Policy & Research Unit - PHN

2.3 - Health Proj./Program Evaluation

WORLD HEALTH ORGANIZATION ORGANISATION MONDIALE DE LA SANTE

Téléphone Central/Exchange: 91 21 11
Direct: 91 34 67

In reply please refer to: PDP
Prière de rappeler la référence:

Dr Gerald Warford
The World Bank
1818 H. Street, N.W.
Washington, D.C., 20433

12 August 1981

Dear Gerry,

It was a pleasure to meet you at the World Bank and have a chance to discuss some of our ideas. I am leaving for London on the evening of 19 August 1981 and thus will be missing you in Geneva, but possibly you might be arriving a day earlier than planned. My contact in London will be Mr Peter Smith who is at the London School of Hygiene and Tropical Medicine. Pat Rosenfield is away until the day I leave, but I will try and tell her of your visit; please do contact her.

I shall be discussing with Dr Lucas my visit to the Bank and some of the ideas we talked about regarding further interaction between the Bank and the Special Programme concerning approaches to evaluating benefits from health programmes.

I will be seeing David Bradley also in London and will discuss these issues with him as well.

When are you likely to get to Geneva again?

Yours sincerely,

Dick

R. H. Morrow, Jnr., M.D.
Secretary, Scientific Working Group
on Epidemiology

DRAFT
KLHall:mlo
April 15, 1981

Selected Issues in Health Development

I. Determining Real Demand for Health Care

A. Defining Health Status

i. Data

1. To what extent, and with how much confidence can we use existing health and related info on incidence and prevalence of disease to ascertain priority health problems?

2. Do typically incomplete, hospital and outpatient clinic-based data hamper assessment of overall health status? Is such data biased toward the curative side?

3. From the preventive perspective, how will need be defined? What risk indicators should be used for identifying vulnerable population groups?

4. Does epidemiological data overstate the real demand for health care, e.g. system will have to respond, in effect, to only a small proportion of total illness as expressed in biological need terms?

5. Do adequate proxies exist to determine key health problems when general health data is inadequate? What techniques exist for rapid appraisal?

6. To what extent are surveys desirable? Is sampling sufficient/sensitive to determine major problems? When are baselines needed?

7. What are tradeoffs between increased specificity of need through community diagnoses and costs (financial and start-up time?)

ii. Analysis

1. How will we set priorities among numerous health problems - which criteria can be used to rank health problems - prevalence, severity, technological, financial and administrative feasibility of redressing?

2. Can distinct levels of health development be defined both between and within countries, based on key indicators and/or patterns of diseases?

3. How do we quantify demand for health care from epidemiological need data?

4. How far can one go in assessing underlying causes of major health problems without detailed epidemiological studies? How are determinants identified to select appropriate interventions (especially important given diverse causes for specific diseases, e.g., TB environmentally or nutritionally related, diarrheas food-borne, water-borne, etc.)

5. Are there certain patterns of illness for which health/medical services are not the best response, e.g., TB?

II. Setting Objectives/Targets

A. Objectives of Health Investments

- What is the principal objective(s) of health sector support?

- Key interrelated objectives of health investments include:

  1. improve health status; specifically to prevent premature deaths and reduce morbidity

  2. meet basic needs/enhance equity

  3. raise productivity/human capital investment

- Which will have priority? Are they consistent; are they inseparable?

- What effect does the choice have on target population? On interventions? On key health status indicators IMR, CDR, life expectancy?

- How far can we go in documenting worker productivity improvements from health interventions? Other forms of productivity, beyond worker - e.g., educational attainment, women's ability to cultivate subsistence plots, obtain sufficient amount of water, cook, including fuel collection.

B. Targets for Health Status Improvement

- What are reasonable targets to set for reduction in rates of key health indicators - (IMR, MMR, CBR) over typical 4 year project life? What has been the experience from developed/developing world to date?

- How will project targets relate to overall sectoral targets?

- To extent that Bank project may often be viewed as "partial equilibrium" model, how will we set realistic performance targets e.g., what can project alone achieve? Or is it so intimately linked to other inputs that it can't be assessed in isolation?

III. Project/Program Design

A. Identifying Appropriate Interventions/Services

- What alternative technologies exist to redress the predominant health problems, especially in under fives in developing countries today?

Typical profile:

for rural and urban poor :
- low birth weight and malnutrition
- neonatal tetanus
- diarrheas
- respiratory infections
- measles and other communicable diseases
- malaria and other endemic diseases

for urban:
- hypertension
- cancers
- respiratory, TB

- By what criteria can the most effective approach within a given country/area setting be identified? How do we optimize our choices?

- What are the specific tasks which should be carried out at each level of care to redress this common health profile?

- Are there logical, cost-minimizing packages of services/interventions which respond to the key problems identified? Is there a minimum set of services which must be instituted concurrently to have a positive health impact?

- What are the tradeoffs between integrated vs. vertical, single-purpose program designs for addressing key problems? Are there certain diseases or categories of health problems that respond most favorably (in terms of impact and time frame) to vertical approaches? horizontal approaches?

- How do fixed site/service vs. outreach/mobile team approaches affect desired outcomes--mortality and morbidity reduction/prevention objectives? What is the effect on utilization of services of these alternative delivery modes?

- What methodologies exist to determine actual number and type of interventions that a given population will need from basic health statistics? Can techniques used by health maintenance organizations/group health insurance schemes be adapted for developing country use?

B. General Project Planning Issues

i. Critical Path Analysis of Inputs

1. Is there optimum, logical phasing or sequencing of inputs to maximize sector development?

   a. does it make sense to concentrate on supply/logistics, service delivery concurrent with manpower training, facility construction/renovation or should they follow?

   b. where do planning, administration/management development fit into the sequence?

   c. is health education most effective before, concurrent with (demonstration effect) provision of services which people should use?

2. Is typical Bank project life of 4-5 years an optimum time frame for health sector development - if not what is a realistic period of time?

3. Can a rolling project design be adopted (phasing) acceptable to Bank standards?

4. Is there an optimum number of discrete activities/components that should be incorporated in design, beyond which efforts are too dissipated to have desired impact?

ii. Targeting Inputs

1. Is there a threshold of health/nutritional status below which health service interventions are necessary but not sufficient to improve health status? Do methodologies exist to identify this threshold?

2. In targeting should greatest weight be given to "need" as epidemiologically defined or are other factors of greater importance - e.g., infrastructure capacity?

3. Do we pick the "winners" (regions/areas with relatively better infrastructure for their greater potential for project success and replication/demonstration effect) or the "losers" to ensure that the model is viable under the worst conditions?

4. What are the tradeoffs in overall program effectiveness between provision of a minimal set of services to the total population vs. intensive inputs to highest risk groups?

C. Health System/Infrastructure

i. Defining the System

1. For sector analysis and project design purposes how will we define the health sector?

2. What are the elements of a health delivery system and how do they interrelate?

3. Is there an optimum model for evolution of a health system? What are the characteristics of each stage? e.g., resource distribution by level of care. Can we define an optimal balance between primary, secondary and tertiary care levels at a given stage of health status and disease pattern? Is the pyramid concept an appropriate one for most developing countries?

4. How does the balance of public and private sectoral resources affect system design, e.g., determine technology choice, skill mix, functions of personnel?

5. What steps need to be taken to ensure system design is dynamic one, responsive and adaptable to changing consumer expectations and health problems (epi transition) over 5-10 years?

6. Where is the Bank's comparative advantage in the system? Should it fill gaps critical to total system operations? In a system-oriented approach, does it make sense to focus on the secondary care level if the primary care level is not in place? Should priority be given to building the base of the pyramid or are there conditions which support a top-down approach?

ii. Assessing the Process

1. By what criteria/objective measures can national health policies be assessed? How can the Bank rapidly assess the impact of existing and proposed legislation on the sector?

2. What measures exist for assessing the efficiency with which a health system operates?

3. How do we assess administrative/managerial capability? Do objective criteria exist to measure infrastructure capacity and monitor project contribution to strengthening such infrastructure?

4. How will we determine planning capability? e.g., existence/ nature of plan, trained planners

5. What are the key ingredients of "institution-building"and how will performance be measured?

D. Resources/Inputs

i. Manpower

1. General

a. What is true cost of use in service delivery of alternative medical and paraprofessional cadres, not just in initial training and salaries but support expenses and additional costs these staff generate over time? e.g., considerable literature on substantial indirect, secondary costs of MDs in US and Canada.

b. To what extent will supply of manpower (coupled with other inputs such as drugs) in previously under - or unserved areas induce demand for health care above levels to which government can respond structurally and financially? Is it desirable to control utilization and what mechanisms exist to preclude overutilization? Can we estimate with confidence the impact on utilization of various personnel inputs?

c. How will optimal staffing patterns for health services responsive to specific health profiles be determined?

d. What types of incentives/disincentives should be incorporated in project design - how is "status" determined and how can projects be sensitive to reward system?

2. Physicians

a. Are present nature and magnitude of medical training cost-effective? What are the distributional effects of current policies and programs?

b. What is the actual expense incurred by the system for placing domestic-trained physicians in rural service? e.g., number of MDs trained that don't go to rural areas and are not needed in urban areas? Can we depend on emerging urban market saturation to effect a more equitable distribution of MDs in the developing countries - is there evidence that market forces actually work in area of manpower location/distribution?

3. Other Staff

a. What are implications of use of minimally trained staff e.g., illiterate CHWs*, over time - are there inherent limitations to hiring vs. creating temporary posts? Can their skills be sufficiently upgraded as health needs change?

b. Is considerable resistance to mobilization of paraprofessional (pp) cadres in many countries, especially strict circumscribing of roles, justifiable? What has been experience with use of pp in health services? performance?

c. What are key causes of high attrition rates (students and personnel) and appropriate incentives to attract and retain qualified staff for health services?

ii. Facilities

1. Are catchment/coverage areas an appropriate approach to health facilities planning? How can realistic coverage areas be defined?

2. By what criteria are facility site location decisions maximized?

3. How do we factor present and projected demand for services into facility design?

4. How do we define productive capacity/source loads to which facilities must respond? What is the logical/optimal balance of facility types through the various levels of care in a health system? Do data exist from a functioning system which would assist in facility planning?

5. What is the role and appropriate proportion of hospitals for PHC oriented systems?

6. What are the comparative advantages of facility over outreach services?

7. Disbursements in construction tend to lag and impede meeting implementation schedules; what are principal constraints in facility construction/renovation activities? How should project designs be adjusted to redress these problems?

8. Assuming community participation is desirable to reduce financial needs for this major cost item, "self-help" clinics have often led to inefficient facility planning/location decisions - how can such inputs be more effectively and productively channeled?

*See Attachment A "Appropriate Skills/Trainings and Functions of CHWs," for detailed questions on CHWs.

9. What mix of facilities (proportions at each level of care) is desirable at various stages of health needs within a country?

iii. Logistics/Supplies/Equipment

Selection of inappropriate commodities/equipment and lack of preventive maintenance and repair capabilities (including budgetary support) adversely affect program/project implementation in many countries - what has worked in redressing these problems?

1. Given the high foreign exchange requirements to meet current sector needs, is domestic production of key commodities a viable alternative? Which ones?

2. With particular reference to vehicles, what lessons for project design can be learned from project experience to date on such critical issues as controlling abuse/accountability, maintenance/repair, product selection criteria?

3. To the extent that training dictates in part health service supply needs, are curricula for key health personnel compatible with what equipment etc. should be used from an economic and appropriate technology standpoint? e.g., microscopes, photocopiers.

4. What is the effect of existing supplies/equipment on service utilization by level of care? What are key commodities essential to support typical primary care level, without which referrals to higher levels of care occur and represent inefficient use of human resources?

5. Can we determine the true economic cost of commodities to guide more appropriate selection e.g, price (foreign exchange) of spare parts?

6. What have been experiences with radio communication systems to link various levels of care and thus help build a true health system?

iv. Drugs

1. Examining the principal problems related to pharmaceuticals in health system, which are top priorities and how should projects respond?

2. What are the actual expenditures on drugs in developing countries and areas for potential cost-savings?

v. Economics and Financing

Amidst government commitments to Health for All, the implications of alternative approaches to financing this goal have assumed increasing importance. Key questions include:

1. Economics of Health

a. What are the principal cost elements in health programs and how well documented are they? Are there areas for potential cost-savings?

b. Examining the true economic costs (including the foreign exchange requirements) of primary health care model, as commonly conceived, is it affordable to the poorest countries? Are primary health care schemes truly "low cost" as promoted?

c. What are the cost- and cost-containment implications of prevalent components of PHC program designs:

- revolving drug funds
- "free" services

d. What pricing mechanisms have been used for health services? Can we derive the comparative costs of unit of given services delivered at various types of institutions/levels of care e.g., dispensary, health center, district /regional hospital, specialized hospitals.

e. What is the recurrent cost implication of expanding health delivery system/health sector coverage? Can we expect economies of scale as the number of users increases? Is there an optimum/equilibrium point?

f. Financial implications of alternative health technologies -- e.g., multipurpose vs. teams of specialized providers?

g. To what extent does a "second economy" exist within the health sector and what is its economic impact e.g., sales of pilfered drugs; physicians use of public health facilities for private patients "after hours"?

h. How does prevalent legislation affect costs? e.g., drug and contraceptive sales prohibitions or limitations of distribution to certain levels of health workers.

2. Sources of Funds

a. How and by which groups are health sector revenues raised? Who pays in the developing world?

b. - What innovative sources of financing have been or could be tapped to supplement public sector revenues and consumer direct payments?

- What are the benefits and drawbacks of alternative financing mechanisms with regard to such considerations as funding levels for services, equity? What impact do alternative financing systems have on the structure, functioning and efficiency of the health services delivered? What options are available to Governments to direct the flow of money to influence achievement of sector objectives e.g., balance between preventive and curative care?

c. Past infrastructural investments have committed many countries to a pattern of high and rising recurrent costs. To what extent can expenditures be shifted down in the health system to extend coverage to under-and unserved population groups, taking into account that closing of hospitals is not realistic in most settings? Will system expansion be limited to incremental revenues or are there potential cost-savings to reduce the existing recurrent cost burden?

d. Are cost-recovery (self-financing) and cost-containment objectives compatible? What is the impact on utilization, demand for supplies, etc. when the consumer is paying the full bill? e.g., Senegal village health committee experience. What guidance does experience provide on phasing in full cost-recovery based systems?

e. What progress has been made to date in cost-recovery of PHC programs within the public sector? What lessons from private sector e.g. missions? How are fee-paying schemes designed and the fee-processing administered? What mechanisms are critical to ensure accountability, maximize revenues generation?

f. What lessons can be gleaned from social security financed health systems? To what extent are allegations valid that they compete with the Ministry of Health system and support inequitable distribution of sector resources?

g. What is the potential for health insurance schemes in developing countries? Is there a threshold of income, health status, population below which they are not viable? When they have worked, what have been the critical conditions (economic, management, etc.) for their effective operation?

h. Does donor/external support of health programs/projects affect the use of resources: 1) are initial capital investments different? 2) do recurrent costs patterns vary from strictly domestically financed activities? 3) does the proportion of capital to operating budget vary?

i. How do tracking/reporting requirements vary between domestic and external finances and what influence, if any, has this on the mix of financing mechanisms used at the country level?

3. Allocations

a. How is revenue allocated within the sector? What kinds of information are used to guide sectoral allocation by level of care, by type of provider, by program area?

b. What budgeting processes have been used and how can they be strengthened to improve revenue requirements estimating and constituency building within the Ministry of Finance?

c. How do we determine the appropriate level of public resources to be allocated to health from general government revenues for a given country? What is the productive function for health vis-a-vis other development needs/priorities?

d. What is the impact of different methods of payment and reimbursement of various levels of health care providers on their geographic distribution?

4. Effective Demand

How do time and money costs related to health care affect utilization e.g., distance travelled, transport mode, queue lengths. What are the true economic costs of utilizing services 1) by alternative types of providers (paraprofessionals/physicians); 2) by demographic groups; 3) by payment methods, by socio-economic groups?

a. What is the actual level of aggregate health expenditures in the developing ocuntries? How much is spent out-of-pocket not only on providers per se but pharmacies, traditional practitioners? What is proportion of private to public expenditures?

vi. Health Education

1. Recognizing difficulty of changing human behavior, have health education efforts worked? Under what circumstances? In what time frame, on average?

2. A broad range of programmatic vehicles exist from one-to-one communication to mass media and telecommunications? What approaches appear most successful and under what conditions are they replicable in other settings, e.g. Chinese experience.

vii. Technical Assistance (TA)

1. How can we optimize contribution of TA/cooperation to achievement of institution building and training objectives?

2. Is TA compatible with objective? To what extent does TA displace/substitute for national inputs?

3. Do common TA procedures (remuneration/supervisory structure/counterparts) and issues (dual allegiance to institutions) affect project performance - how do you ensure TA is cost-effective? Is there an inherent conflict between consultants job-preservation and transfer of technology responsibilities?

E. i. Private Sector

1. How can we assess current and potential role in sector given paucity of information?

2. Is it possible to identify incentives to encourage emphasis on preventive health activities vs. curative orientation?

3. How can private industry involvement be expanded to augment resources available to sector?

4. Are there health activities for which private sector has comparative advantage, e.g., commercial distribution?

5. What is economic effect of common policies permitting government MDs to see private patients in public facilities "after duty hours?"

6. To what extent is HMO feasible delivery model in LDCs - is there a certain level of development necessary e.g., critical mass of people able to pay - to make this viable approach? e.g., Korea, Brazil. What are the successes/limitations?

7. How can successful designs developed in private sector be merged into government designs? Where are, or could be, the linkages e.g., missions role in National Health Councils?

ii. Community

1. What are true opportunity costs of volunteer inputs? Is model viable if workers actually had to be remunerated by system?

2. How can government/Bank projects determine consumer demands/expectations? What documentation of perceived needs exists?

3. How do KAPs modify epidemiological need? Can we focus on few families in which things appear to be "going right" as model for program design?

4. Which health activities can be delegated realistically to family household level - experience to date, e.g., weighing, ORT.

5. When is CP critical to success of project implementation - are there discrete components/activities which can be identified? e.g., maintenance.

iii. Traditional

1. Should national (largely government dominated) health systems displace, complement or integrate traditional health care activities? Does traditional system have comparative advantage in certain areas? Which parts? e.g., TBAs.

2. Are traditional approaches cost-effective?

F. Other Sectors

i. What has been the experience with using other sectors'workers to achieve health status input objectives? What innovative designs have been tested - e.g., agricultural extension, agents in family planning?

ii. Is Ministry of Health structurally suited for conducting/ coordinating multisectoral activities?

iii. Are multisectoral projects feasible within government structure? Bank structure?

iv. Can we determine from existing data an optimum sequence, order of multisectoral activities to maximize health impact?

IV. Outputs/Outcome/Impact

A. Intermediate Measures

i. Utilization

1. What are the principal determinants of utilization - to what extent can we project, with accuracy? Can we predict health status changes from utilization rates - e.g., immunization coverage.

2. How will varying service packages/financing mechanisms influence utilization with regard to volume, equity, cost considerations?

ii. Quality of Care

1. What objective criteria exist for assessing quality of care?

iii. Management/Information System

1. What are the data requirements to assess outcome and to what extent do they routinely exist within developing countries?

2. What are the principal data which should be collected within health projects? Are there successful models to follow?

3. To assess impact, we must know not only indicators but also when change can be expected and thus at what time measurement is desirable. Can we identify when outcomes (prevention or reduction in illness/death) should be measurable for common health interventions e.g., measles immunization?

B. Assessing Impact

i. What is physical effect of alternative interventions - what predictions can we make about reductions in morbidity and mortality from a given intervention?

ii. What mechanisms are needed for long-term evaluation and how can they be built into typical 3-4 year project design?

iii. For specific interventions/components, what are key intermediate variables that let us know we're heading in the right direction, even if outcome can't be measured?

iv. How do we differentiate between management information system and evaluation data requirements - and relate them?

KLHall:mlo
4/15/81

ATTACHMENT A

Appropriate Skills/Training and Functions of Community Health Workers

1. What are the minimum tasks primary care workers must perform at the community level to affect health status? Maximum feasible without overloading?

2. What is the appropriate balance between curative and preventive/promotive functions? How do you avoid creating another basically curative care level and yet maintain system's credibility with the community? Sustain community interest in promotive activities?

3. Are these tasks in which sex differentiation of worker is critical to successful implementation? e. g., must only women do deliveries? family planning, etc.?

4. How will primary care functions interface with traditional practitioners in community already? Does it make sense to recruit workers from these practitioners? How do you avoid competition/conflict?

5. How has financing mechanism affected PHC worker recruitment, performance, retention? How will workers be remunerated? Is volunteerism feasible/sustainable over the long-term? Will (not can) communities pick up the tab, at least in part?

6. What selection criteria have proven most important for such workers? Are there optimum educational levels, age groups, marital status for PC worker cadres?

7. What has experience been with attrition rates? Are levels of staff retention sufficiently high--can they be improved?

8. Are training programs used in other countries "transferable technology?" Can modules be imported or must they be completely indigenous to area? And if domestic internally designed ones have proven to be best, how are they developed? What institution(s) should be involved? Based on what? Can/must training be standardized within a country?

9. A quick review of several PHC programs indicates little if any correlation between tasks to be performed and length of training. In fact, workers performing similar functions have been trained anywhere between a few weeks to a few years. Is there a means by which an "adequate" training period can be defined? What experiences have countries had with too little basic training? too much? What modes of training have been tested? Results? What are common weaknesses of PHC worker basic training? strengths? as perceived by program management? by workers themselves? by communities served? What is appropriate theoretical/practical mix?

10. What are the trade-offs between long initial training and short basic and extensive continuing education? What are in-service training needs? Are there identified areas essential to continuing education? What is the trade-off, if any, between personalized/individualized and general in-service training?

11. What is the current consensus on optimum training site? If secondary care level is to be involved, has it been sufficiently prepared for this role?

12. What is the comparative advantage of public sector/Ministry of Health in PHC worker training, supervision? etc. Is there a role for private sector (including local nongovernmental organizations). Can governments train sufficient number of PC workers to meet population ratios and Health for All goals? Should they do it alone? Since the massive training requirements are for the short term only (until basic cadre developed), will public institutions be able to trim down once needs diminish or will large public training facilities generate a life of their own once created? And if so, what are the alternatives?

13. Lack of adequate supervision has been identified consistently as one of the weakest parts of PHC worker-based systems undertaken to date. What leassons can be drawn from past programs? Who are best supervisors? Why do supervisory functions tend to break down? e. g. transport, morale, lack of accountability? What are the essential ingredients of an adequate and working supervisory structure?

14. Must other training programs for other health/medical workers be adapted as well to operationalize the PC strategy? How? Are certain cadres instrumental for a "team approach"?

15. What types of evaluation have been/should be conducted to assess PC worker performance? How do such cadres compare in diagnosis/ treatment with other levels of health/medical workers? How do you measure effectiveness of polyvalent workers? versus unipurpose? Are there certain functions they have done exceptionally well? poorly?

16. Any additional factors which influence success of PC worker programs?

FORM NO. 75
(9–78)

THE WORLD BANK

ROUTING SLIP	DATE: January 23, 1981

NAME	ROOM NO.
Mr. Clifford Owen, SEC	A-1241

	APPROPRIATE DISPOSITION		NOTE AND RETURN
	APPROVAL		NOTE AND SEND ON
	CLEARANCE		PER OUR CONVERSATION
	COMMENT	X	PER YOUR REQUEST
	FOR ACTION		PREPARE REPLY
	INFORMATION		RECOMMENDATION
	INITIAL		SIGNATURE
	NOTE AND FILE		URGENT

REMARKS:

The attached listing, while not exhaustive, is a good starting point for the University of Bradford to identify documents related to and professionals involved in primary health care program evaluation. If I can be of additional assistance, please call.

FROM: Karen Hall Karen L. Hall	ROOM NO.: N-422	EXTENSION: 61557

I. Evaluation Documents

Evaluating the Impact of Nutrition and Health Programs. Robert E. Klein, S. Read, et. al. New York: Plenum Press, 1979.

Demystifying Evaluation. Noreen Clark and James McCaffery, New York; World Education, 1414 Sixth Avenue, New York, NY 10019, 1979 (African field work orientation)

Council. Food and Agriculture Organization of the United Nations. Seventy-Fifth Session. Rome 11-22, June 1979. UN Joint Inspection Unit Report on Glossary of Evaluation Terms (JIU/REP/78/5) CL 75/6 February 1979.

For UNICEF documents on health program evaluation, contact:
Mr. Howard Dale
Assistant Librarian, Planning, Programming and Evaluation Section
Room A-6507
UNICEF
866 United Nations Plaza
New York, NY 10017

B. Abel-Smith, Value for Money in Health Services; a comparative study.Heinemann, London, 1976.

J.A. Burdette, et al, Primary Care Evaluation. The AAFP - UNC Collaborative Study. JAMA, vol. 230, No. 12, 1974.

A.L. Cochrane, Effectiveness and Efficiency.Nuffield Provincial Hospitals Trust, London, 1972.

B. Cvjetanovic and B. Grab, Rough Determination of the Cost Benefit Balance Point of Sanitation Programmes. Bulletin of WHO, Vol. 54, Part 2, pp.207-215, 1976.

A. Donabedian, Evaluating the Quality of Health Care in Programme Evaluation in the Health Field, SCHULBERG, H.C. et al (eds.) Behavioral Publications, New York, 1969.

J.M. Last, Evaluation of Medical Care. Med. J. of Australia, November, pp.76-785, 1965.

S. Litsios, Developing a Cost and Outcome Evaluation System. Int. J. of Health Services, Vol.6, No. 2, 1976.

I. McDowell and C.J.M. Martini, Problems and New Directions in the Evaluation of Primary Care, Int. J. of Epidem, Vol. 5, pp.247-250, 1976.

J.D. Pole, The Use of Outcome Measures in Health Service Planning. Int. J. of Epidem, Vol. 2, No.1, 1973.

P.S.S. Rao, et al, Methods of Evaluating Health Centres. Brit. J. Prev. Soc. Med., No. 26, pp.46-52, 1972.

M.I. Roemer, Evaluation of Community Health Centre. Public Health Paper, No. 48. WHO, Geneva, 1972.

S. Shapiro, End Result Measurement of Quality of Medical Care. Millbank Memorail Fund Quarterly, Vol. XLV, No. 2, 1967.

A.K. Sprinivas Murthy and R.L. Parker, New Methods for Assessing Health Care Delivery Systems. Proceedings of 12th Annual Conference of Indian Assoc. for the Advancement of Med. Educ., January 12-14. Gujrat. Unpublished, 1973.

M.D. Warren, Process and Methods of Evaluation of Public Health Programmes. Proceedings of WHO Conference. Copenhagen. Unpublished, 1967.

K.L. White, Evaluation of Health Care: How can Nations Cope? Canadian J. of Public Health, Vol. 67, Part 5, pp. 391-396, 1976.

J.K. Wing, Principles of Evaluation. In WING, J.K. and HAFNER, H. (Eds.) Roots of Evaluation, Oxford Univ. Press, 1972.

For copies of U.S. Agency for International Development's Impact Analysis series, with particular relevance to Morocco (nutrition) No. 8 and Senegal (health) No. 9, write directly to: Editor, ARDA/DSDIUDI, Rm. 813, SA-18, U.S. AID, Washington, D.C. 20523.

II. Professionals Involved in Health Evaluation in Developing Countries

Dr. Susan Cole-King (with Institute of Development Studies, Sussex, England)
c/o E. Tarimo
Division of Strengthening Health Services
World Health Organization
1211 Geneva 27
Switzerland

Dr. Peter Knebel -- U.S. Agency for International Development
AID Regional Advisor - Sahel
c/o United States Embassy
Bamako BP 34, Mali

Dr. John Carrier
Department of Social Administration
London School of Economics

Dr. Clive Gray
Harvard Institute of International Development
1737 Cambridge St., Room 618
Cambridge, Massachusetts 02138
Tel. No.: (617) 495-3748

REFERENCES

1. ABEL-SMITH, B (1976) Value for Money in Health Services; a comparative study. Heinemann, London.

2. ACHESON, H.W.K. (1975) Medical Audit and General Practice. Lancet, March 1, pp 511-513

3. ALDERSON, M.R. (1974) Evaluation of Health Information Systems Br. Med. Bull., Vol.30, No.3

4. ALDERSON, M.R. (1973) Objectives and Concept of Health Information Systems
Proceedings of WHO Conference on Health Information Systems, Copenhagen, 18-22 June WHO Geneva.

5. ALEXANDER, C.A. et al Cost Accounting of Health Centre Expenditure Ind.J. Med.Res., vol 60, No.12.

6. ASHFORD, J.R. (1975) How Can quantitative Methods help the Health Services Manager?
In McLachlan, G. (Ed)
Measuring for Management. Quantitative Methods in Health Services Management.
Nuffield Provincial Hospitals Trust.

7. ASHFORD, J.R. and RILEY, V.C. 1975. An Approach to Monitoring the Quality of Health Care.
In McLachlan, G. (Ed.)
Measuring for Management. Quantitative Methods in Health Services Management.
Nuffield Provincial Hospitals Trust.

8. BABSON, J.H. (1973) Disease Costing Manchester Univ. Press.

9. BABSON, J.H. (1971) Hospital Costing in Great Britain
The Hospital, 67, 4, pp 106-111.

10. BARKER, K. (1970) et al (Eds). Health in the Developing World
Cornell Univ.Press, Ithaca, New York.

11. BIOLOGICAL SCIENCES COMMUNICATION PROJECT (1973) Delivery of Health Care in less developed countries with exphasis on integration of family planning with mother and child health. George Washington Univ.
Washington D.C.

12. BLUM, H.L. (1974) Planning for Health. Development and Application of Social Change Theory.
Human Sciences Press. New York.

13. BRADSHAW, J. (1972) A Taxonomy of Social Need. In Problems and Progress in Medical Care, No. 7
Nuffield Provincial Hospitals Trust.

14. BRYANT, J. (1975) Health and the Developing World
Cornell Univ. Press.

15. BURDETTE, J.A. et al (1974) Primary Care Evaluation. The AAFP - UNC Collaborative Study.
JAMA, vol. 230, No.12.

16. COCHRANE, A.L. (1972) Effectiveness and Efficiency Nuffield Provincial Hospitals Trust, London.

17. CVJETANOVIC, B and GEAB, B. (1976)
Rough Determination of the Cost Benefit Balance Point of Sanitation Programmes. Bulletin of W.H.O. Vol. 54, Part 2, pp 207-215.

18. DJUKANOVIC V and MACH, E.P. (Eds) (1975)
Alternative approaches to meeting basic health needs in developing countries. Joint UNICEF/WHO Study. WHO, Geneva.

19. DONABEDIAN, A (1969)
Evaluating the Quality of Health Care in Programme Evaluation in the Health Field, SCHULBERG, H.C. et al (Eds.) Behavioural Publications, New York

20. FELDSTEIN, M.S., PIOT, M.A. and SUNDARESAN, T.K. (1973)
Resource Allocation Model for Public Health Planning. A Case Study of Tuberculosis Control Supplement to Bulletin of W.H.O. Vol.48. WHO. Geneva.

21. FORD FOUNDATION (1976)
The development of health services in Bangladesh Ford Foundation, New York

22. GRIFFITHS, A. (1976)
Health Planning, Management and Training Needs in Underdeveloped Countries Unpublished.

23. GUPTA, M.C. et al (1977)
Effect of Periodic Deworming on Nutritional Status of Ascaris-Infested Preschool Children Receiving Supplementary The Lancet, July 16, pp. 108-110.

24. HARRISON, P. (1977)
Basic Health Delivery in the Third World New Scientist, Feb.17. pp. 41-43

25. IDRISS, A.A. et al (1976)
Sudan; National health programme and primary health care, 1977/78 - 83/84. Bull. World Health Organ, vol.53, part 4, pp 461-471.

26. LAST, J.M. (1965)
Evaluation of Medical Care. Med. J. of Australia, November, pp 781-785.

27. LITSIOS, S. (1976)
Developing a Cost and Outcome Evaluation System.
Int.J. of Health Services, Vol 6, No.2

28. LOGAN, R.F.L. et al (1972)
Dynamics of Medical Care. Memoir No. 14. LSHTM.

29. MASON, A.M.S. et al (1973)
Disease Costing in Hospitals. A review of completed work. Unpublished.

30. MORLEY, D. (1974)
National Nutritional Planning BMJ, October 12, pp 85-88.

31. MORLEY, D (1975)
Paediatric Priorities in the Developing World. Butterworths. London.

32. McDOWELL, I. and MARTINI, C.J.M.(1976)
Problems and New Directions in the Evaluation of Primary Care
Int. J. Of Epidem., Vol.5, No.3 pp.247-250

33. McLACHLAN, G. (Ed.)(1975) — A Question of Quality. Roads to Quality Assurance in Medical Care. Oxford Univ. Press. for Nuffield Provincial Hospitals Trust.

34. NEWELL, K.W. (Ed) (1975) — Health by the People WHO, Geneva.

35. PARKER, R.L. et al (1972) — Relating Health Services to Community Health Needs. Ind. J. of Med. Res., Vol.60, No.12.

36. PIACHAUD, D & WEDDELL, J.M. (1972) — Cost of Treating Varicose Veins. Lancet 2, pp 1191-1193, Dec.2.

37. PIACHAUD, D & WEDDELL, J.M. (1972) — The Economics of Treating Varicose Veins. Int.J. Of Epidem., 1,3, pp 287-294.

38. POLE, J.D. (1973) — The Use of Outcome Measures in Health Service Planning. Int. J. of Epidem., Vol. 2, No.1

39. POPOV, G.A. (1971) — Principles of health planning in the U.S.S.R. Pub.Health Paper No.43. WHO, Geneva.

40. RAMALINGASWAMI, P and RAMALINGASWAMI, V. (1973) — In Health Service Prospects; an International Survey. Ed. by DOUGLAS-WILSON, I. and McLACHLAN, G. The Lancet and Nuffield Provincial Hospitals Trust.

41. RAO, P.S.S. et al (1972) — Methods of Evaluating Health Centres. Brit. J. Prev.Soc.Med., No. 26, pp 46-52.

42. REINKE, W.A. et al (1974) — Functional Analysis of Health Needs and Services. Johns. Hopkins Univ., Baltimore. Unpublished.

43. REINKE, W.A. (1972)(Ed.) — Health Planning: Qualitative aspects and quantitative techniques. Johns Hopkins Univ., Baltimore.

44. ROEMER, M.I. (1972) — Evaluation of Community Health Centre. Public Health Paper, No. 48. WHO, Geneva.

45. ROEMER, M.I. (1976) — Rural Health Care. The C.V. Mosby Company, St.Louis.

46. SANAZARO, P.J. (1974) — Medical Audit. BMJ, February 16, pp 271-274.

47. SANJIVI, K.S. (1971) — Planning India's Health Orient Longman, Bombay.

48. SCOTTISH HOME AND HEALTH DEPARTMENT AND WHO JOINT TEAM (1974) — The Child Health Services. A Systematic Planning Approach. WHO, Geneva.

49. SHAPIRO, S. (1967) — End Result Measurement of Quality of Medical Care. Millbank Memorial Fund Quarterly, Vol. XLV, No.2.

50. SRINIVAS MURTHY, A.K. and PARKER, R.L. (1973)
New Methods for Assessing Health Care Delivery Systems.
Proceedings of 12th Annual Conference of Indian Assoc. for the Advancement of Med. Educ., January 12-14. Gujrat. Unpublished.

51. SUSSER, M.W. and WATSON, M. (1971)
Sociology in Medicine Oxford Univ. Press, London.

52. WARREN, M.D. (1967)
Process and Methods of Evaluation of Public Health Programmes.
Proceedings of WHO Conference. Copenhagen. Unpublished.

53. WHITE, K.L. (1976)
Evaluation of health care; how can nations cope?
Canadian J. of Public Health, vol. 67, part 5, pp 391-396.

54. WHO (1975)
Randomised Trials in Preventive Medicine and Health Service Research Report on a study group. Copenhagen, 8-12 December WHO Reg. Off. for Europe, Copenhagen.

55. WHO (1966)
Sampling Methods in Morbidity Surveys and Public Health Investigations.
WHO Technical Report Series No. 336. WHO, Geneva.

56. WING, J.K. (1972)
Principles of Evaluation.
In WING, J.K. and HAFNER, H. (Eds.) Roots of Evaluation Oxford Univ. Press.

WORLD BANK / INTERNATIONAL FINANCE CORPORATION

OFFICE MEMORANDUM

TO: Dr. J. Evans, Messrs. J. North and H. Messenger, and Ms. I. Husain

DATE: July 24, 1980

FROM: K. Lashman Hall, PHN KLH

SUBJECT: AID Health Evaluation and Survey Activities

1. The AID Health Evaluation Group's paper, "Toward Evaluation of Health Program Impact," has been forwarded to me by the Office of Health, AID for review and comment (Attachment 1). This paper is an expanded and revised version of the discussion paper, "Toward a Framework for Health Project Evaluation," of October 1979 which had generated much interest among Division staff when I disseminated it in late June. Your views on the usefulness of this document as a program development tool are welcomed by AID; since I have promised to share staff recommendations with them, I would appreciate your forwarding any comments that you or Division staff may have directly to me so that I can consolidate the Department response.

2. With regard to the ongoing and proposed AID vital statistics and population surveys outlined in my memo to the files of June 23, 1980, I have attached more detailed information recently received from AID and the National Center for Health Statistics (NCHS), Department of Health and Human Services (DHHS). Mr. Jack Lawson, AID Office of Population, Demography Division, has prepared a current status report for us on the Birth and Death Data Collection Project under contract to the Population Laboratory, University of North Carolina (Attachment 2). Dr. Robert Hartford, Acting Director, Vital Statistics Improvement (VISTIM) Program of the Office of International Statistics, NCHS, DHHS, has forwarded a project summary (Attachment 3) and an evaluation of the first stage of project activities in Peru (Attachment 4). As you will note, the majority of the countries are those in which this Department is directly involved.

3. Mr. Lawson has offered to meet with PHN to discuss both of these activities for which he serves as AID project officer. Please let me know if you would like me to arrange such a meeting.

Attachments

KLHALL:las

July 7, 1980

Ms. Theresa Lukas
Health Planning Adviser
Office of Health, Development Support Bureau
RPE AID, Room 301
Washington, D.C.

Dear Terri:

Thank you for the latest draft of the Health Project/ Program Impact Evaluation paper prepared by Stewart Blumenfeld. I unfortunately could not attend the session yesterday but the Bank was represented by Mr. Rashidur Faruqee from the Development Economics Department who is currently working on the area of health evaluation for the Bank. I hope to be able to continue to participate in AID's informal working group, at least indirectly through sharing views with you and Dave.

As mentioned, the initial draft which I circulated to the Operational Divisions of the Population, Health and Nutrition Department generated much discussion on methodologies for health program evaluation. I will, of course, forward any staff comments which I receive on the latest version for the group's consideration in preparing the final draft.

Thank you again for sharing this paper with us.

Sincerely,

Karen Lashman Hall
Population, Health and Nutrition
Department

WORLD BANK / INTERNATIONAL FINANCE CORPORATION

OFFICE MEMORANDUM

TO: Files

DATE: June 23, 1980

FROM: Karen Lashman Hall, PHN KLH

SUBJECT: AID Informal Health Evaluation Group

1. On June 18, 1980, at AID's request, I attended a working session of the Agency's Informal Health Evaluation Group. This intra-agency group, formed in the Fall of 1979, is comprised of health professionals representing the four regional bureaus, the central offices of Health and Nutrition, the office Program and Policy Coordination (PPC). Its central purpose is to establish a framework to guide AID in assessing the impact of AID-assisted primary care programs in the developing world.

2. Discussions to date have focused on the need within the health sector to move beyond traditional measures of program outcome, as expressed in changes in "final health indicators," e.g., reductions in morbidity, mortality, and fertility. Such measures, it is felt, have little value in the typical AID project with a median three years duration, and in the common developing country setting of inadequate health information systems. Alternatively, the Group is working to identify and promote the systematic use of intermediate measures which can be expected to change in relatively short periods and are valid indicators of progress toward achievement of project outcome objectives. A draft discussion paper, "Toward a Framework for Health Project Evaluation," prepared in October 1979 (Attachment 1) is the only formal product of the taskforce to date. A second draft is to be available on June 25, and will be forwarded to PHN.

3. At the June 18th session, Mr. Jack Lawson of the Demography Division of the Office of Population, AID, provided an overview of his office's ongoing and proposed vital statistics and population surveys. These include: 1) the World Fertility Survey. This extensive country-specific project has developed modules which can be used, as a country deems appropriate, to collect information on health services availability as well as standard fertility data; 2) the Contraceptive Prevalence Survey. This survey of females in the reproductive age groups to determine contraceptive use is to include information on the availability of contraceptives at the community level; the design for this supply component of the survey has not been finalized yet, however; 3) the Birth/Death Survey project. Under contract to the University of North Carolina Population Lab, this survey of fertility and mortality is planned initially for Colombia and Somalia. Data collection and processing will require an estimated one and one-half years; 4) the Civil Registration Project. The National Center for Health Statistics, (NCHS) Public Health Service, Department of Health and Human Services, under

an interagency agreement with AID, will serve as implementing agency for this Vital Statistics Information Management (VISTIM) project to improve national civil registration systems, especially birth and death registrations. In contrast to the other surveys outlined above, of a relatively short-term nature, improving the civil registration systems will be a long-term process of approximately three to four years. The project is currently underway in Jamaica, Peru and Thailand; a grant has been approved for work in Ecuador to commence Fall 1980; and the project is in the planning stage for activities in Brazil and Indonesia. Subsequent to the AID meeting, I contacted Dr. Robert Hartford, Acting Director, VISTIM, NCHS, who is forwarding detailed project information for PHN Department staff use. He can be reached at 436-7039; and 5) 5-Year Multipurpose Household Surveys. These surveys are not population nor health specific but can include questions on these areas.

4. The Group discussed the potential for incorporating health questions in these surveys to assist the Agency in identifying, and evaluating program/project success in meeting, health needs. It was determined that the World Fertility Survey is not an appropriate vehicle for adding-on health questions since the design is relatively fixed at this point in time. Any of the other survey instruments may be effectively utilized, depending on the type of information desired. The survey with the greatest potential for incorporating health components is the new Birth/Death Survey. As in the other surveys, however, this will provide data only on the macro level. If data is needed for smaller geographic areas, e.g., districts, to target population groups "at risk" and guide and monitor project implementation at the local level, the civil registration project is the most appropriate and viable tool.

5. Drawing on the extensive experience of the Office of Population in the conduct of baseline surveys and program evaluations, Mr. Lawson cautioned the Health Group in their work in several areas. He underscored the need to define precisely the health research questions which they want such surveys to address prior to any data system design or implementation to preclude the common mistake of excessive and costly data processing. He highlighted the value of sampling in facilitating, at an adequate level of confidence, decision-making based on more information per person surveyed at the same cost as more people with less data on each one. As an example, he compared the typical national survey which includes 6,000 to 10,000 households, and 250 variables per household, at an average cost of $100,000 to $300,000, with the recent AID-funded non-sample study of a single Morocco District which cost $0.5 million. He stressed the benefits of standardizing surveys (for baseline and evaluation), where feasible, both within and across countries, rather than undertaking studies on a case-by-case basis. Such standardization facilitates tabulation, processing and comparative analyses, while at the same time does not preclude the possibility of adding specific information required in a given project/program setting. He also emphasized the need for evaluation activities to continue beyond the life of the AID population (and health) projects if impact analysis is to be adequately conducted.

6. The participants concurred that the next step is to develop the key health questions which they would like incorporated in the general household surveys for possible field-testing in the forthcoming Birth/Death Surveys in Indonesia and Kenya. A small working group was formed to define the types of information required for intermediate measures, the questions needed to obtain these data and the funding mechanisms. They expect to have a discussion paper completed by September.

Attachment

cc: Dr. Evans
Mr. North
Mr. Messenger
Ms. Husain

KLHall/mlo

WORLD BANK / INTERNATIONAL FINANCE CORPORATION K.L.H.

File: Health Prog Eval

OFFICE MEMORANDUM

There is a group in DED actively collaborating in Household Survey work. You might want to discuss this with Timothy Ray's group i.e Jamieson. JNB.

DATE June 23, 1980

TO: Files

FROM: Karen Lashman Hall, PHN KLH

SUBJECT: AID Informal Health Evaluation Group

1. On June 18, 1980, at AID's request, I attended a working session of the Agency's Informal Health Evaluation Group. This intra-agency group, formed in the Fall of 1979, is comprised of health professionals representing the four regional bureaus, the central offices of Health and Nutrition, the office Program and Policy Coordination (PPC). Its central purpose is to establish a framework to guide AID in assessing the impact of AID-assisted primary care programs in the developing world.

2. Discussions to date have focused on the need within the health sector to move beyond traditional measures of program outcome, as expressed in changes in "final health indicators," e.g., reductions in morbidity, mortality, and fertility. Such measures, it is felt, have little value in the typical AID project with a median three years duration, and in the common developing country setting of inadequate health information systems. Alternatively, the Group is working to identify and promote the systematic use of intermediate measures which can be expected to change in relatively short periods and are valid indicators of progress toward achievement of project outcome objectives. A draft discussion paper, "Toward a Framework for Health Project Evaluation," prepared in October 1979 (Attachment 1) is the only formal product of the taskforce to date. A second draft is to be available on June 25, and will be forwarded to PHN.

3. At the June 18th session, Mr. Jack Lawson of the Demography Division of the Office of Population, AID, provided an overview of his office's ongoing and proposed vital statistics and population surveys. These include: 1) the World Fertility Survey. This extensive country-specific project has developed modules which can be used, as a country deems appropriate, to collect information on health services availability as well as standard fertility data; 2) the Contraceptive Prevalence Survey. This survey of females in the reproductive age groups to determine contraceptive use is to include information on the availability of contraceptives at the community level; the design for this supply component of the survey has not been finalized yet, however; 3) the Birth/Death Survey project. Under contract to the University of North Carolina Population Lab, this survey of fertility and mortality is planned initially for Colombia and Somalia. Data collection and processing will require an estimated one and one-half years; 4) the Civil Registration Project. The National Center for Health Statistics, (NCHS) Public Health Service, Department of Health and Human Services, under

an interagency agreement with AID, will serve as implementing agency for this Vital Statistics Information Management (VISTIM) project to improve national civil registration systems, especially birth and death registrations. In contrast to the other surveys outlined above, of a relatively short-term nature, improving the civil registration systems will be a long-term process of approximately three to four years. The project is currently underway in Jamaica, Peru and Thailand; a grant has been approved for work in Ecuador to commence Fall 1980; and the project is in the planning stage for activities in Brazil and Indonesia. Subsequent to the AID meeting, I contacted Dr. Robert Hartford, Acting Director, VISTIM, NCHS, who is forwarding detailed project information for PHN Department staff use. He can be reached at 436-7039; and 5) 5-Year Multipurpose Household Surveys. These surveys are not population nor health specific but can include questions on these areas.

4. The Group discussed the potential for incorporating health questions in these surveys to assist the Agency in identifying, and evaluating program/project success in meeting, health needs. It was determined that the World Fertility Survey is not an appropriate vehicle for adding-on health questions since the design is relatively fixed at this point in time. Any of the other survey instruments may be effectively utilized, depending on the type of information desired. The survey with the greatest potential for incorporating health components is the new Birth/Death Survey. As in the other surveys, however, this will provide data only on the macro level. If data is needed for smaller geographic areas, e.g., districts, to target population groups "at risk" and guide and monitor project implementation at the local level, the civil registration project is the most appropriate and viable tool.

5. Drawing on the extensive experience of the Office of Population in the conduct of baseline surveys and program evaluations, Mr. Lawson cautioned the Health Group in their work in several areas. He underscored the need to define precisely the health research questions which they want such surveys to address prior to any data system design or implementation to preclude the common mistake of excessive and costly data processing. He highlighted the value of sampling in facilitating, at an adequate level of confidence, decision-making based on more information per person surveyed at the same cost as more people with less data on each one. As an example, he compared the typical national survey which includes 6,000 to 10,000 households, and 250 variables per household, at an average cost of $100,000 to $300,000, with the recent AID-funded non-sample study of a single Morocco District which cost $0.5 million. He stressed the benefits of standardizing surveys (for baseline and evaluation),where feasible, both within and across countries, rather than undertaking studies on a case-by-case basis. Such standardization facilitates tabulation, processing and comparative analyses, while at the same time does not preclude the possibility of adding specific information required in a given project/program setting. He also emphasized the need for evaluation activities to continue beyond the life of the AID population (and health) projects if impact analysis is to be adequately conducted.

6. The participants concurred that the next step is to develop the key health questions which they would like incorporated in the general household surveys for possible field-testing in the forthcoming Birth/Death Surveys in Indonesia and Kenya. A small working group was formed to define the types of information required for intermediate measures, the questions needed to obtain these data and the funding mechanisms. They expect to have a discussion paper completed by September.

Attachment

cc: Dr. Evans
Mr. North
Mr. Messenger
Ms. Husain

KLHall/mlo

K. Lashman

Toward a Framework for Health Project Evaluation

Drafted by a working taskforce* of the
Informal Health Evaluation Group of AID

Washington, D.C.

October 1979

*Taskforce members; defined by active participation include:

Abby Bloom
Stuart Blumenfeld
Charles DeBose
David Dunlop
Katherine Fort
Elizabeth Hunt
Theresa Lukas
John Massey
Maureen Norton
Suzanne Olds
Barbara Pillsbury
Hope Sukin
Mel Thorne
Barbara Turner

Table of Contents

Introduction

AID has been a pioneer among, and cooperating partner with, international donors in promoting primary health care systems. Recently this approach to health care has gained approval as reflected in the WHO Conference on Primary Health Care at Alma Ata where delegates established a global target of "health for all by the year 2000".

Despite the increasing acceptance, there is little hard evidence that primary care involving heavy use of outreach workers, paraprofessionals, and triage is more effective in improving the health status of larger numbers of people - or in assuring more effective use of limited resources allocated to health - than are the previous urban-based, highly-specialized, largely-urban hospital systems. It seems that this should be so; however, the evidence to date is inconclusive at best.

The U.S. Congress also has become increasingly interested in the success of the programs (it funds for the U.S. people) and its support for international health programs has been increasing (FY 1975 - FY 80, $ mil). Some talk in terms of effectiveness rather than success, but basically the Congress is demanding to know in precise terms if its investments are paying off.

Independent of, though in agreement with, Congressional concerns, a number of A.I.D. health professionals have become interested in the evaluation of health programs and projects. They have formed an intra-agency health evaluation group to address this issue. The group is highly interested in determining if primary health care is working as efficiently as hoped. They have become highly dispirited about the possibility of obtaining changes in final health indicators in the short time periods over

which AID projects operate (median = 3 years) or in countries where collection of statistics is unrefined, uncommon, or nonexistent.

The group also has been concerned with the fact that importation of statistical and survey expertise to assess the projects significantly increases the cost of delivery services, especially given that the evidence produced by such expensive inquiry may still not indicate long-term impact, especially in the short-run. Finally, even where changes can be established, it is almost impossible to attribute them to specific causal agents if several things have begun to change at once. If, for example, a road is built which allows farmers to get better prices for their crops due to improved quality or reduced costs of marketing, or if a garden project encourages the local people to broaden their diet at the same time a health promoter starts working in the area - how does one establish what percent of those health improvements are attributable to which intervention?

This group and many other health professionals have begun to ask whether there may be alternatives to final outcome measures, i.e., intermediate measures which can be collected and will show changes in a short time and in less than ideal research conditions.

Theoretical arguments suggest that these intermediate measures, often based on changes in usage of services or delivery of services, can be linked to final outcome measures in a vigorous manner. If these assumptions can stand up to emperical analysis, then the less expensive intermediate measures can be used in a majority of situations with final outcomes being deduced from them. Such extrapolation has been standard procedure for some years in certain types of interventions - most especially in the case of inoculation

campaigns where effectiveness is judged on the basis of the percentage of the population covered and effectiveness of the vaccine administered is assumed.

Therefore, the primary purpose of this paper is to provide a rationale for a set of useful intermediate measures. A secondary purpose is to identify more clearly the various audiences for evaluation results and to delineate what sorts of questions are most relevant to each of those audiences.

This document is envisioned as an introductory one. The AID health evaluation working group plans to develop a series of technical papers. Each of these will focus on a specific aspect of the evaluation process. It is hoped that the series will not only be useful for individual participation in project evaluations, but, in addition, will improve the approach to health care evaluation throughout the world.

Recent AID initiatives have sought to improve the quality of life of the most disenfranchised members of society. Priority has thus been placed on issues of equity and access affecting the poor. For health endeavors it becomes important not to simply increase life expectancy or reduce infant mortality in the aggregate, but to assure that all members of society share in the advances. Clearly, the long-range goal still includes the latter two efforts. However, increased access and greater equity are necessary if not completely adequate first steps in achieving greater life expectancy and reduced infant mortality among those most disadvantaged groups. Furthermore, as a first step it is believed that changes in access and equity measures can be measured earlier than can changes in final outcome indicators and can be more effectively monitored in field settings. Finally,

by concentrating on access and equity, health programs can be incorporated into the larger developmental context and that health can be recognized as only one component of an entire set of activities underway to assist the less advantaged populations of the developing world - multiple activities which often function synergistically.

In considering the usefulness of intermediate measures, it is important to recall that final outcomes of program interventions often require a considerable passage of time before they become measureable. If resource allocations are predicated on significant short-run changes in final outcome measures, many activities which could later demonstrate significant achievements would be terminated long before such changes became apparent. For example, if resource allocations of both domestic and international population programs had been based on final outcome measures, i.e., fertility reduction, they would most likely have been phased out before they had borne fruit. As it was, allocations were made instead of the basis of intermediate performance measures and the programs survived to meet or approach their longer term goals. It is the intent of this paper to suggest a similar set of intermediate performance measures appropriate for health programs.

The main body of the paper first discusses the mechanisms by which health can be improved. Second, the present outcome measures used in health programs are re-assessed and the family planning experience is more thoroughly reviewed for lessons applicable to health programming. Third, a theoretical case for an alternative approach to final outcome measures is made where consumption rather than investment is considered to be the more appropriate goal. Finally, a set of practical evaluation measures is developed along

several dimensions. These measures are evaluated against timing and data requirements, and against audience considerations. Recommended evaluation guidelines conclude the paper.

A Reassessment of Impact Measures for Health Programs

There has been a long and continuing search for measures of final outcome that can be utilized across all types of health program activities in order to make comparative analyses of the relative success of each endeavor. Thus it is not surprising that decision makers concerned with resource allocation have come to accept measured changes in various vital event rates such as infant mortality and longevity as yardsticks for success in the health field. Such measures were utilized in part because information networks had been developed to monitor vital events in general in the U.S. Another reason is that in the 1960s there was considerable interest in being able to conduct benefit-cost analysis for all human service programs with the benefits of each program being easily identifiable and comparable. Since the infant mortality rate has generally been assumed to provide a relatively sensitive measure of the outcome of a number of types of health-related programmatic interventions, e.g., nutrition, sanitation, and MCH services, it is considered particularly useful.

Health: Alternative Approaches to Its Attainment

As in any situation, a simple measure is desired as it facilitates comparability across programs. However, in the case of health, the effort to define output in simple terms yields considerable problems.

One of the reasons it is difficult to find an adequate simple measure is the very complexity of health. A positive state of good health is obtained via a complex interweaving of many factors and many inputs. Health care services comprise only one of these factors, and may be in many respects a relatively minor one when compared to agricultural production and family incomes. This is not to say the availability of health services is not extremely important to a population but rather to point out such services operate in a context and as part of a complex system.

At the present time in LDCs, many of the inputs that can improve the health of a population are provided via categorical delivery systems, e.g., separate immunization campaigns for individual diseases. Health planners currently are attempting to use the concept of primary health care to weld the present disparate delivery systems with outreach and education services into a unified and thus more productive strategy for health improvement.

Many of the services embodied in this broader primary health care approach are not limited to the health sector and traditionally may not have been considered part of that sector at all. Among these are food production and distribution, provision of safer and more abundant water for household use, sanitation and housing. Even though a particular health program may well not choose to tackle all these in themselves multifaceted problems, it is useful to incorporate them into the concept of primary health care. This incorporation helps policy-makers account for the large number of variables influencing the health status of any

given population, understand the wide range of options available for improving that status, and not lose sight of the possible complementarities or synergism between those options.

It may be useful for persons outside the health sector to consider how the various inputs to health complement those in sectors even further removed than those just mentioned. For example, there is the economic question of the ways in which inputs can complement each other. Something is known of the positive synergism of various health inputs in improving health status. Less is known about how much health inputs directly or indirectly improve health status and thus improve returns to investment in education and other areas producing human capital by increasing peoples ability to concentrate, to learn, and to labor. Similar questions can be pursued in many sectors. Thus, a broader, more complex view of health yields definite benefits.

However, such complexity presents serious problems with the use of final impact measures. One is that it is doubtful that truly simple measures exist to adequately assess changes in such a complex system; if they do we certainly have not identified them.

Attribution

Another, basic problem is that of attribution - which is perhaps a corollary of the simple measure one. If a change of y in measure x is achieved, but if a, b, c, d - h have all been changing, how does one decide if the change in measure x is attributable to changes in "a" or in "b" -- and in what proportion?

Furthermore, it is important to distinguish between statistical correlation and ascribed causality. In many instances, it is not clear, particularly if other activities are under way in the same area or locale, that the efforts made by one program or intervention were not attributable in an indirect way to another intervention or set of interventions established in the same locale. While the occurrence of such multiple effects can potentially be statistically disentangled, the program or the project information system is generally not designed to accumulate information about the larger environment and the changes in that environment which may be the actual reason for the success of the program. In two cases that have been systematically studied in recent years, the changes in the infant mortality and subsequent birth rate in both the Kerala stateof India and the country of Sri Lanka have generally been attributed, not solely to health programs, but also to a commitment by both governments to minimize fluctuations in food consumption.[1/] At the same time, these countries (or parts thereof) have made the political commitment to explicitly address the issues of distribution of wealth and income. Thus, to attribute causality or direct impact to a particular programmatic intervention, e.g., health, may not be valid given the larger context in which the program operates.

Feedback and Indirect Impacts

The problem opposite attribution is that of indirect effects. If simple measures in a complex system make it difficult to determine what has caused a given change, they also make it difficult to assess all the effects a given intervention had.

The term program impact generally connotes a change in a direct outcome measure which is attributable to a particular intervention. As in most human resource programs, however, often there are not just direct effects. There are many indirect effects as well and these may be either desirable or undesirable.[2/] For example, in the case of education, it has generally been assumed that increased education improves labor productivity. However, the benefits of increased labor productivity not only accrue to specific individuals but also to society.

At the same time, it has been statistically demonstrated that increased education is associated with household decisions to restrain family size in most countries, including the higher income countries; there is a high correlation between educational status and the probability of migration from rural to urban areas. In certain situations this may be a negative externality. This is particularly true of those areas where economic growth is slow or stagnating and rural-urban migration may exacerbate poor living conditions for both urban and rural dwellers.

While the purpose of this discussion is not designed to trace all potential effects of education, the above examples are suggestive of the multiple impacts attributable to improved education. In the case of health, programmatic outcomes may manifest themselves in multiple ways, both in direct improvements in health status irrespective of the measure used, and in human resource measures, e.g., rate of learning, attentiveness, anthropometric measures, and changes in desired family

size. Thus the impact of a health intervention may take many forms. A person's or a program's own creativity and ability to monitor (including a budget constraint) are the only reins on the possibilities.

Timing of Evaluation

A third generic problem in using any measure of outcome, including vital events is determining when to make the assessment, i.e., the timing problem. Without a continuous long-term monitoring system in place, when to assess program impact becomes a very significant judgement which may well seal the fate of the evaluation endeavor. There may be programs that have a long gestation period before any measured change occurs. The impact may be cumulative and may peak subsequent to the assessment. In such a case, the evidence may indicate "failure," yet were the assessment to have been conducted at a time when the peak impact of a particular intervention had occurred, the program would be considered successful. The opposite can also occur. A program in the short-run may appear more successful than its longer term effects would warrant.[3/]

Operations research studies may be required to decide on appropriate timing of evaluation. A set of studies to monitor both the flow (amount and timing) of benefits and of costs may be necessary. However, another conflict can arise. The Agency, in order to maintain any comparability, establishes standard cycles. Current project design standards require that impact analyses be conducted within four years of project inception and projects are rarely evaluated after termination.

Thus, there is general underreporting of long-run program impacts. This is, of course, not a direct problem with vital events or other final impact measures. It does, however, skew the results and virtually guarantees that changes which require periods of time of more than three years will not be observed.

4 Benefit Disaggregation

A fourth problem is that even if a change in vital events occurs, there may be no clear indication of which part of the population benefitted. The effects must be appropriately disaggregated. If all of the positive changes in vital events for a particular project occurred among a segment of the population that was relatively well-off prior to project initiation, the project may be greatly successful in terms of overall statistics and an utter disaster in terms of improved equity. A number of multi and bilateral donors are interested in improving the living standards of the poorest groups in the population. Thus, it is important to obtain a disaggregated distribution of the measured changes and measure initial events at least according to income group. Perhaps in rural areas, the best approach would be according to size of land holding and/or whether there are off-farm income sources.

5 Baseline Information

A fifth problem with present measures is that frequently, prior to the initiation of the proposed health activity, there is no measurement taken of the vital event rate of the target population. This is true whether the target population is described in geographical terms or

via some other taxonomy, e.g., age, sex, or income level. The present lack of such data can be corrected in future program efforts by conducting baseline surveys and establishing other information systems prior to the project's inception. Nevertheless, at the moment there are a number of projects for which little or no baseline data are available and which may or may not have been collecting data appropriate to determining changes in vital events.

In some cases where baseline data is lacking for smaller areas, adjustments of country-wide estimates of vital rates are used as substitutes. However, such adjustments may grossly under or over represent the conditions among the target population which may deviate more or less than expected from the national figures. Furthermore, this assumes that the countrywide vital events statistics available are accurate in the first place. This may be untrue for several reasons. For example, political and other similar reasons may dictate that a country's or region's vital event rates, such as infant mortality or the crude birth rate be set at some predefined level. Few political leaders would be willing to allow official documents under their control to reveal that the "true" infant mortality rate in their country approached 200. Irrespective of the reason, if the vital event rate recorded prior to the health or other program intervention makes the situation look better than it in fact is, a project may have to achieve monumental success in order to appear to show minor gains when measured against the apparent "before" rate. This is not just a theoretical concern. A number of instances exist, particularly in rural areas, where the infant mortality

rate is generally underestimated for a variety of reasons. If a rural primary care program were developed to address some of the reasons for the high vital rate, i.e., infant mortality, the program may inappropriately be deemed a failure. Another reason for doubtful vital events rates is that data are not always recorded correctly nor are all events incorporated into the rate numerator. There is mounting evidence that underreporting of such events are common. For example, research conducted in Honduras indicated that there was a 60 percent underregistration of infant deaths.[4/] Since there may be considerable variance in underreporting throughout a country because of differences in ethnicity, income, and other social and cultural factors, it is difficult to derive appropriate data from national norms. Thus the data concerning even an event as absolute as death are often "soft" and may not provide the policy guidance desired.

Cost of Information

(b) Finally, the cost of obtaining accurate vital event information is of serious concern. Epidemiologists, and other survey research experts have repeatedly pointed out the problems of collecting such information.[5/] They have ponted out that without expensive survey research procedures and careful records maintenance (which also has costs) such information is unattainable.

Several other points must be considered when addressing the cost-of-information issue. First, any organization, be it private or voluntary, systematically gathers and assesses information about its performance for review by decision makers. The information may be

organized as an income statement or in the form of some other production statement. In all cases, significant resources are used for the accounting, data processing and auditing functions needed to collect such information. At present, it is not clear that the collection of vital events information in health projects or related human resources projects is necessarily more costly than the record-keeping for any other standard production activities.

However, information does have a significant cost, and significant costs are always an issue. In considering whether information is worth the price, the opportunity cost of not collecting that information must be taken into account. The past opportunity costs stand out more clearly than do future costs which can be anticipated only in part. In the case of AID, the past opportunity costs are quite clear. Quite simply, Agency evaluation efforts would not be in their present bind if more attention had been paid and more investment made in good information collection, analysis, and storage in the past. With more information the ambiguity of what constitutes appropriate technology for a given service under given conditions would be reduced. The decade of the 1960s, both in developing countries as well as in the U.S., should have taught an important lesson: do not set social expectations too high. However, it is also important not to set them too low. Successes and failures of the past will not improve the future without evaluation efforts which are based on both sound information and a framework for its interpretation. Those who make resource allocation decisions must be fully appraised of the fact that they must be "willing to pay" in order to obtain the information required to either continually or

periodically reassess the portfolio of program activities.

For all of the above reasons, even though a given health program may indeed save lives and reduce mortality rates of specific segments of the population, it is not clear that a measurement of changes in vital events rates which can be linked to the program is the best measurement of program success. It is certainly not the only possible type of measurement, and it is a clearly unappropriate measure of short-term change.

need for alternative measures

The Rationale For Measurement Choice in Health Programs

While "resource allocation decision makers" have taken the position that the primary measure of impact of a health or health-related project is changes in vital event measures, it is important to understand the context in which health programs were designed and implemented in the past and the extent to which the primary purpose of health programs has been undergoing change in the present development context. In particular the focus of development efforts has shifted in the last decade from one of growth maximization as measured by changes in per capita income, to one where basic human needs and "quality of life" have crept into the limelight. The Alma Ata slogan, "health for all by the year 2000", is only interpretable in this larger context.

Health: A Consumption Good

The globally changing raison d'etre of development activity reflects itself in the current concern with use of indicators based on changes in vital events to measure final outcomes of health programs. To measure a change in vital events attributable to a health program is primarily to measure the investment output of such efforts. For example, changes in the infant mortality rate or in age and sex-specific death and morbidity rates can be used in a classical human capital benefit-cost framework to estimate the benefits and thus the internal rate of return to the activity. 6/ The question is whether or not this is the most appropriate way to look at health care. If the U.S., in concert with WHO and the United Nations, is seriously concerned with addressing the goal of quality of life and

the issue of access to basic necessities, then health care, particularly that provided through a primary care delivery mechanism, must be viewed primarily as a consumption rather than an investment good. Given that consumption is the underlying raison d'etre, it is more appropriate to look at measures of medical care and health services utilization and measures of consumer satisfaction particularly among important target groups, i.e., the poor, in order to measure impact. In the case of changes in utilization patterns, it is also important to distinguish between an increase in total consumption of medical and/or health services and a shift in consumer selection among alternative providers: public and private, or "traditional" and "modern". Here, both the intent of the intervention and the actual outcome must be examined. Further, if the intervention is meant to increase awareness and consumption of preventive health practices (family planning represents one special subset), it may be appropriate to monitor changes in the knowledge, attitude, and practices (KAP) related to that particular type of service. An example of such an approach taken in the health area is the evaluation of the Tanzanian mass health education program conducted in 1973. 7/

There are several approaches to measuring changes in satisfaction resulting from changes in consumption. First, household-based health interview surveys (HIS) which have been systematically conducted in the U.S. for the last ten to fifteen years, can be utilized to obtain population-based medical care or health service utilization/consumption rates. 8/ More precise estimates of this relationship are possible if the resulting change information is either related to changes in consumer satisfaction established

using survey instruments developed by John Ware and others or to changes in "willingness-to-pay" responses elicited by the methods suggested by Ed Clark, Joseph Lipscomb, and others.[10/] Other inferential information about changes in consumer satisfaction can be obtained by monitoring changes in household activities particularly the allocation of time among consumption, production and leisure activities which can reveal changes in the quality of life.

Health: An Investment

Perhaps the investment component of health programs itself can be more easily monitored and/or verified by certain other measures than via vital events. First, if development interventions are occurring simultaneously in other sectors, particularly agriculture, it is possible to obtain information about the allocation of scarce household time when the demand for labor is the greatest (planting, weeding, and harvesting).

There have been a number of farm management studies that have obtained fairly precise information about time allocation from season to season for various activities,[11/] and there is information available from certain selected primary health care facilities that indicates utilization patterns (particularly amongst adults) are inversely related to peak agriculture demands for labor, i.e., when there is an increased demand for labor at planting, weeding, and harvesting time the number of visits to primary health care facilities falls.[12/] It may also be that evidence can be obtained about the rate and timing of planting which can be related to instances of reported morbidity in

the population. This information could then be related to agricultural output fluctations. These data could also be supplemented by labor information obtained from migration studies (disaggregated on the basis of age and sex) and by information indicating the degree to which households go beyond their own labor resource base to hire additional labor.

Another interesting set of measures is associated with seasonality. Particularly in rural areas, patterns of abundance or scarcity of both time (as has been discussed) and money (or barter goods) follow the patterns of planting and harvest. Measurement of the number of people suffering from certain health problems at various times of year, can be highly illuminating especially when scarcity patterns are also measured and compared.[13/] Likewise, corollating the extent to which various traditional health practices are employed or "modern" health services are sought with both the calendar and ritual years can produce important information. It should be emphasized that such information can be critical in designing program or project appropriate to a given region. Later changes in these patterns can be evaluated together with investment (more/less morbidity) and consumption (more/less or different services sought) impact data to give a more complete analysis of total impact as compared to one employing only one subset thereof.

The above discussion has implied that consumption rather than investment measures be given more priority in project evaluation than has been the case in the past. The case for such a reallocation is fully consistent with AID's recent congressionally approved strategy

of supporting ... "The achievement of self-sustaining equitable growth oriented toward the establishment of basic human needs." 14/ However, in order to achieve self-sustaining equitable growth, many, if not all programs must yield a positive return. Thus in health, as well as in other sectors, the relationships of increased consumption of basic goods and services (e.g. health care when sick) and investment returns (e.g. changes in vital events rates, including morbidity and functional health status) must be investigated over time. The intent of AID is not to provide unending support to programs or countries solely to increase consumption. However, evidence from LDCs increasingly suggests that unless a minimum level of consumption of basic goods and services is reached, investment returns will not be forthcoming due to the lack of changes in behavioral and motivational factors related to increased consumption. Thus, as greater investment in consumption of human resource services such as health and education is both linked to and necessary for greater productive output, the increased use of short-range measures, which monitor changes in consumption brought about by the human resources investment programs of both countries and donors (including AID), becomes more respectable in economic analysis of development progress.

rel's/p C and I outputs of h services t/away

Situational and Programmatic Constraints

The outcome of a health program is defined in a situational and programmatic context. It is important to determine the characteristics of that context. When possible, they should be measured.

First there are non-programmatic constraints: the political, socio-economic and cultural realities of a socity. The way wealth and power are distributed in a society, the ways the status quo is threatened, and the ways the threatened parties will react are all critical. What happens in the Ministry of Defense logically should not effect a health program, but in reality it may. The religion of a program director should not affect the program, but it may. To ignore these factors is to court the fate of the ostrich, and to not make them explicit from the program's inception leads to unrealistic rigorous evaluation efforts.

An equally important set of constraints are intra-programmatic.[15/] For example, these constraints include the way the managers manage themselves or how they use other inputs. Other examples include how the critical design or technology is chosen. What is the process of program implementation? What are the administrative process constraints, e.g., logistics, personnel supervision and the regularity of supply of the necessary inputs of the program. For primary health care programs, perhaps the extent to which there is a regular supply of efficacious drugs is one of the most critical variables affecting "success", particularly if the country is having a foreign exchange problem. Finally, it is often easier to look at the constraints in the foreign society when those intrinsic to our own bureaucracy or society may be causing equal or greater problems.

Currently, data systems developed for purposes of evaluation and managerial control rarely gather information on these program and

non-program constraints. Furthermore there is often little documentation of why certain strategies were chosen or certain technologies were incorporated into a project. This lack of documentation is particularly true when design changes occur in a project over time. As a consequence, it is difficult to determine whether the program's "success", or lack thereof, is attributable to the program's management or largely due to environmental variables outside the control of the project itself or rather to the project per se. For example, if there has been a recent coup, or if the price of the primary export product of the country has dropped by fifty percent, e.g., copper, many social economic difficulties would intervene and thus thwart the successful implementation of the program. Evaluations which do not measure and analyze these non-program constraints will never find the cause of program failure.

In the area of initial project design and technology definition, the Agency has considerable internal and external resources upon which it can draw. Nonetheless, in many health care programs the set of necessary inputs, their complementarity and their substitutability, have not been systematically defined. Some operations research has been conducted in this area, e.g., on the differential outcomes between para-professional and professional personnel.[16/] Continual experimentation with the technology of health care provision under alternative types of programs is necessary to know how a given technology may likely perform in a given environment. A systematic applied research agenda is implied and it may be useful to follow the Office of Population's lead in developing an operations research contract.

A Proposed Set of Evaluation Measures for Health Programs

As our discussion grows more specific, it becomes useful to step aside a moment to review the definition of terms currently common in evaluation. One important definitional distinction is between the terms "effect" and "impact". A WHO panel in 1975 equated "impact" with what has been referred to in this paper as a "final outcome measure".[17] "Impact" was further defined as an induced and human specific effect of an intervention. The programmatic output measures were termed direct effects. An example of a direct effect health indicator would be the number of physician or clinic visits. Further, the direct effects are "chained" into a relationship with the induced effects (human specific) and these are termed impact measures. In short, by this definition, an impact is a sub-category of an effect and specifically refers to the changes wrought in the target population.

The term "program effectiveness" has at least two important components. First, it connotes comparability; a particular impact measure has been altered in some desired way and in accordance with some a priori expectation. In this context the term has an evaluative component which the

term impact does not. Second, the term effectiveness connotes economy of resource use in attaining the desired or intended effect. This attribute of the term applies when the technology of a health program for example, is analyzed in the context of its efficiency of resource use relative to what it produces. Both of the above senses apply when a program is analyzed in the context of its efficiency of resource use relative both to what it produces and to what other programs would have produced.

Since both terms -- program impact and effectiveness -- have several meanings and can be interpreted in various ways, it is important that they be used only when they can be used quite generally. In instances where evaluation requires more precision and specificity of intent other terms are required.

Levels of Evaluation

P.S. Mohapatra has developed a useful taxonomy for categorizing human resource program effects via a typology of objectives. He suggests: "All ... programs involve a hierarchy of objectives, conviently grouped into three classes - ultimate, intermediate, and program execution objectives."[18/] He defines the term ultimate objective similarly to the way the term "final outcome measure" was defined earlier. Intermediate objectives in his parlance is synonymous to measures of consumption, and/or attitudes or behavioral change with respect to health sector intervention. Finally, he says that "program-execution objectives refer to the performance of the specific activities carried out in pursuit of the intermediate goals. The mobilization of resources

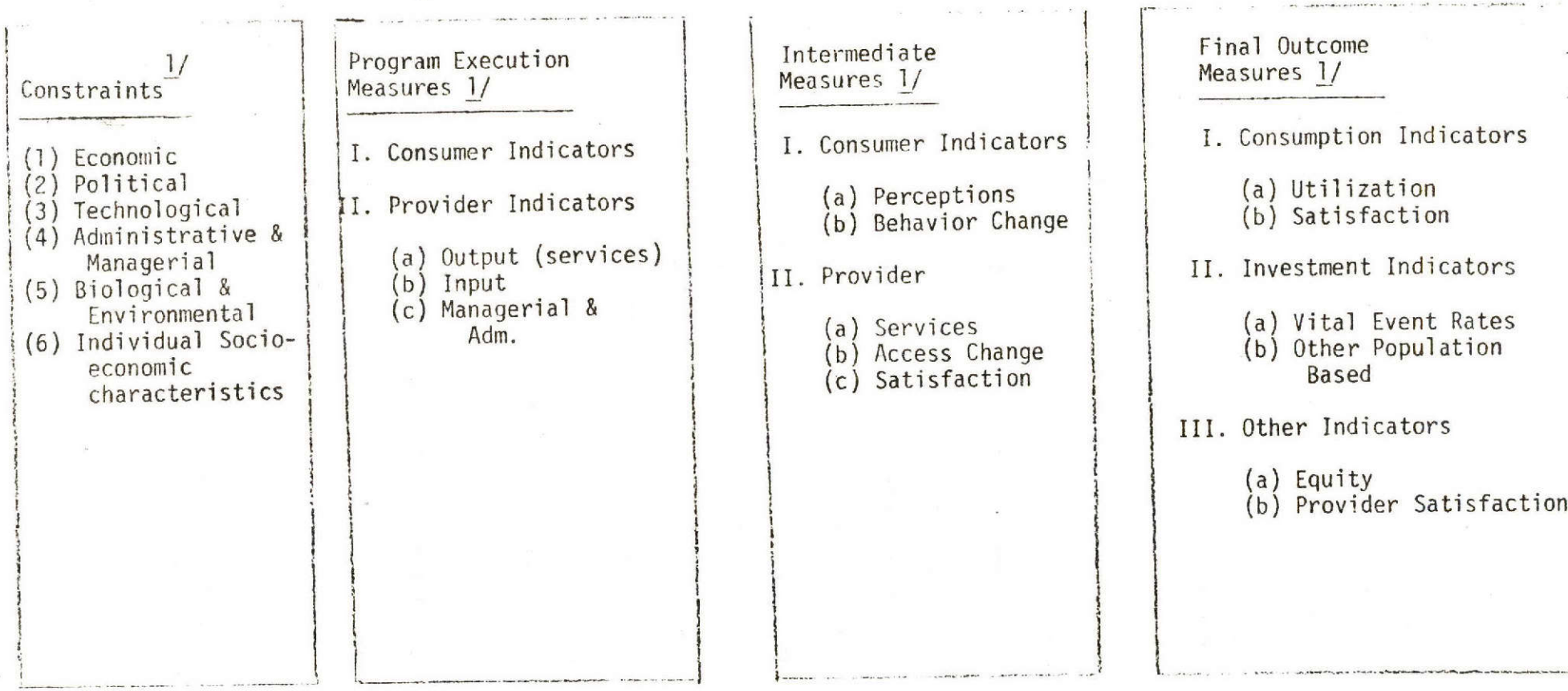
(program inputs) ... and the actual provision of services (program outputs) ... would be an example ..."[19] He goes on to say:

> "The heirarchy of objectives is linked together in a series of input-output chains in which lower-order outputs ... become inputs in higher-order activities. Any program can be analyzed in terms of the large number of input-output chains of which it is composed; indeed, the soundness of a program can be judged by the realism of the assumptions ... used to construct these linkages."

For purposes of AID health program evaluation, Mohapatra's taxonomy can be a useful point of departure. In Figure 1, the level of program evaluation is identified in conjunction with the direction of chained linkages. This is done for each level of evaluation. Within each level of evaluation a suggested but non-exhaustive list of indicators has been developed and presented in Tables 1A-1D at the end of the paper. For both the program-execution and intermediate evaluation levels, the indicators are disaggregated into two subsets: consumer and provider oriented. To conform with the guidance provided in the previous section, the final outcome indicators or ultimate objectives are disaggregated into consumption and investment indicators, with the investment indicators disaggregated into (a) vital events, (b) other population-based indicators, and (c) program-based indicators.

The set of at least 6 different types of constraints on program performance, irrespective of level or objective is defined in the following categories: economic, political, technological, administrative and managerial, biological and/or environmental, and socio-economic target population characteristics. Within each of the generic constraint

Figure 1: Towards a Health Program Evaluation Framework



Notes: (1) The specific measures are included in Tables IA-ID

categories, a set of indicators is suggested. The set is not meant to be exhaustive. A minimum set of administrative and managerial constraints might be composed of: measures of logistics management, information system development and information flow, personnel development, supervision and management, financial control and budgeting procedures, evaluation and planning procedures, leadership ability and commitment, and organizational and structural characteristics.

From the perspective of integrating intermediate or process evaluation efforts, it is important to evaluate programs on the extent to which, both non-program and intra-program constraints have aided or impeded the project's development. Besides basic lack of program resources, other intra-program constraints - such as managerial problems, the initial design, and the technology embodied in a program are examples of topics for "process evaluation". Evaluations which analyze political, socio-economic, and income distribution variables assist in determining the extent to which such constraints altered the course of the program, and thus, define the extent to which the program in and of itself could have been expected to achieve its final outcome goals.

It is important to point out at this point that figure 1 implies a set of functional relationships between final outcome indicators and intermediate outcome indicators, as well as sets of relationships with inputs and with constraints. These relationships are indicated by the arrow from one set to the other. For each set of indicators, it is possible to hypothesize the sign of the relationship between two variables or indicators, holding all other things constant. In practice, however,

certain hypothetical relationships are not supported by the available evidence. Depending on the decisions to be made, it may become vitally important to ascertain why a divergence occurred between expectations and actual results. Such findings can generate a useful applied operations-research agenda.

Why These Indicators?

There are several reasons for recommending the use of this taxonomy and the suggested indicators outlined in Figure 1:

First, by developing a set of measures a greater range of programmatic activities and types of impact are monitored. By using a larger number of indicators, the subtleties of causal chains and the relative importance of constraints can be investigated more systematically and simplistic ascertions of cause and effect can be avoided.

Second the provider indicators are, for the most part integrally tied to the successful performance of the administrative and managerial functions. The systematic monitoring of administrative and managerial constraint indicators can provide those persons making program decisions with improved knowledge of problems of program execution and better understanding of their cause.

Third, each evaluation level proposes one or more indicators of consumer response. These measures give an additional set of criteria for evaluating the newer AID activities armed specifically at meeting basic human needs in which consumption outcomes have been given more weight.

Fourth, with this set of measures, programs can be monitored at an earlier point. Thus, earlier and more accurate mid-course corrections could occur (see section outlining below). Since the set of indicators measures a number of different attributes of program activity, it is also possible to better ascertain where a potential problem exists. Without such breadth, it is difficult to pinpoint difficulties and generally only the broad generic problem is noted.

Fifth, data gathered from household surveys are not generally required for many of the proposed measures. A well-designed, program-specific information system can provide virtually all of the monitoring and intermediate outcome measures on both the supply and demand side as well as some of the consumer-oriented final outcome measures. Much of the constraint information can be obtained through existing documents. Thus only a minor amount of necessary information requires collection via a household survey. Thus, the cost of short-run evaluation efforts can be significantly reduced.

If specific case studies are strategically picked, household-based data may already exist, e.g. the Bicol region in Philippines. Finally, by virtue of having a multiple set of measures and indicators which can be used respectively in short, intermediate and long-term evaluation contexts, various reporting and evaluation requirements can be met using appropriate measures without attempting to conduct impossible exercises simply because changes in final outcome measures were established as the long-range raison d'etre of the endeavor. More focused, and therefore, less expensive, evaluation efforts can take place throughout the

life of the project thus increasing the feedback usefulness of such activities.

While all suggested measures suffer from a certain amount of measurement bias and under-reporting, many problems can be circumvented or minimized earlier in the life of the project by virtue of having the program specific information system more closely integrated into the evaluation process.

Other Evaluation Considerations

There are four other issues or questions pertaining to the development and use of these evaluation indicators: (a) for whom (audience) is the evaluation being conducted; (b) at what point is it possible to obtain information about each measure; (c) is data available (information flow issues); (d) and what does it cost to obtain and use information, especially in terms of manpower and computer time. These issues are addressed in light of the proposed set of measures indicated in Figure 1. The judgement of the working group on how each measure addresses the four general issues raised in the discussion to follow is summarized in Table 1 at the end of the paper.

Audiences.

No one indicator is appropriate for all audiences nor does any one audience normally require all indicators to answer the questions relevant to its interest. As one evaluator has commented, "even when there is a clear commitment to evaluation, per se, there must be a clear understanding

of why (a given) evaluation is being carried out. ... the decisions which are to be made on the basis of information to be collected must be known for the right information to be collected." 21/

In the context of health, there are a number of specific audiences which request evaluation. In many instances these varying audiences define the scope of inquiry and the extent to which different indicators have relevance in terms of the decisions they must make. In the case of health project evaluation exercises, there are at least six distinct audiences which may request an evaluation report. These distinct audiences include (a) external donor organizations -- sometimes for themselves (policy) and sometimes for their individual constituencies, i.e. AID for Congress (justification); (b) the national government or the relevant ministry with jurisdiction over the project; (c) the regional (provincial) subset of that ministry; (d) the local administrative office or person responsible to the ministry, i.e. the district medical officer; (e) the project or program director; and (f) the workers involved in the development and running of the project.

In general, those most removed from the daily operation of the activity are interested in knowing if and how the project achieves final outcome targets, e.g., lowered infant mortality. They are concerned with large resource allocations. Program administrators and service-providers are more interested in measures of direct programmatic activity. They require information for supervisory purposes. Given these differences, it is likely that the different audiences will require different evaluations. It is not surprising that conflict

may result between the levels, nor should it be surprising that more than one evaluation may be required or that multi-purpose and independent evaluations must be conducted on a regular and periodic basis. Given that virtually all health and related projects are not financed from the personal resources of a single independent individual who is unconcerned with what happened, evaluation activity and related management and analysis of information must be an integral part of program activity. Operations research is necessary to ascertain what is the minimum information necessary to meet the multiplicity of evaluation requirements by all audiences such that correlations between changes in one indicator imply change of a particular nature in other, longer run indicators.

Timing: When to Measure

Just as there are many different audiences for evaluation and just as each proposed indicator is more useful to certain audiences than others, certain phenomena are more appropriately measured earlier or later than others. It is not appropriate to attempt to find changes in vital events measures within the first six months of a project. It may also be inappropriate to measure other indicators of programmatic activity at certain points throughout the life of a project. Further, as has been referred to above (see ftn 3) the potential for measurement error is great if an unappropriate moment for monitoring is chosen and is greatly reduced with a more continuous monitoring of all indicators.

If data on final outcome measures or other measures are demanded too early, the information yielded most likely will be inaccurate and

will lead to uninformed and thus possibly poor decisions.

In Table IA-ID each indicator is categorized by when it should be measured for evaluation and related decision-making purposes. Clearly, in order to measure the changes in many aspects of a project, the conditions which exist prior to the beginning of the project intervention must be ascertained by "baseline" survey or other means. Further, it is suggested that most measures be monitored more than once over the life of the project in order to obtain a better understanding of the nature of change through time. If many measures of short-run effects or of program execution are to be used for on-going administrative or managerial decision-making as well as used for proxies of final outcome measures (assuming appropriate research is conducted on these possible linkages), it is critical that they be monitored on a frequent and regular basis.

Finally, the issue of timing in measuring change in each indicator implies a certain minimum resource base be made available within the design of projects. Moreover, it is critical that the data gathering process be integrated into project planning from the point of inception.

Data Availability

While much of the information desired for measure/indicator development (particularly provider as opposed to consumer measures) can be obtained from a project's managerial information system, there are many other sources available. These other sources must be systematically reviewed before launching new efforts. Finally, it is important to realize that some information is only obtainable via periodic population-based survey techniques. In such cases, the cost of obtaining the

information must be reviewed in financial, real resource (e.g. manpower), and time dimension terms (see next section below).

Instructive, multi-purpose, household-based information systems are already in place in certain selected sites throughout the world. For example, there is a multi-purpose data collection effort funded by AID underway in the Bicol region of the Philippines which has incorporated health status and time allocation questions amongst many others.[22/] Further, there are similar systems being readied for use in Bolivia and other areas.

With respect to sources of information other than the project-specific information system, several ideas may be helpful. First, there are many government reports and documents available at the national, regional and local level. Second many non-governmental agencies and/or institutions may have similar reports and studies available. The universities and related research institutes, bureaus, or centers may have conducted studies on the issue under consideration. A particularly neglected source of information is student papers, theses and dissertations. Finally, in many countries ongoing consumer-based, household-interview surveys exist. Whether new survey organizations require development, or whether existing mechanisms can be tapped for use in obtaining consumer based program impact data, is a question requiring case by case determination. But it is clear that the creation of a new survey mechanism is not the only option.

Cost of Data Collection and Use

While it is often desirable to have as much information as possible, basic resource constraints require that every information

gathering and analyzing endeavor be subject to its own cost-benefit analysis. These costs are disaggregated into three basic categories: (a) estimate of financial resources required as a proportion of total project costs; (b) estimates of minimum manpower skills required to obtain reliable and timely information; and (c) time required to obtain the data and make it available for evaluation and decision-making uses.

Of particular interest is the extent to which (a) additional staff must be employed to gather data rather than direct service providers, (b) expatriate experts (researchers) must be involved in the design, management and analysis of the data; (c) specialized computer facilities are required; and (d) data can be timely gathered and analyzed for both programmatic/administrative decision-making and evaluation purposes related to the attainment of final outcome goals.

Summary

Our discussion implies that the evaluation process is one that is subtle, long-term -- one that requires integration with other health and development activities. It involves the management and development of information-gathering, coordinating, analyzing and disseminating systems. This document has not only reviewed the evolution of evaluation efforts in a similar programmatic activity, e.g., family planning, but it has also analyzed the past health project evaluation focus on vital event change and has suggested other foci, which are theoretically more consistent with the present socio-economic development goals as defined by the basic human needs strategy of development. It further outlines a multiple-purpose evaluation strategy which can be used as a guideline for many types of health evaluation efforts, depending on time, personnel, audience, data availability and other constraints. The guidelines are proposed as being realistic, i.e., meeting the test of field implementation requirements, as well as providing guidance for addressing, to the extent feasible, more general outcome issues and concerns.

Recommendations

(1) The field of population and family planning has engaged in program evaluation for many years. There are many lessons to be learned for health program evaluation from that program's past experience. It is recommended that a series of state-of-the-art papers reviewing the population field's experience be made available to

those working on health program development and evaluation on such subjects as evaluation methodology, field work strategy, indicators, and program impact.

(2) There is currently too much emphasis on defining success on the basis of vital events change while many other impacts of such projects go unnoticed. It is recommended that an alternative, more pluralistic approach to the definition of success be used. Such an approach has been developed in this paper focusing more on changes in consumer and less on investment measures. A number of alternative indicators have been proposed as additions to, or, in some cases, as substitutes for vital-events changes. A full appraisal of success in the health sector requires their incorporation.

(3) Many lessons can be learned from evaluative studies of past and ongoing health endeavors which can improve future activities. In order to improve such evaluative studies it is recommended:

a) that the management information systems in health projects should be made as compatible as possible in the design phase with short and intermediate run evaluation needs;

b) that multi-purpose, population-based survey instruments be designed and implemented in selected primary health care delivery projects to monitor the general welfare impact of basic human needs investments as measured from various perspectives on target populations; and

c) that measurement and evaluation activity be undertaken periodically, i.e., that several repeated observations be taken on each measure over a sufficiently long period that the dynamic nature of the changes can be ascertained.

(4) A complementary recommendation to (3) is that a systematic applied operations research agenda be developed which would focus its attention on cost-effectiveness analyses of alternative technologies for low-cost, primary-health care delivery systems. This recommendation follows the lead of the recent operations research contract lead by the Office of Population.

(5) Evaluation, analytical work, and information flow all imply additional resources be made available and that a long-run commitment be forthcoming not only from AID but also from its funder, the U.S. Congress. If the Congress wants to know what is successful and if research activities on health services are to be undertaken in as systematic a way as has been done in hopes of finding a cure for malaria, schistosomiasis, heart disease or cancer, there must be a willingness to pay. Such activities can be cost-effective. It is recommended that the necessary commitments be made.

(6) The log frame used in virtually all AID projects presently has several important flaws as it applies to health projects. In particular there is no way to incorporate: a) intermediate measures of impact, b) realistic expectations for achieving such changes, c) important assumptions and/or constraints on project activities, and d) an integration of effects/outcomes to define a "critical path" toward final outcome achievement. It is recommended that appropriate variations on the log frame be developed.

Table 1 A: An Analysis of Program Execution Measures for Audience, Timing, Data Sources, and Cost of Obtaining Data, Including Manpower and Time to Develop

Health Program Measures I Program Execution	I Audience	II Timing (When to Measure)	III Data Sources	IV Cost of Obtaining Data (A)	(B) Manpower	(C) Time to Develop Measure
	(a) Donor Orgs. (b) Host Cntry National (c) Host Cntry Regional (d) Host Cntry Local (e) Project Adm. (f) Project workers	(a) Prior to Project (b) Beginning of Project (c) Up to 6 mths (d) " " 1 year (e) " " 2 years (f) 3-5 years (g) Longer (Specify)	(a) Project Mgt. Info Sys. (b) Govt. Docs. (c) Govt. Files (d) Special Studies (e) Journal Articles (f) Survey Data	(a) less than $10 (b) $10 - $1,999 (c) $2,000 - $4,999 (d) $5,000 - $9,999 (e) $10,000 - $24,999 (f) $25,000 - $99,999 (g) > $100,000	(a) Local Clerk (b) Local Supervisory (c) Local Cntgr (d) Local Research (e) Expatriate Expert	(a) < 1 mth. (b) 1 - 3 mths (c) 3 - 6 mths (d) 6 - 12 mths (e) > 12 mths
(A) Consumer Indicators						
(1) Community Participation in project identification, design & impl. or involvement						
(B) Provider Indicators (1) Output Measures (i) Services provided (by type)						
(ii)						
(2) Input Measures (i) Inputs available for Service provision (ii) Personnel trained & Recruited.						
(iii) Transport available (iv) Job tasks & technology understood by personnel						
(v) protocols in existance (vi) (Administrative & Managerial Systems & Personnel in Place						

Table 1B: An Analysis of Intermediate Measures for Audience, Timing, Data Sources, and Cost of Obtaining Data, Including Manpower and Time to Develop

Health Program Measures / II Intermediate Measures	I Audience (a) Donor Orgs (b) Host Country National (c) Host Country Regional (d) Host Country Local (e) Project Aim (f) Project Workers	II Timing (Phasing Measure) (a) Prior to Project (b) Beginning of Project (c) Upto 6 mths (d) 6 mths - 1 year (e) 1 - 2 years (f) 3 - 5 years (g) Long term (Ex-post)	III Data Sources (a) Project Mgt Info Sys (b) Govt. Docs. (c) Govt. Files (d) Special Studies (e) Journals (Secondary) (f) Survey Data	IV. Cost of Obtaining Data. (A) (a) Less than $10 (b) $10 - $49 (c) $50 - $99 (d) $100 - $499 (e) $500 - $999 (f) $1000 - $4999 (g) $5000 - $9999 (h) $10000 and above (B) Manpower (a) Clerical (b) Supervisory (c) Local College (d) Local Research (e) Expert, International (C) Time to Develop Measure (a) < 1 mth. (b) 1 - 3 mths (c) 4 - 6 mths (d) 6 - 12 mths (e) > 12 mths
(A) Consumer Indicators				
(1) Perceived Reduction in Mortality & Morbidity in Target Pop				
(2) Satisfaction Measures				
(3) Changes in Health Behavior (Accessibility, Confidence in Providers, Preventive Orientation, Status)				
(B) Provider Indicators				
(1) No. of contacts in Community By Outreach Workers (disag. by type)				
(2) Access of Target Pop. to Clinics				
(i) Financial				
(ii) Distance				
(iii) Time				
(3) Geographical Coverage per stds.				
(4) Provider Satisfaction				

Table I C: An Analysis of Final Outcome Measures for Audience, Timing, Data Sources, and Cost of Obtaining Data, Including Manpower and Time to Develop

Health Program Measures — III Final Outcome	I Audience						II Timing (When to Measure)							III Data Sources						IV Cost of Obtaining (A) Data							(B) Manpower					(C) Time to Develop Measure				
	(a) Donor Orgs	(b) Host Cntry National	(c) Host Cntry Provincial	(d) Host Cntry Local	(e) Project Adm.	(f) Project Workers	(a) Prior to Project	(b) Beginning of Project	(c) Upto Ends	(d) in 1 year	(e) in 2 years	(f) 3-5 years	(g) Longer (Specify)	(a) Project Mgt. Info. Sys.	(b) Govt. Docs.	(c) Hosp. files	(d) Special Studies	(e) Journal/Studies	(f) Survey Data	(a) Less than $10	(b) $10 - $499	(c) $500 - $999	(d) $1,000 - $4,999	(e) $5,000 - $9,999	(f) $[illegible] - $[illegible]	(g) > $100,000	(a) Local Clerk	(b) Local Supervisory	(c) Local Cntrs	(d) Local Research	(e) Expatriate Expert	(a) < 1 mth.	(b) 1-3 mths	(c) 4-6 mths	(d) 6-12 mths	(e) > 12 mths
(A) Consumer Indicators																																				
(1) Utilization Rates Disag. by: (i) User charges (ii) program &/or Service																																				
(iii) Alternative Delivery Systems																																				
(2) Consumer Satisfaction																																				
(B) Investment Indicators (1) Changes in Vital Events																																				
(i) Infant Mortality (ii) Death Rate. (iii) Life Expectancy (iv) Morbidity, Rates (prevalence/incidence)																																				
- Restricted Activity Days - Bed Disability Days - Illness Episodes.																																				
(2) Other Indicators. (i) Employment status (ii) Work Impairment.																																				
(C) Other Indicators																																				
(1) Provider Satisfaction																																				

Table 1D: An Analysis of Constraint Measures for Audience, Timing, Data Sources, and Cost of Obtaining Data, Including Manpower and Time to Develop

Health Program Measures / III Constraints	I Audience: (a) Donor Orgs; (b) Host Country National; (c) Host Country Parliament; (d) Host Country Local; (e) Project Adm.; (f) Project workers	II Timing (When with Measure): (a) Priority Project; (b) Beginning of Project; (c) Upto 6 mths.; (d) 6 mths - 1 year; (e) 1 yr - 2 years; (f) 3 - 5 years; (g) Long run (Capacity)	III Data Sources: (a) Project Info Sys.; (b) Govt. Docs.; (c) Govt. Files; (d) Special Studies; (e) Univ. research/study; (f) Survey Data	IV Cost of Obtaining Data: (A) Data: (a) less than $10; (b) $10 - $499; (c) $500 - $999; (d) $1,000 - $4,999; (e) $5,000 - $9,999; (f) $10,000 - $49,999; (g) $50,000 - $99,999; (h) $100,000 >; (B) Manpower: (a) Local Clerk; (b) Local Supervisory; (c) Local College; (d) Local Research; (e) Expatriate Expert; (C) Time to Develop Measure: (a) < 1 mth.; (b) 1 - 3 mths; (c) 4 - 6 mths; (d) 6 - 12 mths; (e) > 12 mths
(A) Economic				
(1) Balance of Payments				
(2) Import Quotas				
(3) Foreign Exchange				
(4) Employment Growth				
(5) Govt. Financial Picture				
(6) Distribution of Income				
(7) Sectoral Commitments eg. Defense				
(B) Political				
(1) Distribution of BS & CS				
(2) Ideology Consistency				
(3) Political Leadership Capacity				
(4) Political Commitment (Budget)				
(5) Govt. Structure				
(C) Technological				
(1) Capital or Labor Bias (Recurrent cost)				
(2) Key imported inputs				
(3) Timing of Input Combination				
(D) Adm. & Managerial				
(1) Info. System				
(2) Logistics				
(3) Personnel				
(4) Financing & Budgeting				
(5) Leadership				
(6) Structural/Organization				
(7) Evaluation/Planning				

Pg. 2

Table I D: An Analysis of Constraint Measures for Audience, Timing, Data Sources, and Cost of Obtaining Data, Including Manpower and Time to Develop

Health Program Measures — IV Constraints (cont.)	I Audience: (a) Donor Orgs	(b) Host Country National	(c) Host Country Provincial	(d) Host Country Local	(e) Project Adm	(f) Project workers	II Timing (When to Measure): (a) Prior to Project	(b) Beginning of Project	(c) Up to 6 mths	(d) 1 in 1 year	(e) 1 in 2 years	(f) 3-5 years	(g) Longer (Specify)	III Data Sources: (a) Project Mgt. Info Sys	(b) Govt. Docs.	(c) Hosp. files	(d) Special Studies	(e) Journals/Articles	(f) Survey Data	IV Cost of Obtaining (A) Data: (a) Less than $10	(b) $10-$4,999	(c) $5,000-$49,999	(d) $50,000-$99,999	(e) $100,000-$499,999	(f) $500,000-$999,999	(g) $1,000,000	(B) Manpower: (a) Local Clerk	(b) Local Supervisory	(c) Local College	(d) Local Research	(e) Expatriate [illegible]	(C) Time to Develop Measure: (a) < 1 mth.	(b) 1-3 mths	(c) 3-6 mths	(d) 6-12 mths	(e) > 12 mths
(E) Biological/Environmental																																				
(1) Vector Prevalence																																				
(2) Initial Health Status of Pop.																																				
(3) Seasonality considerations																																				
(4) Cultural (specify)																																				
(F) Socio Economic Characteristics of Individuals in Target Pop																																				
(1) Education																																				
(2) Income																																				
(3) Access to Land.																																				
(4) Age																																				
(5) Ethnic origin																																				
(6) Sex.																																				

Footnotes

(1) See for example the paper by David Gwatkin, "Food Policy, Nutrition Planning and Survival: The cases of Kerala and Sri Lanka," Food Policy, (Nov. 1979) forthcoming.

(2) For a more complete analysis of the direct and indirect effects of human resource programs, see Vol. 2, pp. 20-93 or Family Health Care, Planning for Health and Development: A Strategic Perspective for Technical Cooperation, Report to USAID (Washington, D.C.; Family Health Care, July 1979).

(3) For a more complete treatment of such timing issues, see Lawrence W. Green, "Evaluation of Measurement: Some Dilemmas for Health Education, American Journal of Public Health, 67, 2 (February 1977), 155-161.

(4) Page 1 and 2, Larry Heligman, et. al., Measurement of Infant Mortality in Less Developed Countries, International Research Document #5 (Washington, D.C.; U.S. Department of Commerce, U.S. Bureau of the Census, August 1978).

(5) See for example, W. Brass, A. J. Coale, P. Demeny, D. F. Heisel, F. Lorimer, A. Romanick and E. Van de Walle, The Demography of Tropical Africa, (Princeton, New Jersey: Princeton University Press, 1968).

(6) For a useful review and critique of the human capital approach. See Jan Acton, "Valuing Life Saving Alternatives and Some Measurements", Journal of Law and Contemporary Problems, (Winter 1976); Jan Acton, Evaluating Public Programs to Save Lives: The Case of Heart Attacks, (Santa Monica, California: Rand Corp., 1973) for a constructive analysis of alternative approachs to benefit measurement and a more consumer-oriented approach to the issues raised in this section. See Selma Mushkin and David Dunlop, eds., Health: What's it Worth? (New York: Pergammon Press, 1979).

(7) Budd Hall, "Mtu Ni Afya" (Health Education Campaign in Tanzania), Paper 4. (Institute of Adult Education, University of Dar Es Salaam, Dar Es Salaam, Tanzania, January, 1973).

(8) For an explanation of the purpose, scope, and methodology of the U.S. National Health Interview Survey, See NCHS, Health Survey Procedure: Concepts, Questionnaire Development and Definitions in the Health Interview Survey, Series 1, No. 2, (Washington, D.C., U.S. Dept. of HEW, Public Health Service, 1958), 66 pp.

(9) See John Ware and Joanne Young, "Issues in the Conceptualization and Measurement of Value Placed on Health," pp. 141-167, in Selma Mushkin and David W. Dunlop, eds., Health: What is it Worth? (New York: Pergammon Press, 1979).

(10) Ed Clark, "Social Valuation of Life-and Health-Saving Activities by the Demand-Revealing Process", pp. 69-91, in Selma Mushkin and David W. Dunlop, eds., Health: What is it Worth? (New York: Pergammon Press, 1979).

(11) See for example John Cleave, Africa Farmers: Labor Use in the Development of Smallholder Agriculture, (New York Praeger Publishers, 1974); H. A. Oluwasanmi, et.al., Uboma: A Socio-economic and Nutritional Survey of a Rural Community in Eastern Nigeria, Occasional Paper of the World Land Use Survey, No. 6 (Bude, England, 1966); and D. W. Norman, Economic Analysis of Agricultural Production and Labour Utilization Among the Hausa in the North of Nigeria, Rural Employment Paper No. 4, African Rural Employment Study (East Lansing, Michigan; Michigan State University 1973).

(12) Unpublished Monthly Data From Uganda Health Facilities available from David W. Dunlop. See also paper by _____ on the relationship of seasonal changes and disease incidence in human populations, from the Institute for Development Studies, Sussex University, the U.K. 1979; and the study entitled Health in Regionville.

(13) A highly illuminating study of this phenomena was conducted by John Hunter, "Seasonal Hunger in a Part of the West African Savanna. A Survey of Bodyweights in Nangodi North East Ghana," Transactions and Papers 41 (1967) 1671-185.

(14) PP. 2, AID, A Strategy for a More Effective Bilateral Development Assistance Program: An AID Policy Paper (Washington, D.C.: USAID, March 1978).

(15) For an instructive elaboration on the distinction between non-programmatic or situational constraints and intra-programmatic or policy constraints, see James S. Coleman, Policy Research in the Social Sciences (Morristown, N.J.: General Learning Press, 1972) pp. 4 and following.

(16) See for example David W. Dunlop, The Economics of Uganda's Health Service System: Implications for Health and Economic Planning, unpublished Ph.D. Dissertation, Michigan State University, East Lansing Michigan, 1973; and F. Grundy and W. Reinke, Health Practice Research and Formalized Managerial Methods, Public Health Papers, #51, (Geneva: WHO, 1973).

(17) WHO, The Evaluation of Family Planning Activities. Conducted in Health Services, Technical Report Series #569 (Geneva: WHO, 1975).

(18) P.S. Mohapatra, "Measuring the Performance of Family Planning Programs: Three Essays," pg. iii, N.D. for an earlier conceptualization of the same ideas, see C. Chandrasekaran and Moye W. Freymann, "Evaluating Community Family Planning Programs," pp. 266-286, in Mindel Sheps and Jeanne Ridley, eds., Public Health and Population Change, (Pittsburgh: University of Pittsburgh Press, 1965).

(19) P.S. Mohapatra, ibid pg. iii, N.D. for a similar view about evaluation chain linkages see also WHO, The Evaluation of Family Planning Activities ..., (1975), op. cit.

(20) P.S. Mohapatra, ibid, pg. iii, N.D.

(21) James Veney, "Health Program Evaluation in a Developing Nation: An Evaluation Methodology Case Study," unpublished paper prepared for discussion, Department of Health Administration School of Public Health, University of North Carolina, 1979, pg. 31.

(22) For an introduction to potential uses, importance, and methodology of the Bicol multipurpose survey, See Barry Popkin, et. al., 1978 Bicol Multipurpose Survey Vol I: Survey Design and Implementation, (Chapel Hill, North Carolina: Carolina Population Center, 1979).

File: Health Proj' Evaluation

June 18, 1980

To: Mr. John Evans
Room N 437

I talked to you about working on this paper in May (when we were attending the Inter-American Development Bank's seminar on Health Project Impacts). Various things diverted me from the paper, but when I finally concentrated on it, I felt the need of slightly extending the scope of the paper from what we talked about. Anyway, this is the very first draft; I need to do a lot more work on it. I shall appreciate receiving your comments and suggestions.

From: R. Faruqee
Room I-8-1234
Ext. 61261

FIRST DRAFT
For Comments Only

IMPACT ANALYSIS OF HEALTH SERVICES: THE NARANGWAL EXPERIENCE

CONTENTS

FIRST DRAFT
R. Faruqee

IMPACT ANALYSIS OF HEALTH SERVICES: THE NARANGWAL EXPERIENCE

Introduction

Impact analysis of any project in social sectors, such as health and education, is difficult because the project operates in an environment affected by many nonproject forces. 1/ Observed changes, if measurable, cannot often be precisely ascribed to the various forces causing the changes.

Traditionally, two types of impact analyses are done: ex ante and ex post. Ex ante analysis is needed to determine the overall desirability of the project and to compare its merit with other projects. Ex ante analysis constitutes the key element in the project preparation and appraisal. Ex post analysis is needed to monitor and assess whether anticipated benefit of the project matches actual benefit. Ex post analysis constitutes the key element in the project evaluation.

In physical sector projects, ex ante impact analysis is often straight forward: economic analyses of internal rate of return, benefit/cost ratios and cost effectiveness criterion are used. These calculations are, of course, made on the basis of certain assumptions and definitions, which often prove inadequate or inaccurate by experience. However, the overall conceptualization is straight forward. The cost of the project is considered investment for the expected outcome in the form of output (income) growth. The analysis has to focus on (a) inter-sectoral efficiency (for the scale of the project/program), (b) intra-sectoral efficiency (for the selection of design and objectives of the project), and (c) intra-project efficiency (for the selection of the particular components of the project). Since all these dimensions are analytically identifiable and measurable under given assumptions, an economic analysis proves useful for investment decisions.

1/ In a recent two-day workshop on Health Project Impacts, organised by the Inter-American Development Bank, this point became abundantly clear. The workshop concluded that while considerable research is focusing on ideological and methodological issues, a workable guideline does not exist for health project preparation and appraisal.

In the social sector projects, certain crippling problems arise. First, most social sector projects have both consumption and investment aspects; this creates a problem for impact analysis, because consumption benefits from social services are subjective and hard to measure. Education, health, and family planning services are desired by consumers for the utility they derive from the use of these. Yet, these services also improve productivity or reduce fertility, contributing to output growth. In some sense, the objective function of social sector projects is to increase welfare, and not merely to increase output or income. This gives rise to the problems of measuring nontangible aspects of welfare and comparing welfare of one individual or group with another.

The second problem flows from the first and involves the political considerations inherent in the decisions about social sector projects, making impact analysis inapplicable. Because there are more undefined project goals, political decisions often determines the size and distribution of the investment funds in social sectors. Although politics enter into physical sector projects--the difference is in the nature and magnitude. In physical sectors, the degree of freedom for the politicians is much more limited than the social sectors. For example, the location of a hydroelectric project is limited by the nature of the river's course Conversely, there is more freedom involved in the decision to locate a hospital. On the other issues a project design such as the size and components--the restrictions in a physical sector project imposed by economic analysis is even more binding than in a social sector.

Third, although a few social sector projects have a single goal, such as malaria eradication programs, social sector projects usually have multiple goals. The situation is just the opposite in most physical sector projects.

The multiplicity of goals in social sector makes both ex ante and ex post analysis highly complex.

This short paper provides a general overview and simple guideline to those who are concerned about preparing, appraising and evaluating health projects. The paper reviews the state of the art, examines the experience of a research project based on field experiments at Narangwal, Punjab, India and summarizes the results to help the design of health projects both for (ex ante) appraisal and (ex post) exaluation.

The State of the Art

Use of Economic Analysis: Dunlop (1980) has recently reviewed the problems of using economic analytical techniques in health projects and also has noted the meager progress, mostly in the context of the developed societies, in economic analysis of health project impact. Major techniques of economic analysis of projects--benefit-cost analysis, cost effectiveness analysis, linear programming, and macro simulation -- have been used in health projects. The problems of using the analytical techniques are discussed in this section.

Benefit-cost analysis is a common technique used in any resource allocation exercise. There are, however, serious problems in using it in the health sector projects, similar to any other social sector. The most important problem is the conceptualization and measurement of health benefit and, to a limited extent, measuring problems relating to cost. Since this is at the heart of the problem, the next section focuses on the issues of health status indicators. However, the work of Katz et al. (1963) in the past, and Densen (1978) and Mushkin (1979) more recently, have attempted to come in grips with the theoretical issues of measurement and provide guidelines for continuation of functional health index. These measures only relate to the United States and are illustratively applied to specific cases, of such as ambulatory and nursing home care.

While indicators of health outcome are many and complex, valuation of the outcome for benefit-cost analysis is also a serious issue. Without valuation, benefit cannot be compared with the cost in one health project or with benefits across several projects. The most common way to place a value on health benefits so far has been the human capital approach, which measures economic benefit by the discounted value of output foregone. The foregone output has either been due to premature death or to a lower level of productivity resulting from illness. One obvious problem with this approach is that life is valued equal to the marketable output one can produce.

Clark (1979) and Tullock (1976) have, therefore, attempted to approach valuation by "the-willingness-to-pay" criterion. This is essentially a process to reveal demand. Surveys have been used for this purpose with limited success. In fact, no viable method exists to get to "the-willingness-to-pay" criterion of valuation.

Because of intractable problems of valuation of the health outcome, cost effectiveness is often advocated in place of benefit-cost analysis. Cost effectiveness is only a partial substitute for benefit-cost analysis, because cost effectiveness is only a search for cost numinimization for a given output. This does not allow a comparison with cost, or a comparison across projects of different kinds of outcome. In addition to the valuation problem, benefit-cost and cost effectiveness analyses also have the limitation of a micro focus, which may become unsuitable when the health sector is dealt with in a macro framework. Also, benefit-cost analysis is limited in finding proper weights for objectives like equity on the benefit side, and in shadow pricing unpaid family labor and nursing services of family members on the cost side.

Other techniques like linear programming and macro simulation require considerable data. Since measurement of output remains a serious problem in

case of health program evaluation, the data problem is even more serious for health projects and programs for these techniques to be used. Data of dubious quality, if used, can produce meaningless results. Barlow (1968) and Sheldon et al. (1970) have attempted to use macro simulation and operations research techniques to health program evaluation with only limited success.

Health Status Indicators and their Change: Since the international conference on primary health care in Alma-Ata in 1978, the declared goal, "health for all by the year 2000", is being actively pursued by the World Health Organization (WHO). The declaration commits developed and developing countries to improve health equity by broadening health and economic development. However, many issues remain unresolved.

WHO has generated a debate on the definition of health. What exactly is meant by health or what aspects of health should be emphasized are the issues being raised. A definition for the term "for all" is also being sought. The goal has emphasized the role of primary health care. The idea is to reach more people than has been possible in the past with a curative health structure, but the discussions of the concept of "health for all" have underscored the need of suitable indicators of health and monitoring their change to see improvement. These discussions have relevance for the impact analysis, because the concept of "health for all" has to be operationalized in terms of the indicators and their change over time. Several issues require discussion.

First, the debates and controversies have widened our understanding of health indicators and of the ways to monitor the progress in the indicators. But developing an understanding of these indicators has hardly helped the project people because no attempt has been made to relate the knowledge to the operational use in project preparation. So, there is a significant advancement in conceptualizing health status indicators and even in developing approaches to measure these indicators. Yet, few practical guidelines have

for one who is struggling with ways of justifying and evaluating health projects.

Second, the nature of controversies indicates that there are essentially two points of view. Some favor simple index of health; and others point out that such simple index are not only hard to construct, but can be counterproductive because they can be misleading about the true status of health.

Third, most of the discussion on the health indicators arising from the declared goal of, "health for all by the year 2000," actually focuses at the national level within the framework of intercountry comparison. For a health project constituting only a part of the overall health program of a country these discussions may not totally apply. However, the basic points regarding the problems and complexities of constructing health indicators and monitoring their changes remain relevant for the analysis of impact of health services as a part of health project preparation.

Fourth, it has been recognized that health services constitute only some of the factors determining health. That is, health system is a broader concept than health services, or the health delivery system.

From the discussions that have taken place thus far and the papers that have been prepared on the concepts of health indicators and their monitoring, four important categories of indicators to measure health and its improvement stand out.1/

The first category includes indicators reflecting health environment--what WHO has termed as social and economic conditions relating to health--and health policy measures. The second category includes inputs on health care services that are rendered to change health status. The third category

1/ WHO in a recent paper, 'Indicators for Monitoring Progress Towards Health for All' has proposed four categories of health indicators: 1. Health Policy Indicators; 2. Health Status Indicators; 3. Social and Economic Indicators Relative to Health; and 4. Indicators of the Provision of Health Care.

includes the output indicators that are actual changes in the utilization rates of health services. The last category includes outcome indicators reflecting long-run changes in health status. The distinction between the third and fourth categories lie not only in the time framework but also in what each implies for the health status. The third category has more of those indicators that reflect the change in the utilization of health services rather than the benefit from the utilization whereas the fourth category more directly reflects changes in the improvement in health status.1/

Categorizing important variables for impact analysis of health projects suggest that a few environmental factors should be considered in evaluating the changes resulting from the health services.2/ These are predisposing variables that would include the socioeconomic conditions affecting health at the beginning of the project or the program. Then, there are policy variables that may not be totally related to the particular health project being evaluated but would nevertheless effect the total outcome that would follow from such a project. These are also predisposing variables and will fall in the category of health environmental variables.

The second category includes measures of the magnitude of services offered and utilized that help understanding of the size and nature of the intervention offered. Since all health services are not provider initiated, it is at the same time an indicator of the level of health or lack of it. Health projects can be of different types. First, health projects can have a

---

1/ In a recent paper, Dunlop (1980) has offered a somewhat similar framework for evaluating health program. He has suggested that the evaluation framework should consider four categories of variables: 1. constraint, 2. program execution measures, 3. intermediate outcome measures, and 4. final outcome measures.

2/ When the discussion focuses on a country situation about health, these environmental factors themselves become an indicator about the health situation in the country. However, when evaluating impact of a health project or a program, these factors are to be examined as predisposing variables.

specific objective like eradication of malaria, immunization against measles, and so on. This can be country-wide in scope, but since it has a specific objective, monitoring its progress is straighforward and involves following up of key indicators relating to the specific objective. The second type of health projects may be national health programs and projects to pursue the overall health goals. In this case, the different kinds of indicators have to be used. The third type is a specific health project for a specific area covering different kinds of services often using specific target groups, multiple health and welfare objectives and so on.

The scope of the health indicators has to be carefully defined considering the type of health project being examined. The health environmental variables for an overall health program should include socioeconomic variables such as per capita gross national product, income distribution level, availability of food, proportion of population with access to adequate water and sanitation, literacy rate, female education, and so on. In considering a specific area-based health project area socioeconomic indicators that are part of the overall system of the economy should be examined. These indicators should include a political commitment to health, the resource allocation level, the degree of equity of distributional health resources, and the nature of community involvement in the program. In the Narangwal experience the predisposing variables included income, occupation, education, and so on. In India, caste is a stable indicator of socioeconomic status and was used as a predisposing variable to determine how it affected health services.

Two aspects of the health care services are important in examining the input variables. One is the access. For example, the average distance of households from service centers is an important consideration. The second aspect is the availability, which is somewhat different from access. Even with a clinic within a certain distance, the facilities may not always be available.

Physical accessibility of services is important, but socioeconomic and cultural aspects also are involved in accessibility. Physical accessibility is often defined in terms of distance or time needed to travel. Again, this would depend on particular environment being considered. For example, in certain areas distance may be short, but physical accessibility may still be difficult. Economic accessibility includes the ability of the individual or the community to cover the cost of services. If the service is available but the community or the individual cannot afford to pay for it, then it is not economically accessible. Cultural accessibility implies that using the available services is affected by cultural reasons. For example, in certain cultures, health services offered by male workers are not utilized by female clients. In this situation, only health services offered by female workers would ensure accessibility for female clients.

Output variables constitute the third category of impact analysis variables for health services. They relate to the utilization of services, which actually represents the effective coverage of services. These are represented by the proportion of people in need of service who actually receive it within a given time, often 1 year. The indicators for such utilization are the proportion of children at risk immunized and the proportion of pregnant women who received ante-natal care or have their children under the supervision of a trained attendant.

Finally, outcome variables reflect the final impact of health services. This variable may reflect overall health status in a population or may include the disease specific variables, which may or may not reflect the overall health status. Since the conference in Alma-Ata, the discussions about health indicators have brought out the importance of one particular variable, namely infant mortality rate as a key indicator of health status.1/

1/ James Grant (1980) has suggested an index of the quality of life, in which infant mortality rate figures most prominently.

The Narangwal Experience

At Narangwal, Punjab, India, a field experiment was carried out by the Department of International Health, Johns Hopkins University, between 1968 and 1974. In the experiment, groups of villages were provided with various combinations of health, family planning, and nutrition services; the households in each group were observed over time. The World Bank is collaborating with Johns Hopkins University to analyse the Narangwal data to study outcomes of alternative intervention strategies. Two monographs on the results of the Narangwal research results are available. Although the focus of these monographs are different from this paper, the monographs are relevant to the question of measuring the impact of health services.1/

The Narangwal field experiment actually consisted of two action research projects: one was the Narangwal population study, which was an in-depth examination of the outcome of integrating health with family planning, and the other was the nutrition study, which considered the interaction between malnutrition and infections in weaning-age chidlren.

In the population study, there were five experimental groups of villages matched as closely as possible for comparability. Each group of villages received a different service package, as follows:

1. Family planning, women's services, and child care services (FPWSCC).
2. Family planning and women's services (FPWS).
3. Family planning and child care services (FPCC).
4. Family planning education (FPed), and
5. Control group (CONT-P).

In the nutrition study, which covered all children under three years of age, there were four experimental groups of villages, receiving the

1/ This section is based on the two monographs: (1) Integration of Family Planning and Health: The Narangwal Experience, and (2) Malnutrition, Growth and Development: The Narangwal Experience. For details of the Narangwal research results, see these monographs.

following service inputs:

1. Child care services: nutritional supplementation and health care (NUTHC).
2. Nutritional supplementation (NUT).
3. Health care (HC), and
4. Control group (CONT-N).

One experimental group, namely FPCC or NUTHC overlapped in the two studies (See Figure 1).

The Population Study: The primary focus of the Narangwal population study has been to understand whether and how integration of family planning with a different component of health care will increase the acceptance of family planning and lead to a decline in fertility. However, the effects of services on health status were also analyzed.

When prospects of measuring outcome such as changes in the health status are limited, inputs including numbers of services provided or worker/facility and target population ratio are often used. Since the Narangwal experiment included control groups that provided reference points for comparing outcomes, the impact analysis was designed to focus on output and outcome variables. However, as a part of another study project,1/ a systematic observation and recording of workers' activities, using an appropriate sampling framework, was made. Types of information collected included the amount of time spent by each category of worker carrying out his or her activities and the distribution of various activities and tasks among workers, or an overall basis for the service package as a whole. Table 1 presents the overall time inputs of the health worker in the various experimental groups and by the types of services offered. Although such results can be useful to show trends, they do not indicate the impact of services. Moreover, collecting

1/ This is known as a Functional Analysis Study and involves data collection about activities of workers, analysis of service records, and surveys of utilization of health services.

Figure 1

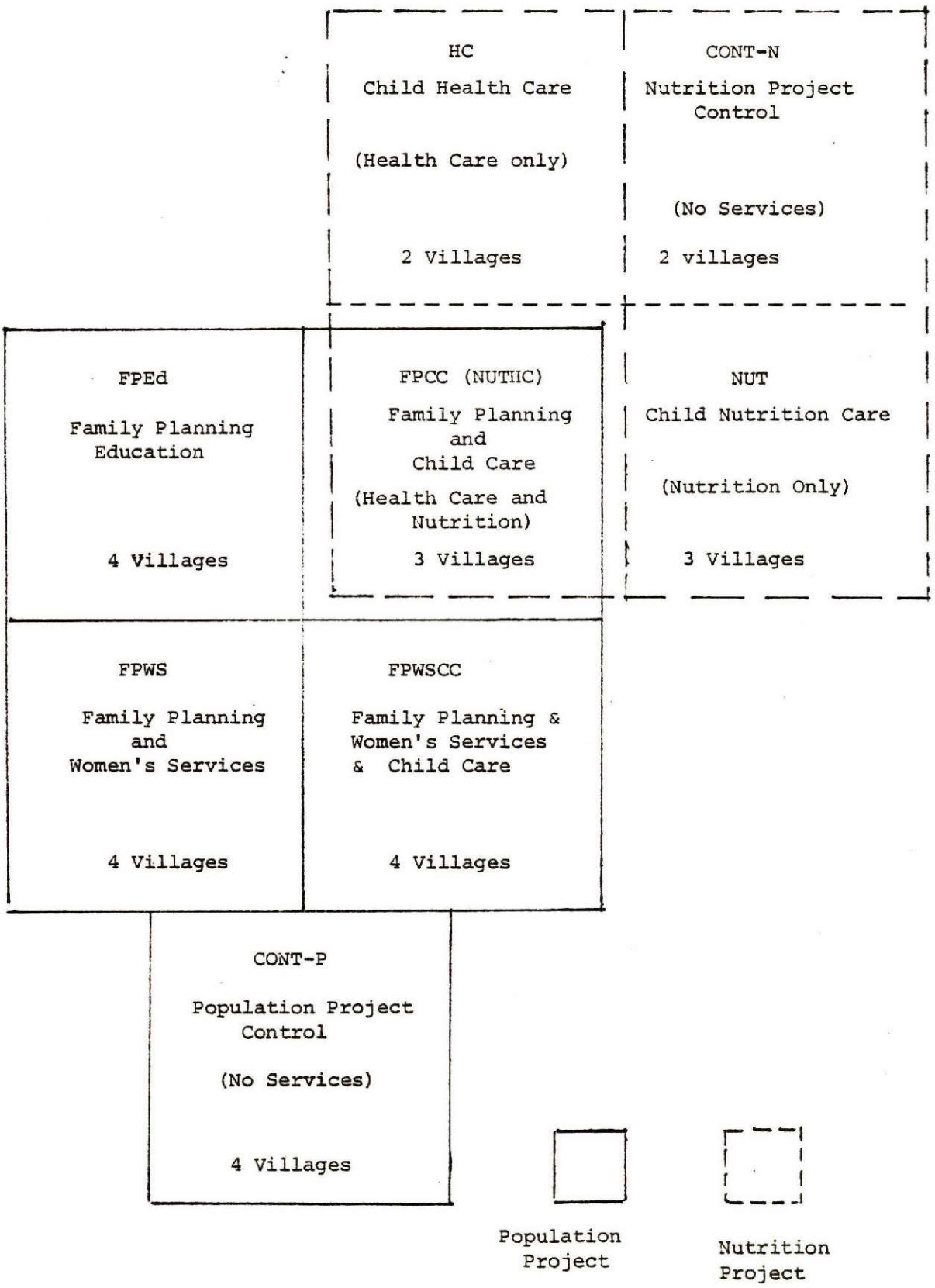

HC
Child Health Care
(Health Care only)
2 Villages
CONT-N
Nutrition Project Control
(No Services)
2 villages
FPEd
Family Planning Education
4 Villages
FPCC (NUTHC)
Family Planning and Child Care
(Health Care and Nutrition)
3 Villages
NUT
Child Nutrition Care
(Nutrition Only)
3 Villages
FPWS
Family Planning and Women's Services
4 Villages
FPWSCC
Family Planning & Women's Services & Child Care
4 Villages
CONT-P
Population Project Control
(No Services)
4 Villages
Population Project
Nutrition Project

Table 1

TIME SPENT PERFORMING DIRECT SERVICE ACTIVITIES OF HEALTH WORKER IN EXPERIMENTAL GROUPS, 1974
(Minutes per Week per 1000 Population)

	DIRECT SERVICES		FPWSCC	FPWS	FPCC
CHILDREN	Illness	1 yr.	25	4	116
		1-2 yrs.	21	-	73
		3+ yrs.	46	3	65
	Preventive	1 yr.	16	3	42
		1-2 yrs.	12	-	22
		3+ yrs.	1	-	2
	Nutrition		13	-	1
	Subtotal		134	10	321
WOMEN	Illness	50 yrs.	127	114	33
		50+ yrs.	46	29	3
	Preventive		11	34	15
	Antenatal		24	82	15
	Labor and Delivery		2	5	-
	Postnatal		17	9	3
	Subtotal		227	273	69
MEN-Illness			11	4	4
RAPPORT			32	45	39
TOTAL			441	415	490

this type of information accurately is difficult except by having observers follow workers over a long enough time to get a representative sample of the total work pattern. Detailed information of this sort is ruled out except in the case of a special field study.

Five output and outcome measures for health and nutrition services were used: Service contracts or coverage, the volume (intensity) of use, mortality effects, morbidity effects, and children's nutritional status. First, service contacts or coverage, both at home or clinic, were obtained from the detailed individual patient or service records. These were then standardized by expressing these visits per week per 1000 population. The paramedical workers were responsible for nearly 90 to 95 percent of health service contacts, which fell broadly in two categories -- children's health services and women's health services. In a project area, the availability of nonproject services add a complication to the determination of project impact. In the Narangwal analysis of service contacts, the proportion of individuals who received some form of health care was included and the source of such care was identified to indicate the relative role of the project source.

Table 2 summarizes the contacts made by all staff in FPWSCC, FPWS and FPCC villages from 1969 through 1973. The table shows that the volume of services peaked in 1970-71. Such time series of volume of services can be helpful for understanding the general trend and provide insights about certain shifts. In the table, child care contacts drastically dropped from 1971 to 1972, indicating either a shift in emphasizing of service coverage or a decline in the need for services.

The emerging pattern of service coverage in relation to needs shows that more intensive service villages (FPWSCC) achieved generally more coverage. Percentage differences of coverage are as follows: 60 percent of the ill children

Table 2

AVERAGE NUMBER OF RECORDED TOTAL SERVICE CONTACTS IN EACH EXPERIMENTAL GROUP PER 1000 POPULATION PER WEEK

Experimental Groups/ By Services	Years 1969	1970	1971	1972	1973
1. FPWSCC					
a. Women's Services	23	49	52	51	49
b. Child Services	57	70	73	47	36
2. FPWS					
a. Women's Services	27	63	61	56	57
b. Child Services	-	-	-	-	-
3. FPCC					
a. Women's Services*	2	11	12	13	14
b. Child Services	81	109	100	94	66

*These are mainly pregnancy surveillance visits.

in the FPWSCC villages and about 50 percent in FPCC received some type of care, as against 30 to 40 percent of ill children in control villages or in the villages with no child care services. The percent of ill women who were treated ranged from 26 percent in control villages, 35 percent in FPEd, 38 percent in FPCC, 42 percent in FPWSCC, to 47 percent in FPWS. These percentages also help to see the total coverage effect of project services, compared to nonproject area. But they do not fully capture the relative contribution of the project services, which often substituted for the non-project services. If the project services are expected to be better in quality, the substitution effect should indicate an improvement.

To probe the substitution effect, how much of the total actual coverages came from project sources was estimated. This involved examining what proportion of the children and women already receiving health care used the project services. Project services constituted the primary source for the child health care in FPWSCC and FPCC villages, where Narangwal staff accounted for more than two-thirds of the care provided. It appears that alternate sources of care, primarily indigenous private practitioners, continued to be provided to one-third of the children. This underlines the continued significance of traditional sources of care, despite the availability of modern, easily accessbile services. In FPCC villages about half of the child care services earlier used was substituted by project sources.

The effects of project services on overall use of services by women in FPWSCC and FPWS were very similar. About 30 percent of ill women received care from the project in both of these experimental groups; this care represented between 60 and 70 percent of the total care received. Again, other sources of care were used at about half of the level seen in control and FPEd villages. Providing some care to ill women in the FPCC groups (mostly symtomatic treatments to maintain the project) produced an intermediate picture, but the few project-sponsored child care services in FPWS villages

did not modify the pattern of using other child care sources.

Second, the volume or intensity of use was examined. The average number of "illness" visits for each woman using such services in 1969 was 4.0 in FPWSCC villages and 5.1 in FPWS villages. The results show on average that each woman required more visits, when women's health services were provided without child care than when combined with child care. In 1970 women in FPWS averaged 5.9 contacts compared with 5.4 per woman in FPWSCC. The average number of contacts per woman receiving other care--including monitoring of fertility and pregnancy, pre- and post-natal care, and supervision of deliveries--was almost identical in FPWS and FPWSCC.

The average number of children illness care (clinic) visits per year was about 10 in FPWSCC and 7 in PFCC, while "other care" (both home and clinic) visits (for periodic check up, immunization, nutrition supplementation, and so on) were slightly less than 20 per year in FPWSCC and 40 per year in FPCC. These show different focus of child care depending on whether child care is combined with women's services (see Table 3). We find that the average number of contact for other care for children in FPCC group is almost double in FPCC group that of the integrated service group. (FPWSCC).

Third, mortality effects were analysed. Although we had data on all deaths, we only used the death rate of children under 3 for our analysis. We did this because only in the case of children under 3 were deaths of sufficient number to make analysis possible and the children's health services largely concentrated on children under 3. An important measure of the effect of women's services would have been changes in maternal mortality rates, but the analysis of the maternal mortality indicator was not feasible because there were few cases of maternal deaths.

Table 3

AVERAGE NUMBER OF VISITS OR CONTACTS AMONG WOMEN WHO UTILIZED WOMEN'S AND CHILDREN'S HEALTH SERVICES

Type of Visit	Experiemntal Group	Average No. of Visits 1969	1970
Women's Illness Services	FPWSCC	4.0	7.7
	FPWS	5.1	10.8
Women's Other Services	FPWSCC	4.9	5.4
	FPWS	4.1	5.9

Type of Visit	Experimental Group	Average No. of Visits During 1969-71
Children's Illness Services	FPWSCC	10
	FPCC	7
Children's Other Services	FPWSCC	20
	FPCC	40

Mortality rates were calculated for children by age -- still birth, infant mortality (disaggregated into infant deaths under 1 month and those from 1 to 12 months), and children's deaths between 1 to 3 years (see Table 4). Overall, the still birth rate was 57 percent per 1000 live and still births in the control villages and was 23 to 35 percent lower in the services villages. This roughly measures the probable effect on the fetus of prenatal care for the mother, especially the effect of providing iron and folic acid to all mothers and nutritional supplementation to poorly nourished mothers.

The mortality effects of services were seriously affected by the caste variable. Comparable effects were produced in both FPWSCC and FPCC in the low caste group, but only in FPCC group, did some effect occur among the high caste families. This implies that more intensive child care services of FPCC could cover all caste groups, while FPWSCC workers, with less available time, may have concentrated their efforts involving prenatal care and nutritional supplementation on low caste mothers.

Infant mortality rates were disaggregated into neonatal death rates (first month of life) and post-neonatal death rates (1 to 12 months). The difference in the neonatal death rates was significant; in control villages the rate was 78 per 1000 live births, as against 64 in FPWSC and 47 in FPCC. It is interesting to note that this variable was affected differently by the caste variable. In terms of neonatal survival rates, the high caste children in FPWSCC group benefitted from the services, but not the low caste group. This was in contrast to FPCC, where the low caste groups appeared to receive the maximum benefit, a reduction of 43 percent. Apparently, in villages where services were less intensive in their outreach (FPWSCC), high caste families actively sought care from the project or other sources and achieved moderate reductions in neonatal mortality. However, low caste children benefitted only in those villages (FPCC) where home visits were much more intensive (weekly home visits in FPCC compared to monthly visits in FPWSCC

Table 4

EFFECTS OF CHILDREN'S SERVICES ON MORTALITY RATES BY AGE, CASTE AND EXPERIMENTAL GROUP (1970-73)

Experimental Group	Caste	Stillbirths*		Infant Mortality **				Child (1-3) mortality***	
				< 1 month		1-12 months			
		High	Low	High	Low	High	Low	High	Low
a. Control Villages †		47	71	81	84	52	52	7	26
		(57)		(78)		(51)		(19)	
b. FPWSCC Villages		52	39	58	86	40	66	3	10
		(44)		(64)		(54)		(7)	
% Difference ††		+10.6	-45.1	-28.4	+2.4	-23.1	+26.9	-57.1	-61.5
		(-22.8)		(-18.0)		(+5.9)		(-63.2)	
c. FPCC Villages		27	37	67	48	28	33	12	18
		(37)		(47)		(34)		(13)	
% Difference ††		-63.8	-47.9	-17.3	-42.9	-46.2	-36.5	+71.4	-30.8
		(-35.1)		(-39.7)		(-33.3)		(-31.6)	

Note: "High" caste were the Jat Sikhs, the landowning farmers, and "Low" caste were the scheduled Sikhs, predominately landless laborers. These castes made up between 75-85 percent of the population. Total mortality rates including other castes are shown in parentheses.

* Rates are per thousand live and stillbirths combined.

** Rates are per thousand live births

*** Rates are per 1000 children 1-3 years of age.

†† $\frac{b(\text{or } c) - a}{a} \times 100 = \%\text{-Difference.}$

† Combined rates from the control villages of the Population and Nutrition Studies.

after the immediate post partum period). Basically the same pattern held for children in the post-neonatal period, with high caste children in FPWSCC and all children in FPCC receiving significant benefits from the child care services.

For evaluating project impact, there was trade off between the qualitative and quantitative impact of child care services when combined with women's health services or offered separately. If, combined with women's health services, less intensive home visits are made, the higher caste group benefit more than the lower. If more intensive home visits are made with less extensive services, then more coverage of the low caste is possible. This is different from the other indicator, still birth, where substantial improvement is done in the high caste group in FPCC villages. In the case of infant mortality, the impact is felt among the low caste group only in FPCC (and not in FPWSCC). This implies that still birth as an indicator is more difficult to influence than infant mortality, given the caste situation. The improvement in still birth incidence comes slowly; only the high caste group with concentrated child care services attained a significant fall in that incidence. Both neonatal and post neonatal rates can be more significantly influenced by intensive services. If, however, services are more extensive with less home visits, higher caste groups benefit more.

Although deaths among 1 to 3 years old were fewer in number, the impact of services on such deaths was equally important as in the case of deaths at an earlier age. The 1-3 year mortality rates in FPWSCC and FPCC were 7 and 13 per 1000 children compared with 19 in control villages. In this the impact was greatest in FPWSCC villages (63 percent lower than controls) and services had equal impact among all castes. Interestingly enough, services for older children in FPCC had no effect on high caste children.

This may be reflecting the fact that high caste older children did not avail the nutrition suplemental program. Nutrition care was most effective in the 1-3 years of age group.

It is clear from the discussion above that in the Narangwal context caste proved to be an important variable in understanding the impact of services measured through mortality rates, still births, infant mortality (both neonatal and post-neonatal), and child mortality (1 - 3 years). Caste differences indicate the relative socioeconomic status of the households, which is a strong predisposing variable influencing the impact of child health services.

To illustrate the point that predisposing variables are important, we note here that in-depth interviews of mothers whose children died provided interesting insights that possibly explain some of the differential impacts of services on death rates among different castes. A few children of either high or low caste did not receive treatment for illness that led to death. However, high caste families tended to seek care earlier than low caste families. Sixty eight % of high caste children dying received care in the first 24 hours of their illness compared with about 50% of low caste children. These differential rates in seeking services explain why accessible services have impact on those who take advantage of their own initiative. In those cases where services are rendered through home visits in intensive coverage, the lower caste groups benefitted relatively more. The impact of services seemed to have worked differently in the infant vs. child groups. In this sense, age also is an important pre-disposing variable to measure the impact of services.

The fourth important indicator used in our analysis is the <u>morbidity effect</u>. For some of the groups in the Narangwal experiment weekly morbidity surveillance was carried out. We, therefore, have data that measure the

effect of child care services on the illness incidence of children in the FPCC experimental group that can be compared with the children in the control group of villages. The morbidity indicator used is the average duration of episodes of 7 specific illness in control and the other villages, where child care services were offered. These illnesses are fever, cough, pneumonia, diarrhea, vomiting, eye and skin infections. These were selected for their frequency and also their importance in the zero to three year age group. For each illness, the average duration was less in the villages of FPCC services in comparison to the control villages (see Table 5). The differences range from 14 to 33%.

We further refined the measures. Using the above durations and average incidence rates for each illness, the total annual days of illness per child were estimated for infants under 1 year of age and for children 1-3 years of age. Our analysis show that services in FPCC reduced the amount of illness by 22 days per year in each age group (Table 5). A reduction of 16% under 1 year of age and 21% from 1 to 3 years of age. This variable has proved to be a stable impact indicator to show the difference in the health status of children under 3 from the use of services. However, such an indicator requires weekly morbidity surveillance, which will be expensive, if provided for in the normal health projects. So, even if this particular indicator appears to be most desirable in an experiment of the Narangwal type, it cannot be expected to be replicated in other health project areas without an intensive system of longitudinal data collection on regular basis. If instead of weekly surveillance, these particular indicators were constructed through cross sectional surveys in two time periods, the results would not have been so easily discernible, because, when the illness episodes are narrated from memory, lapses tend to confuse the picture.

Fifth, the nutritional status indicator, growth of children, was also used. Analysis of the effect of services on the growth of children is available

Table 5

EFFECTS OF CHILDREN'S SERVICES ON MORBIDITY LEVELS
(1970 - 1973)

			Control Villages (days)	FPCC Villages (days)	Percent Differences
Average Duration of Specific Conditions in Days per Episode	Fever		3.9	2.9	- 25.6
	Cough		11.4	8.5	- 25.4
	Pneumonia		3.6	3.1	- 13.9
	Diarrhea		6.3	5.1	- 19.1
	Vomiting		5.2	3.5	- 32.7
	Eye Infection		8.3	6.3	- 24.1
	Skin Infection		8.7	7.2	- 17.2
Annual Average days of Illness per Child	For All Above Conditions	< 1 year	135	113	- 16.3
		1-3 years	105	83	- 21.0

only in the FPCC group of villages, because this group formed a part of the nutrition study as well. As in the morbidity analysis, the data came from the nutrition project and can be compared with the nutrition control data to show how child care services affected the growth of children. Significant differences were observed between FPCC villages and the controls in average weights and heights at ages 17 months through 36 months. Children in FPCC villages exhibited significantly higher average weights and heights. The pattern was consistent for both males and females and for high and low caste. The differences between the surveys and control villages were on the order of 3 to 4% points corresponding to 0.4 and 0.6 kilograms beyond 2 years of age. Differences in average weight at ages younger than 13 months were not statistically significant. At ages 21 months and older the average height was significantly greater in the service villages than in control villages. Differences between heights in the FPCC and control groups averaged about 1.5 to 2.5 percentage points or 1 to 2 centimerer starting a little before the second year of life. Like some other mortality indicators, the nutrition indicator was influenced by caste differences. The data show that high caste children weighed on average approximately 0.7 kilogram higher than the low caste children on average. Similarly, caste had a strong influence upon heights of children below three years of age. The differences in average height between high and low caste increased from 1.4 centimeter at age 9 months to approximately 2.5 centimeters after two years of age. Such important differences across castes point the need of the monitoring of the impact of services separately for the castes. If that is not done the influence of services on this kind of measures is muddled by the intercaste differences.

The Nutrition Study: The outcome indicators used in the nutrition study were essentially three: child's nutritional status and child growth, morbidity

experience and mortality. In the design of outcome analysis it was assumed that socio-economic status afffects child growth, child development and morbidity mainly through three intermediate variables, namely availability and quality of mother care, quality and quantity of diet and housing and environmental conditions. We shall record the results of analysis regarding these indicators.

First, we used nutritional status and growth indicators and found that nutritional care alone or in combination with health care improved weight and height of the children under study beyond 17 months of age. As noted before the nutrition study of the Narangwal experiment included four groups: namely, nutrition service villages (NUT) health care (HC), mainly concentrating on infection control, integrated services of nutrition and health care (NUTHC) and the control group (CONT-N). Our research results show that nutrition care alone or when combined with health care improved both height and weight of the study children at 36 months; children from nutrition care villages weighed on the average 560 grams more and were 1.3 centimeters taller than children in control villages. Children in HC villages had the mean weights and heights intermediate between those in the NUT and control villages.

Among the many socioeconomic and demographic variables tested, sex and caste were shown to have an especially pronounced independent effect which averaged .60 to .75 kilograms in weight and about 2 centimeters in height. Beyond 13 months of age the proportion of underweight children in nutrition care villages was consistently lower than in other villages. The difference between the proportion of underweight children in nutrition care and control villages was expectedly not impressive for high caste children. This only shows that caste proves to be an important variable in terms of explaining the impact of services on height or weight of children. In such cases, caste-specific indicators are important. For example, the difference became

highly significant for low caste children, especially females, suggesting that children whose undernutrition had resulted primarily from lack of care and poverty profited most from the program.

Normally, the framework of experimental health research does not produce clearcut examples about the impact of services on nutritional status. Control groups tend to improve along with the experimental groups, because as control group children are being weighed and measured mothers spontaneously provide the extra nutrition care that they need. In the Narangwal experiment, it was, however, possible to show statistically significant differences between control and experimental groups by carefully controlling and systematically measuring the program inputs.

The other method used to study the impact of nutritional services on nutritional status of children was the technique of regression analysis on a subsample of 180 children on whom exact dietary measurements were obtained. Regression analysis showed a strong relationship between dietary intake and achieved anthropometric status. This is an indirect method of impact analysis. If the services were affecting dietary intake, they could be expected to have an important effect on the anthropometric status achieved. Here, again, the intervening variable of socio-economic status proved significant. The regression technique is useful for controlling for the effect of such intervening variables.

Second, a morbidity indicator was used. Health care caused a significant reduction in the average duration of infectious diseases as compared with villages without health care. Each episode of diarrheal disease was reduced on the average by two days, lower respiratory tract infections by 1-1/2 days, fever by 1 day, cough by 2-1/2 days and skin infections by 1-1/2 days, in comparison with villages without health care. The only condition for which the combination of nutrition and health care exerted a larger effect than health care alone was eye infection. There the mean duration in NUTHC villages

was 6.3 days comparad to 7.1 in HC villages and 8 days in NUT and 8.3 days in control villages. This differential in average length of disease episodes indicate the differential impact of services. This is of course a complex design in which packages producing different results in terms of the selected indicator of morbidity. The results have shown that there is no evidence for synergism in infection control and nutrition among children in the study villages.

Third, mortality rates were used and we found that prenatal mortality was significantly reduced in NUT villages (31 per 1,000 lives and still births) as compared to HC villages (45 per 1,000 lives and still births) or control villages (57 per 1,000 lives and still births). This decline in mortality probably resulted mostly from better nutrition of all mothers due to the iron and folic acid supplement. However, it was not possible to distinguish this effect clearly from another influence. The latter influence came from food that was provided during the pregnancy of the mothers, which was done on the basis of the judgment of the workers whether the mother was nutritionally at risk or not. This illustrates the problem of the cases where differential treatment based on health workers' discretion led to complications for deciding what component of services had caused a better result, if any.

Neonatal, postneonatal, and 1 to 2 year old child mortality were reduced by 1/3 to 1/2 in villages where infectious disease control services (HC or NUTHC) were provided as compared with control villages. NUT villages had an intermediate effect for under 1 year of age, and an equivalent effect on mortality among 1 to 2 year old children. It is interesting to note that when we see the impact on the concentrated group for the health services in the parallel population study (FPWSCC), the intensity of home visits was much lower than disease control villages, namely HC or NUTHC. It is also found that in those cases where there is less intensive activity, the effect on child mortality below 1 year of age was negligible. However, beyond 1

year of age mortality improvement was very similar to that in HC or NUT villages where intensity of surveillance is much more than the population project of concentrated health activities.

Cost Analysis: In ex-ante analysis, the cost of the project components is an important task. Here we review analysis about cost of services in the Narangwal experiment. The overall cost of Narangwal integrated services (FPWSCC)was $2.2 per capita per year in comparison to that less integrated services. FPWS was $1.8; FPCC, $2.6 and FPed, $1.2 per capita per year. Because of the marked differences in output, these costs should be related to particular benefits in order to obtain a more accurate picture of the comparative efficiency. All calculations about cost showed the greater cost effectiveness of integrating services.

First, we separated out the cost for various types of services. The salary component varied greatly, ranging from 45% in the integrated package (FPWSCC) to 71% in FPed. This is in contrast to government primary health centers where the salary component was 75%. The drug component was 10 to 15% and supplies accounted for 6% in the integrated package, which really means if integrated services were to be provided, the essential expenses would be relatively small and could be readily funded. Seven to 11% of costs were for transportation. Finally the allocation needed to amortize building costs was 1 to 2.5% of the total cost.

A functional classification from the detailed cost analysis gives even better insight into the options. In each instance, the integrated services cost 1/3 to 1/2 of the particular service cost when a service was given in isolation. On an annual per capita basis child care was $1.56 in every FPWSCC, but $1.07 in FPCC and nutrition care was $0.60 and $0.87. Women services where $0.60 in FPWCC and $0.88 in FPWS, and maternity care was $0.23 and $0.43. Family planning cost shows a particularly wider range from FPWSCC, $0.24, FPWS, 0.44, FPCC 0.67 and FPed 1.08. These estimates

are on the basis of clients receiving services. In the case of family planning services these were estimated on the basis of number of family planning acceptors. In the case of children services these are on the basis of number of children covered and similarly for women services the number of women receiving services.

When costs were related to the number of service contact the cost per contact ranged from $0.9 to $1.31 per women's or child's contact from $0.79 to $1.31 for maternity care contacts. In government primary health centers the cost per curative contact was estimated at $0.20, giving an idea of the difference in overall cost in government services as compared with Narangwal. The cost per family planning contact showed an even wider range. FPWSCC $0.53, APWS $0.80, FPCC $1.87 and FPed $1.47. The government cost per family planning contact was estimated at $0.47.

A comparison of the project cost with costs of all the services gives the project cost not very high. Private annual per capita expenditure ranged from $1.73 in FPWSCC to $2.13 in FPWS, FPCC and control and up to 3.07 in FPed. The Punjab government expenditures for health services used by village people range from 0.20 per capita in FPWSCC villages and FPCC to $0.40 in FPWS villages,$0.45 in FPed villages and $0.84 in control villages. Use of project services therefore produced a saving from other expenses of nearly $1.55 per capita in FPWSCC, compared with FPed, and $1.0 as compared with the control villages. Combined costs were lowest in control villages next in FPWSCC and highest in FPed, which fits other evidence that the latter villages had the most affluent population and were actively seeking health care.

Cost per new family planning acceptors was $12.27 in FPWSCC, twice of that in FPWS and three times of that in FPCC and FPed. Cost per couple of family planning used was $10.27 in FPWSCC and 1.7 times of that amount in FPWS, 2-1/2 times in FPCC and 3 times in FPed.

When health costs were attributed to the outcome of mortatlity prevention, the total child care cost per death averted came to about $800 in FPWSCC and slightly more in FPCC. The cost per neonatal death averted was $9.87, for infant death averted, was $37.33, and for 1 to 3 year old child death averted, it was $101.46. The portion of the cost attributed to morbidity reduction resulted in calculations of the cost of the day of illness averted which came to $0.53 for infant and $0.40 for a child 1 to 3 years of age.

Detailed measurement of service inputs showed clear differences between experimental groups in terms of staff time service service contacts and costs. In the nutrition part of the Narangwal experiment, the largest amount of service time and service contact provided per child were in NUTHC villages. However the costs per year of service were very similar in all experimental groups. The NUT villages where the most costly per child, primarily because of the higher average number of child feedings provided per child under three. Cost in NUTHC villages for combined nutrition health care were about $21.0 per year per child under 3, or less than $2 per capita of the total population. The average costs per service contact of 0.20 was about equal to the cost per patient visit in government primary health centers in Punjab in 1969. Cost per child feeding per session average about $ 0.04. Because project services partially replaced use of private and government services, the combined child care and nutrition program increased overall health care expenditures in those villages receiving the program by only 40% above expenditures in control villages or private and government care.

Cost effectiveness estimates show that prenatal child care cost per prenatal care averted were the lowest in NUT villages as for the most favorable cost for averting an infant death of $25 was found in HC villages. The HC villages also produced the lowest cost per child death averted. ($31 in 1 to 3 years of age). Costs per day of illness averted in children under

1 year of age were $0.40 in HC villages and $0.56 in NUTHC villages. Similar estimates for children 1 to 3 years of age were $0.35 and $0.39 respectively. Nutrition costs per additional centimeter of growth attained by 3 years of age was $26 in NUTHC and $30 in NUT villages.

Summary and Conclusion

Impact analysis of a project, either ex ante (for appraisal) or ex post (for evaluation) is much more complex in the social sectors like health or education than in the physical sectors like agriculture or industry. The complexities for the social sector projects arise from the problems of:

- measuring and valuing the outcome of the projects,
- separating investment and consumption aspects of the project,
- making interpersonal comparison of utility or welfare,
- allowing for political considerations and influences, and
- assigning weights to multiple goals of a project.

Essentially there have been two major approaches to analytical development for impact analysis of health projects or programs. The first approach -- dominated by health economists -- attempts to apply conventional economic techniques to health project analysis. The analytical development so far is confined to the measurement and valuation of outcome for doing benefit cost analysis of health projects. Investment approach which values life by one's production has been frequently used. Recently some economists have attempted to apply "the willingness-to-pay" criterion of valuating the output of a health program. All these analytical developments about the measurement and valuation of health outcome relate to the North American context.

Another economic analytic technique, used more frequently, is the cost-effectiveness analysis. This technique bypasses the problem of output valuation and attempts to evaluate the efficiency of a project to produce certain

output or outcome. However, the problem of quantifying output remains, and what is more serious, cost-effectiveness gives no clue to compare two projects when output and outcome are supposedly different.

Linear programming and macro simulation have sometimes been used, though very rarely, to analyze efficiency and benefits of a health program. Since measurement remains a serious problem, these techniques have even less applicability, data of dubious quality, if used in an interdependent system (which human programming and macro simulation use), can produce meaningless results.

The second approach to analytical development -- dominated by the public health and health specialists -- concentrates on measuring health and its progress. There have been numerous attempts to develop health indicators. Several generated from a recent International conference on primary health care held at Alma-Ata and sponsored by WHO. Discussions have focused on the goal, 'health for all by 2000', which was declared at the Alma-Ata conference. What the goal means and how it can be reached are the two major issues. From the discussions that have taken place so far and the papers that are available, we see that 4 categories of indicators have been suggested: (a) health environment, (b) health inputs, (c) health outputs, and (d) health outcomes.

The first category of indicator would include socioeconomic conditions and health policy measures, which affect health. In a macro setting, this indicator also reflects health status and any changes that are clearly relevant to measure progress in health. For a given project, this indicator is the predisposing variable; its role is important to understand to correctly assess the impact of a health project.

The second category of indicator measures the magnitude of services offered and utilized. These are not impact of the services per se, but

actually help us to understand the size and nature of the intervention offered.

The third category of indicator relates to the utilization of services, which actually represents the effective coverage of services. The effective coverage has to show the provision of services in relation to the need for the services. Often this is represented by the proportion of people in need of the service who actually received treatment. The concepts of population groups at risk (e.g. children needing immunization) and the target population (e.g. all expectant mothers needing antenatal care) are used to convey the effective coverage of services.

The fourth category of indicator includes outcome variables, reflecting the final impact of health services. These variables are to reflect either the general change in the health status or a particular change such as reduction of a particular disease incidence.

At Narangwal, Punjab, India, a field experiment was carried out between 1968 and 1974. In the experiment, groups of villages were provided with combinations of health, family planning, and nutrition services, and the households in each group were observed over time. There were two parts of the experiment--population study and nutrition study.

In the population study, there were five experimental groups of villages, with four of them with various combinations of services and one control group. Since this was a micro setting, no environmental indicator was used except as a predisposing variable. Also, because this was an experimental design with a control group, the use of input variables for impact analysis was not extensive. The focus was an output and outcome measures.

For the input variables, time used by the main categories of workers was measured. Collecting the date is difficult when an intensive study is not planned.

Five types of output and outcome variables were used in the population study with varying results and insights about their uses in other settings.

(1) Service contacts or coverage. The total number of service contacts was measured, but the contacts were then standardized by estimating the total number of visits per week per 1000 population. The time trend of these contacts showed that they were increasing up to a point and then declined in all experimental groups. These are like input indicators, because they do not relate to how much of the need was met by these services.

It was, therefore, necessary to relate the visits to the need--showing what proportion in need were covered by service contacts. It was clear that the intensive services group achieved generally more coverage.

It was also necessary to compare the project coverage to the total health coverage from both project and nonproject sources. We find that in the intensive service groups, project services constituted the primary source for the child health care in the FPWSCC villages which had comprehensive care--including child health care, women's health care and family planning--and in the family planning and child care villages (FPCC). This shows that with a comprehensive service package and with certain types of services (here child care services), project sources substituted for other services to a greater extent.

(2) Volume or Intensity of Use. The average number of visits per client using such services was used. This helped to show which of the services needed more contacts per client than other services. The results show that on average more visits were required per woman, when women's health services were provided without child care than when they are combined with child care. The volume or intensity of use is a helpful indicator, in addition to the general use indicator, in those situations where various types of services offered and where the different modes of delivery, as for example, home and clinic visits are used. In the Narangwal population study, it was

found that the average number of contacts for the prevntive health care of children--for check-ups, immunization and so on--is almost double in the child care group compared with the comprehensive group. This shows that in the child care group village health workers had more time for preventive health care than those in the comprehensive villages.

(3) Mortality Indicators. Mortality rates were calculated for children by age -- still birth, infant mortality under 1 month and infant deaths between 1 and 12 months. We found that the still birth rate 23-35% lower in the service villages than in the control villages. For neonatal deaths, the difference was 20 to 40%. For deaths among children between 1 and 3 years, the difference was 30 to 60%.

In this variable, the influence of the predisposing variable, caste, is most significant. Comparable effects were produced in both the FPWSCC and FPCC villages in the low caste group, but only in the FPCC group was there some effect among the high caste families. This implies that more intensive child care services in the FPCC villages could cover all caste groups while FPWSCC workers, because of less available time, may have concentrated on the prenatal care and nutrition supplementation affecting more low caste families. Also, in terms of neonatal survival rates, the high caste children in the FPWSCC group benefitted from the services, but the low caste group did not. In the FPCC villages the low caste groups appeared to receive the maximum benefit, a reduction over 40% in survival rates. Apparently in villages where other services were less intensive in their outreach, the high caste group actively sought care from the project or other services and achieved moderate reductions in neonatal mortality. However, low caste children benefitted only in those villages (FPCC) where home visits were more intensive.

(4) Morbidity Effects. The morbidity indicator used is the average duration of episodes of 7 specific illness in control and other villages where child care services were offered. The frequency of fever, cough, pneumonia, diarrhea, vomitting, eye infection, and skin infection among children 3 months to three years was reduced in the present villages. The differences among the various services groups ranged from 14 to 33% less than the episodes experienced by the control villages. Also, the total annual days of illness per child were estimated and proved to be a useful indicator to show the difference in the health status of children under 3 as a result of child care services.

(5) Growth of Children Indicator. Analysis of the effect services on the growth of children showed that there were interesting differences between FPCC villages and the controls in average weights and heights for children at ages 17 through 36 months. The differences were between 0.4 to 0.6 kilograms beyond 2 years of age and between 1 to 2 centimeters starting a little before the second year of life.

In the nutrition study, three indicators were used.

(1) Nutrition Status and Growth. The results were similar on this indicator as those revealed by the growth of chidren indicator, for the population study. Caste proved to be an important predisposing variable. The impact was most significant among the low caste children, suggesting that children whose undernutrition had resulted primarily from lack of care and poverty profited most from the program.

(2) Morbidity Indicators. In this case also, the average duration of most common diseases were used. Health care villages in the nutrition study had achieved significant reduction in the average duration of the disease episodes, by 1 day to 3 days.

(3) Mortality Rates. Child mortality rates, same as the population study, were used and the results obtained were similar.

Some cost analysis, which is an important ex ante analysis of projects, was done for the Narangwal experiment. An overall cost of $2.2. per capita per year was calculated for the most comprehensive package (FPWSCC). We disaggregated the cost. The salary component varied widely, ranging from 45% in the comprehensive package to 71% in FP ed, the Indian government and primary health care centers have a salary component of 75%. A functional classification from the cost analysis gives some insights about alternative strategies of service packaging. In each instance, the integrated services cost 1/3 to 1/2 of the particular cost when a service is provided in isolation.