Development Program (Colombia)

The goal of this project was to design, implement and evaluate a rural health care delivery system designed to improve the health status of rural communities. (39) The health care delivery program was one constituent element of a larger multidisciplinary development program conducted by the Centro de Investigaciones Multidisciplinarias en Desarrollo (CIMDER), located in Cali, Colombia. A wide range of services were offered, including family planning, nutritional surveillance and education, and the promotion of community participation in health-related activities. Although the project monitored nutritional status, the project did not include a supplemental feeding program. Neither did the project involve the construction of improved wells and latrines, although the community health workers promoted their construction and utilization. The project period extended over five years (1974-1979). The project area was located in Cauca state, and contained a population of approximately 10,600. The study area was divided into seventeen subareas, with a trained health promoter assigned to each subarea. The objectives of the health care project were to make primary health care available to 100% of the experimental population, to improve accessibility to primary health care services for at least 80% of this
population, and to reduce the magnitude of their health problems.

To evaluate the performance and impact of the project, the researchers used a one group pretest-posttest evaluative research design. Baseline demographic and health data were collected in a household survey of 1,200 families (50% sample). The health promoters also collected health status information monthly on all individuals residing in their area of assignment. This information was used to plan their activities during the project, and to supplement the data collected in the baseline survey. Information was also collected on community acceptance of and satisfaction with the promoters, and their participation in health-related activities. The promoters were also asked their opinions on the attitude of the recipient population to their activities, and on their workload. The household surveys were repeated in 1978 and 1979.

In general, it appears that the project objectives were attained according to the results of the evaluation. For example, the project claimed to have achieved an 84% level of coverage, based on a rate of 2.52 contacts per person per year after twelve months, and a rate of 3.76 after twenty four months. The researchers were unable to calculate the degree of decline in maternal mortality as the benchmark rate had not been established at the commencement of the project. However, the evaluation study reported that the incidence of
complications at delivery and during the postpartum period had decreased from 75/1000 births to 67/1000 births over the final two years of the project.

Infant mortality reportedly decreased over the project period as well, from 151/1000 infants to 39/1000 infants over a twenty-four month period. A significant decrease in the incidence of diarrheal-related diseases was recorded between 1977 and 1979 (from 295/1000 children aged less than 14 years to 105/1000 children). The evaluation also noted an increase in the number of latrines constructed and that the number of persons drawing water from improved wells increased by 66% over the study period. Whether this was a result of the construction and utilization of improved wells and latrines was not established.

The prevalence of incapacitating diseases in children, as measured by the number of missed schooldays, reportedly decreased from 52/1000 school children to 21/1000 children over the five year study period. Immunization-controllable diseases also were reported to have decreased over the same period, from 24.7 cases per 1000 children to 15.8/1000, following an intensive study area-wide immunization program. Improvements in general nutritional status of children were also reported. It was estimated that the percentage of children classified as being in a normal state of nutrition increased by 40% over the study period. Dental care records were not analyzed. The
evaluation also reported that the morbidity rate for adults decreased significantly over the study period. A decrease in the general mortality rate was also noted, from 107/1000 persons to 75/1000 persons. The researchers believed this to be of significance, since the national mortality rate for the same period was 90/1000 persons. The researchers concluded in the evaluation report that since the socio-economic conditions of the study population had not changed over the study period, yet the level of coverage and availability of health care services had, the decreases recorded in the mortality and morbidity rates were a direct consequence of the effects of the health program.

iv. Rural Health and Family Planning Project (Haiti)

This project, which was conducted by personnel from the Public Health and Population Department of the Haitian Ministry of Public Health, was part of an experimental primary health care program which had as its goals the provision of nutrition and general community-based preventive services, the delivery of family planning, maternal-child health services. Community health workers (CHWs) were trained and deployed to deliver these services. The program also included the training of local research personnel to conduct a process and impact evaluation of the health care project. The project was implemented in
a rural area in southwest Haiti, with a total study population of approximately 34,000 people. The Canadian International Development Agency (CIDA) was financing a major integrated rural development program in the study area, in cooperation with several Haitian government ministries. The primary health care project extended over five years (1975-1980). The evaluation study covers the first three years of project activities.

A case-control study, using three experimental sites and one control, was originally planned to evaluate the PHC project. However, a time series analysis which compares outcome data from the three experimental sites was ultimately used. No reason for this change in design structure was provided by the researchers. Indicators to measure the impact and performance of the project were mortality, morbidity and fertility rates. Baseline demographic and vital statistics data were collected at the three experimental sites for all households. A sample of 750 women aged 15-44 years (approximately a 20% sample of this cohort) was selected to be interviewed and followed concerning births and contraceptive practices. An epidemiological survey on the prevalence of intestinal helminths was conducted at the start of the project, as well as a socio-economic survey of the study population. Morbidity and mortality data and records of vaccinations were collected from records maintained by the CHWs during the course of the project. A follow-up demographic
survey was conducted in 1978. Specially trained enumerators were trained to collect the demographic data.

The evaluation provided details on some process and outcome measures. For example, the evaluation reported that over 90% of children aged 0-4 years received three doses of DTP vaccine, and over 85% of children aged 0-14 years received TB immunizations. The evaluation also reported that as a result of the campaign to eradicate neo-natal tetanus, the infant mortality rate at one of the study sites decreased from 117/1000 infants to 106/1000 infants over the first year of the project. This trend continued through the second year, and by 1978, it was reported to be as low as the infant mortality rates in the other two study sites.

The nutritional education program for mothers offered by the CHWs also was reported as being very successful when compared to the same program offered at local nutrition rehabilitation centres. This conclusion was based on statistics which indicated that while the percentage of children who gained weight over the study period was higher for those attending the rehabilitation centre program, the weight gain of children whose mothers participated in the community-based program was higher, and extended to siblings.

The impact of the program in affecting mortality rates for major disorders is unclear from the information provided in the evaluation. For example, over the three year period that the
evaluation covers, the mortality rate for adults aged 15-49 years for deaths related to TB remained unchanged, as did the infant mortality rate for deaths due to gastroenteritis. In fact, the latter increased over the study period.

The research team concluded that the project had been a success as it demonstrated that the nutritional assessment methods used had a greater impact on infant growth than the rehabilitation centre-based program; that mobile teams of health auxiliaries had increased access to health care services for the study population; and, that neonatal tetanus had been virtually eradicated in the study area over the project period.

v. Volunteer Health Workers (Thailand)

This thirty month project (1976-1978) was conducted by staff of the Faculty of Medicine, Khon Kaen University, located in Khon Kaen, Thailand. The project took place in a rural area of northeast Thailand. The objectives of the project were to train local villagers as primary health care workers, use them as contacts with the existing health care system, and evaluate them in their tasks. The volunteer health workers (VHWs) were trained in such tasks as advising villagers on nutrition and prevention of communicable diseases, and in providing first aid, malaria treatment, and delivering babies at home.(41)

The researchers had originally intended to use a one group
pretest-posttest evaluative research design. However, this approach had to be abandoned as the quality of data collected by the VHWs was inadequate and not of a quality which allowed comparison with baseline data. As a result, the evaluation used a one-shot case study design.

Demographic data and vital statistics were collected from a sample of 282 households. This figure represented 50% of all households in four study villages (the study was to include nine study villages initially). The VHWs were also expected to maintain records of activities performed, to be turned into the researchers on a monthly basis. The VHWs were supposed to conduct baseline surveys in the study villages. They were to collect mortality and morbidity, demographic, and socio-economic data for each study village. They were also to conduct a survey of environmental conditions, and the availability of improved wells and latrines. However, in the end, the researchers took a random sample of 325 families in two of the study villages (out of a total of 9 study villages). These households were interviewed to collect information on knowledge, attitudes and practices with respect to sanitary behaviour, and morbidity data on the prevalence of parasitic infections. Information on mother and child health was solicited from 311 of these same households, while an unspecified number of study households were asked questions pertaining to their knowledge, attitudes and practices about
family planning.

The results of the evaluation indicate that less than 50% of the households interviewed stated that they had used the VHW's services. For those who did utilize the VHW, they stated that they were generally satisfied with the services rendered. At the time of the baseline survey, 42% of the respondents stated that they used self-medication and drugstore services for their health needs. The follow-up survey indicated that this level had decreased to 16% of the respondents. However, there was no indication whether the same households were interviewed, nor what affected the decision of villagers to use these methods. The researchers claimed that the VHWs were responsible for increasing the use of family planning methods. While the number of acceptors did increase, according to the data collected (from 2,061 to 2,620 people), there is no corroborative data to support this claim. For example, it was never established whether this trend was symptomatic of a similar trend throughout Thailand.

The monthly reports submitted by the VHWs indicated that they most frequently treated colds, headaches, dysentry, and burns. However, the quality and reliability of these data are questionable, since report preparation and submission was very inconsistent. For example, the initial response rate for submitting activity records at the start of the project period was very low. The researchers discovered that the VHWs did not
know how to complete the forms. Following a short training course in the villages, the response rate for submitting the activity reports did increase to 58% (67/105). This increased during the subsequent months, reaching an 80% response rate at one point. But after one year, the proportion of VHWs submitting activity reports had declined to below 50%. Consequently, the evaluation was inconclusive as to the effect of the VHWs on the health status of the study population. The researchers, nevertheless, concluded that the VHW program was a success, although "it is too early to conclude on its total impact".

vi. Rural Teachers Health Program (Paraguay)

The aim of this thirty-two month study (1978-1980) was to improve the health status of the rural population by implementing a primary health care program utilizing teachers trained in and assigned specific health duties. The project was designed, implemented and evaluated by senior staff of the Ministry of Public Health and Welfare in Paraguay. The project was based on two fundamental assumptions: that the teacher, in his capacity as a community leader, has a significant influence on children and parents, and could therefore be used to mobilize people to solve their own problems; and, that the school children could serve as a link between the teacher and
their families to promote improvements in environmental conditions and health behaviour in the home. The teachers were expected to provide preventive health education in the primary schools, make home visits to promote better health and environmental conditions, to provide basic curative services (primarily first aid), and to diagnose and refer more serious cases to local health clinics. (42)

To evaluate the impact and performance of the teachers, the researchers utilized a one group pretest-posttest evaluative research design. The project would be evaluated according to the following indicators: changes in key mortality and morbidity data; change in environmental conditions; utilization of the teachers for health problems; and, acceptance of the teachers by the community. A baseline demographic and morbidity survey was conducted for all households in the study area. A permanent record system was established, to be maintained by the teachers, to gather demographic, morbidity and mortality data during the course of the project. An end-of-project survey to assess acceptance of the program was also conducted.

The evaluation reported that the rural health promoters were referring cases to the regional health clinic, and that the referred cases were following referral instructions they had been given. However, it was difficult to establish the actual number of cases since some of the patients referred did not go to the health centre to which they were referred. In
some cases the health clinic did not retain or file the referral card, or though the patients did travel to the health clinic, they received no treatment.

The evaluation also noted that accessibility to primary health care personnel had increased in the study area, and that a modification in demand patterns for health personnel had also occurred, with 13.7% of demand for primary contacts directed towards the rural health promoters by the end of the study period. An increase in the utilization of trained midwives for deliveries were also noted over the study period, although there was no evidence presented whether this was a continuance of a pre-study trend. Immunization rates of 60% were reported, with 95% of those immunized reporting that they had attended talks on the subject prior to the campaign.

While the morbidity pattern of the study population apparently changed over the study period, the research failed to demonstrate a causal link between the availability of rural health promoters and this trend. However, the program appears to have had a significant impact on the installation of latrines. There also appears to be some association between this and the lower frequency of gastroenteric disorders. During the study period, the percentage increase in the number of latrines constructed was 375%, as compared for an increase of 146% for the entire health district over the same period. During this same period, morbidity due to gastroenteric
diseases dropped from 260/1000 to 170/1000 persons.

Improvements in the sanitary conditions of village homes were also recorded, as illustrated by the fact that the proportion of homes where garbage was burned rose from 53% to 68% over the study period. An increase in households using water drawn from the village wells was also noted. As the schoolchildren had been the recipients of a major health education program, it was assumed that this program was responsible for the changes. However, historical data which may have provided support was unavailable to corroborate this conclusion.

There was an indication that the increased responsibilities given to the teachers affected the teachers' capacity to function effectively. Involvement in the program was voluntary and required the teachers to donate their free time to health activities. Over the course of the project, a significant proportion of the teachers initially involved in the project (20%) found this unacceptable and resigned from it. How this affected the overall impact of the program was not recorded in the evaluation study.

Although it was not a specific objective of the project, the evaluation indicated that the children are important change agents in the communities. It was noted that they exercised a significant effect on the decisions of their parents to seek preventive health care and improve environmental and sanitary
conditions. Based on the evaluative research findings, the researchers concluded that the rural teachers could be trained to provide certain health-related activities, and that this was an acceptable arrangement to the communities. However, the evaluation failed to provide conclusive evidence that the retrained rural teachers were directly responsible for changes in mortality and morbidity rates.

Analysis of the Evaluation Reports

The foregoing provides a descriptive overview of the objectives, activities and indicators used to evaluate the five case studies selected for this review. In summary, the evaluations claimed that the deployment of community-based health workers affected positively the health of the target populations. Yet even a superficial analysis of the case studies seems to reveal that such a claim is unsubstantiated. The remainder of this chapter analyzes the five case studies in order to identify those factors which limit the validity and utility of the evaluation results.

Perhaps the primary limiting factor relates to the nature of the measures used to assess project effectiveness. Indices based on mortality (such as crude and age-adjusted death rates, infant mortality rates, and the expectation of life) and on morbidity (prevalence and incidence of major disease
categories) have traditionally been used as measures of improvement in health status. Control of mortality and morbidity have usually been the paramount goals of health programs. Variations in death and illness rates have been considered as direct measures of progress toward that goal. (43)

The supposed importance of these indices as criteria of success for health care programs is illustrated in the following statement:

The ultimate goal of a health program is to reduce morbidity and mortality in the population covered. All descriptions about acceptability, coverage, costs, or feasibility, and quality control process are meaningless to health officials unless some concrete outcome in reduced illness and death is demonstrated. (44)

But while changes in health status may be the best indicator of impact, such data are expensive, difficult, and often time-consuming to collect, analyze and interpret. They often yield ambiguous results. This may be due in part to the fact that there is no fully satisfactory definition of health that permits the use of comprehensive measurements, or the establishment of conclusive relationships between health-related interventions and change in health status. The concept of health as the 'absence of disease' is unsatisfactory for several reasons.

First, health is more than just the absence of disease. Considerations of consumer satisfaction or quality of life are
also important considerations. Additionally, the interrelationship between health care programs and health status is complex. Statistical indicators such as morbidity and mortality may reflect changes that have not directly resulted from health service delivery programs, but rather from improvements in other sectors. (45) Due to these shortcomings, more adequate and appropriate measures to evaluate the impact and performance of health care programs were proposed.

Second, the definition of health is a function of culture. Both the concept of disease and its causation are culture-bound. What is perceived as ill-health in one community or segment of the population may not be similarly perceived in another community or population subgroup. This may be due to differences in culture.

Third, to measure whether, or to what degree, a particular disease or condition is present in a population ignores the macro-level relationships between health and ill-health. Focusing on one particular condition may be impractical or unrealistic when other diseases exist which may have an equally debilitating effect. As Schaefer points out, "to be saved from malaria so as to suffer or die from malnutrition offers a poor choice". (46) The lesson here is that an evaluation, if it intends to document the relationship between interventions and health status, must be comprehensive. This entails a thorough knowledge of the inter-relationships among identified
conditions of ill-health, an their cumulative effect on health status.

It is also difficult to 'prove' that health interventions lead to improvements in health. It is not simply a matter of documenting the availability and utilization of technical interventions. If these interventions are expected to be effective, then prevailing behaviour patterns must also be affected. But to document the process of behavioural change is both a time-consuming and expensive exercise.

Four of the five case studies cited above used mortality and morbidity data to evaluate the impact of the PHC projects. Yet the validity of these measures is put into doubt when examined closely. For example, in the Iran project, the researchers claimed that the effectiveness of the auxiliary health workers was demonstrated. They based this claim on evidence presented to show that the infant mortality rate in the study villages was less than in the control villages at the end of the project (84/1000 infants versus 138/1000 infants respectively). However, the actual trend for infant mortality statistics in the study and control regions was the following:

<table>
<thead>
<tr>
<th>Year</th>
<th>Study Villages</th>
<th>Control Villages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972</td>
<td>36</td>
<td>134</td>
</tr>
<tr>
<td>1975</td>
<td>64</td>
<td>128</td>
</tr>
<tr>
<td>1978</td>
<td>84</td>
<td>138</td>
</tr>
</tbody>
</table>

(47)
If one accepts this data as valid, then it denotes a different interpretation from that reported by the Iranian researchers in their evaluation report. While the infant mortality rate in the study villages was lower than that in the control villages at the end of the project period, this was not a recent phenomenon. In fact, the infant mortality rate in the study villages was historically inferior to that recorded in the control villages. Yet no comparative analysis of this trend was reported in the project evaluation.

A second major flaw in the evidence presented in the project evaluation is that over the study period, the infant mortality rate recorded in the study villages increased significantly. However, the infant mortality rate in the control villages actually decreased over the first three years of the project, before increasing towards the end. Given that the total population for the study villages was approximately 10,000 people, and assuming that the proportion of infants aged less than five years of age constituted approximately 20% of this population, then it could be estimated that there were approximately 2,000 infants resident in the study villages. Based on the recorded infant mortality data, the actual number of infants to die in 1972 and 1978 increased from 52 to 168 (assuming a constant population for this age group over the study period). This constitutes a 223% increase in infant mortality. By comparison, infant mortality increased by 3% in
the control villages over the same period. This would seem to refute the researchers' claim that the auxiliaries were effective in improving health status.

Using infant mortality as an indicator of community health status was not practical or valid in this case. Additionally, the researchers did not provide statistical information for other ill-health conditions in the study and control villages. One is left with the impression that infant mortality represents all ill-health conditions in the study area, an unrealistic assumption at best. By restricting the range of conditions to be measured, the researchers in effect limit the validity of their research.

The nature and means of using morbidity and mortality data also limited the conclusions which could be drawn from the evaluation of the Colombian project. In this case, the claim concerning the success of the project was based on evidence drawn from several different measures (see Table 2). However, while the morbidity and mortality data used would seem to indicate that improvements in health occurred for specific population groups over the study period, it is difficult to base the impact of the program on the evidence presented. First, as in the Iran project evaluation, historical morbidity and mortality rates were not established or reported. Therefore, the morbidity and mortality data cited in the report could not be seen in perspective of historical trends or
patterns. Perhaps the decrease in the infant mortality rate cited in the project evaluation was simply an extension of a trend, and not the result of the impact of project activities. Second, no data for control villages were recorded or analyzed. It is difficult to claim success of an intervention without the presence of comparative data for study populations which did not receive the intervention. While certain patterns of infant mortality were recorded in the study villages, the researchers provided no data with which to determine if such trends were unique to the area. Third, while the Colombian researchers did use a wider range of disease categories to calculate morbidity data than the Iran project did, these data were representative only of specific population groups. They do not necessarily reflect the ill-health conditions which characterize the majority of the study population. The researchers use a narrow band of ill-health conditions to generalize about community health status. Consequently, the validity and the utility of the research results are thrown into doubt.

In the Rural Health and Family Planning Project (Haiti), the morbidity and mortality data recorded were neither comprehensive, nor consistent in the format in which they were reported. For example, data on the prevalence of tuberculosis in the three study villages was reported as an average over the study period. The lack of disaggregated statistics does not allow for comparison among these villages, nor across time.
Additionally, the data recorded related to ill-health conditions for specific population subgroups. There were no comparative statistics from populations outside the study area, nor was there an examination of ill-health conditions representative of the entire study population. The statistics reported in the evaluation report do not provide sufficient evidence to suggest that the PHC approach used in this particular project had a positive impact on the health of the intended beneficiaries.

The Rural Teachers Health Program (Paraguay) recorded morbidity data for those diseases most common to the study population. These statistics were collected at the start and at the termination of the project. Over the study period, the researchers found that the prevalence of these diseases tended to decrease in the study communities. However, as with the previously cited studies, no comparative data was collected. Consequently, the researchers did not establish whether the recorded morbidity trends and patterns were unique to the study area, or were consistent with morbidity trends and patterns elsewhere in Paraguay.

Another shortcoming which characterized the project evaluations which used general morbidity rates for measuring impact is that these measures tend to subsume the existence of differential rates for various population groups and across diseases. Not all members of the study population have the same
disease or ill-health condition, at the same level of severity, or at the same point in the disease cycle. Likewise, those who are ill may be the victims of multiple ill-health conditions simultaneously, although they are demonstrating symptoms of only one disease. Therefore, it is difficult to determine whether or not appropriate and adequate health care services are being provided by collecting morbidity data across an entire study population at a unique point in time.

The indices used to measure project performance or output usually focus on the availability of health services and personnel, and on the nature and frequency of tasks performed. But this information provides little if any indication of the benefits accruing to the recipient population. Even when both impact and performance indices are used, little effort is made to identify the nature or the degree of the relationship between the characteristics of the health care delivery system and health. As one author states, "we do not know very much about how or whether most health personnel service procedures are related to improvements in the health of individuals [or] of populations". (48) One is usually left to assume that if positive changes in health status indices occur over the project period, that the health care services delivered were appropriate, and the quality of care adequate. Yet such is not necessarily the case.

This analysis has demonstrated that the use of morbidity
and mortality data in the case evaluations presented have not supported claims that the PHC approaches used have had either a positive or significant effect on health. First, the range of ill-health conditions cited represent only a subset of the conditions which affect any community. Second, some of the measures were recorded only at one point in time. Third, comparative information was not recorded for populations not receiving the interventions. Standing on their own merit, such data provide little substantive information. Additionally, the lack of a historical perspective of morbidity and mortality data also limits the validity and utility of the recorded data.

Another factor which limits the credibility of the morbidity and mortality data recorded in these evaluation studies is the time interval over which the data are compared, particularly the time interval between the introduction of the intervention and the ultimate collection of data. There appears to be an expectation that changes in health status will occur shortly after the introduction of health-related interventions. Yet before one can expect health to change, attitudes and beliefs which affect behaviour must be changed. This is a long and complex process. Also, assuming that project activities will proceed as planned is unrealistic. Most projects incur delays, many due to unexpected events. Typically, projects fall behind schedule because of delays incurred in the transfer and availability of funds; delays in the training and deployment of
community-level health workers; delays in receiving equipment and supplies; and, delays in planned activities due to inclement weather conditions, unavailability of transportation, or more seriously, political or civil unrest in the study area. The cumulative effect of such factors can have a deleterious effect on the expectations of project staff with respect to the impact of project activities on health.

The longest project period was five years (Colombia and Iran). But the actual period between the introduction of the health-related interventions and the point in time when the final surveys were carried out were much shorter. For example, the community health workers in the Haiti and Iranian projects were not deployed at the start of these projects, nor to all study villages simultaneously. In other cases, equipment and supplies were late in arriving at project sites, thereby causing significant delays in the start-up of project activities (Colombia and Paraguay). In the Paraguay project, inclement weather and poor roads were responsible for reducing the number of supervisory and monitoring visits by the research team to several schools, particularly those more inaccessible locations in the study area. Due to their already heavy workload, and the fact that their health activities were voluntary and performed in their spare time, the teachers involved in the program required close supervision and encouragement. As the final evaluation report points out, the
supervision provided was inadequate to ensure a continuous and comprehensive program. This could conceivably affect the intensity of their health-related activities, and therefore their impact on the community. Had the project period been longer, more supervision might have been provided. As it was, the project period was too short to demonstrate any significant change in health. The effective project time period was significantly reduced in these projects. The USAID report suggests that five to ten years, rather than three, be adopted as a more realistic timeframe in which to implement project activities satisfactorily, and in which to measure health indices. (49)

Another factor which limited the usefulness and validity of the five case study evaluations was the evaluative research design used. Two of the five cases cited (Iran and Haiti) originally intended to use a true experimental design (a pretest-posttest control group design). However, for a variety of reasons, none of which were explained in the evaluations, the actual research design used was non-experimental. The rationale for using a true experimental design is to demonstrate that the intervention caused a significant difference in selected measures between the study and comparison (control) populations over time. Such paradigms require two conditions: that subjects be randomly allocated to the treatment and control groups; and, that pretreatment
measures be taken in both populations.

The random allocation of subjects to treatment and control groups constitutes the cornerstone of the experimental paradigm. Randomization is a powerful tool in that it increases the probability that treatment and control groups will be similar so that alternative explanations for program results may be ruled out. In random selection the experimental group and control groups are both drawn from the same population. Selection of persons for the two groups is carried out in such a way that each person, or segment of the population (dependent on the selection criteria) has the same probability of being selected in either group.

A second factor required for the correct comparative analysis of the treatment (experimental) and non-treatment (control) groups is the use of pretreatment measures. These measures compare the experimental and the control groups before the experiment and define those variables considered to be relevant to the criterion that will be used to measure the change. The utilization of measurement before the experiment serves two purposes. First, it permits demonstration that the experimental and control groups are equivalent to begin with, and second, it provides the reference point for change.

Despite the methodological rigour and the objectivity offered by the experimental designs, their shortcomings, particularly when applied to a field setting, have placed them
under increasing scrutiny. Many of the criticisms are related to the practical problems which arise when using them, particularly in field settings. Researchers have found that many of the key requirements can rarely be met: the subjects cannot often be randomly assigned to groups; equivalence of the experimental and control groups cannot be guaranteed; intervention and outcome commonly have only tenuous theoretical links; many programs simultaneously employ multiple interventions; and, key sources of variance cannot be controlled. In the Iran and Haiti cases, the evaluative design chosen was inappropriate since none of the key requirements could be met. Consequently, the effectiveness of the proposed evaluative research designs was severely limited.

The three remaining case studies used non-experimental research designs. While these approaches are less rigorous, and require fewer controlling assumptions, they are also largely ineffective in demonstrating causality. For example, in both the Colombian and Paraguayan case studies, the researchers used a one-group pretest-posttest design. However, without external comparative populations, the value of the data collected is significantly limited. Simply comparing before-after measures within one study population without reference to trends and patterns in comparative groups which did not receive the intervention is not sufficient to claim that the interventions were effective in affecting health.
Another factor found to limit the validity of the evaluation studies presented in this study is the failure on the part of the researchers to account for the existence and possible effect on health status of other development projects which occurred simultaneously with the primary health care projects. Other development activities could have a profound effect on health, although they are not health specific. For example, the Rural Health and Family Planning (Haiti) project evaluation claimed that the community-based nutrition education program was largely responsible for improvements in the nutritional status of children in the study area. Yet simultaneous to the deployment of the project community health workers, a major integrated development program was being implemented in the study area. Changes in the availability of local foodstuffs and better marketing procedures resulting from development program activities may have been more influential in affecting the nutritional status of Haitian children than the community-based nutrition education program. However, this possibility was not discussed in the project evaluation.

In the Colombian project, the health-related activities were only one component of a wider spectrum of community development activities. Again, the researchers did not examine the potential effect of the other project activities on the health of the study population. Therefore, claims that the health program was uniquely responsible for noted changes in
health indices are spurious. The same can be said about the results reported for the Paraguayan project, as no information was provided concerning the existence of parallel health promotion campaigns in the study area.

The lack of adequate and reliable data collection techniques is another factor which limits the validity of the evaluation exercises and the utility of the data collected. This criticism relates as well to the instruments used to gather data and the manner in which they are administered. The results obtained may not be valid if the questionnaires administered before and after the program were not identical, or if interview techniques differed, including those responsible for conducting the interviews. While follow-up surveys were conducted in most projects, there was no indication that the content of the questionnaires was the same as that used in the baseline survey. For example, it was unclear in the majority of the studies whether the same families were interviewed during both the baseline and follow-up surveys. No indication was provided as to the rate of absenteeism among the population during the follow-up surveys, a situation which could significantly bias the validity and interpretation of the data collected. As well, while some of the studies utilized specially trained personnel to conduct interviews, it can only be assumed that the same people conducted the follow-up surveys in order to save time and
expense related to training.

A case in point is the information collected for the Volunteer Health Worker (Thailand) project. The methodology used to collect the data severely limited the usefulness and significance of the entire evaluation exercise. The collection of baseline data on study villages was based upon the VHW's recall abilities. They were expected to recall specific conditions in the villages prior to their training, even though over two years had lapsed in many cases. There was also bias in reporting health statistics due to the high rate of non-response of VHWs in submitting their activity reports.

In the Kavar study, the research team intended to collect and analyse data related to improvements in sanitary facilities and changes in sanitation behaviour over the course of the project. However, as the evaluation noted, these data were collected too infrequently to be of much use. This is also the result of those expected to deliver health care services also being expected to collect and compile statistical data. The greater the amount of data to be collected by field workers, particularly if they are the same people delivering the services, the greater the probability of data errors. The same is also true for questionnaires or forms which are complex and difficult to administer and interpret. A case in point was the Paraguay project, where the researchers reported that even those people given prior training (including the rural
teachers) experienced difficulty in understanding and using the data sheets correctly. Despite the training sessions where it was explained how the forms were to be completed and despite the emphasis placed on the importance of the forms in the evaluation of the project, the on-going training of teachers and supervision of project activities, a significant number of the teachers made virtually no use of them. Of those forms which were completed and submitted to project administration, it was reported that approximately 40% were incorrectly or inadequately completed. Consequently, the research team was unable to measure changes in some of the indices which the program originally sought to examine.

Another shortcoming of many of these evaluations is the lack of verification of the data collected. For example, while some of the projects used specially trained enumerators to collect data, no details were provided of the methods used to verify the validity of the data collected. An exception was the Colombia project which undertook to ensure a high degree of reliability and completeness for the information collected by the rural health promoters. In a separate survey, the data collected in the initial surveys were checked against the responses of a random sample of one third of all the individuals served by them. All family members were interviewed and the prevailing sanitary conditions of the household compound recorded. Information on sickness and incapacitating
illness was also collected over a period of two weeks, while mortality and parasitic infection rates were collected over a period of one year. The evaluation found that the information collected by the rural health promoters was consistently of a high quality, and differed little from the data collected using the comparative survey.

The use of quantitative data, almost to the exclusion of qualitative data, is another factor which seems to be responsible for limiting the significance and utility of the case study evaluation results. Researchers tend to concentrate on collecting quantitative data by objective survey methods, to the neglect of other data which may be more salient to the performance and effect of the innovative approach being used. They disregard such data because it is 'subjective', 'anecdotal', or 'impressionistic'. (50) Cochran (1980) argues that quantification tends to add elements that are unwanted and overlooks dynamics that are part of the process of social change and interaction. (51) She warns that quantitative data may have an aura of authenticity even when they are inaccurate or false, especially if they come from a reputable source. (52)

For the most part, the five case studies relied on quantitative data as evidence of the impact of their respective project activities on health. But the validity of such data is significantly decreased due to the absence of comparative and corroborating data, as has been previously discussed.
Qualitative data could have served as corroborative evidence to support the quantitative data collected in these studies.

For example, in the Paraguay PHC program evaluation, interviews on reactions to the program were conducted with both the teachers and the villagers. While the qualitative data collected in the study did not conclusively link the activities of the teachers with apparent changes in the behaviour of villagers and their health status, the data collected during the interviews lent support to the association suggested by the quantitative data. The qualitative information collected suggested that the schoolchildren assumed the role of community change agents, playing an extremely important role in disseminating information relating to health activities. As a result, the vaccination campaign was very successful and included their pre-school aged siblings as participants, an activity which had not been foreseen in the original project goals. It was also reported that due to the influence of the children on changing behavioural attitudes within the family, a significant number of households decided to modify their hygiene and sanitation habits. Had such information been collected in the other studies, perhaps the results drawn from the quantitative data would have been of greater significance and usefulness.

The foregoing analysis would seem to indicate that the factors which affect the validity and utility of evaluations of
primary health care programs designed and carried out by expatriate agencies and organizations also characterized those designed and implemented by indigenous researchers. The major problems identified in this study can be summarized as follows: evaluative research designs which are either too complex for the field situation at hand, or are over-simplified such that they are incapable of demonstrating even a strong association between the intervention and program performance and impact; an insufficient timeframe over which the impact on health of health-related interventions can be adequately measured; the nature and quantity of data collected; the reliance on morbidity and mortality data to demonstrate causality; the lack of a global perspective with respect to activities in other economic and social sectors which may affect health; and, the lack of qualitative data to substantiate the claims made using quantitative data. These factors seem to fall into two major categories: design deficiencies and weaknesses in methodology and technique. The task now becomes that of identifying strategies to overcome these limiting factors, so that field-based evaluations are credible.

Summary

The chapter has presented an analysis of the major factors which limit the validity, and ultimately the utility of
evaluative research on primary health care in developing countries. The analysis has illustrated that these are not restricted to evaluative research design faults. The indices chosen to measure program performance and impact, and the techniques used to conduct the evaluation, particularly the data collection methods, also have a significant effect on the validity and reliability of evaluation results. The issue is therefore not so much whether one particular research design is better than another. Rather, the aim of the exercise is to ensure that the construct deficiencies inherent in the design and the techniques chosen are adequately dealt with. The next chapter presents a number of suggestions to strengthen the validity and utility of field-based evaluations conducted in developing countries.
REFERENCES


15. American Public Health Association, Delivering Health Services, p. 53.

17. American Public Health Association, Delivering Health Services, p. 53.

18. Parlato and Favin, Primary Health Care, p. 79.


21. Ibid., p. 79.


26. Parlato and Favin, Primary Health Care, p. 81.


35. Parlato and Favin, Primary Health Care, p. 80.

36. Ibid., p. 87.

37. Ibid; discussions with CIDA personnel.


42. R.H. Malloquin et al., Informe Final y Conclusiones del Programa Extension de los Servicios Salud a traves de Maestros Rurales [Final report and conclusions concerning the extension of health services through rural school teachers] (Asuncion, Paraguay: Ministerio de Sanidad y Asistencia Social, 1980).


49. Parlato and Favin, Primary Health Care, p. 88.


CHAPTER THREE

ALTERNATIVE APPROACHES FOR CONDUCTING FIELD EVALUATIONS

Introduction

Before further investments are made by governments and funding agencies in the development, implementation and expansion of primary health care programs, valid and useful information about program performance and impact are required. But as the thesis has pointed out, the results of many, if not most, of the evaluations conducted to date are not credible. The previous chapter identified several factors which have limited the validity and utility of the results of evaluative research on primary health care in developing countries. While some are related to design considerations, most are associated with the techniques and methods used in conducting an evaluation. Therefore, if the quality of evaluations is to be improved, then strategies to address these shortcomings must be found.

The purpose of this chapter is to identify alternative designs and techniques for evaluating primary health care programs in developing countries which may be appropriate in addressing these limitations. The first issue to be dealt with concerns the need for demonstrating causality. The chapter therefore begins with a discussion of the level of certainty
required for field evaluations of social service programs conducted in developing countries. It then presents some alternative strategies and techniques for conducting valid and useful evaluations of primary health care projects in field settings. Although the case studies reviewed refer particularly to PHC projects carried out in developing countries, the suggestions made are meant to apply as well to field-based social service evaluations in general.

The Level of Certainty Required in Evaluative Research

Many of the evaluation studies referred to in the previous chapter attempted to establish a causal link between a set of services provided by primary health care workers and improvements in the health status of the study community. The evaluations sought to produce conclusive evidence about the relationship between primary health care activities and health status. An evaluation study which provides conclusive evidence provides the soundest basis for action, as it "establishes that . . . an event, process, or outcome [was] possible and probable, [and] that indeed it did occur". (1)

For Shortell and Richardson (1978), rigorous experimental program evaluations are necessary when beliefs about cause/effect relationships are certain. In such cases,

rigorous experimental evaluations [are] . . . important in program evaluation research . . . as a
defense against critics [of the program] who will attempt to refute the scientific evidence by criticizing the study design, methodology and analysis. The stronger the design used . . . the greater the ammunition available . . . to defend the program's operation, assuming, of course, that the results are basically favorable. (2)

However, while rigorous, experimental evaluative research designs may provide the strongest form of evidence, they are the most expensive, difficult and time-consuming to carry out. However, as the previous chapters has pointed out, there are numerous practical problems related to conducting program evaluations in developing countries. As a result, what are originally intended to be scientific, conclusion-oriented studies often failed to produce the expected results. In many cases, they simply report that village health workers had been trained and deployed, and in some cases, that primary health care services were utilized.

Yet, the need to prove, to demonstrate conclusively that a causal relationship exists, may not be necessary when one is concerned with the production of information for administrative decision-making. Several authors have discussed this issue at length. (3) House concludes that expecting evaluation to provide conclusive and compelling evidence is "hoping for more than evaluation can deliver".

But if [evaluation] cannot produce the necessary, it can provide the credible, the plausible, and the probable. The results are less than certain but still may be useful . . . evaluation persuades rather than
is credible rather than certain. . . .

Once the burden of certainty is lifted, the possibilities for informed action are increased rather than decreased. (4)

While there is a need to consider establishing conclusive evidence, one cannot afford to be overly zealous in attempting to ensure the highest degree of confidence and credibility in the evaluation results. As has been previously stated, too rigorous an evaluation framework, particularly in a non-supportive environment, will most likely lead to erroneous results. As Smith explains, there are many levels of certainty, and that in general, the level of certainty of an evaluation study is dependent on the following criteria:

1. that alternative conclusions have been eliminated;
2. that systematic bias or error has been eliminated;
3. that contextual factors have been appropriately accounted for;
4. that assumptions made are accurate;
5. that the most relevant variables and relationships have been examined;
6. that all salient audience perspectives have been considered; and,
7. that evidence was first-hand and material rather than circumstantial and hearsay. (5)
Instead, demonstrating a strong association between several fundamental characteristics of primary health care services and personnel (i.e. their availability, accessibility, performance, etc.) and changes in quality of life and health indices may be sufficient. By defining evaluation as an administrative tool rather than as a science, less than conclusive evidence may be both appropriate and sufficient to satisfy the information needs of those for whom the evaluative research is undertaken. Where there are limited resources for evaluative research, less than conclusive evidence may be the only possibility. (6)

Smith (1981) proposes that where treatment effects are not expected to be harmful, where the incremental benefit to individuals is high, where the potential risk to participants is low, and where the program is developmental, less certainty is needed for establishing relationships between the project interventions and demonstrated impact. He defines two other levels of evidence which may be appropriate to the needs of program administrators of health care programs. The first of these is what he calls suggestive evidence, wherein it can be stated that 'X [the hypothesis] is possibly true'. As Smith explains, an evaluation study that provides only suggestive evidence:

merely establishes that certain events, processes, or outcomes are plausible . . . [it] is often the only type of evidence available in formative stages of programs . . . [and] is the weakest form of evidence, the easiest to obtain . . . (7)
The second level of certainty in health program evaluative research proposed by Smith is one for which it can be stated that 'X is probably true', called preponderant evidence. An evaluation study which provides preponderant evidence establishes that . . . an event, process or outcome . . . more likely occurred than not. A preponderant claim indicates that the weight of the information collected to date supports the claim as probably being true.(8)

Most 'information-based' decisions are actually made on the basis of preponderant evidence due to the lack of available data to support conclusive evidence. This level of evaluation provides a stronger claim than suggestive evidence on which to base decisions. Where resources are limited, and the need for conclusive evidence may not be required or appropriate. In such cases, the use of either a preponderant or suggestive level evaluation study is advocated on practical grounds. Although these factors may suggest what level of certainty is most desirable, determining what level of certainty is attainable will require balancing these factors against the human and financial resources and the time available as well as the prevailing field conditions.(9) Using professional judgement in balancing these criteria will enable the evaluator to justify whether a higher or lower order of proof is required.
Of critical importance is choosing the approach most suitable for the situation and context in which the evaluation is to be conducted. The method should fit the program and the problem rather than the other way around. The research approach generally should be dynamic and adaptable, and concentrate on providing the most relevant and comprehensive information. Less experimental, open-ended approaches allowing for serendipitous findings, may be the most appropriate. (10)

Alternative Approaches in Evaluative Research

Proving conclusively that a particular intervention has a particular outcome is a difficult task even under the best of circumstances. The problems related to carrying out valid and meaningful evaluations in field settings has prompted the search for alternative approaches. As the analysis carried out in the previous chapter points out, there are several factors which tend to limit the validity and value of developing country-based primary health care evaluations. These can be grouped into two broad categories: evaluative research design; and, evaluation methodology and techniques. The remainder of this chapter will address each of these and suggest alternative approaches which are expected to address these limitations.
i. Evaluative Research Designs

The evaluative research designs used by the research teams to evaluate the primary health care projects cited in the previous chapter were insufficient to demonstrate any significant relationship between the project interventions and changes in health status of the recipient populations. The factors responsible for limiting the validity of these designs were identified and reviewed in the previous chapter. Basically, the rigorous experimental designs, besides being costly and time-consuming, require so many controlling assumptions that the program being studied becomes fragmented. In the end, the research results are unrealistic and irrelevant to programmatic decision-making. (11)

For Hall (1979), the experimental evaluative research design tends to oversimplify social reality by employing survey research methods which extract information from individuals in isolation of one another, and then aggregate these data into a single set of figures. According to Hall, this serves to reduce the complexity and richness of human experience. Social responses to problems by groups of people are not necessarily the same as the total of individual responses of people acting alone. To facilitate the analysis of this information, responses are often quantified. However, what often results is
an illusion of accuracy, resulting in often unclear and mystifying research.\(^{(12)}\)

Boruch and Gomez (1979) argue that given the realities which affect the implementation and evaluation of field-based programs, much of the effort involved in ensuring precise experimental research and testing is in actuality a futile exercise. They contend that statistical tests of less than perfect reliability and sensitivity, measures of less than perfect accuracy, and programs which are less than perfectly implemented combine to compound the possibility (and probability) that these experiment-based models will not demonstrate significant program effects, even if they occur. Given the rarity of conducting a successful true experiment, and that quasi-experiments require additional statistical manipulations, then the program impact would have to be enormous to be noticed.\(^{(13)}\)

The non-experimental approach used by some of the evaluations cited in the previous chapter, on the other hand, are structurally weak. The lack of control or comparative populations, historical data, or investigations into parallel program activities and their possible effect on the health of the study population constitute serious shortcomings. While they are easier to conduct, their results provide insignificant and irrelevant data.

The evidence presented suggests that neither the rigorous
and complex experimental research designs, nor the simplistic non-experimental designs are appropriate for conducting field-based evaluations. What are required are research designs which allow researchers to demonstrate strong associations between the presence and utilization of primary health care services and personnel and changes in health status, under often difficult and demanding field conditions, using the human and capital resources at their disposal. Quasi-experimental research designs perhaps offer an appropriate and realistic solution. Although the term was not coined until later, Stouffer (1950) and Campbell (1957) were first to introduce the concept of quasi-experimentation.(14) The classic study of quasi-experimental models, which evolved from the field of educational research, was produced by Campbell and Stanley (1980).(15)

Quasi-experimentation is a research strategy designed to investigate causal relationships in situations in which true experimental designs are inappropriate or difficult to apply. As Cook (1983) states, the strategy is to test whether a particular pattern of results is obtained despite the operation of potential validity threats.(16) As Campbell and Stanley (1980) explain, the quasi-experimental designs are applicable in situations where the researcher wants to introduce something akin to a true experimental design into the scheduling of data collection procedures (eg., the when and to whom of
measurement), even though full control over the scheduling of experimental stimuli (the ability to randomize exposures) which makes a true experiment possible, is lacking. The strength of the quasi-experimental paradigm is that it permits probes of causal connections in field settings when or where the true experimental designs are not feasible, suitable, or desirable.

The quasi-experimental paradigm utilizes a before/after design, in which the status of the outcome, or dependent variable, is measured in either a single group, or in both the study and comparative group, both before and after the introduction of the intervention. These designs are intended for use in settings in which the researcher has partial but not complete control over the assignment process. Like the conventional experimental paradigm, quasi-experiments have interventions, outcome measures, and experimental populations. However, they do not use random assignment to create the comparison from which change caused by the intervention is inferred. Rather, assignment to the treatment group is by self-selection or administrator fiat rather than by chance. The comparison depends upon possibly nonequivalent groups that differ from each other in several characteristics other than the effect of some intervention. The task confronting those who interpret the results from quasi-experiments is basically one of assessing the extent of differences and separating the effects of a treatment from those due to the initial
noncomparability between the average units in each experiment group; only the effects of the intervention are of interest. To achieve this separation of effects, the researcher has to explain the specific threats to valid causal inference that random assignment rules out and then in some way deal with these threats. The irrelevant causal factors must therefore be identified and explained within the ceteris paribus of random assignment. (18)

The goal of the quasi-experimental paradigm is to identify the factors that are presumed to make a difference to specific outcomes, even if these involve only a small part of a larger system. Quasi-experimentation is moot on the issue of the multivariate complexity required for complete causal understanding, complete causal explanation, and total prediction. It seeks only to discover those parts of the real world in which causal relationships exist and in which manipulating one thing predictably causes another to change. The testing of causal laws, the generation of causal understanding are not its strong points. (19)

There are four basic quasi-experimental research designs which are appropriate to the present discussion: time series; multiple time series; nonequivalent comparison group; and, recurrent institutional cycle. (20) The time series design is perhaps the most unsophisticated of the quasi-experimental designs. It allows researchers to calculate multiple measures
both before and after the introduction of the intervention. The primary advantage of this particular design over the single pretest-posttest are its control of maturational effects and partial control of historical effects (due to the presence of multiple measures prior to the introduction of the intervention). The time series design is most useful when evaluating programs which are not subject to seasonal variations. If such is the case, then the observations periods must be more frequent.

The multiple time series design is similar to the previous design, except for the addition of a comparison group. This population must be as similar as possible to the study population (maximization of equivalency). The greater the similarity between the treatment and the comparison group, the greater the degree of confidence which can be placed in the research results. Assuming that the treatment and comparison populations are matched, then posttest differences between these two populations which are not due to the effect of the program under study must be related to selection differences. Multiple observations are collected for both groups prior and subsequent to the introduction of the intervention. The major advantage of this design over the single time series is that it allows for greater control for historical effects.

The nonequivalent comparison group design is similar to the experimental pretest-posttest control group design. However, in
this case, the subjects are not randomly allocated to the treatment population and the control group. Instead, a second population is selected which is as similar to the treatment population as possible. While comparable to the multiple time series analysis design, the nonequivalent comparison group design has only one pretest and one posttest observation, instead of multiple observations.

The recurrent institutional cycle design is a composite approach which compares a one-shot case study (posttest observation) for the treatment population with a single pretest-posttest design for the comparison population. It allows the researchers to compare the characteristics of a population which has been involved in the innovative program for some time with those of a population which is about to be exposed to the same program. Singularity, the individual components of this design are subject to several threats to validity. But when combined, this design controls for a number of threats to validity, particularly historical, testing, instrumentation and attrition effects. However, maturation and regression effects remain as potential alternative explanations for changes in population characteristics.

The evaluative research designs used in the five primary health care project evaluation case studies cited in the previous chapter could have been adapted to a quasi-experimental design format. For example, the original
intent of the researchers in the Iran and Haiti case studies was to use an experimental case-control design. But due to several factors not identified in the project documents, these designs were abandoned. However, even the modified designs were faulty, since the treatment and comparison populations were not matched. Consequently, differences in posttest measures between these two populations can not be attributed to the effects of the project, as too many other factors which could conceivably affect program impact were unaccounted for. Had the researchers matched the study and comparison populations, a nonequivalent comparison group design would have resulted (assuming, of course, that pretest and posttest observations were made for both population groups).

The addition of multiple pretest and posttest observation points in the other case study could have transformed the evaluative research designs used from single pretest-posttest designs to single time series designs. As was pointed out in the Colombia and Paraguay case studies, the absence of historical data for the selected measures limited the relevance of the study findings. The addition of comparison groups would have resulted in a multiple time series design. If the comparison group was matched with the treatment population, then a nonequivalent comparison group design would have resulted. In any case, the use of additional observation points and/or matched comparison groups would have served to
strengthen the evaluative research designs of these particular studies.

The extension of the time interval over which the observations are taken is also an important consideration. As pointed out in the preceding analysis, the project durations tended to be of insufficient length to adequately demonstrate a significant or positive link between the primary health care interventions and changes in health in the study community. Change in health is a complex and often lengthy process. Nor is it a linear process, wherein changes in health status are a direct function of the availability of preventive and curative health-related activities. If a realistic timetable for assessing changes in health is to be adopted, then a more practical and realistic project period than normally used is required. But a longer period between the point in time when the interventions are introduced and the point in time when the final project surveys are made will necessarily increase the number of observation points and the tasks of those involved in data collection and analysis. The costs associated with the project evaluation will also increase. However, by increasing the timeframe for the evaluation, the validity and significance of the research design will be enhanced.

But even a 'good' design can guarantee the validity, significance, and relevance of research results. As the next section of this chapter will demonstrate, the methods and
techniques used to collect data and the nature of the data can must also be addressed.

ii. Evaluation Methods and Techniques

Besides ensuring that the evaluative research designs are appropriate for demonstrating program impact, and conducive to the realities of the prevailing field conditions, the researchers must also ensure that the methods and techniques used, particularly with respect to the nature and means of collecting data, are also appropriate. The aim should be to ensure that the methods and techniques used serve to strengthen the capacity of the evaluative research design to demonstrate program impact.

The nature and number of measures to be used in the evaluation are an important consideration. The selected case studies did not establish any linkage between the activities of the community-level health workers and the measures selected. For example, while infant mortality may be considered a sensitive indicator, it is affected by more than just the actions of the community health worker or the primary health care project activities. Although a PHC program may include an immunization campaign to combat certain diseases, or nutrition education classes, unsanitary environmental conditions and/or the lack of adequate nutritious food may have a greater and
opposive effect on infant mortality.

Additionally, health cannot be perceived as a singular activity or sector. Health is an intrinsic component of the development process, the ultimate objective of which is the eradication of poverty. The nature of the interrelationship among the various manifestations of poverty is such that one factor influences all others in both positive and negative synergistic effects. (21) As Hendratta (1979) states:

The implications on the development of a health program are clear. Health is just one element in the total development picture; accordingly, health problems cannot be tackled in isolation from the other elements. In other words, the attack on the problems of health should be mounted as an integral part of a broad frontal attack directed against the multifarious forms of poverty. There can be no hope for a significant and lasting improvement in health in the absence of a simultaneous improvement in all other sectors. (22)

The lack of consideration of development activities in other sectors which parallel those reported in the case studies also serve to limit the relevance of the evaluation results. For example, in the Colombian case, the health activities were one component of a larger community development project. Yet no effort was made to examine the potential effect of development activities in other economic sectors with changes in health.

There are also factors internal to the process of delivering health care services which may affect outcome, but which are not necessarily identified or accounted for during
the course of the evaluation. These include consumer attitudes to project activities and personnel, or health-specific behaviours which affect practices, and ultimately, health. This broader frame of reference is necessary in order to account and control for factors which could affect the relationship between the project's dependent and independent variables. However, simply increasing the number of measures will not necessarily redress the issue, as this may only serve to increase the complexity and the costs associated with the evaluation. Care must be exercised to ensure that the most relevant and sensitive measures are selected. Such an approach requires a redefinition of the focus of the exercise.

One method which could be used to highlight the existence of these factors, as well as the direction and degree of their relationship to changes in health status, is the case study. The case study approach concentrates on the program processes and on how people view the program. This approach provides an understanding of the perceptions of those who are familiar with the program. It is naturalistic investigation. The aim is to establish credibility and confidence in the reported outcomes. Stake calls case studies a responsive approach to evaluation, which elicits opinion, finding out what is of value to the audience.(23) Case studies provide rich and persuasive information that is not available from other approaches. A strength of this approach is that it allows the representation
of diverse points of view, and can account for factors which affect outcome, but which would not necessarily be accounted for in the evaluation.

An equally important element of this approach is that it provides an opportunity for service providers and community members who use the services to become directly involved in an assessment of the performance and acceptability of the primary health care services, an important consideration if one of the aims of the program is to implicate the community in the development and operation of these services. The evaluator makes no attempt to manipulate, control, or eliminate situational variables. The existing and complex socio-economic, cultural and ecological environment is taken as a given. The case study approach is based on the following assumptions: the importance of understanding people and programs in context; a commitment to study naturally occurring phenomena without introducing external controls or manipulation; and the assumption that understanding emerges most meaningfully from an inductive analysis of open-ended, detailed, descriptive, and at times anecdotal data gathered through direct contact with the program and its participants. As House states, the case study approach "allows one to come to grips with the inner workings of the program". It is a strategy which advocates 'progressive focussing'. The task of the evaluator is to provide a comprehensive understanding of the complex realities
surrounding, an possibly affecting, both the process of implementation and the impact of the program. The evaluator tries to sharpen discussion, disentangle complexities, and isolate the significant from the trivial.(26) As Foster (1969) states, a more holistic approach to evaluation is more than seeking and explaining univariate causes and effects. The use of case studies in evaluation is an iterative process, one which requires an unusually broad and flexible field research methodology based on a holistic view of society and culture. Where the technical, social, cultural, economic, psychological, and other pertinent factors are not recognized in advance, this exploratory quality is enormously advantageous. It vastly increases the investigator's chances of hitting upon the critical element in any situation. (27)

The integration of qualitative and quantitative methods in evaluation can provide a more comprehensive analysis of the relationship between program activities and their impact. But as Patton (1980) points out, there is a need to carefully balance the proportion of quantitative and qualitative data in evaluating programs. As he states:

In many ways the real trade-off between quantitative methods and qualitative methods is a trade-off between breadth and depth. Qualitative methods permit the evaluator to study selected issues in depth and detail; the fact that data collection is not constrained by predetermined categories of analysis contributes to the depth and detail of quantitative data. Quantitative methods, on the other hand, require the use of standardized stimulus so that all experiences of people are limited to certain response
categories. The advantage of the quantitative approach is that it is possible to measure the reactions of many subjects in a limited set of questions, thus facilitating comparison and statistical aggregation of the data. By contrast, qualitative methods typically produce a wealth of detailed data about a smaller number of people and cases. (28)

The balance between the 'right amount' and 'right type' of quantitative and qualitative methods is an important issue. As Patton (1980) points out, it is qualitative data that are most often made to bear the larger burden of proof. Qualitative data tend to be rejected if they do not support the quantitative data. (29) The successful integration of qualitative and quantitative methods can be a difficult and complex task. However, the richness of detail with respect to explanations of 'how' and 'why' program outputs and processes affect program outcome which result from a combination of these two methodologies may be worth the price.

The use of case studies as supplements to the quantitative data collected may have enhanced the credibility of the results of the five case study evaluations. While some of the case studies cited in the previous chapter did collect and analyse some qualitative data, it tended to be related to surveys of the quality of primary healthcare services provided and personnel, based on the perception of users. Such surveys tended to be closed-ended, and rarely allowed for an expanded explanation. The significance of the information collected is also questionable. Expecting anything other than a positive
answer in reply to a question as to whether villagers are satisfied that PHC personnel are available in a situation where no health personnel previously existed would be naive. This also applies to the responses given in reply to questions about whether the community-based personnel should remain. Given that their primary purpose appears to have been to provide curative emergency care, their absence would mean having to walk long distances to health clinics to obtain treatment. Such data added relatively little to the significance and utility of the quantitative data which was collected.

The analysis of data which examined the relationship between the range of community development activities being implemented in the study areas and the effect (both real and perceived) of these activities on the health and welfare of the inhabitants of the study communities would have constituted a valuable addition to the quantitative data collected. For example, in the Colombian and Haitian cases, regional community development programs were being carried out in the study areas simultaneously to the primary health care field activities. Yet neither evaluation examined the relationship between these activities, nor their singular or cumulative effect on the well-being of the PHC study population. In the Thai project, case studies of some of the study villages may have provided a more comprehensive perspective about historical economic, social, cultural and physical environmental factors which
affect the health of the inhabitants. A linkage between these factors and the activities of the village-level health workers could have established whether, and how their health-related activities were responding to the manifestations of these factors. The same holds true for the Paraguayan and Iranian cases.

There are a myriad of survey techniques for collecting both quantitative and qualitative data. Direct interview techniques (questionnaires), counting surveys, and the examination of established record systems (precollected data) are probably the two most extensively used techniques. As previously discussed, the five case studies cited in the previous chapter tended to collect quantitative data through these traditional approaches. Direct interview techniques were used extensively for collecting demographic data and information on the presence of latrines, improved wells, and general environmental conditions in the study areas. Data drawn from clinic records were used to calculate rates of disease and to verify the nature and frequency of referrals made by the community health workers. The objective of these surveys was to collect information to calculate rates and frequencies.

But the collection of qualitative data may require a somewhat different approach. There are instances where the collection of data on certain activities may be culturally unacceptable. In such cases, it would be expected that the
evaluation would not collect such data on ethical grounds. However, an explanation should be made in the final report to this effect. For other activities, the direct interview techniques may be totally inappropriate. Examples include eating and personal hygiene behaviour. In such cases, an alternative approach is prescribed.

Participant observation is useful in case studies where one is interested in determining the attitudes and views of particular individuals within a population, but direct interview techniques may not be practical or appropriate. The assumption behind this research method is that in order to understand the people under study, one must get as close as possible to them in their natural setting. Only by extensive participation with the group under study can one understand how they view themselves, the program being evaluated, and their interrelationships and attitudes towards the program. Participant observation includes a variety of data collection methods such as indirect observation, informal interviewing, formal interviewing and document analysis. Good field work uses several of these techniques.

For the purposes of evaluative research on primary health care programs, participant observation may be useful for determining the following:
1. How the program actually operates at the local level, particularly with respect to how local environmental factors affected its operation;

2. How the program has been implemented. A detailed description of program characteristics such as its structural organization, activities, and the views of health care providers and both users and non-users of services;

3. How individuals feel about the program. As opposed to listing satisfaction according to an interval or ordinal scale, detailed analyses can be provided as to consumer satisfaction with the services provided, for instance; and,

4. How individuals have been changed by the program. Participant observation can provide a descriptive account of how service user's health status was prior to entering the health care delivery system, how particular services affected them, and how their lives changed following treatment.(31)

Participant observation offers a means of viewing the dynamic interrelationship of people and their environment. It provides the opportunity to the researcher to view first-hand the structures and processes which affect, and which are
affected by, the program interventions. None of the five case studies used participant observation as a means of collecting village-level data. This technique would have been useful to verify behaviour and to identify non-users of primary health care services. For example, although the Paraguay and Colombian evaluation teams collected information on the number of latrines constructed in the study areas over the project period, there are no data to indicate that they were being used. The success of the PHC programs should be measured relative to the use of such facilities, not simply with respect to their being constructed.

But traditional data collection techniques, such as those referred to above, place the research team as observers. Those being evaluated play a passive role throughout the exercise. The information needs of the evaluation are defined external to the study area. Feuerstein (1980) refers to this as the 'study of specimens' approach, whereby program participants are expected to play only a minimal role in the evaluatory process. This approach consists essentially of a brief explanation of the study's objectives, wherein the participants agree to be counted, examined and even questioned. These results are then removed from the study area for analysis. There is no feedback of the findings to the participants - neither do they expect it. The utilizers of the services are divorced from the evaluation results. The product of the exercise is a
conventional evaluator-focused evaluation. (32) This was the approach used in the five case studies.

Hall (1979) contends that conventional data collection methods tend to divorce evaluators from the realities of the situation they are meant to investigate. Questionnaires or interviews designed in an office of a university are by nature one-sided. The study population is relegated to being a source of information. People are regarded as having bits of isolated knowledge. However, they are never expected or apparently assumed able to analyze a given social reality. At the extreme, researchers take up people's time with badly formulated questions, and make interpretations based on little experience in the area.

Hall also proposes that rigorous, highly technical, and scientifically 'pure' research creates an illusion that only the university-trained are capable of collecting data. The abilities of people to investigate their own objective realities are not stimulated, and the pool of human creativity is kept within narrow confines. As he states:

Those most familiar with the problems and those whose daily existence is affected by poor health, poor nutrition, and low levels of production are effectively taken out of the active process of making the changes which might lead to improvements. Control is left to those who, by definition and levels of training, are outside the experiences within which change is sought. (33)
Hall also criticizes the use of experimental designs for failing to provide easy links to possible subsequent action. Although shortcomings in stated program objectives and in program activities are often cited, strategies designed to overcome these limitations and improve program impact are rarely identified. Yet the purpose of research is to promote action. It is expected that when changes are made, the people in the community or the participants in the program will participate more actively, more efficiently, or will gain increased benefits over what had existed before. However, this often is not the case. In addition to resulting in poor sources of information, such research methods tend to alienate the respondents, or at best treat them as sources of primitive information. Consequently, there is little likelihood of creating the active and supportive environment essential for change. (34)

Such techniques assume that individuals do not place value on the intervention, that they react objectively to its introduction. Yet, as previously discussed, the value of qualitative data lies in its ability to identify and explain interrelationships overlooked or ignored by quantitative data collection and analysis techniques. An alternative approach to collecting and analyzing such data is one element of the participatory research concept.
Participatory research is defined as a method of social investigation involving the full participation of the study community. The aim of this approach is to create an awareness on the part of the community of their own resources, and mobilize them to take responsibility for their situation. The evaluator becomes a participant and learner in the evaluatory process. The community takes a full and active role in the evaluation process, and have priority in decision making with respect to the design and implementation of the evaluative research exercise and in the dissemination of research findings. Because members of the study population have been intimately involved in all aspects of the program and the evaluation process, they may then require only minimal help in initiating and carrying out future evaluation studies. It is this latter approach which is critical to the success and relevance of the case study strategy.

The strength of participatory research is that it recognizes that the evaluator and subjects inhabit different realities and that it is presumptuous for the investigator alone to determine what is to be investigated. Participatory evaluation describes and analyzes the process by which health care is administered and the personal meanings attributed to this process by the service users. In this regard, it challenges the conventional systems-oriented approaches of input-output analysis and quantified results. Moreover, the
participatory evaluation approach attempts to connect the individual case to the larger economic and social context in which the program exists. (36)

Proponents of the participatory approach agree that there is no definitive methodology for conducting an evaluative research study of this nature. The basic element of this approach is observation and description. Qualitative rather than quantitative information is collected and analyzed to a greater extent. What ensues as a result of the active participation of the community in the research process is a more accurate and authentic analysis of the program and its functional capabilities, strengths and weaknesses within the context of the prevailing social and economic reality. Referring back to Smith's discussion of the level of certainty required to ensure that the research results are utilized, the participatory research approach provides more than suggestive or preponderant evidence. The first-hand experiences of individuals, combined into the whole of the community, provide a real and holistic analysis of the performance and impact of program activities, equal in validity, sensitivity and significance to quantitative data.

The active participation of field workers and the community in research on primary health care is advocated by the World Health Organization. (37) In fact, most of the five case studies used village-level health workers to gather
information. While the data collected in some instances was valid and relevant, in other cases it was not. As was pointed out in the previous chapter, the factors responsible for limiting the validity and relevance of the data collected by the community health workers were related primarily to inadequate training and supervision, and an expectation that they would be capable of completing large numbers of forms in addition to all the other tasks they were expected to perform. However, these tend to be logistical problems, largely unrelated to the ability of the community health worker, or other community members for that matter, to collect data for the program. The problem is to ensure that the techniques used to collect the data are within the abilities of those engaged to carry out this task. This entails designing surveys and other data collection techniques which are understood and relevant to those who will be completing them. Additionally, a realistic and practical interview/observation schedule should be designed, taking into consideration the daily workload of those responsible for collecting the data. But they are not confined to collecting the data. They also collaborate in their compilation and analysis. One study found that a vital part of the evaluatory process was the discussion of the findings of the data collection exercise among the village health workers and the villagers responsible for collecting the data. The potential feedback from such a discussion has important
implications for the future success and impact of the program. One of the community health workers involved in the project observed that

We did not want an evaluation that we would not understand and that would not have helped us to understand our problems - like just answering questionnaires. Our evaluation is very important because in all the years we have worked we have never so clearly seen the value of our work. (38)

Allowing the community to become involved in the data collection, analysis and discussion activities adds another dimension to the evaluation, not only in logistic terms, but more importantly, with respect to the comprehensiveness of the findings of the entire exercise. Qualitative information on attitudes, perceptions, relationships, fears, motivations, communication barriers, reasons for priorities in carrying out program activities are critical in understanding the factors which affect the performance and impact of the program. (39)

The active participation of community members and field staff in the evaluation can provide detailed information as to 'why' they react to the availability of program services. Through this approach, the evaluator(s) actively seek to understand user and provider concerns. In this way, program goals and activities can be modified to adapt to changing attitudes and context in which the program exists, thereby making the program more relevant and significant to the needs of the intended beneficiary population.
Summary

The results of an evaluation are location-specific, and cannot necessarily be generalized. Attempting to apply any one evaluative research design or method universally is impractical. Such an approach will fail to identify important local factors which affect the performance and impact of the program. Consequently, the validity and relevance of the results of the evaluation will be questionable.

What is required is a more flexible approach to conducting evaluations in field settings. As this chapter has shown, quasi-experimental designs offer a practical means of overcoming many of the factors which limit the validity of the more experimental evaluative research designs in a field setting. Additionally, the use of measures other than morbidity and mortality may provide the researcher with a more credible basis upon which to base research findings. The use of research survey techniques, such as participant observation, as alternatives and/or adjuncts to the more conventional approaches should allow researchers to collect a wider range of data.

The use of qualitative data as an adjunct to quantitative data should also serve to strengthen the validity and relevance of the evaluation results. Lastly, consideration should be
given to seek active community and service provider participation in the evaluation exercise. In this way, important information which might otherwise have been overlooked, particularly with respect to their perceptions of the health services and its potential for benefiting the community. What the chapter proposes is a more holistic approach to evaluative research, one which should serve to increase the credibility of the evaluation exercise, but which will also be sensitive to those factors which tend to adversely affect the capacity to conduct field-based research.
REFERENCES


7. Ibid., p. 274.

8. Ibid.


12. B. Hall, "Participatory Research: Breaking the Academic Monopoly," p. 51 (Mimeographed.)


17. Campbell and Stanley, Designs for Research, p. 34.

18. Cook and Campbell, Quasi-Experimentation, p. 6.


31. This description is based on a similar presentation provided in Kurz, "Participant observation," pp. 94,95.


34. Ibid., p. 52.

35. Ibid., p. 54.


CHAPTER FOUR

CONCLUSION AND RECOMMENDATIONS

The primary health care approach is proffered as a more effective and lower-cost alternative to conventional health care delivery approaches for developing countries. Over the past two decades, millions of dollars have been invested in the development and implementation of primary health care projects throughout the Third World. The common assumption of these projects is that the availability and utilization of the primary health care services and personnel will necessarily result in a significant improvement in the health status of the recipient population.

However, to date there is little evidence to support this hypothesis. For the most part, few projects have been evaluated. Of those which have been, the validity of the results of the evaluation is debatable. Yet despite this lack of substantive evidence of a significant impact on health, donor and host national governments and foreign aid organizations continue to invest heavily in this strategy. This thesis has identified several factors which tend to limit the credibility, validity and utility of evaluations of primary health care projects in the Third World. These factors can be grouped into two categories: evaluative research designs; and evaluatory methods and techniques.
In some cases, the evaluative research designs are simply too complex and rigorous for field situations and for the capabilities of those responsible for carrying out the evaluations. Many of the criticisms levelled at the true experimental research designs relate to the practical problems which arise when applying them to a field setting. Many of the fundamental prerequisites can rarely be met: the subjects cannot be randomly assigned to study and control groups; equivalence of the experimental and control groups cannot be guaranteed; intervention and outcome commonly have only tenuous links; many programs simultaneously employ multiple interventions; and lastly, key sources of variance cannot be controlled. Additionally, the expense and manpower qualification required to implement well-monitored and tightly supervised experiment-based evaluations are often beyond the capabilities of the vast majority of those involved in implementing primary health care projects. Such designs can also be politically unacceptable, especially when control areas where services are not to be introduced are contiguous or nearby the experimental communities.

In other cases, the designs are too weak. The absence of historical data and comparative populations limits the ability of evaluation reports to claim that the intervention has had any significant positive effect on health in the study population. In order to enhance the credibility of the
evaluative research design, additional observation points and comparative populations should be considered.

For the most part, though, the limiting factors were related to practical methodological issues, such as data collection instruments and techniques. In many cases, quantitative data were of an inferior quality due to a lack of training of personnel. In other cases, the type of data collected before and after the intervention varied, thereby making comparative analyses impossible. The amount of data collected was also found to be a contributing factor. In some cases, those responsible for delivering health care services were also expected to complete long and complex questionnaires and reports. Related to this was the format and exigencies of the questionnaire or data collection instrument. The more numerous, complex, or difficult to administer and interpret, the higher the probability of errors in data reporting and compilation.

An important element which contributed to limiting the validity and utility of the evaluation studies is the type of indices used to measure program performance and impact. In many cases, the indices chosen were unsuitable for the purpose of measuring program impact. Much of the data related to the quantity and nature of services provided, the number of health personnel trained and deployed, and in some cases, consumer perceptions about the services and those delivering the
health care services. While such indices provide useful information concerning program output, they provide no indication of how the services or personnel affected the health status of the recipient community.

Another concern is the reliance on morbidity and mortality data as outcome measures. Often, the data collected pertains to a subgroup of the study population. Yet generalizations concerning the success of the project's PHC approach are made for the entire population, based on this narrowly defined measure. Additionally, health is not affected solely through health-related activities. By restricting the scope of measures used, and the range of activities which may affect health, the evaluation becomes limited in its approximation of real interrelationships.

Another factor found to influence the credibility and utility of the evaluations is the dependence on quantitative data, almost to the exclusion of qualitative data. Part of this reluctance to collect and analyze qualitative data may be due to the extra expense and manpower requirements involved, as well as a perception that such data are generally of an inferior quality to quantitative data. While qualitative data may provide some corroborative support to the findings of quantitative data, it is widely held that the use of qualitative data alone is unscientific, and therefore insufficient. However, qualitative data can provide unique
insights into the reasons why people choose to (or choose not to) use primary health care services. They can also provide researchers with important information on the nature and reasons for unanticipated side effects of the program, and can assist in identifying factors which impede the effectiveness and efficiency of the program.

Perhaps one of the more important limiting factors is the preoccupation with demonstrating conclusive evidence of a link between the availability/utilization of primary health care services and improvements in the health status of the recipient community. In that evaluations are used primarily for decision-making purposes, the present study suggests that less than conclusive evidence may be appropriate in such cases. This is especially true when one considers that to demonstrate an improvement in health requires a long time period and before/after differentials of significant magnitude. However, program managers require informative feedback quickly. What may be required, and may be a more appropriate approach given these requirements, is a trade-off between the need for a high degree of proof of program effectiveness and the need for a long evaluative research time period. Demonstrating a strong association between the dependent and independent variables may be sufficient if an equally important consideration is to decrease the time interval between the commencement of project activities and the point in time when the evaluation of its
impact and performance takes place.

Based on this analysis, alternative approaches for evaluating primary health care programs which addresses many of the identified limiting factor are proposed. The thesis suggests the use of quasi-experimental evaluative research designs may serve to strengthen the capacity of the evaluation to demonstrate a strong relationship between the availability and utilization of primary health care services and a change in health of the study population. While these designs do require careful monitoring, they are more adaptable to field settings. They are not as complex as the true experimental designs. Nor do their underlying assumptions restrict their flexibility. A review of the evaluative research designs used in the case studies examined illustrated that slight modifications to these designs would have resulted in acceptable quasi-experimental designs.

As a means of increasing the credibility and utility of the research results, it is proposed that qualitative data be used as an adjunct to quantitative data. To this end, a case study approach is recommended. By examining particular population groups or PHC activities, researchers may be able to identify factors which either increase or decrease the capacity of the PHC approach to affect health. Qualitative information about various aspects of the program from service providers, consumers and non-consumers can identify both the strong and
weak points of the program. Such information is not restricted to questions concerning the availability and utilization of services and health care workers. Rather, such data can provide important insight into the perceptions, attitudes and behaviour of these groups. The researchers, therefore, not only come to know the 'when', 'where' and 'how much' of the program, but also the 'why'.

Conventional data collection techniques also serve to restrict the value of the data. Structured direct questionnaires which are usually constructed externally to the study area and population may overlook important variables which affect program performance and potential impact. Alternative data collection methods, such as participant observation, are suggested as a means of increasing the importance of the data to be collected.

Another approach which could strengthen the potential for program success is the active involvement of service providers and community members in the evaluation process. All too often, providers and beneficiaries are perceived as sources of information, rather than elements critical to the program's success. Researchers, either knowingly or unknowingly, tend to relegate their concerns to a negligible position. They are not viewed as collaborators in the project, but as subjects. Consequently, they often do not believe that they have any personal stake in the success of the program. For the
community, this can only serve to affect their attitudes and perceptions of the program, and ultimately their behaviour and utilization of the services. Health workers view their role as a job, believing that they have no input to contribute concerning the future development of the program. This is unfortunate, since they can usually identify factors responsible for limiting the effectiveness and efficiency of a program. By integrating these groups into the evaluation, they will be provided with feedback. This should increase their understanding of the nature and dynamics of their problems, and in some cases, the means to resolve them, or at least to minimize their effects.

Based on the findings of this study, the following recommendations are presented. They are designed to ensure that evaluative research for primary health care projects, and for social development programs in general, in developing countries is credible and useful. First, evaluative research in the Third World must be action-oriented. In other words, it must be a process for clarifying and responding to actual problems of the rural and peri-urban poor. This can be achieved by ensuring that it is a bottom-up approach, actively incorporating the entire program community in the evaluation process. In this way, the potential of success for the program will be increased, since community members will be involved in the development of future program decisions.
Second, the design and methodologies used must be appropriate to the situation at hand. This pertains not only to prevailing environmental factors, but also the needs of the audiences for the evaluative results. Development projects, including primary health care programs, do not exist in a vacuum. They affect, and are affected by, a wide range of ecological and socio-cultural factors. Research designs should be flexible to allow researchers the opportunity to use a variety of methodologies which may assist them to identify the nature of these factors and to measure their impact on health and well-being. As was stated previously, no one methodology is appropriate for every situation; but neither does every situation require the use of every methodology available.

Third, it is evident that there is a lack of developing country people trained in evaluative research methods. Few have the skills to design or manage an evaluation, or to train others in interview techniques. What is required are short-term, but intensive, training programs in a variety of evaluative research approaches. By building the capacity of Third World researchers to conduct evaluations of social service programs, the effect of cultural insensitivity which often characterizes evaluations conducted by expatriate researchers should be significantly reduced.

Finally, more funds should be made available by donor agencies to field-test alternative evaluative research
methodologies, such as that proposed in this study. Donor agencies must become more aware of the need to evaluate health and social programs in terms of their real effect on the intended recipient population. Simply parachuting conventional evaluative research approaches into developing countries does not guarantee useful or valid evaluation results. Evaluations should be more than simple accounting exercises.

The goal of development is to improve the quality of life of people. Therefore, evaluations must be people-focused. Measuring changes in impersonal indices without a consideration and understanding of the relationship between the program and the people it is intended to benefit severely limits the significance and validity of an evaluation. In the end, the evaluation becomes an academic exercise instead of a practical, problem-solving exercise. Additional research is required to identify other alternative approaches which may be appropriate for conducting field-based evaluative research in developing countries to produce valid and meaningful information.
SELECTED BIBLIOGRAPHY


Canada, Canadian International Development Agency (CIDA). "Health: CIDA's Involvement." Ottawa, April 1980. (Mimeographed.)


Echeverri, O., and Saravia, J. Sistema de servicios de salud rural como un componente basico de los proyectos de desarrollo rural en Colombia [Rural health services system as a basic component of rural development projects in Colombia]. Cali, Colombia: Universidad del Valle, Centro de Investigaciones Multidisciplinarias en Desarrollo Rural (CIMDER), 1974.


Fendall, N.R.E. "Primary Medical Care in Developing Countries." Internation Journal of Health Services 2(2) (1972):297-315.


Hall, B. "Participatory Research: Breaking the Academic Monopoly." Toronto, n.d. (Mimeographed.)


King, M. Medical Care in Developing Countries. London: Oxford University Press, 1966.


Malloquin, R.H.; Rolon, P.A.; and Kunzle, C. *Informe Final y Conclusiones del Programa Extension de los Servicios Salud a través de Maestros Rurales* [Final report and conclusions concerning the extension of health services through rural school teachers]. Asuncion, Paraguay: Ministerio de Sanidad y Asistencia Social, 1980.


"Planners' approaches to community participation in health programmes: Theory and reality." *Contact* 75 (October 1983): 3-16.


Tungsubutra, K. "Volunteer Health Workers: A Research Project carried out by the Faculty of Medicine, Khon Kaen University," Khon Kaen, Thailand: Khon Kaen University, 1979.


Williams, W. "Implementation and Analysis and Assessment." 


"Rural Basic Health Services: The Lardin Gabas Way."


