TW
HEL
027

## Health and Nutrition Survey

# *Training Manual*

CCRDA

Prepared By: Community Health Department
Faculty of Medicine
Addis Ababa University

Edited by: Yemane Berhane, MD, MPH
Abeba Bekele, MD, MPH

*This health and nutrition survey training is organized and sponsored by the*
*CHRISTIAN RELIEF & DEVELOPMENT ASSOCIATION (CRDA)*

September 1995

## Contents


| | | Pages |
|---|---|---|
| 1. | Basic Concepts of Epidemiology | 1 |
| 2. | Major health problems | 8 |
| 3. | Major Nutritional problems | 12 |
| 4. | Community Survey | 15 |
| 5. | Community nutritional assessment methods | 21 |
| 6. | How to conduct a survey | 31 |
| 7. | Data Collection and Questionnaire design | 39 |
| 8. | Processing and Analysing data | 43 |

Appendixes

1. Ethical guidelines for epidemiological investigations
2. Introduction to microcomputer use in data processing
3. Anthropometric Measurements
4. Sample proposal form


CCRDA

## I. Basic Epidemiologic Concepts

Epidemiology is the study of the distribution, frequency and determinants of health problems and disease in human populations. The purpose of epidemiology is to obtain, interpret and use health information to promote health and reduce disease. The basic epidemiological concepts are highly practical and are relevant for all health workers.

There are four phases in the use of the epidemiological approach. The first is **descriptive epidemiology**. It asks: what is the problem and its frequency, who is involved, where and when? In the above definition it is the part concerned with disease distribution and frequency.

The second phase is often called **analytical epidemiology** because it attempts to analyze the causes, or determinants, of diseases by testing hypotheses to answer such questions as: how is the disease caused and why is it continuing?

The third phase is **intervention or experimental epidemiology** in which clinical and community trials are used to answer questions about the effectiveness of new methods for controlling diseases or for improving underlying conditions.

The fourth phase may be called **evaluation epidemiology** because it attempts to measure the effectiveness of different health services and programmes and to answer the very important question: so what? Have there been any improvements in health status?

The analytical and experimental phases rely on complicated research methods, such as case-control and cohort studies and controlled trials, and are unlikely to be used in the health workers day-to-day activities. However, descriptive and evaluation epidemiology can be extremely useful and should be widely used.

Epidemiology is about information: the information needed for planning, managing and evaluating all the activities required to promote health and to prevent and control disease. The key components of the data needed to provide this information can be approached through a series of questions:

**What** is the health problem, disease or condition, and what are its manifestations and characteristics?

**Who** is affected, with reference to age, sex, social class, ethnic group, occupation, heredity and personal habits?



CCRDA

| | |
|---|---|
| **Where** | does the problem occur, in relation to place of residence, geographical distribution and place of exposure? |
| **When** | does it happen, in terms of days, months, seasons or years? |
| **How** | does the health problem, disease or condition occur, and what is its association with specific conditions, agents, vectors, sources of infection, susceptible groups and other contributing factors? |
| **Why** | does it occur, in terms of the reasons for its persistence or occurrence? |
| **So What** | interventions have been implemented as a result of the information gained and what was their effectiveness? Have there been any improvements in health status? |

Although members of a health team are not full-time epidemiologists, they need to know how to base their health work on the needs of the population and of special subgroups. To do this they require health information, not only on those who come to the health services but also on those who do not attend.

A most important concept in epidemiology is that of the total **population at risk**, also called the **denominator population**. It is vital to know about all the people at risk of developing a disease or having a health problem, as well as those who are currently suffering from it. The concept of the denominator is the key to developing the information needed for health planning, management of health programmes and the evaluation of health promotion and disease control activities. For instance, the health team should not only know how many children are immunized each month, or how many patients attended for tuberculosis treatment, or how many houses were sprayed with residual insecticide; they also need to know the total number of children, tuberculosis patients or households that should have received the immunization, treatment or insecticide. A comparison of those who actually received the service with those who should have received it is called **coverage**. Achievement of high coverage by the main programmes is the single most important managerial task for a health team. An assessment of the coverage being achieved is the starting point for improvements in health planning.

## Descriptive epidemiology

The first stage in understanding a health or disease problem from an epidemiological perspective is to describe it by the characteristics or variables of Who? Where? and When? After all the information has been assembled, the second stage is an attempt to explain all the facts.

CORDA

***Who?*** The most important variables are age, sex, education, occupation, income, cultural and religious group, family size, nutritional state and immune status. Other groupings might be by such characteristics as clinic attendance and non-attendance, those with latrines and those without, or normal and low-birth-weight infants. Any relevant variable may be used, provided that subjects can be clearly placed in one category or another.

***Where?*** The place where people live or work may partly determine which health and disease problems they suffer from and what use they make of the available health services. For example, the variables might be:

- Town, village or isolated dwelling.
- High or low altitude.
- Proximity to rivers, forests, wild animals or sources of toxic substances.
- Distance from dispensary, health centre or hospital.

***When?*** It is important to know when health problems are most severe, or when the incidence of new cases is greatest. To show this, cases, episodes or events can be grouped according to new cases per day, week, month or year. The time period depends on what is being analyzed, for instance:

- New cases of cholera per day.
- New cases of measles per week.
- New pregnant mothers registering per month.
- New cases of yellow-fever in one year.

**Measuring frequency**

The two main measures of the frequency of diseases, health problems and the use made of health services are **incidence** and **prevalence**. It is most important to be clear about which is being used.

**Incidence** measures the number of **new** cases, episodes or events occurring over a **defined period of time**, commonly one year. Incidence is the most basic measure of frequency and is the best indicator of whether a condition is decreasing, increasing or remaining static. It is, therefore, the best measure to use in evaluating the effectiveness of health programmes. It is also the measure used in surveillance systems and for Analysing how people are using the health services.

Examples include the births and deaths occurring in a district in one year, cases of neonatal tetanus diagnosed per year, number of women attending antenatal clinic for the first time per month, and number of trypanosomiasis cases diagnosed per year.

CCRDA

**Prevalence** measures the total number of **existing** cases, episodes or events occurring at **one point in time**, commonly on a particular day. Prevalence may be more complicated to interpret than incidence because it depends upon the number of people who have developed their illness in the past and have continued to be ill to the present time. It is a combination of the previous incidence of a condition and its duration. Examples of frequency measured by prevalence are the total number of leprosy patients on a register at the beginning of each month or the number of hospital beds occupied per day. Whereas prevalence is also very useful for chronic conditions, incidence is more useful for those diseases with a short average duration like measles, diarrhoea, and pneumonia. Cross-sectional surveys generally provide information about prevalence and are particularly useful in establishing information about chronic diseases such as leprosy or schistosomiasis.

Under stable conditions, incidence and prevalence are related by the following formula:

prevalence = incidence x average duration of the condition.

Thus for those conditions with a long average duration, such as leprosy and tuberculosis, the incidence per year is much lower than the prevalence. For example, the prevalence rate of pulmonary tuberculosis is commonly between 0.5% and 1.0% (or 5-10 per 1000 people) and the average duration of untreated illness is estimated to be about 4 to 5 years. This means that the incidence of new cases of pulmonary tuberculosis is 0.1% to 0.2% or 1-2 cases per 1000 people per year. In countries with good diagnostic and reporting systems the incidence of new cases of tuberculosis may be used. However, in many developing countries without such systems, reliable information can usually be obtained from cross-sectional surveys which provide prevalence data.

CCRDA

**Numbers and rate**

Incidence and prevalence are used in reporting health information and may be given as a whole number or a calculated rate.

The most readily available data will be in absolute **numbers**. These are often used in monitoring the occurrence of important infectious diseases, especially in outbreaks, when the populations involved are restricted in time and locality and the population structure can be assumed to be stable.

When we have to look at trends over a period of time, or compare the frequency of diseases between subgroups or communities, using the total number of cases can lead to invalid conclusions. The population size and age-sex structure of each group must also be considered before the groups can be compared and the information should then be expressed in terms of **incidence or prevalence rates**.

When calculating rates, the events or cases are related to the population which has given rise to them. This is the **population at risk** and refers to the group of people who have the potential to get the disease and thus may contribute to the total number of cases. For example, such a "population" may refer to the whole population of a district or a part of the district or all the people in a particular age-sex group.

The number of cases (made up by counting people, episodes or attendances) is called the **numerator** and the total population at risk the **denominator**. All people in the denominator must be, by definition, at risk of becoming a part of the numerator population. Each rate must have a time period or a date attached to it and this should always be stated.

Incidence and prevalence rates are defined as follows:

$$\text{Incidence} = \frac{\text{new cases in specified period of time}}{\text{total population at risk}} \times K$$

$$\text{Prevalence} = \frac{\text{existing cases at specified point of time}}{\text{total population at risk}} \times K$$

The rate is multiplied by a factor(K), the size of which is chosen so as to enable the rate to be expressed as a suitable whole number; the factor used is commonly 100,1000, or 10 000.

**Example 1**: In a district with an estimated mid-year population of 200,000 people, there were 40 cases of tuberculosis reported during 1987.

$$\text{Incidence rate} = \frac{40}{200{,}000} = 0.0002 \text{ cases per person per year.}$$

Depending on which factor is used, this can be expressed as:

0.02 cases per 100 people per year, or
0.2 cases per 1,000 people per year, or
2.0 cases per 10,000 people per year

CCRDA

**Example 2:** On 1 June 1987 there were 120 registered leprosy case in district with a estimated population of 200,000 people.

$$\text{Prevalence rate} = \frac{120}{200{,}000} \times K$$

= 0.06 cases per 100 people on 1 June, or
= 0.6 cases per 1,000 people on 1 June.

The importance of using the appropriate denominator in the calculation of various rates needs to be emphasized. If we are estimating the prevalence rate for a particular disease, then the denominator used should be the total number of individuals who may be a risk of contracting the disease in question. In the case of a sample survey, this denominator may comprise all individuals in the sample. If, on the other hand, we wish to estimate the "positive" rate for Plasmodium vivax from the blood samples that have been collected for the same survey, then the denominator used generally should be the total number of individuals from whom blood samples have been taken and the slides read-not the total sample size.

For age-and sex-specific rates the denominator includes only the people in the relevant age or sex groups. For example, in the age-specific fertility rate for women aged 20-24, the denominator should comprise only females in the study sample who are between the ages of 20 and 24 years.

**Making use of rates**

There are two main reasons for using rates as opposed to whole numbers.

- **To make comparisons** between two different populations that may have different numbers of people at risk, by standardizing for population size. For example, it may be important to make comparisons between several districts or between what is happening now in the district compared with 10 years ago.

- **To calculate the number of expected cases.** By using a known rate (e.g. the national prevalence of leprosy may be 1 per 100 people, or the infant mortality rate 120 per 1000 infants per year), the approximate number of cases that are expected to occur in the district in one year (total leprosy patients and infant deaths) can be calculated.

CCRDA

## Health Status assessment

An assessment of the health status of the community, based on information about health problems, is necessary for planning and evaluating the health services. However, since the cost of obtaining such information, in terms of resources and time, can be considerable, it is very important that the health team selects the data that are most feasible to collect and that will be most useful for its work in the area. In addition to this information, the team will also need information on the health resources of the area and how they are being used.

## Sources of epidemiological information

The routine health information system commonly has information on the frequency and distribution of the locally important causes of morbidity and mortality. However, often this information is not presented in a way that is easy to understand and use.

Morbidity information gives the overall picture of ill health in the community. Although this information is often deficient in quality and quantity in developing countries, the sources include the following:

- Hospital inpatient records
- Outpatient records
- Disease notifications
- Workplace records
- Schools
- Special surveys

CCRDA

## II. Major health problems in Ethiopia

As in many other underdeveloped countries in the world, the available morbidity and mortality data in Ethiopia clearly indicate that communicable diseases are the leading public health problems besides nutritional health problems. These include a wide range of vector borne and non-vectored diseases and by far most are preventable. The climate and topography of the country favour the existence of various disease vectors such as mosquitoes, flies and snails. The problem is further compounded by the low socioeconomic status of the population, poor personal and environmental hygiene and lack of adequate health services to the majority of the rural population. Among the many communicable diseases prevalent in the country, the following are recognized as the most important public health problems.

I. Infectious

A. Tropical: Malaria, Onchocerciasis, Trypanosomiasis, leishmaniasis, schistosomiasis, leprosy, Yellow fever

B. Non tropical: Tuberculosis, STDs, VPD (measles, diphtheria, pertussis, poliomyelitis, tetanus, tuberculosis), Trachoma, Diarrhoeal diseases, Hepatitis, Epidemic diseases (meningococcal meningitis, typhus, relapsing fever, cholera, bacillary dysentery, yellow fever) and intestinal parasite.

II. Non-infectious

A. Nutritional problems:
- Deficiency diseases (Vitamin A, Iodine)
- Protein energy malnutrition

B. Chronic diseases: Diabetes mellitus, Hypertension

The overall impact of these endemic and epidemic communicable diseases on the health status of the population is reflected in the high mortality rates and very low life expectancy at birth for the whole population. The mortality from communicable diseases is very much increased due to the poor nutritional status of the population. The burden of morbidity and mortality is however high in children, the under served rural population and urban poor.

CCRDA

***Malaria***

About 75% of the total area if Ethiopia is below 2000 meters of altitude and is estimated to be malarious; and approximately 65% of the country's population is at risk of infection.

The most common malarial parasite in Ethiopia is Plasmodium Falciparum, followed by Plasmodium Vivax. The other two parasites Plasmodium Ovale & Plasmodium Malare, are rare in Ethiopia

The major and important mode of transmission for malaria is through the bite of female mosquito. Important vectors in malaria transmission in Ethiopia are the Anopheles species.

***Tuberculosis***

In Ethiopia prevalence of tuberculosis among different age groups as determined by tuberculin reaction ranged between 22 to 71.8% by region; urban regions had the highest rates. Annual risk of infection, that is, the probability that an individual who has not been previously infected will be infected during the ensuing period of 1 year, is found to be 3%, the highest in Africa.

Mycobacterium tuberculosis and Mycobacterium bovis are the bacteria responsible for the disease. The disease is transmitted mainly through respiratory route, followed by alimentary route.

***Leprosy***

Prevalence of leprosy in Ethiopia differs by area; the overall prevalence is estimated to be 3/1000 population. Higher rates of the disease are found in densely populated areas particularly in central and northern Ethiopia.

The infectious agent responsible for causing the disease is a bacteria, Mycobacterium leprae. The infection is transmitted through the nasal and respiratory mucosa or through the skin from a prolonged close contact with a sick person.

***Diarrhoeal diseases***

Diarrhoeal diseases are first or second leading cause of under five year children morbidity and leading cause of under five year children death in Ethiopia (average annual mortality rate in '75 is 11/ 1000). They are also a major contributor of malnutrition.

Infectious agents include viruses, bacteria, and intestinal parasites. The mode of transmission for these agents is mainly faecal-oral.



CCRDA

***Vaccine preventable diseases***

In general, the VPDs are major causes of morbidity and mortality in children. The diseas this category can be prevented by providing the appropriate immunization at the right time i life of a child. The magnitude of the specific diseases are discussed below:

**Measles:** Incidence in Ethiopia is very high in age group below 15, particularly in the 1-4 y according to the routine reports of health facilities, but the annual incidence is not ac fatality rate in Africa is 26%, and among malnourished population it raises to 80%.

**Pertussis & Diphtheria:** Incidence of these diseases is not known in Ethiopia. In Africa case fatality rate of pertussis is estimated to be 2.3-6.3%.

**Polio:** It is a major cause of lameness in developing countries. Residual paralysis in s children due to polio in Ethiopia was found to range between 2.9-13.1/1000 in two surveys. western world is free of this disease and eradication programme is currently undergoing i developing world.

**Neonatal tetanus:** A commonly fatal diseases of newborns with nearly two thirds of all te deaths occurring among neonates. Community based surveys in Ethiopia demonstra mortality rate 4.5-6.7/1000 LB, and an estimated incidence rate of 8.4/1000 LB.

**Tuberculosis:** Is the most common EPI disease in Ethiopian children and causes high mort

***Meningitis***

Ethiopia lies within the "meningitis belt" of sub Saharan Africa which extends be 5-16° North of the equator. Seasonal epidemic occurs from January through April, and de in May and June.

The disease is caused by a bacteria, Nisserie meningitides through a direct contac droplets and discharges from nose and throat of infected persons.

*Trachoma*

It is a major cause of blindness in Ethiopia. The prevalence differs from area to ranges from 62- 83% as seen from different community based studies.

The infectious agent is a bacteria, Chlamydia trachomatis. The disease is trans through direct contact with infectious agent.

CCRDA

***Louse borne diseases***

In this group are included typhus and relapsing fever and Ethiopia is one of he few countries where these diseases continue to be reported. Of the 18,359 cases notified to WHO in 1979, 95% were from Ethiopia. These diseases occur in colder areas where people live under unhygienic condition and are louse infected. Several thousand cases were reported to MOH between 1981 and 1990, with the largest number in 1983 (43,727) where an epidemic occurred in Wolayita Awraja. The disease is transmitted by body louse (Pediculus humanus corporis) from person to person.

***Sexually transmitted diseases***

Over 20 STDs are recognized and the commonest are: syphilis, gonorrhoea, lymphogranuloma venereum, soft chancre, non-gonococcal urethritis, herpes genitalis, granuloma inguinale and AIDS. Prevalence of the classical STDs is high in Ethiopia. They ranked sixth among 15 most frequently reported diseases by MOH health facilities, constituting 4.7% of all cases of communicable diseases in 1986-87. The exact prevalence of these diseases is not known because of lack of diagnosing facilities, under reporting, self treatment, use of traditional healers, and non use of health resources for fear of stigmatization. Infectious agents include viruses, bacteria, protozoa, fungus, and ectoparasites. The main mode of transmission is via genital contact.

***Intestinal parasitism***

Dysenteries & gastroenteritis, and helminthiasis are the leading cause of outpatient morbidity. They accounted for 8.5% of male and 10.4% of female outpatient visits in Ethiopia in 1988-89. Prevalence rates higher than 70% and high rates of multiple infection have been reported from many urban and rural communities. Infectious agents include protozoa and helminths which are mainly transmitted through faecal-oral route.

***Schistosomiasis***

Surveys in 1978-82 in selected areas of all the 14 administrative regions found a prevalence rate of schistosome infection of about 15%. A national survey in 1988-89 reported an overall prevalence of 25%. Geographically, S.mansoni infection and its intermediate host are distributed where small rivers, streams and irrigation canals are present, and in the altitude range of 1300-2000 meters. Parasites responsible for the disease are Schistosoma mansoni (endemic in Ethiopia), Schistosoma hematobium (endemic in Ethiopia), and Schistosoma japonicum (non-existent in Ethiopia).

CCRDA

## III. Major Nutritional Problems in Ethiopia

The diets of the numerous ethnic groups in Ethiopia vary considerably, and there are differences in the diets of rural and urban people and different socioeconomic classes. Cer mainly *teff*, maize, barley, sorghum, wheat, and millet form the main staple in most parts of country outside the areas of pastoral nomadism. Legumes are also an important part of the in most parts of the country. Fruits and vegetables are not commonly eaten by Ethiopians ex in some of the larger towns, and even in these towns vegetables are consumed mainly du periods of religious fasting.

The average daily per capita consumption of the various food items, according to results of the extensive ICNND (Interdepartmental Committee on Nutrition For Nati Defence) survey, was 2500 calories, or 47 calories per kilogram of body weight. The thiam calcium, and iron intake was over the recommended daily allowance (RDA); amounts of pro and niacin were acceptable, but those of riboflavin, vitamin A, and vitamin C were low.

Dietary studies in different parts of the country showed that the percentage of ani protein within total protein consumption to be extremely low in all areas. Surveys reveal 65 to 80% of calories are derived from carbohydrates, 10 to 15% from protein, and 15 to 2 from fat. This pattern is typical of developing countries.

### Nutritional problems of public health importance

The prevalence and incidence of specific nutritional deficiency diseases vary am geographical areas, ecological zones, and population groups.

#### *Protein-Energy Malnutrition*

PEM describes a range of clinical disorders. At one end, marasmus is due to a contin restriction of both dietary energy and protein, as well as other nutrients. At the other en kwashiorkor, due to a quantitative and qualitative deficiency of protein, but in which ene intake may be adequate. PEM is the most important public health problem in underdevelo countries today, and is largely responsible for the fact that in some areas up to half the child born do not survive to the age of 5 years.

PEM occurs characteristically in children under 5 years, wherever the diet is poo protein and energy. Typically the marasmic form of the syndrome occurs in infants under 1 y of age and kwashiorkor in the second year of life.

The most serious health effects of PEM are stunted growth, body wasting, retar mental growth, and high mortality in younger children. Mortality rates are influenced by

CRDA

occurrence of measles, diarrhoea and other communicable diseases.

***Goitre***

The term goitre is used to denote enlargement of the thyroid gland of whatever kind. It is endemic in many parts of the world including Ethiopia. Environmental rather than dietary factors determine the prevalence of most simple goitres - especially dietary factors of which iodine deficiency is the major one. Other than iodine deficiency, goitreogenic substances in food were also found to be causes of simple goitre. Most of the goitreogens identified in plants like cabbages, turnips and particularly the seeds of cabbage, mustard and rape are organic compounds containing sulphur.

The national goitre survey in Ethiopia in 1981 found out the prevalence of goitre to range from 0.4 to 66.3% in different communities, with a mean of 25.1%. Mean prevalence was higher in the highlands above 2000 meters (33.9%) than the lower elevations (19.2%). The specific cause of goitre in Ethiopia has not been determined, although it was found that the soil, particularly in the highlands, contains insignificant amounts of iodine, apparently due to high erosion rate.

***Xerophthalmia***

Xerophthalmia (*Greek Seros, dry; ophtalmus, eye*) is a condition caused by vitamin A deficiency. In its mild form it is confined to the conjunctiva and this is very common in many areas. At this stage there is no disability, but it is a clear warning of the probability of vitamin A deficiency. When it spreads to the cornea there is danger of corneal ulceration and a permanent defect in vision. In severe cases there is softening of the cornea, keratomalacia, which, if not immediately treated, soon leads to permanent blindness. Keratomalacia is frequently associated in young children with PEM. When this occurs, the mortality is high even with the best treatment.

Xerophthalmia arises when the diet contains practically no whole milk and butter and very limited amounts of fresh vegetables and fruit and so lacks both retinol and carotenes. Xerophthalmia and keratomalacia both occur in the first year of life amongst artificially fed infants but are rare amongst the breast fed. Children in poorly nourished communities are born to mothers who have had small intakes of vitamin A and consequently their liver stores are small at birth. PEM further aggravates the partial deficiency because absorption and plasma transport of vitamin A are impaired.

In Ethiopia vitamin A deficiency prevails in much of the country and attains public health importance particularly in children and pregnant women. The ICNND survey found about 10% of the pregnant women with vitamin A levels of less than 20 micrograms per 100 ml of serum.

CCRDA

### *Rickets and Osteomalacia*

Rickets, a word derived from the Angle-Saxon *wrikken* to twist, is a disease of child in which the bones are softened and deformed. It arises as a result of deficiency of vitamin and a failure to absorb calcium from the small intestine. Osteomalacia, which means soften of bone, arises when there is vitamin D deficiency in adults. The resultant calcium deficie leads to demineralization of the bones.

### *Nutritional Anaemia*

Iron deficiency anaemia is by far the most common variety of anaemia through out world affecting mainly women in their reproductive years, infants and children. In both r and urban areas in the tropics, this type of anaemia is extremely common. Losses of iron oc not only in menstruation and pregnancy but also as a result of infection with hookworms. M tropical diets do not contain sufficient absorbable iron to make good these losses.

Dietary deficiency of folate and vitamin B12 leads to megaloblastic anaemia. In countries, megaloblastic anaemia is much less common than iron deficiency. These oc commonly in underdeveloped countries. It may occur at any age, but adult women, infants young children are affected more frequently than men. It is often made manifest by pregna

In tropical countries, demands for iron may be greatly increased by the los haemoglobin in the faeces and urine resulting from haemorrhage due to infection schistosome or hookworm. Other protozoal and helminthic infections as well as bact infections which are common in the tropics may be responsible for anaemia. When t additional factors operate in persons who live on diets deficient in iron and protein, it is surprising that anaemia in tropical countries is both frequent and severe.

CORDA

## IV. Community Survey

Surveys measure the prevalence of disease & are called prevalence studies. Measurements of exposure and effect are made at the same time.

A community survey consists of a series of activities. These activities can be broadly categorized into two:

1. What we need to **know**
   - Why a survey is necessary
   - Types of survey that can be done
   - Who should do the survey
   - When the survey should be done
2. What we need to **do**
   - Prepare materials
   - Train the survey team
   - Carry out the survey
   - Use the results

### I. What we need to know

#### A. Why a survey is necessary

There are several reasons for doing surveys:

1. To discover the main health needs of the **community**.

   Accurate, baseline information on the community's health and population structure is obtained **before** we help to bring solutions. Then by resurveying **after** a period of time, say 3 or 5 years later, we can estimate how much progress has been made.

2. To discover those **individuals** and **families** with the greatest health needs.

> In hospitals and clinics we treat those who come to us. In CBHC we identify and help those who needs are greatest.

3. To build relationships.

   A survey done in a relaxed and friendly manner will help to build friendships and create trust. It will prove a good 'starter subject' for community participation.

4. To teach and sensitise.
   We will have a chance to discuss health problems with individual families and create awareness about how they can be solved.

CGRDA

## A. Types of survey that can be done

There are several types of survey that can be carried out in CBHC. These include:

1. **Comprehensive** surveys
   Every home is visited and questions are asked concerning all family mem
   The advantage is all individuals at risk can be discovered and comprehensiv
   can be started without delay. Information obtained is also more complete
   it is time consuming.

2. **Sample** surveys
   Some, but not all households are visited. Every 5th or 10th house can be c
   or houses selected at random. Sample surveys are used either if numbers
   project are very large or if a quick, initial survey is needed.

3. **Mixed** surveys
   Here we may visit **each** house to record certain important information such
   weight or upper arm circumference of the children, but only **some** hou
   record other details such as socioeconomic data.

4. **Pilot** surveys
   These are small scale surveys carried out at the start of a project, eit
   estimate the needs of the people, obtain an approximate census or to pre
   surveying technique.

## C. Who should do the survey?

Project and community members should work together. For example: a project m
can **record** answers, a community member **find out** information.

The **project member** will need to be:

- open and friendly in manner,
- exact and neat in recording.

The **community member** could be one of the following:

- a community health worker,
- a health committee member or
- any community member.

CCRDA

### D. When the survey should be done

This will need to be:

1. When the **community** is ready for it.

   We should make sure the community understands the reasons for the survey and is ready to participate. A suspicious or unwilling community may give inaccurate answers.

> We can reduce suspicion by always working with the community - in planning the survey, carrying it out and acting on results.

2. When the **project** has enough resources.

   At the very least we will need:

   - Team member(s) who have been trained in surveying.
   - Survey materials including forms or family folders.
   - Time to do the job properly.
   - The ability to work with the community in response to the needs discovered.

3. At a **time of day** and a **time of year** when most people are at home, and not too busy with other activities, such as harvesting or seasonal employment.

## III. What we need to do

### A. Preparing for the survey

1. Decide what **information** needs to be gathered.

   Information should be useful and easy to collect. Consider:

   Will it help in making plans?
   Can it be used to bring about improvements?
   Is it simple to ask and record?

2. Study and adapt existing **survey forms**.

   Find out if the government or voluntary health association have materials available. See what nearby projects are using. Forms already in use can probably be used or adapted.

CCRDA

If there is a plan to do research or a detailed evaluation later it is worth consu a specialist, statistician or an epidemiologist at this stage.

3. Collect the **materials** needed for the survey.

- print questionnaire
- collect other instruments required for the survey

4. **Train the survey team**

Surveyors should be trained in two basic skills:

a. **Relating** to the social skills:

The surveyor must be friendly, tactful, persistent and discreet.

Before asking questions the surveyor will need to explain **who** he is, **where** he is from and **why** he is doing the survey.

b. Related to **Obtaining** the information (technical skills):

the surveyor will need to learn **questioning**, **Record** **Measuring** as specified in the interview manual required.

**B. Carry out the survey**

Stages to follow:

1. First arrange a day well in advance, inform the people and co-ordinate community partners.

2. Work in teams, a team member and community volunteer working together

3. Decide on a numbering system for the houses with the survey teams be **starting**.

Numbering often causes problems. Frequently houses have no num or else numbers change as extra houses are built. If a perma numbering system exists, use that. If there is no such system devise and made sure each family records and remembers its health su number. Mark the number on the outside of the house for the conven of visiting health workers.

4. Budget time

Estimate the time needed to complete an interview. Allow adequate to avoid time pressure on both the respondent and interviewer.

CCRDA

5. Conduct the survey

   a. Data collection
   b. Quality control - monitoring and supervision
   c. Data analysis
      Tally and perform tabulation and rates calculations.
   d. Interpret results - understand the results

C. **Use the information for action**

> Spend time with all the surveyors explaining what has been discovered about the health of the community. Help them to understand the significance of the findings. This helps each person to be informed, involved and interested.

Having completed the survey the results must be used to plan your activities, not stored away and forgotten. Note that the main objective of doing surveys is to direct actions/activities.

1. What are the major health problems of the community?
   List major health problems as observed in the surveyed community.

2. Set priorities

   Rewrite the health problems according to their importance. This tells you what must be done first. It is **not a list to be imposed on the community rather this exercise must be done preferably with the presence of community members.** In doing so consider:

   . How **serious** is the problem?
   . How **widespread** is the problem?
   . How **important** is the problem to the community?
   . How **suitable** is it for joint project/community action?

The team can now rate answers to each of the above questions from 1 to 3 making a maximum possible total of 12. Problems are then prioritised in order of their score see the following table. In the example shown the resulting order of priority is: water 11, tetanus 10, drinking 8, malaria 6.

CCRDA

3. Based on the agreed priorities develop plan of action

   a. Formulate objectives
   b. Establish monitoring and evaluation indicators
   c. Decide what activities needs to be performed to achieve the objectives
   d. Identify responsible persons/institutions to carry out the activities
   e. Outline time table for the activities
   f. Allocate adequate resources
   g. Identify possible sources of funding/resources
   h. Plan monitoring and evaluation activities.

Table 1 Prioritising problems

| Problem | How Serious | How widespread | How important | How suitable to community joint action | Total |
|---|---|---|---|---|---|
| Lack of clean | 3 | 3 | 2 | 3 | 11 |
| Occasional malaria in wet season | 2 | 1 | 1 | 2 | 6 |
| Many deaths from neonatal tetanus | 3 | 2 | 2 | 3 | 10 |
| Heavy drinking among men | 2 | 3 | 1 | 2 | 8 |

CCRDA

## V. Community Nutritional Assessment Methods

Nutritional assessment in a community can be direct or indirect. The term 'direct assessment' refers to methods in which individuals or communities are investigated directly, while indirect assessment uses a variety of vital and health statistics. Methods available for the direct assessment of nutritional status are clinical signs, laboratory tests (biochemical, haematological, parasitological), biophysical methods and anthropometry.

### *Clinical signs*

Clinical examination has always been, and remains, a widely used practical, direct method for trying to assess the nutritional status of individuals and communities. Clinical signs are obviously more useful in community studies where the nutritional situation is seriously abnormal. Essentially, assessment by clinical signs is based on physical examination of the individual for changes believed to be related to inadequate or excessive nutritional intake. In addition, clinical examination can include the observation and recording of more obvious signs of some nutritionally important conditions or contributory infections (measles, diarrhoea, tuberculosis, etc.). Table 1 indicates some commonly recorded clinical signs suggestive of dietary deficiency or excess.

Table 1. Commonly recorded clinical signs 'strongly suggestive' of dietary deficiency or excess

| signs | suggested nutrient abnormality |
|---|---|
| **Deficiency** | |
| Pale congectiva | Anaemia |
| Bitot's spots | Vitamin A |
| Angular stomatitis | Riboflavin |
| Spongy/bleeding gums | Vitamin C |
| Thyroid enlargement | Iodine |
| Oedema (both lower legs) | PEM |
| Epiphyseal enlargement | Vitamin D |
| **Excess** | |
| Mottled enamel | Fluoride |
| Dental caries (cavities, etc.) | Sugar |

CCRDA

The following are some of the advantages and disadvantages of clinical signs -compared w other methods- as a community nutritional assessment method in a survey.

Advantages:
- relatively low cost (no equipment needed)
- highly qualified personnel not needed (junior personnel and even parents trained)
- fast to do and easy to organize
- non invasive

Disadvantages:
- subjective (less marked signs)
- limited diagnostic certainty (nutrient or non-nutrient)
- well trained staff needed
- many signs are seen only in severe cases, and there may be n early on

***Laboratory tests***

The main laboratory tests of nutritional significance can be considered in three g biochemical, haematological, and parasitological.

Advantages:
- findings may precede other evidence
- objective
- gradable

Difficulties:
- expensive
- difficulties with specimen collection, preservation and transportation
- absence of laboratories
- lack of reference value for different age groups

***Biophysical tests***

This is a method in which non-biochemical methods are used to measure functional of biological tissues, or alteration of structures inaccessible to clinical examination, esp the bones. These include tests to assess work capacity and energy expenditure, dark adap immune response, hair root morphology, and radiologic examination.

CCRDA

*Anthropometry*

Anthropometry is the most commonly used direct method for the assessment of two of the most wide-spread nutrition problems in the world: PEM and obesity. They can be used to monitor normal or abnormal growth in childhood, pregnancy, and fetal life, as well as the pathological 'overgrowth ' of obesity and the pathophysiological alteration of body dimensions in the elderly,

The measurements usually employed are weight and height. The arm circumference and triceps fatfold are also commonly measured, and can give an indication of change in calorie stores (subcutaneous fat) and/or protein stores (muscle).

Advantages:
- relatively inexpensive (basic instruments are used)
- objective
- understandable by the population at large
- give results which can be numerically graded
- supply information concerning adequate growth, early PEM and obesity not easily or economically obtainable by other methods

Disadvantages:
- have a considerable potential for inaccuracy of measurements
- the need for reasonably precise age in young children
- limited nutritional diagnostic relevance (mainly to PEM)
- problems with the selection of appropriate reference data and of cut-off points above or below levels suggesting abnormality

**Anthropometric measurement methods**

The methods and measurements employed in anthropometry can vary greatly in number and complexity. Obviously, those for children will depend on purpose, objectives, and level of accuracy needed in the particular survey, study, or investigation, and on the resources available

1. Equipment

The selection of suitable instruments is obviously important. For survey work, they must be sufficiently accurate for the particular study, simple to use, easy to read, affordable, readily portable and rugged enough to withstand severe handling. The design and type of anthropometric equipment needs to be related to its proposed use.

CCRDA

## 2. Techniques

Of the different possible techniques of body measurements, those selected should be t simplest and quickest to take, the most acceptable to the subjects concerned, and the easies replicate that will, nevertheless, give the maximum information concerning the particu nutritional problem under investigation. They should employ affordable equipment and developing countries should be easy to use by primary health care workers, including train members of the community.

If many people are to be examined in surveys, sufficient personnel should be availa to permit rest periods. Otherwise the repetition, monotony, and physical strain of making la numbers of careful measurements, especially on uncooperative young children, lead to fatig boredom, carelessness, and inevitable error. The average time required to carry out e procedure must be also known in order to determine how many subjects can be dealt with at e session.

The investigator can record the readings immediately, or can dictate each measurem to a recorder, who repeats aloud the figure as it is given. It should be appreciated that ma errors in taking measurements are caused by mistakes in reading the scales and writing down results.

### Major measurements

The most usual measurements are those made to assess:

1. body mass, as judged by weight,
2. linear dimensions, especially height or length,
3. body composition and reserves of calories (energy) and protein, as reflected the principal accessible soft tissues -subcutaneous fat and muscle.

CCRDA

### 1. Body mass

Weight is the most used anthropometric measurement and is a basic indicator in clini practice, especially paediatrics. Although instrument and observer errors easily occur, potential value of weighing, especially for children, is appreciated not only by health personn but by patients and parents, for whom weight changes can also be useful as a part of health a nutrition education.

Weight estimations can be made on single occasions, as in cross-sectional prevalen surveys, or repeated at intervals, as at under five clinics, prenatal clinics, hospitals, schools, the community, or in individuals. Serial measurements also form an integral part of longitudi surveys

The evaluation of the significance of measurements of body weight also must take into account the influences of body build and composition, especially height, breadth (in adults), and the relative proportions of fat, muscle, and bone, and the possible presence of pathological weight due, for example, to oedema. Thus, it should be combined with visual inspection of the body and measurement of height.

*Equipment*

Points to consider in choosing equipment are:
- 'sufficient' accuracy for particular purposes,
- clear readability,
- appropriate weight range, and weighing surface suitable for age,
- cost, simplicity, possible transportability,
- a capacity for reliable and reproducible recordings under different conditions, and
- not frightening for young children.

Two types of measuring instrument are suggested:

-For children below 6 years: a Salter spring balance with the scale measuring up to maximum of 25 kg with increments of 100 g. With this type of balance the child hangs in a specially designed 'bag'. The model is sturdy, compact and easily transported (Annex 3).

-For those above 6 years: a bathroom scale on which the child is made to stand. The apparatus usually reads to a maximum of 100 kg with increments of 100g.

In both age groups the readings are taken to the nearest 100g.

*Techniques*

As a principle, all clothing should be removed for taking weight. But this may not be practical when mothers are reluctant to undress their children due to fear of 'evil eye', or in weighing adults. Light weight clothing, with known weight should be used in these cases, and subtracted from the individual's weight. Details should be included in the final report.

**2. Linear measurements**

The height of an individual is principally a measure of skeletal bony tissue. For nutritional assessment, usually only the total height is measured routinely.



CCRDA

*Equipment*

For older children and adults, a vertical measuring rod of a scale or a non-stretch t
fixed to a vertical flat wall can be employed for measurement height. In very young childr
recumbent length (crown-heel length) has to be employed, as a measurement of standing hei
is either impossible or very inaccurate with an uncooperative child. It is usually measured w
a length-board, often made of wood (Annex 3).

*Technique*

After removing the shoes, the subject should stand up as straight as possible on a
floor, or board, by the scale with feet together and with heels, buttocks, shoulders, and back
the head touching the upright. The feet should be held, flat on the floor, and with the he
almost together. The legs should be straight and the shoulders relaxed. The head should
looking straight forward and heels comfortably erect, with the lower border of the orbit in
same horizontal plane as the external auditory meatus (the'Frunkfort plane'). The arms sho
be hanging by the sides in a natural manner ('stand tall and look straight ahead') (Annex 3

The head piece is gently lowered, compressing the hair, and making contact with the
of the head. The presence of unusually thick hair requires to be taken into account. When
assistant is available, this person should place one hand against the child's knees to detect
flexion of the lower limbs and the other hand on the upper surfaces of his feet to detect
lifting of the heels from firm contact with the floor or platform. The eyes of the measurer sho
be as level as possible with the reading to avoid optical errors. The measuring scale should
175 cm high and capable of measuring to an accuracy of 0.1 cm.

3. **Body composition**

*Skinfold thickness*

Body composition studies, including information concerning the amount and distribut
of fat, and hence of calorie reserves, can be carried out by various technically complica
methods; but anthropometry using calipers is convenient and reasonably practicable in f
circumstances as an indirect but objective indicator of complete body stores of fat.

*Equipment*

Various skinfold calipers have been devised, but experience suggests that the instrum
used should have a standard contact surface or 'pinch' area (20-40 $mm^2$), should read thro
a range of 2-40 mm, with gradations of 1 mm (or less), and exert a constant pressure (10 g/m
through the whole range of skinfold thicknesses at all distances of separation of the jaws.



CCRDA

caliper dial should be set to zero before each day's work, or allowance made for zero error by subtraction (Annex 3).

Although precision calipers are designed so that their jaws exert a constant tension, in practice, this has different results according to fatfold compressibility, which in turn varies with age, sex, hydration, and fatfold site. The subcutaneous fat in babies and young infants has been shown to be very compressible, while the obese appear to have decreased compressibility.

*Techniques*

Despite the fact that an increase or decrease of subcutaneous fat stores is not uniform all over the body, the essence of the problem is to select one or two easily accessible sites that may be expected to give an approximate practical indication of calorie reserves. For this purpose, the triceps fatfold is the most practical measurement for all age-groups. This can be combined, if needed, with the subscapular fatfold as the preferred second site. The triceps fatfold also has the practical advantage of not altering substantially during the pre-school period, and so from 12-60 months it is relatively independent of precise age in well-nourished populations.

The fatfold at the mid point (half way between the tip of the acromial process of the scapula and the olecranon process of the ulna) of the left upper arm is measured, arm hanging relaxed at the side. The fatfold at the back of the upper arm parallel to the posterior long axis is picked up between the thumb and forefinger of the left hand, away from the underlying muscle, and measured at this point. The triceps fatfold is easily accessible and no undressing is needed. The subscapular fatfold is measured just below and laterally to the angle of the left scapula.

***Arm circumference measurement***

As a measurement it is relatively simple to take and can be done accurately, provided the tape is snugly fitted and not squeezed too tight. A tape is very cheap and easily carried. It has special value when weighing scales and/or a length-board are not available, because of cost, convenience, and portability.

*Equipment*

Tape measures generally should not be wider than 12 mm nor thicker than 0.2 mm. Metal and firm plastic tapes are more likely to retain a circular shape during measurement, while less rigid tapes tend to follow the actual surface of the arm. The numbers on the tape should be clearly visible and understandable by the measurer.

CCRDA

***Technique***

The mid upper arm circumference (MUAC) of the left arm, the arm hanging free measured to the nearest 0.1mm.

**4. Assessment of age**

Assessment of exact age is difficult particularly in areas where age is not recorded exact date of birth of a child/person is not known, the approximate age could be use introducing the time when major events has occurred in the area or the country. But, this c be still very difficult and in this case age independent methods should be used.

**Selection of measurements**

The most important nutritional problem in the world today is that of PEM particu in the less developed countries. It is also a deficiency which, in general, supplementary fee is intended to correct. For this reason, it is suggested that the following three measurem should be chosen. These are age, weight, and height (length supine for children aged less 2 years; height standing for children aged 2 years or more).

These three measurements are combined to form three indicators of nutritional stat
-weight for age,
-weight for height, and
-height for age.

These indicators are compared with those obtained from an international reference popula

In analysing data collected with this indices, the percentage of children in the sample indicators at low levels as compared with the reference children should be determined. determine these proportions, cut-off points have to be chosen beyond which the indicators (W H/A, W/H) are considered to be "low" or "not low". This choice is an arbitrary one which vary from one programme to another.

**Statistical distribution**

Almost all anthropometric measurements have a 'normal' Gaussian distribution i ethnically homogeneous population. In other words, some healthy, well nourished adults wi smaller or taller than average. However, the chances of their height being abnormal incre the further they are form the median. The most usual statistical limits of the 'normal' range from +2 to -2 standard deviations, or from the 97$^{th}$ percentile to the 3$^{rd}$ percentile (Annex

It follows, therefore, that individuals above or below reference levels may be classi statistically with regard to their distribution in relation to percentiles or SDs from the med For example, X percent of a population may be below the 3$^{rd}$ percentile (or -2 SD) for W/A,



CCRDA

have Y percent of school children above the 97th percentile (+2 SD)for height. This type system, using + or _ 2 SD has been recommended by WHO (Waterlow et as. 1977). Very approximate comparisons between statistical and mathematical levels for commonly used measurements are given in the next table.

Comparison of percentage below median (reference) corresponding approximately to third centile (-2 SD) and to 97th centile (+2 SD) in young children up to 60 months of age.

| | 97th centile (+2 SD) (percent) | Median ('reference') (percent) | 3rd centile (-2 SD) (percent) |
|---|---|---|---|
| ...ght for age | 120 | 100 | 80 |
| ...ght for height | 120 | 100 | 80 |
| ...ght for age | 110 | 100 | 90 |
| ...m circumference for age | 115 | 100 | 85 |

CCRDA

## Selection of Reference data

It is suggested that selection of reference data can be helped by bearing in mind six practical principles.

1. Appreciate that all anthropometric reference data are imperfect.
   Flaws may be biological (e.g. genetically mixed population, or communities with widespread disease, or with nutritional disorders, including obesity), statistical (e.g. inadequate and/or biased sample), technical (e.g. measurements inaccurately taken), or a mixture of all or some of these.

2. Differentiate practical utility from comparability
   Practical utility: e.g. determining the prevalence of PEM in a community.
   Comparability: e.g. determining the relative commonness of PEM in different countries at different times, by comparing results obtained using the same reference data

3. Relate reference data used to 'optimal health'.

4. Use most genetically appropriate reference data.

5. Avoid unnecessary confusion.

6. Recognize the possibility of non-nutritional nationalistic misinterpretations of reference

Table 3. Recently advocated and previously used internationally available reference data: se examples.

| | Recent | Previous |
|---|---|---|
| Weight for height for children | NCHS(1979) | Harvard |
| Arm circumference for children | NCHS(1977) | Wolanski(1966) |
| Weight, height for adults | Frisancho(1984) | Metropolitan Life Insurance(1 |
| Arm circumference for adults | Frisancho(1981) | Jelliffe(1966) |

**Standardization and quality control**

Anthropometric measurements, though they look simple to carry out, are subject t and errors in recording if they are not properly standardized. Four steps should be ta obtain good standardized records:

- train the health personnel in the proper methods of using the measuring apparatu scales;
- adjust the scales regularly before each measuring session;
- check for observer error; and
- whenever possible, rotate field workers among groups of subjects to be measured to reduce the effect of individual bias.

Scale adjustment of a spring balance is easily done by bringing the scale needle b the exact zero mark using the side screw provided for this purpose. A similar adjusting exists on bathroom scales. Accuracy should be checked before each measuring sessi comparing the scale reading with a known weight.

ECRDA

## VI. How to Conduct a Survey

Many health care professionals wish to conduct a modest clinical or epidemiologic study (surveys constitute the majority of the epidemiological studies). Hoping to answer one or more interesting questions, they find themselves in a good position to collect and analyze some appropriate data. However, to someone without previous research experience, the task often appears awesome, and it is not at all clear how to proceed.

This chapter is written as a general guide for the novice who wishes to carry out such a study(survey). Obviously, each research project and each study setting present unique problems that cannot be dealt with here. What will be presented is a general approach that emphasizes the practical difficulties that are frequently troublesome to the beginner.

### Defining the Problem

The first step-and one of the most difficult one -is defining the problem and choosing the question or questions to be answered. There is a tendency for the novice at research to ask questions that are diffuse or vague. Instead, the problem must be stated in terms of clear, simple answerable questions.

Intimately involved in the asking of vague, overly broad questions is the tendency to be too ambitious. The new researcher wishes to make important discoveries and solve big problems. These unrealistic expectations can lead only to failure and disappointment. For the most part, medical science progresses gradually by very small steps. So much of health care is based on tenuous evidence and incomplete knowledge that a careful study of even a simple question will be a worthwhile contribution, of which any scientist should be proud.

### Reviewing the Relevant Literature

Once a problem has been selected, the scientific papers describing previous related work should be read carefully. In addition to learning what is already known about the question, the investigator will become familiar with Problems that other researchers have faced, using various study methods.

The usual result of a literature review is a realization of how little is known about the particular topic one wants to investigate. Seemingly authoritative statements and accepted medical doctrines, perpetuated through textbooks and lecturers, often turn out to be supported by the most meager evidence, if any can be found at all!

Many others are based on little or no evidence or on the results of poorly conducted studies. In these instances a literature review will provide encouragement for proceeding with



CCRDA

the proposed study. However, if it is evident that the question has already been well answ a related problem may come to mind-one that can be studied just as well.

**Preparing a Protocol**

The next essential step is the preparation of a study protocol. Eventhough the begin investigator may have a plan clearly in mind, it is extremely important to set it down in wri

A written protocol serves three major purposes. First of all, when one writes protocol, ideas and procedures must be clearly defined and spelled out. Usually the plan in o mind is not as clear and logical as was hoped, and the gaps and flaws are easier to recognize correct when the plan is seen on paper. Second, a written protocol can be studied by an whose advice is desired or whose approval is required. Third, any person involved in carr out the study, even the investigator, may forget some method or procedure to be followed. written protocol constitutes a permanent record that can be referred to, so that methods do change unnecessarily during the conduct of the study.

Some persons have such an abhorrence of writing that the preparation of a protocol i almost insurmountable obstacle to performing a study. If so, it is probably better just to qu this point, since even if some data are collected and analyzed, the results will probably neve written up, and no one else can adequately study the findings. Another alternative fo nonwriter with a good idea for a research project is to team up with a co-investigator wh willing and able to write the protocol and the final report.

**Contents of the Protocol.** Research grant applications may require strict adherenc prescribed contents arranged in a particular order. However, a protocol prepared for local may be shorter and similar but should contain at least the following elements, unless ther good reason for omitting any.

1. A brief statement of the specific question(s) to be answered and/or the spec aim(s) of the study.
2. Background and significance of the study. This should be a pertinent n rambling discussion of what is known and not known about the problem and w the proposed study would be worthwhile or important.
3. Methods, included should be a description of the study subjects-how they are be selected and how many there are likely to be. The data to be collected and methods for collecting them should be described. Uniform criteria for diagno of disease and for decisions as to the presence or absence of a characteristic outcome should be listed. Data analysis methods should also be present preferably with some sample blank tables showing how the data will be organiz

Plans for safeguarding the rights and welfare of the subjects and the method of obtaining their informed consent (if needed) should be explained.

4. An approximate time schedule for carrying but the various aspects of the study.
5. A budget, if financial support is being requested, with explanation of any personnel and other costs whose requirement is not obvious.

### Consultation

After a draft of the protocol has been written it is wise to seek some expert consultation before proceeding further. Many potential problems and difficulties will be quickly spotted by knowledgeable persons reviewing the protocol and discussing the proposed research.

Help can come from persons in a number of disciplines. An experienced investigator who has worked in the area to be studied can perhaps provide the most comprehensive view of the problem. A clinician who specializes in the area of study will often provide some fresh insights into the subject matter derived from the experience with patients and from familiarity with the current literature. Epidemiologists and biostatisticians are professionally concerned with study design and data analysis and can provide guidance on these aspects of the study. The choice of appropriate statistical tests and the determination of whether the proposed sample size is adequate to obtain meaningful information are of particular concern to the biostatistician.

The protocol should now be revised, taking into account the suggestions of the consultants.

### Presenting the Study Plan to Other Key Individuals

At this time the investigator should inform all the responsible persons whose approval or cooperation is either required or desirable.

CCRDA

### Data-Collection Methods

The data to be collected-whether by observation or interview of subjects, by chart review, laboratory tests, or other means-must be recorded in a systematic and orderly manner. The usual way of bringing order into the data-collecting process is by the use of standard forms. Careful attention to preparation of a form, even if only a few items need to be recorded on it, will save the investigator much trouble and grief later on.

One or more forms will be used for each study subject. Each form should provide space for identification of the subject and for recording necessary data. If computers are to be used

for analysis, the format for recording data on the form should meet their requirements. unit of information must be recorded in a particular space on each form.

**Pretesting of Data Collection**

No matter how carefully the data collection is planned, problems will come to light starting. That is why it is important to pretest procedures and forms before the study form begins.

Suppose, for example, that data for a study of cardiovascular disease are to be colle in a mobile facility in which volunteer subjects are scheduled to pass from station to station e 5 minutes for a series of procedures. It may turn out that the electrocardiogram takes 8 min on the average, instead of the planned 5 minutes. As a result, subjects may pile up at ea stations if there is not provision for a waiting area in case of delays. It may therefor necessary to slow down the examination schedule, or provide two electrocardiographic stat or set up a waiting area. This problem should be uncovered and solved during pretestin not, and the subjects have to wait or get the impression that the study is disorgan cooperation may be seriously impaired.

Similarly, a series of interview questions may seem perfectly clear and appropriate they are written down. Yet when study subjects are actually asked these questions they ma understand, or be offended, or give responses that were not anticipated. In a study of radi exposure, for example, the investigator may consider it perfectly reasonable to ask, "Have ever received radiation therapy?" It will undoubtedly turn out that some subjects answer because they misinterpret the question to mean x-ray examinations. The question will ha be reworded and supplemented with additional clarifying questions in case of a "yes" resp Problems such as these quickly become apparent when an interview is tried out on friend associates first and then on some persons similar to the potential study subjects but not offi part of the study.

**Data Collection**

If the investigator is relying on others to collect and record the data, he or she s supervise this aspect of the study closely, especially during the early stages. The wo persons collecting the data should be observed and completed data-collection forms shou checked carefully. In this way, the investigator can ensure that the study plan is being follo Where possible, data collection and recording procedures should be duplicated throughou study on at least a sample of subjects to measure the error rate.

Not all problems will have been discovered during pretesting. Further changes

procedures and forms may have to be made after the study officially begins. These modifications should be kept to a minimum to avoid inconsistencies in the data. Any changes or new rules to be followed should be recorded as additions to the protocol.

CCRDA

## Data Analysis

For most epidemiologic studies, data analysis mainly involves sorting into categ counting and computing proportions, rates, means, and other group characteristics.

To proceed in an orderly fashion and end up with the answers that were desired first place, it is often helpful to draw up some blank tables showing the format for displayi results of data analysis as they would be presented in a final report. These tables are then in with the appropriate counts, rates, and so on when these results become available.

For the novice, preparing blank tables is often quite difficult, requiring a good d patience and self-discipline. However, the results are well worth the effort and, with experi subsequent table-making becomes much easier.

Data analysis tables should show the results broken down by age, sex, and other per variables. In addition to showing the key results that one is after, that should show the nu upon which the results are based. For example, Table 13-1, showing just incidence rat inadequate. The counts upon which these rates are based should also be listed, as in Table Similarly, when means and standard deviations are shown, the number of persons enterin each of these computations should be given.

The actual work of data processing can be carried out in a variety of ways. The c of a proper method depends on how many subjects are involved and the complexity analysis. If only a few counts and proportions are to be determined for a few dozen sub manual counting of items on the data collection forms will be quite adequate and analys be performed with the help of electronic calculators. In fairly large studies investigator wish to use microcomputers and appropriate data-base and statistical analysis progra Statistical program packages such as EPI INFO, SAS, Dbase, and SPSS can usually be ma by nonprogrammers.

CCRDA

## Preparing the Final Report

The difficult job of preparing the scientific paper describing the study becomes a less imposing task if the investigator writes portions of it during the course of the investig

The introductory section of the paper briefly outlines the problem and the purpose study. Note that this material has already been set down in preparing the written protocol that has to be done is to make any modifications that seen necessary for a final paper o come to mind once data collection and analysis are completed.

Similarly, the methods section of the paper can be readily adapted from the protocol. It should describe exactly what was done and, in addition, inform the reader on how subjects were selected for study, how many subjects were included or excluded, and the reasons for exclusion. Criteria for classifying subjects and for decisions as to outcomes should be spelled out. Methods of measurement should be clearly described.

Writing the Results and Discussion sections is simplified if the investigator takes advantage of the fact that the results usually appear in stages. Tables of data are usually completed one at a time. When each table is prepared or received, one should immediately write a paragraph or two discussing the implication of this results for the Discussion section. Then, by the time the data are completely analyzed, most of the writing will be done. The Results and Discussion sections will still have to be organized and edited, but the task of writing them will not have to be faced all at once.

In the discussion section the implications of the study and its relation to previous work should be described. In addition, the difficulties, problems, and potential errors and biases of the present investigation should be reviewed. All investigations have some obvious limitations and others that are not so obvious. It is best if the investigator recognizes and points these out, before someone else does.

## Importance of Good Communication

Science is a social process. Each investigation is related to previous work, either attempting to confirm it or, usually, to build upon it. Investigators need to know and understand what others have done and are doing. It is, therefore, an important responsibility to present a study as fully and clearly as possible.

A paper or an oral presentation at a meeting should be clear and simple. Jargon and unnecessarily complex or obscure terminology should be avoided. Although tables of data in a written paper should be complete, they need not to be repetitious. Thus, after the basic numbers have been shown once (as in Table ), they do not need to be repeated over and over again. Some relationships may be communicated most clearly by means of graphs and figures; these visual aids should be used freely.

In preparing slides or charts to accompany an oral presentation, the temptation to crowd a lot of information into one slide should be resisted. Each slide should have only a few lines or numbers, displayed with large characters, easily seen by those in the back of the room.

CCRDA

**Importance of Investigator Worry**

Many things can go wrong and many errors can occur during a study. Theref essential that, preferably, the investigator, or a conscientious person responsible investigator, worry about details. The careful investigator might well adopt a questio even a suspicious attitude toward each phase of the effort.

In addition to observing the process of data collection, as recommended earl investigator should see to it that every data-recording form is checked carefully by someo than the person who filled it out to detect and correct omissions and obvious errors. should be made of all completed data-collection forms so that the original information be available if any forms are lost. Complete lists and counts of all study subjects sh maintained to provide a check against lost forms. It is surprising how often forms misplaced or fall behind a desk. Key-entry of data should be verified, which involves r the key-entry on a machine that detects discrepancies.

All mathematical calculations should be done twice by two different person investigator's own work should be double-checked by a conscientious individual. Newly computer programs should be tested on small samples of data and the results compared w calculations.

Data tables should be checked to make sure that all the numbers are correct and to the totals. Surprising or inconsistent results should provoke redoubled efforts t whether something has gone wrong.

Finally, it would be sad, indeed, if after all this work the resulting paper were to misleading typographical or printing errors. The manuscript and galley proofs sh proofread carefully.

CCRDA

## VII. Questionnaire Design

For a field researcher there is no satisfactory substitute for experience to know how to formulate and administer a questionnaire satisfactorily.

### Introduction

*Why do we use questionnaires?*

simplicity and convenience

the only best way of obtaining information

Conceptual definitions of variables

- Clearly formulate the purpose of the questionnaire
- Requires a well-defined objectives of the study
- Conceptual problems must be solved before the questionnaire is prepared

### The questions

- Complexity
- Skipping
- Face validity
- Content validity

- Phrasing of questions
  - clear and unambiguous
  - simple and contain only one idea
  - Closed Vs open response
  - Memory aid (probing)

- Embarrassing questions
  - be nonjudgemental
  - use introductory phrases
  - use open-ended questions
  - use words familiar to the respondent
  - Ask past experience before asking about current behaviour

- How Many questions for a single variable

CCRDA

## III. The questionnaire

- **Translation**
  - selection of 'equivalent' word
  - Back-translation
- **Length of the questionnaire**
  - allow adequate time to complete fully
  - avoid unduly time constraint
  - avoid paying per interview completed
  - limit interview period
- Order of questions
  - ♣ follow natural orders
  - ♣ ask in their order of importance
  - ♣ avoid long successive questions
  - ♣ avoid awkward sequencing
  - ♣ place sensitive questions at the end
- Testing questionnaire

> *"if you do not have the resources to pilot-test your questionnaire, don't do the study."*

  - Pre-testing helps to check:
    - acceptability and intelligibility of the questions.
    - identify difficult or unsatisfactory questions and awkward sequences.
    - usefulness of open-ended questions
    - reasons for "don't know" and skipped answers.

==> revised version should submitted to a similar test.

- Validity of the questionnaire
  - Appraising Validity can be done:
    - ♣ inspect the answers
    - ♣ compar results with criterion(gold standard)
    - ♣ Reliability/reproducibility

CCRDA

- **Layout**
  - must be self explanatory/self sufficient
  - should not be too bulky a document
    - time it will take to complete
    - problem of storage
  - instruction to interviewers must be written differently
  - give clear and consistent instruction
  - appearance should be pleasing to look at:
    - paper size and quality
    - allow enough space to write on responses
    - indicate optional questions and jump explanations clearly
    - assign a number to each question

**The Interviewer**

**Selection**

- persons accorded an appropriate level of respect by the study subjects.
- trustworthy
- must have pleasant disposition, be well-mannered, well dressed, reliable and punctual
- must not make promises to respondents that they do not honour
- probationary period and firing procedures
- interviewers should speak the local language
- consider sex, age, and normal place of residency
- Other considerations:
  - health workers
  - extremely educated interviewers
  - young interviewers
  - stranger interviewers

**Training**

- provide both theoretical and practical exercises
- use interviewers manual
  - conduct 'post-mortems'
- monitor performance of the interviewers through out

**Standardization**

- standardization of interviewers is necessary for reproducibility
  - interviewers must be instructed and trained to follow instructions exactly.
  - the place and time of interview must be standardize to obtain similar responses.
  - investigate reasons for obtaining different responses
  - deploy them in a 'balanced' way

CCRDA

**Interviewer manuals**

- ♦ provides detailed instructions regarding how individuals are select approached
- ♦ details any special instruction regarding each question in the questi and the responses that may be given
- ♦ provide guidance on how to deal with unusual situations and how unusual responses
- ♦ outlines checks to be conducted on completed questionnaires an should be submitted for data processing and analysis

Regular updating of the manual is necessary:

♠ based on interviewers report on irregularities

♠ by conducting regular meetings with interviewers to discuss progr queries.

## V. The Interview (Who, Where, and When)

**Who**
- ♥ The mother is the best respondent in studies of children
- ♥ informants should be avoided if possible, other than children

**Where**
- ♥ choice is mostly influenced by logistic considerations
- ♥ chose places for the convenience of the respondent rather t that of the interviewer
- ♥ ensure privacy

**When**
- ♥ chose the time to fit the convenience of the respondent
- ♥ consult or inform in advance respondents regarding the in time.

**Non-response**

♦ high response rate is important to keep the validity of the study. High respo may be achieved in one of the following ways:

- ♣ weed out the questions which can not be answered by sig proportion of the respondents during the pre-test
- ♣ give appropriate and detailed explanation of the study purpo procedure in advance of the main survey to the respondents
- ♣ pay good attention to the selection of interviewers
- ♣ instruct interviewers to call back repeatedly before abando interview.

CCRDA

## VIII. Processing and Analysing data

Data only become really useful when they have been processed and analyzed. Health data may come from routine information systems, surveillance system, investigation of epidemics or special surveys.

Once the data have been collected and coded, the next step is to process them to provide information on all the important questions or variables. Data processing can be undertaken by hand-tallying, hand-sorting or micro-computing.

For simple descriptive epidemiological data on up to 300 individuals, data processing and analysis can be done easily, and quite quickly, by two or three people using hand-tallying and/or hand-sorting. Two advantages of these methods are that recording and coding mistakes can be corrected immediately and that they give a very good "feel" for the data and their meaning. After the initial tallying, the total, percentages and rates can easily be calculated using a small electronic calculator. The use of microcomputers, with the coded data entered directly via the keyboard, is recommended only for large and complicated surveys. Many developing countries are introducing the use of microcomputers for handling health information, but as yet few have good and reliable micro-computing facilities.

Hand-tallying involves counting the number of times a particular category of information appears in all the record forms. If the data have not been pre-coded, all the possible answers must first be listed. For example, when processing data on the symptoms reported for a particular disease, first list all the possible symptoms and then go through all the record forms. Each time a particular symptom is encountered, count it or record it by making a vertical mark against that particular symptom on the tally sheet, (figure 1):

Fig 1. Example of a tally sheet

| Symptoms | Tally | Total |
|---|---|---|
| Nausea | // | 2 |
| Diarrhoea | 𝍸 // | 7 |
| Constipation | 𝍸 𝍸 / | 11 |
| Lethargia | 𝍸 / | 6 |

For easy counting, every fifth mark is drawn diagonally across the preceding four marks to produce the notation (𝍸), which indicates group of five items. In the example just given, we can quite readily see the totals for each symptom.



CCRDA

When the number of forms to be analyzed is larger than about 100, tallying prone to error, often as a result of fatigue on the part of the person doing it. Common tallying include the misclassification of an item, e.g. "nausea" when it should be di double-counting an item or missing one altogether. The risk errors is higher when tal a two-way tabulation, as in Fig.2.

Fig. 2. Hand-tallying for a two-way table, showing the age and birth order of childre

| Age (years) | Birth order of child 1 | 2 | 3 | 4 | 5 or more | Total |
|---|---|---|---|---|---|---|
| 0-4 | 𝍸 | 𝍸 / | /// | | | 14 |
| 5-9 | / | | 𝍸 // | /// | | 11 |
| 10-14 | | / | | // | | 3 |
| 15-19 | | | /// | 𝍸 | // | 10 |
| 20-24 | 𝍸 𝍸 / | // | //// | / | | 18 |
| Total | 17 | 9 | 14 | 11 | 5 | 56 |

CCRDA

## Steps in data analysis

When the data are processed, each variable (or question) for each subject can b and these counts summarized as tables. Graphs and diagrams cannot be drawn witho of tables.

The use of tables for the expected end results (called "dummy" tables) should developed early on in the investigation when the objectives of the study were agreed up dummy tables act as a guide in the analysis, and may be reviewed and modified as th progresses.

When constructing tables three conditions must be satisfied:

- All data should be in a form that can be classified into categories:
  e.g. spleen palpable: Yes or No
  age or haemoglobin: actual numbers or values.

- All table categories must be mutually exclusive
  i.e. no individual can be classified twice in one table.

- Each table should include all the "raw" data

The three essential steps in data analysis are:

- Analyze each variable separately, one after the other, for the distribution of counts in each category. This is called simple or one-way tabulation. What do the distributions look like? What do they tell us?

- Analyze pairs of variables that are relevant, such as incidence of cases by age group, sex or month of the year. This is called two-way or cross-tabulation. Are there any findings or associations that stand out clearly?

- Calculate the appropriate means, percentages and standard deviations. These are called the summary statistics.

By the time these three stages of the analysis have been completed the person doing the analysis should have a good "feel" for the data. In drawing conclusions it is important to look for associations between variables and to check whether any apparent differences are real or not, that is whether they are significant or not. Tests of statistical significance may be needed at this stage, the results of which are commonly expressed in terms of the probability that such an association could have occurred by chance. Often an association is arbitrarily considered as significant if the probability (P) is less than 5% ($P<0.05$), i.e., it would occur by chance less than 1 time in 20.

CORDA

## Simple tabulations

Table 1 shows a set of raw data obtained from a cross-sectional survey of a random sample of 100 villages for haemoglobin levels and hookworm infection. Additional information was also collected on the age and sex of the villagers. This set of raw data will be used to illustrate the various kinds of statistical tables that can be constructed.

Table 1. Raw data on haemoglobin levels (grams per 100 ml) and hookworm infection (p
or negative) for 100 villagers.

| Individual | Age (years) | Sex | Haemoglobin (g/100 ml) | Hookworm infection |
|---|---|---|---|---|
| 1 | 46 | F | 10.3 | + |
| 2 | 9 | F | 8.4 | + |
| 3 | 8 | F | 9.0 | - |
| 4 | 30 | F | 9.4 | - |
| 5 | 45 | F | 8.4 | + |
| 6 | 2 | M | 10.0 | + |
| 7 | 12 | M | 8.3 | - |
| 8 | 33 | F | 10.9 | + |
| 9 | 4 | F | 8.8 | + |
| 10 | 57 | M | 10.9 | + |
| 11 | 38 | M | 10.9 | + |
| 12 | 51 | F | 7.3 | - |
| 13 | 42 | M | 8.3 | - |
| 14 | 38 | M | 9.0 | + |
| 15 | 41 | F | 9.8 | + |
| 16 | 8 | F | 9.3 | + |
| 17 | 16 | F | 9.7 | - |
| 18 | 0 | M | 13.8 | + |
| 19 | 1 | M | 8.9 | - |
| 20 | 4 | M | 7.5 | + |
| 21 | 2 | F | 10.9 | + |
| 22 | 13 | F | 10.7 | + |
| 23 | 8 | F | 10.5 | + |
| 24 | 12 | M | 9.6 | + |
| 25 | 59 | F | 6.8 | + |
| 26 | 19 | M | 11.6 | - |
| 27 | 3 | F | 9.2 | - |
| 28 | 4 | F | 8.4 | + |
| 29 | 32 | F | 10.9 | - |
| 30 | 27 | M | 11.5 | - |
| 31 | 43 | F | 11.4 | + |
| 32 | 4 | M | 9.4 | + |
| 33 | 62 | F | 10.3 | + |

CCRDA

Table 1 (Continued)

| Individual | Age (years) | Sex | Haemoglobin (g/100 ml) | Hookworm infection |
|---|---|---|---|---|
| | 33 | M | 12.7 | - |
| | 8 | M | 10.2 | + |
| | 37 | M | 11.2 | - |
| | 56 | M | 12.2 | - |
| | 39 | F | 10.0 | - |
| | 37 | M | 12.2 | - |
| | 6 | M | 10.1 | + |
| | 2 | F | 11.0 | - |
| | 14 | M | 10.3 | + |
| | 48 | F | 8.4 | + |
| | 62 | F | 8.9 | + |
| | 0 | M | 13.1 | - |
| | 16 | F | 9.3 | + |
| | 9 | F | 6.4 | + |
| | 24 | M | 11.7 | - |
| | 20 | F | 10.0 | - |
| | 68 | F | 9.1 | + |
| | 12 | M | 10.8 | - |
| | 24 | M | 11.8 | - |
| | 71 | M | 6.6 | + |
| | 45 | M | 8.7 | + |
| | 3 | M | 9.8 | - |
| | 52 | M | 12.4 | - |
| | 21 | F | 12.0 | + |
| | 33 | F | 10.6 | - |
| | 7 | F | 10.8 | - |
| | 42 | F | 11.0 | - |
| | 5 | M | 10.0 | + |
| | 36 | F | 10.9 | - |
| | 1 | F | 7.4 | + |
| | 8 | F | 11.6 | - |
| | 29 | F | 10.8 | - |
| | 4 | F | 9.8 | - |

ECRDA

Table 1 (Continued)

| Individual | Age (years) | Sex | Haemoglobin (g/100 ml) | Hookworm infection |
|---|---|---|---|---|
| 67 | 0 | F | 8.6 | + |
| 68 | 2 | M | 10.2 | + |
| 69 | 13 | F | 10.4 | + |
| 70 | 35 | M | 10.8 | + |
| 71 | 63 | M | 7.0 | + |
| 72 | 35 | M | 11.0 | + |
| 73 | 42 | M | 10.2 | + |
| 74 | 3 | F | 10.2 | + |
| 75 | 4 | M | 9.4 | + |
| 76 | 23 | M | 6.8 | + |
| 77 | 34 | M | 12.2 | - |
| 78 | 41 | F | 10.0 | + |
| 79 | 1 | M | 9.4 | - |
| 80 | 51 | M | 9.6 | + |
| 81 | 36 | M | 11.2 | + |
| 82 | 8 | F | 7.9 | + |
| 83 | 11 | M | 9.2 | + |
| 84 | 14 | M | 12.6 | + |
| 85 | 38 | F | 10.2 | - |
| 86 | 56 | F | 9.1 | + |
| 87 | 19 | F | 9.4 | + |
| 88 | 27 | F | 10.8 | - |
| 89 | 9 | M | 12.7 | - |
| 90 | 36 | F | 6.2 | + |
| 91 | 14 | F | 10.1 | + |
| 92 | 42 | M | 12.9 | - |
| 93 | 10 | M | 11.2 | - |
| 94 | 39 | M | 8.4 | + |
| 95 | 11 | F | 11.6 | - |
| 96 | 8 | F | 9.7 | + |
| 97 | 23 | M | 11.9 | - |
| 98 | 13 | M | 7.8 | + |
| 99 | 17 | M | 9.8 | + |
| 100 | 41 | F | 12.1 | |

CCRDA

All tables must have a proper title which should be complete, concise and self-explanatory. The title should give information on the sample being analyzed and on the variables being presented.

Table 2 shows the age distribution by broad age groups of the sample of 100 villagers. As the tabulation involves only one variable, namely age, we have a simple or one-way table. In this example, age is treated as discrete measurement and broad age groups have used because the number of subjects is fairly small. To group the villagers, the two opposite extreme values are found from the set or raw data, (in this case, less than 1 year old and 71 years old). The first category in the table includes the smallest value and the last category the highest. Having formed the classes, we can then process the raw data by hand tallying or sorting.

If a larger sample were being analyzed, the use of a larger number of age classes might be necessary. In that case, the age groups might be: under 1 year old, 1-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64 and 65+ years.

Table 2. Age distribution of the sample of 100 villagers

| Age group | Numbers |
|---|---|
| [illegible] | 19 |
| [illegible] | 25 |
| [illegible] | 40 |
| [illegible] | 16 |
| Total | 100 |

Table 3 shows the age and sex distribution of the sample of 100 villagers as an example of a two-way table. The age-sex distribution of the sample can now be compared to that for the district. If the two are similar, the sample is probably reasonably representative of the whole district population.

CCRDA

Table 3. Age and sex distribution of the sample of 100 villagers

| Age group (years) | Male | Female | Total |
|---|---|---|---|
| 0-4 | 10 | 9 | 19 |
| 5-14 | 12 | 13 | 25 |
| 15-44 | 20 | 20 | 40 |
| 45+ | 7 | 9 | 16 |
| Total | 49 | 51 | 100 |

For continuous measurement variable, such as haemoglobin level, a simple tabu
starts by grouping the data in categories, a shown in Table 4.

Table 4. Distribution of haemoglobin levels in 100 villagers

| Haemoglobin level (g/100 ml) | Number of villagers |
|---|---|
| 6.0-6.9 | 5 |
| 7.0-7.9 | 6 |
| 8.0-8.9 | 13 |
| 9.0-9.9 | 22 |
| 10.0-10.9 | 29 |
| 11.0-11.9 | 14 |
| 12.0-12.9 | 9 |
| 13.0-13.9 | 2 |
| Total | 100 |

CCRDA

To make up the categories, the two extreme haemoglobin values (i.e. the low
highest) are first identified; the difference between them gives the range. In this exam
lowest reading in the sample is 6.2 and the highest 13.8; the range is therefore 7.6 gra
ml. The range is then divided by 10 to obtain a crude value for the class interval. (Sinc
10 categories or groups usually gives a good distribution, dividing the range by 10 is
starting point.) This would give an interval of about 0.8 g, but as a whole number is pr
intervals of 1 g are used. If the number of categories is too small or too large, it may be
to make sense of the distribution of the variable.

## Cross-tabulation

Cross-tabulations involve the use of at least two variables. In its simplest form, a cross-tabulation consists of 2 rows and 2 columns (excluding the row and column totals). Table 5 gives an example of a two-way table, which is often used when both variables are classified as "present" or "absent".

Table 5. The distribution of hookworm-infected subjects by haemoglobin levels of less or more than 10 g/100 ml.

| Haemoglobin level | Hookworm-infection present | Hookworm-infection absent | Total | % with hookworm |
|---|---|---|---|---|
| Less than 10 g/100 ml | 35 | 11 | 46 | 76.1% |
| 10 g/100 ml or more | 24 | 30 | 54 | 44.4% |
| Total | 59 | 41 | 100 | |
| % with anaemia | 59.3% | 26.8% | | |

A haemoglobin value of less than 10 g/100 ml can be used as an arbitrary cut-off point for the diagnosis of anaemia, which affected 46 out of the 100 villagers. Over half of the villagers, 59 in all, were found to be infected with hookworm. Of those villagers with hookworm infection, 59.3% were anaemic compared with 26.8% of those with no hookworm. a further analysis by age and sex could be made using the Chi Squared test, which is a useful test of statistical significance to apply to "2 by 2" tables. Use of the test on these data would show that in these villages anaemia was statistically associated with hookworm infection (P, or probability,

Table 6 gives a further example of a cross-tabulation involving the two variables haemoglobin and sex, but this time with more than two categories for haemoglobin. The grouping of the data for haemoglobin has been done in the same manner as in the simple or one-way tabulation shown in Table 4.

An example of a three-way tabulation is illustrated by Table 7. The three-way table is constructed by processing all the data on the three variables, namely haemoglobin, age and sex, given in the set of raw data. From this three-way table, one-way or two-way tables can be obtained by looking at the totals. For example, in the three-way tabulation given in Table 7, the row totals give the simple tabulation for haemoglobin levels.

CCRDA

Table 6. Distribution of the 100 villagers by haemoglobin levels and sex

| Haemoglobin level (g/100 ml) | Sex | | Total (both sexes) |
|---|---|---|---|
| | Males | Females | |
| 6.0-6.9 | 2 | 3 | 5 |
| 7.0-7.9 | 3 | 3 | 6 |
| 8.0-8.9 | 6 | 7 | 13 |
| 9.0-9.9 | 10 | 12 | 22 |
| 10.0-10.9 | 10 | 19 | 29 |
| 11.0-11.9 | 9 | 5 | 14 |
| 12.0-12.9 | 7 | 2 | 9 |
| 13.0-13.9 | 2 | 0 | 2 |
| Total | 49 | 51 | 100 |

Table 7. Distribution of the 100 villagers by age, sex and haemoglobin level

| Haemoglobin level (g/100 ml) | Age group (years) | | | | | | | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0-4 | | 5-14 | | 15-45 | | 45+ | | | |
| | M | F | M | F | M | F | M | F | M | F |
| 6.0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 3 |
| 9.0 | 4 | 2 | 2 | 3 | 3 | 5 | 1 | 2 | 10 | 12 |
| 10.0 | 2 | 2 | 5 | 5 | 2 | 10 | 1 | 2 | 10 | 19 |
| 11.0 | 0 | 1 | 1 | 2 | 8 | 2 | 0 | 0 | 9 | 5 |
| 12.0 | 0 | 0 | 2 | 0 | 4 | 2 | 1 | 0 | 7 | 2 |
| 13.0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| Total | 10 | 9 | 12 | 13 | 20 | 20 | 7 | 9 | 49 | 51 |

CCRDA

## Summarizing Statistics

All the data on a variable can often be expressed in a much briefer way by using summarizing statistics or indices.

The most useful indices are:

- subjects infected,
  e.g. schoolchildren infected with S. haematobium or malaria.
- Percentage of subjects above or below a certain cut-off point,
  e.g. adults with a systolic blood pressure of 160 mm Hg or more in hypertension and pregnant women attending antenatal clinic at least once.
- The average is another name for the mean and the range expresses the difference between the lowest and highest values in the raw data.
  e.g. mean birth weight and average number of visits made by children to child health clinics.
- (often called SD) can be used when there is a reasonably normal or bell-shaped distribution about the mean. The SD can easily be calculated using some scientific electronic calculators. A large SD indicates a wide scatter of individual values on either side of the mean or average while a small SD value indicates a narrow distribution of values.
  e.g. severe malnutrition defined as those children more than two SDs below standard weight-for-age; normal range for many laboratory tests is often expressed as the mean plus or minus 2 x SD.

## Correlation

When two quantitative variables are associated or correlated with each other, an increase or decrease in one is associated with an increase or decrease in the other. For example, the newborn baby's weight is positively correlated with mother's weight.

Whenever there are two continuous variables, often measured on the same individuals, a very useful technique is to draw a scatter diagram, or scatter gram. The more closely correlated the two variables are, the closer the dots will appear to be on a straight line. This correlation can also be calculated using statistical formulae and it is then expressed as a correlation coefficient.

Whenever two variables are thought to be correlated it does not necessarily mean that one variable is the cause of the other. Great care must be taken in interpreting the relationship between variables. For instance, a mother's height may partly determine the baby's weight but the reverse is biologically not true.

CCRDA

**Standardization**

To make comparisons, for example between the prevalence of malnutrition in the di 10 years ago and at present, or between disease prevalence rates in the district and in the co as a whole, it is necessary to standardize for the different populations. This avoids compari between two populations with different shaped population pyramids and age-sex distributi

To handle increases in population size, we commonly use a percentage or a rate o 1000 people, which can be made more accurate by using an age-, sex- and disease-spe percentage or rate, e.g. 20% of male children 5-9 years old were infected with Bancro filariasis in 1986 compared with 10% in 1976.

Where the age-sex structure or pyramid of the district is noticeably different fro national one, statistical adjustments should be made before comparing district with national r A technique called direct standardization can be used to do this. It is important to standa for the age structure of the district population where a variable, such as the risk of dea unevenly distributed according to age and for the sex structure where the variable, su number of births, is associated with one sex. The crude death rate (CDR) could be much h in one district because there is a higher percentage of children living (and dying) in that di compared with another one nearby. In this case comparing the CDRs for districts would n valid; the same is true if the district rates for today are compared with those of 10 to 20 ago.

CCRDA

# APPENDIX 1 Ethical guidelines for epidemiological investigations

These guidelines were developed by the Scientific Working Group on Epidemiology of the UNDP/World Bank/WHO Special Programme on Research and Training in Tropical Diseases, World Health Organization, Geneva.

## Background

These guidelines focus specifically upon ethical aspects that have posed particular problems in epidemiological studies of communities in developing countries. They are intended as a supplement to the Declaration of Helsinki[a] and the WHO/CIOMS guidelines on research involving human subjects,[b] and should not be viewed in isolation. They represent provisional guidelines for consideration, not rules for execution.

The international declaration concerning ethics as set forth in the Declaration of Helsinki, adopted by the 18th World Medical Assembly in 1964 and revised by the 29th World Medical Assembly in 1975, covers well most of the important ethical considerations concerned with clinical research in human subjects. The WHO/CIOMS proposed international guidelines for biomedical research involving human subjects were framed with special reference to developing countries and focus on problems of informed consent from certain disadvantaged subjects, such as children, the mentally ill or prisoners, and broadly ethical problems of developing countries and community-based research. Further, the guidelines discuss ethical review procedures at institutional and national levels and touch upon the need to inform communities in which research will be undertaken and upon the need for compensation from accidental personal injury.

The ethical concerns of controlled clinical trials, which are well outlined in these documents, include some that are also important to epidemiological studies. In particular the ethical issues surrounding three matters—informed consent, use of control or comparison (e.g. placebo) procedures, and trials of drugs or other materials that have not been approved for use in humans in the country of origin—all may be of great importance in epidemiological studies in developing countries, but are not specifically reviewed here.

---

[a] *WHO Chronicle*, 30: 360-362 (1976).

[b] *Proposed international guidelines for biomedical research involving human subjects*. Geneva, CIOMS, 1982.

## *Level of review*

Proposals for research activities should undergo several levels of procedural review, as described below.

### *Individual*

The fundamental level of review is the individual person; every participant in a study must provide informed consent to take part in the research. The many complex issues that informed consent may involve are generally well covered in the above-mentioned documents. Illiteracy and differing cultural concepts of health and disease do not alter the basic principles of informed consent. Written consent may be a legal requirement in some countries, but from the ethical viewpoint it is neither necessary nor sufficient. Written consent *per se* is not the requirement; the critical need is that the person consenting has an understanding of the procedures, including the benefits and hazards. Informed consent should not be considered as an inhibiting factor in research: indeed, a properly informed subject may be a much more useful participant.

### *Community*

The second level of review, which may be of particular importance for some types of epidemiological research, is a review by the community. It is not simply a matter of consent; it is a matter of understanding. For a research programme involving a community the community and the researchers should be working in partnership since the purpose of the research should be mutually rewarding. A well informed community will not only be more cooperative, but also more useful in its collaboration.

### *National level*

The third level is that of the responsible institutional or national review body and is well covered in the WHO/CIOMS guidelines.

### *International*

The final level of review for externally sponsored projects (i.e. those funded by an external national or international agency) should ensure that the research protocol conforms to the requirements of the Declaration of Helsinki and that the approaches to the three prior levels are documented.

Detailed procedures for obtaining individual consent or community understanding will vary considerably and must be left to the responsible institution and/or national review bodies and to the principal investigator.

Obtaining approval at all levels and, in particular, obtaining the informed consent of individuals in no way reduces the responsibility of the investigator to these individuals and their communities. A checklist of suggested areas for review at each level is given at the end of this appendix.

CCRDA

# Issues of particular concern to epidemiological studies in developing countries

## Health care expectations of the community

Epidemiological studies made in a community or other defined population in developing countries may involve important considerations that are not encountered in clinical research. In many areas, the only medical care available to the population is that of traditional practitioners, and modern health care with the infrastructure required for its delivery may be minimal or absent. The conduct of a health-related research programme brings with it the quite reasonable expectations on the part of the people that some kind of health care will be provided. One-shot, "bleed and fly" safari-type studies have often produced unfulfilled expectations in the people, leading to disappointment and reduced cooperation for future studies. Although one-shot studies have had their usefulness – and may not raise ethical issues when directed at solving a specific health problem such as an epidemic – they are generally unacceptable today. Such studies, usually carried out by outsiders, in addition to raising expectations among the people, make no long-term contribution to the health infrastructure or the research capabilities of the country. Almost always the same information can be better obtained through direct cooperation with institutions in the host country.

Longitudinal community-based studies inevitably raise the expectations of the people for improved health care. There are no clear answers to the vexing questions raised by these expectations. The first step is that a reasonably full understanding must be achieved between the community and the researchers before, during and after the research. Frequently, during a longitudinal, community-based study, the health infrastructure must be strengthened in order to obtain information about the health of the people. When this is the case, every attempt should be made to strengthen it on a continuing basis by such means as training, improved record-keeping, development of sampling procedures for future use. In general, it is understood that the actual provision of health care is the responsibility of the national health services and community: therefore, the researcher must work closely with both and resolve these issues as fully as possible.

## Involvement of local personnel

Related to this is the need to strengthen the research infrastructure on a continuing basis whenever possible. At the least, this means as much involvement as possible of local personnel in the early stages of design of the protocol, its implementation, the collection of data, and the analysis and feedback of the information to the community and the government.

COPY

### *Control or placebo groups*

In establishing the effectiveness of virtually any intervention procedure (vaccine trial, mass drug distribution, etc.) on a community basis, the need for control or placebo groups may create special problems. The basic ethical issues are the same as those in clinical trials, but the procedures used may be different. Important principles are (1) that the control group receive the best currently established form of intervention, if one exists, and (2) that if the new intervention procedure is demonstrated to be better (e.g. higher benefit/cost ratio) then it should be provided to the control population as soon as possible.

All the people concerned should be fully informed about the proposed research. They should be provided with a clear explanation of the investigation, the reasons for undertaking it and the possible implications. Any coercion is unacceptable.

Sometimes there may be practical problems in determining who truly represents the community; occasionally it is necessary to work through dual channels where there may be both legal and traditional authorities. Generally full discussions with community leaders – before, during and after the completion of the research – are both necessary and very helpful. Again to be emphasized is that the ultimate objective of research is to improve the health of those in the community; hence a cooperative partnership arrangement to ensure full understanding on the part of both the researchers and the community is vital.

### *Use of past medical records*

Epidemiological studies carried out from a hospital or clinical base rarely raise ethical issues that are different from those of any other clinical research investigation. Although there has been some concern in the United Kingdom and the USA concerning the use of past medical records without specific individual consent of both the doctor and the patient as a potential "invasion of privacy", it is no longer considered an issue provided that there is no identification of individual patients and that complete confidentiality is assured. As long as these principles are adhered to, the issue seems unlikely to become a problem in developing countries.

### *Anonymity of the community*

The anonymity of the community, in general, should be protected with the same degree of concern as that given to the protection of the individual. If, as often happens, sensitive data emerge in the course of the studies of a community, it will be the responsibility of the investigator to use the utmost discretion in relating these data. However, sometimes the location and the circumstances are important to an understanding of the research that has taken place. Very often no purpose is served in not revealing the identity of the community. Indeed sometimes a community may take pride in being

CCPDA

associated with the research. The best approach is to have continuing communication with the community as suggested above in the community review procedures (level 2 review).

### *Behavioural research*

It has sometimes been forgotten that informed consent may apply to observation of personal behaviour just as with any other research method affecting human subjects. The use of hidden observers and secret observation procedures (including photography) is rarely justifiable in behavioural research. When observers are required, as in research into human contact with water, their existence should be made known, but they can, of course, be stationed inconspicuously without being hidden. Experience has shown that such observers are eventually accepted—and even ignored.

### *Environmental effect*

In wide-scale application of measures to control vectors or intermediate hosts of disease organisms, the effect on the community and the environment of the methods used must be anticipated and carefully monitored.

### *Ethics in training*

Training at all levels should include consideration of the ethical issues that will be involved in the trainees' future duties.

## Checklist

The following is a checklist of suggested areas for review at each level:

- **Parent institution; funding organization**
  - scientific merit of the study
  - consistency with ethical guidelines
  - monitoring of results
  - implementation of recommendations.
- **National review committee**
  - scientific merit of the study
  - consistency with ethical guideline
  - use of controls
  - confidentiality of records
  - anonymity of subjects
  - use of non-intervening observers
  - cost to the community under study
  - means of communication to the community of the nature of the research

CCRDA

- nature of informed consent
- acceptance of the research by the study community
- potential conflict and competition between proposals
- termination of controlled trials
- end points in the research process
- implementation of recommendations
- communication of results to the community.

- **Local community**
  - nature and necessity for research
  - possible harm and possible benefits to the community
  - communication of issues to the people
  - obtaining of consensus and the non-desirability of coercion
  - monitoring of results
  - implementation of recommendations
  - end points.
- **Individual**
  - informed consent is required from all individuals.

## *Acknowledgements*

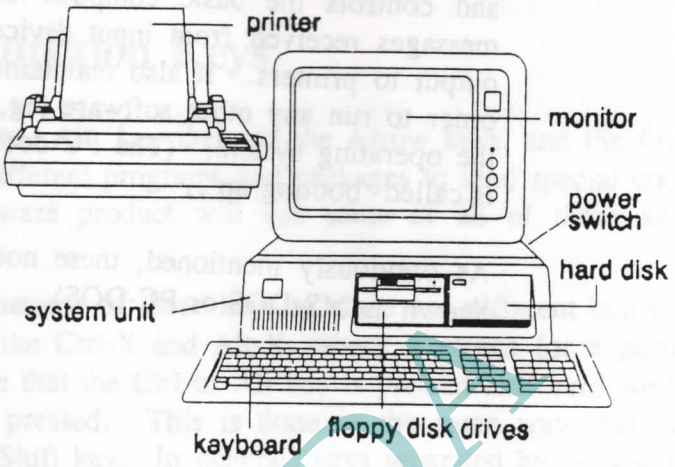
These guidelines were originally drafted by Professor H. M. Gilles, Dean, Liverpool School of Hygiene and Tropical Medicine, with assistance from Professor F. Dunn, University of California, Berkeley. A number of useful suggestions from members of the Scientific Working Group on Epidemiology and the secretariat of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases have been incorporated into the present version. Special thanks are due to Professor I. Riley, formerly Professor of Community Health, University of Papua New Guinea, for his comprehensive review and for the checklist of suggested areas for review and to Dr P. Rosenfield, formerly Secretary, Scientific Working Group on Social and Economic Research, for her compilation and analysis of the many useful suggestions submitted.

CRDA

Appendix 2: Introduction to microcomputer use in data processing

# HARDWARE vs SOFTWARE

## Computer Hardware




Computer hardware is anything you can touch—the machinery. The hardware that makes up a basic PC includes a monitor, a keyboard and a system unit. Among other things the system unit contains the computer chip, the hard disk drive and the floppy disk (or diskette) drive. Other items of computer hardware include a printer, a mouse and a modem.

CCRDA

# Computer Software

Software is the term applied to programs (lists of computer instructions) the hardware to do something. STATA and WordPerfect are ex software.

The operating system is a special type of software. It gets the compute and controls the basic computer activities. For example, it co-or messages received from input devices (eg, the keyboard or mouse) a output to printers. It also maintains an orderly system of files on th order to run any other software (eg, STATA, WordPerfect) you first the operating system. (The PC does this automatically when you turn is called 'booting up'.)

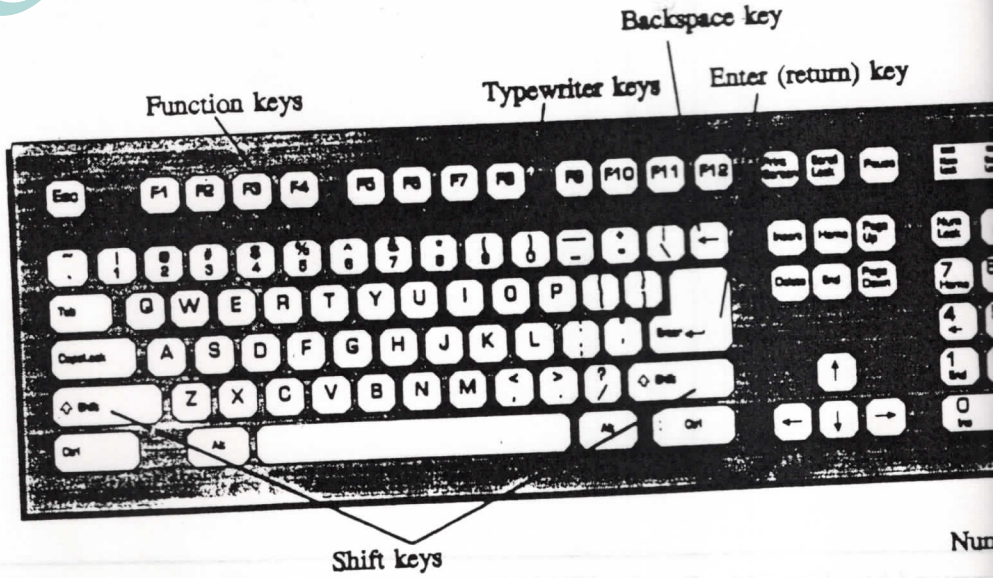
As previously mentioned, these notes focus on the operating system known as MS-DOS or PC-DOS).

CCRDA

# THE KEYBOARD

## Keyboard Layout

The keyboard layout of a typical personal computer




(Some keyboards have only ten function keys, positioned on the lef the arrows and insert/delete group of keys are included only keypad.)

## Standard Typing Keys

The standard typing keys are the same as those used on a conventional typewriter. These include letter and number keys, the ***Tab*** key, and the ***Shift*** keys. Your ***Shift*** key is used to type uppercase letters and the symbols above the upper row of number keys, just as on a typewriter. The ***Caps Lock*** key lets you type alphabet characters in upper case without pressing a ***Shift*** key. When the ***Caps Lock*** is on, pressing ***Shift*** produces lowercase characters.

## Special Control and Function Keys

Special control keys like Ctrl, Alt, Ins, Del and the Arrow keys, and the function keys F1-F12 are used by different programs and packages to send special codes to the computer. Each software product will use some or all of these keys to perform different tasks.

Many packages use combinations of these keys to stand for different commands. For example, expressions like Ctrl-X and Alt-X, where X stands for a particular keyboard character, indicate that the Ctrl or Alt key is pressed ***and held*** while the indicated character key is pressed. This is done in the same way that capital letters are typed using the Shift key. In general, keys separated by '-' are meant to be pressed at the same time to get the indicated function.

CCRDA

# THE COMPUTER CHIP

The computer chip is located inside the system unit. The two important components of the computer chip are:

1. central processing unit (CPU), and
2. the memory of the computer.

The CPU can be thought of as the brain of the computer. It interprets and processes information. Memory is discussed in the section below.

## Memory

To be executed, programs (software) must be loaded from the disk (see later) into memory (sometimes referred to as Random Access Memory, or RAM). In order for programs to access a data file, this must also be loaded into memory.

Any data stored in memory will be lost once you exit STATA. If you want to the data at some later stage you must save it to disk before exiting STATA. N also that if the computer loses power even for a fraction of a second, any data memory is lost. Therefore, during a long STATA session where you have mad number of changes to your data set it is advisable to save your data to disk regular intervals.

The important thing to remember about memory is that it is temporary. permanently store data you need to save it to disk.

## Measuring Memory

The commonly used unit of measurement for describing computer memory byte. A byte can store a single character (eg, the letter 'a', the digit '3' o punctuation mark '?'). A kilobyte (KB) is about 1,000 bytes and a mega (MB) is about 1,000 kilobytes.

Computer chips typically come with 1 MB of memory. Extra chips can be to expand the memory of the computer to 2, 4, 10 or more MB.

CCRDA

# DISKS and DISK DRIVES

The hardware that permanently stores information such as programs (softwa data are called disks. Disk drives read (or write) information from (or to) (or from) memory.

As with memory, the space available on disks to store information is meas bytes. Unlike memory the information stored on disk is not lost wh computer is turned off.

There are two kinds of disks—hard disks and floppy disks (also called diske

The hard disk is a non-removable disk that is usually built into the systen the time of manufacture. Hard disks store very large amounts of inf (anything from 20 MB to 500 MB or more). Hard disk drives are usually 'C:'.

Diskettes (floppy disks) on the other hand store much less information portable. They come in various sizes. Commonly used sizes are 5¼ inc disk that can store 1.2 MB of information and 3½ inch diskette that can s MB. Floppy disk drives are usually denoted 'A:' or 'B:'.

## FILES

Computer programs (software) and data are stored in files. These files are located on the hard disk or a floppy disk. Each file has a filename. DOS has rules for naming files.

- they must be 1 to 8 characters in length,
- they must start with a letter but can include numbers,
- they are not case sensitive—it does not matter whether you use uppercase or lowercase letters when you type them,
- they can have an optional extension of up to 3 characters which is always separated from the rest of the filename by a full stop.

eg, **`myfile1.ext`**

## DIRECTORIES AND SUBDIRECTORIES

A disk can store several hundreds (or even thousands) of files. This can become very confusing and muddled. To overcome this problem, DOS allows you to group related files into directories.

For example, the STATA program files are stored in a directory called 'STATA'. The STATA manual advises users not to use this directory to store your data and graphics files because:

1. you could confuse your data sets with the files that are used to run STATA, and
2. you could accidentally delete an important file and STATA would stop working.

You should therefore create a separate directory (or directories). For instance, you might want to create a directory called 'BIOS' to store the data files for the Biostatistics course. How to create directories is discussed later in this Module.

Note that sometimes directories that are within other directories are called subdirectories.

### The Root Directory

Every disk (or drive) (eg, A: or C:) has at least one directory—the root directory. All other directories branch out from the root directory. In DOS commands, the

CCRDA

root directory is not referred to by its name. Instead, it is represented backslash (\).

You should store as few files as possible in the root directory. The best pra is to make directories off the root directory to store files.

## Directory Tree

The organisation of directories is called a directory tree. Perhaps more correctly it should be described as an upside down tree as the root directory is always placed at the top of the diagram.

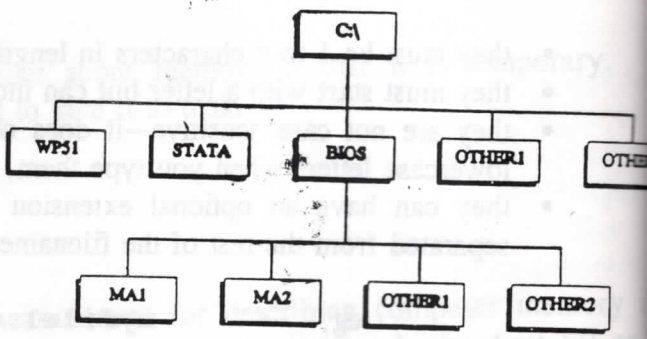



Illustration of a directory structure

# THE DOS PROMPT

When you first turn your PC on, it runs through a number of set procedures include starting the operating system. This is called 'booting up'. This usu takes one or two minutes. When this is complete, you should see the following

```
C:\>
```

This is called the DOS prompt (or command prompt). The flashing unders next to the DOS prompt is called the cursor. The cursor shows where the character you type from the keyboard will appear.

Note that only DOS commands will work w you are at the DOS prompt. common mistake is to try to type STATA co nds at the DOS prompt.

# THE 'ENTER' KEY

After you type a command you must press the ENTER key to send the comma to the computer for processing. That is, after every DOS or STATA comma you must press the ENTER key.

CORDA

## UPPER or LOWER CASE?

DOS is not case sensitive. That is, you can use either uppercase or lowercase letters when you type DOS commands. (In contrast, STATA is case sensitive. In STATA you must take care as to whether or not you should use upper or lower case characters.)

## COMMANDS for MAKING and MOVING BETWEEN DIRECTORIES

| Command | Action |
|---|---|
| `md` | Makes a new directory. |
| `cd` | Allows the user to move between directories. |

The following exercise gives you practice at using the above commands.

### EXERCISE 1

To make a directory called 'BIOS' off the root directory:

a) Ensure that you are in the root directory. That is, the DOS prompt should look like `C:\>` (If you are not at the root directory of the 'C:' drive, type: `cd\ <enter>`)

b) Type the following DOS command:

```
md bios <enter>
```

(You have now created a directory called 'BIOS'.)

To change from the root directory to the 'BIOS' directory:

a) At the root directory type: `cd\bios <enter>`

b) The DOS prompt should now look like `C:\BIOS>` indicating that you are now in the 'BIOS' directory.

*cont'd*

CCRDA

*EXERCISE 1 (cont'd)*

To make the directory 'MA2' off the 'BIOS' directory:

e) Ensure that you are in the 'BIOS' directory. That is, the DOS p should look like `C:\BIOS>`

f) Type the following DOS command:

```
md ma2 <enter>
```

(You have now created a directory called 'MA2' off the 'BIOS' directo

To change from the 'BIOS' directory to the 'MA2' directory:

g) At the 'BIOS' directory type:

```
cd\bios\ma2 <enter>
```

h) The DOS prompt should now like like: `C:\BIOS\MA2>` indicating tha are now in the 'MA2' directory off the 'BIOS' directory.

To change back to the root directory from the 'MA2' directory:

i) At the DOS prompt `C:\BIOS\MA2>` type:

```
cd\ <enter>
```

j) The DOS prompt should now look like `C:\>`

CCRDA

*Notes:*

1. The DOS prompt indicates your location in the directory tree. example, if the DOS prompt is `C:\BIOS>` then you know you are 'BIOS' directory off the root directory. The directory you are in is the current directory. Another name for the current directory is the directory.

2. Directory names can be from 1 to 8 characters in length and shoul with a letter. They are not case sensitive.

3. To change to the root directory from any other directory type:

```
cd\ <enter>
```

4. To change to other directories specify the 'path' (paths are explained in the section *Using Paths to Specify the Location of Files*) you would need to take from the root directory. For example, to change to the 'MA2' directory above you would type:

```
cd\bios\ma2 <enter>
```

5. STATA can access files in the directory that you start STATA from without you having to specify a path. It is therefore good practice to start STATA from the directory that contains the STATA data files you want to use (see Module C2, 'An Introduction to STATA').

## CHANGING DRIVES

Typically, PCs have at least two drives:

1. a hard disk drive, usually denoted 'C:', and
2. a diskette (or floppy disk) drive, usually denoted 'A:'.

You can change between drives by typing the drive letter followed by a colon at the DOS prompt. To change from the 'C:' drive to the 'A:' drive, you would type:

```
a: <enter>
```

You must have a diskette in the 'A:' drive for this command to work. After issuing this command, the DOS prompt should show `A:\>` indicating that you are now in the 'A:' drive. To get back to the 'C:' drive you would type:

```
c: <enter>
```

## VIEWING THE CONTENTS OF A DIRECTORY

To view the contents of the current (or default) directory use the `dir` command. The `dir` command stands for 'directory'. This command lists all the files and directories contained within the current directory.

Sometimes there are so many files and other directories in the current directory that they scroll off the screen. You can view the contents of a directory one screen at a time by typing:

```
dir/p <enter>
```

CCRDA

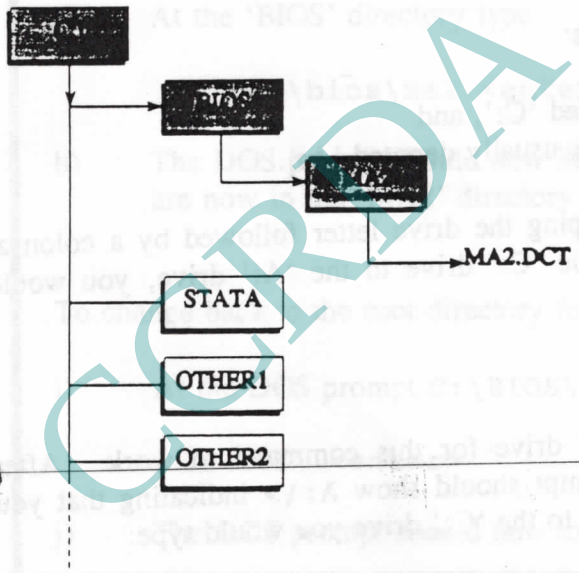
Another option is to type:

```
dir/w <enter>
```

which shows the contents of a directory in wide format.

Note that the list produced by the `dir` command shows a space rather tha
stop between the filename and extension (eg, `myfile1 ext`).

## USING PATHS TO SPECIFY THE LOCATION OF FILES




The term path refers to the cour
DOS must take, starting at th
directory, to get to files in
directory. For example, to access
MA2.DCT in the 'MA2' directory
must go from the root directory
the 'BIOS' directory to the
directory. This is shown in the
on the left.

You would specify the path as:

```
c:\BIOS\MA2\MA2.DCT
```

This is the path to the file MA
The first letter and colon (C:) rep
the drive that the file is on. Th
backslash (\) represents the root directory. The second backslash separa
'BIOS" directory from the 'MA2' directory. The third backlash separa
'MA2' directory from the filename, MA2.DCT.

## COPYING FILES

To copy a file you use the DOS command `copy`. When you use th
command you must specify two pieces of information:

1. the location and name of the file you want to copy (the source), and
2. where you want to copy the file to (the destination).

CCRDA

You separate the source and destination with a space. The `copy` command follows this pattern:

```
copy [source] [destination] <enter>
```

*EXERCISE 2*

To copy the file YOURDATA.TUT from the 'STATA' directory to the directory 'MA2' type:

```
copy    c:\stata\yourdata.tut    c:\bios\ma2\yourdata.tut
<enter>
```

Depending on which directory you are in you could shorten the command. For example, if you were in the 'STATA' directory you could type:

```
copy yourdata.tut c:\bios\ma2\yourdata.tut <enter>
```

That is, you don't need to specify the path of the directory that you are currently in.

Consequently, if you were in the 'MA2' directory you could type:

```
copy c:\stata\yourdata.tut yourdata.tut <enter>
```

To check that the file had been successfully copied you would first change to the 'MA2' directory (if this was not already the default directory) by typing:

```
cd\bios\ma2 <enter>
```

You would then type:

```
dir <enter>
```

to view the contents of the 'MA2' directory.

## DELETING FILES

To delete or remove a file you no longer want on your disk you use the `del` command. The `del` command stands for 'delete'.

Assuming the current directory is 'MA2', to delete the file 'YOURDATA.DTA' from this directory (and only this directory) you type:

CCRDA

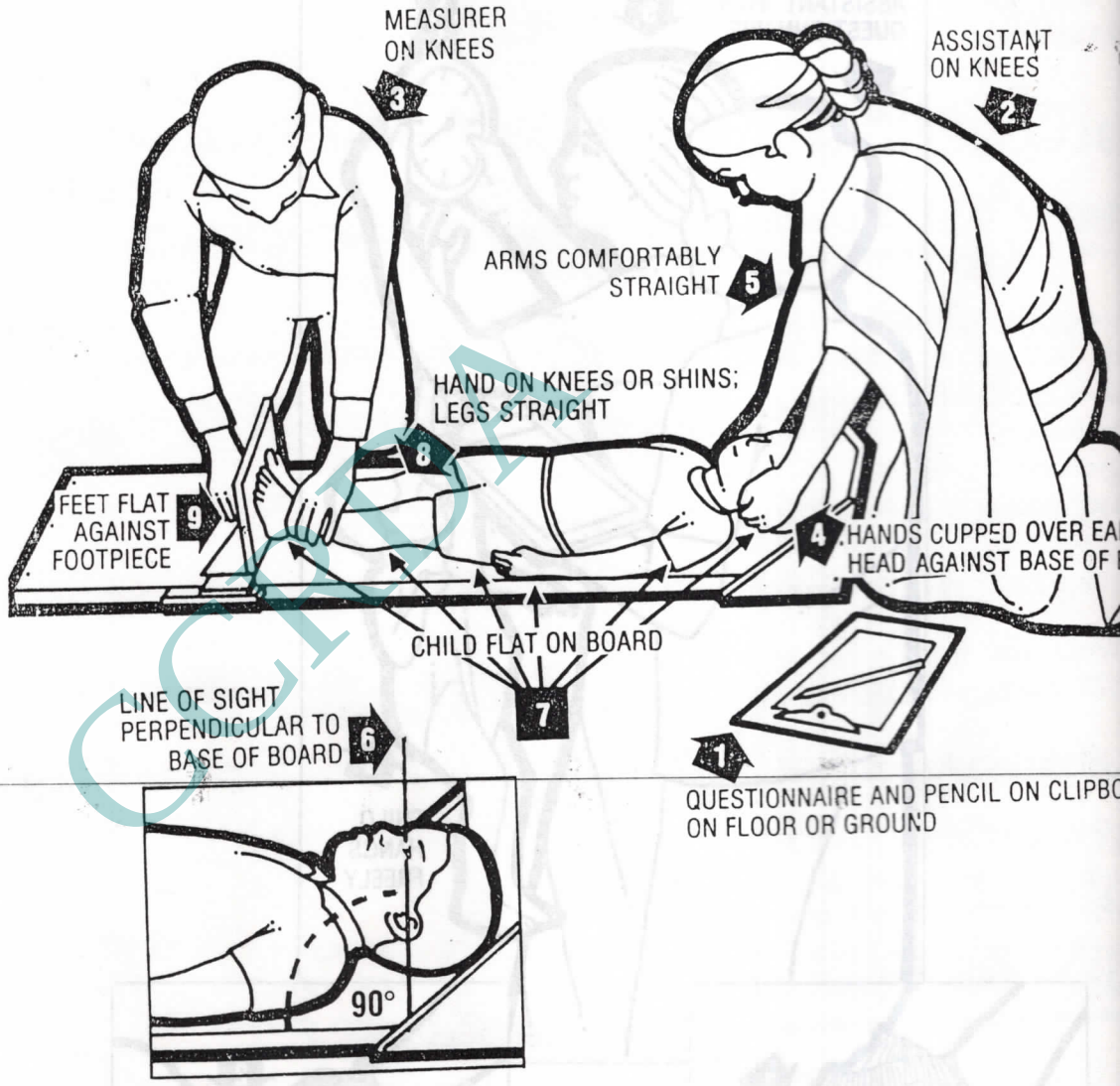
## Appendix 3c Anthropometeric Measurements

### Child Weight

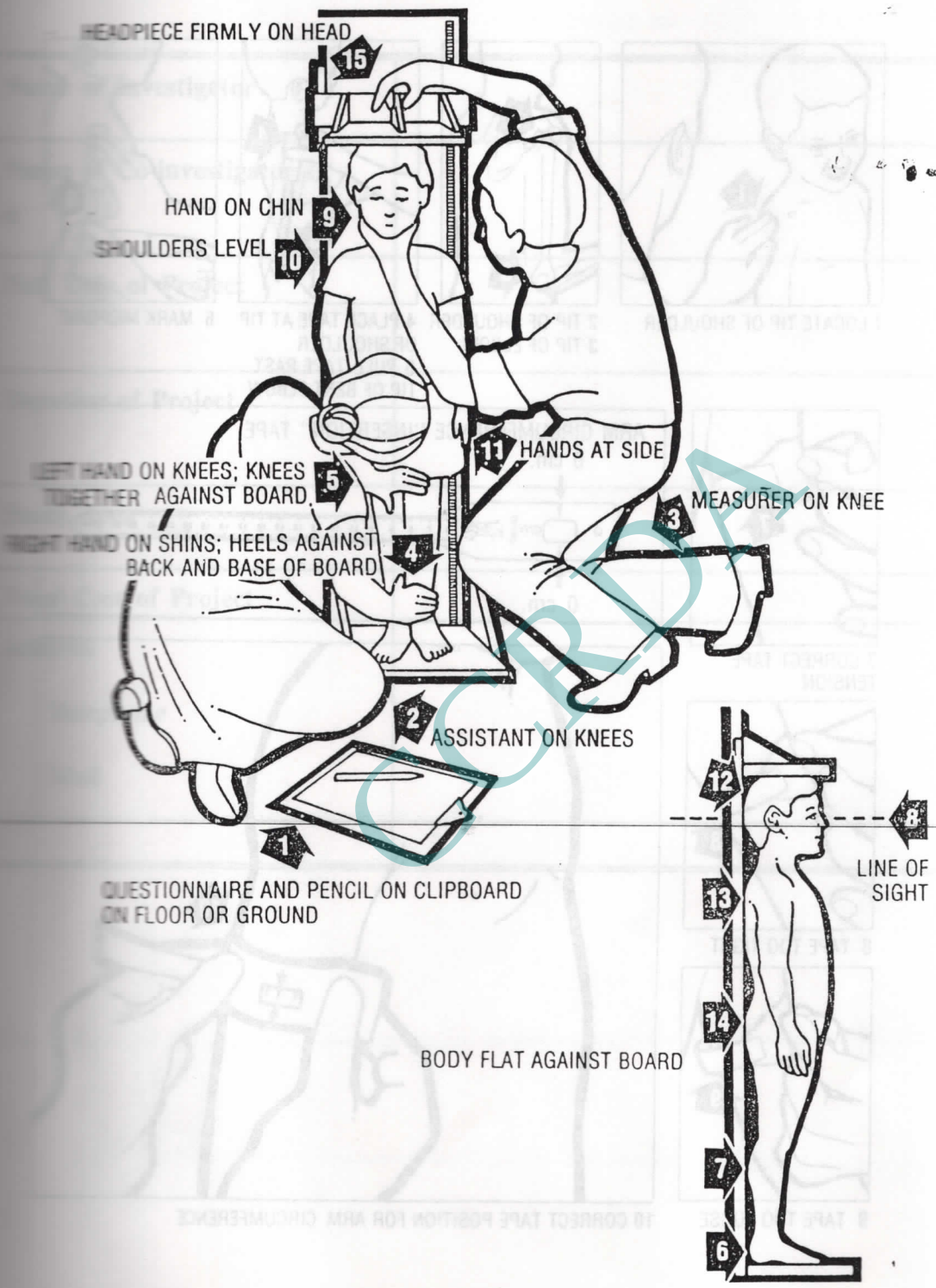



CORDA

## Child Length Measurement



## Child Height Measurement

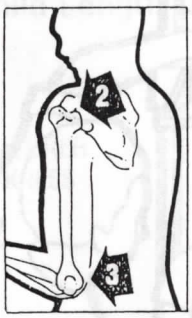



CCRDA

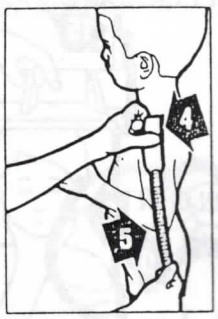
## Child Mid-Upper Arm Circumference Measurement



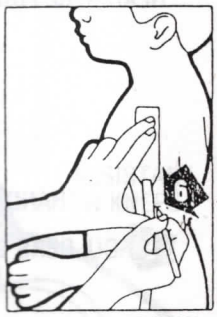

1 LOCATE TIP OF SHOULDER



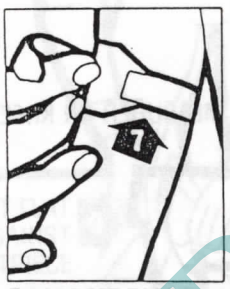

2 TIP OF SHOULDER
3 TIP OF ELBOW



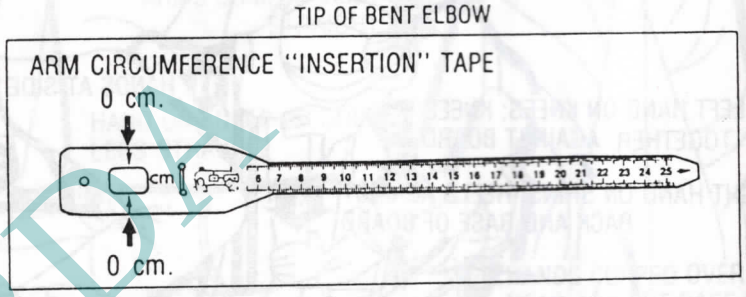

4 PLACE TAPE AT TIP OF SHOULDER
5 PULL TAPE PAST TIP OF BENT ELBOW

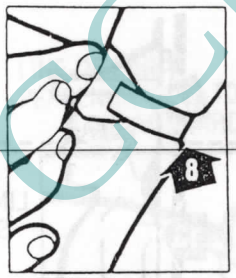



6 MARK MIDPOINT

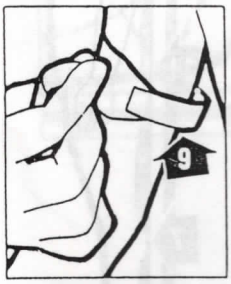



7 CORRECT TAPE TENSION

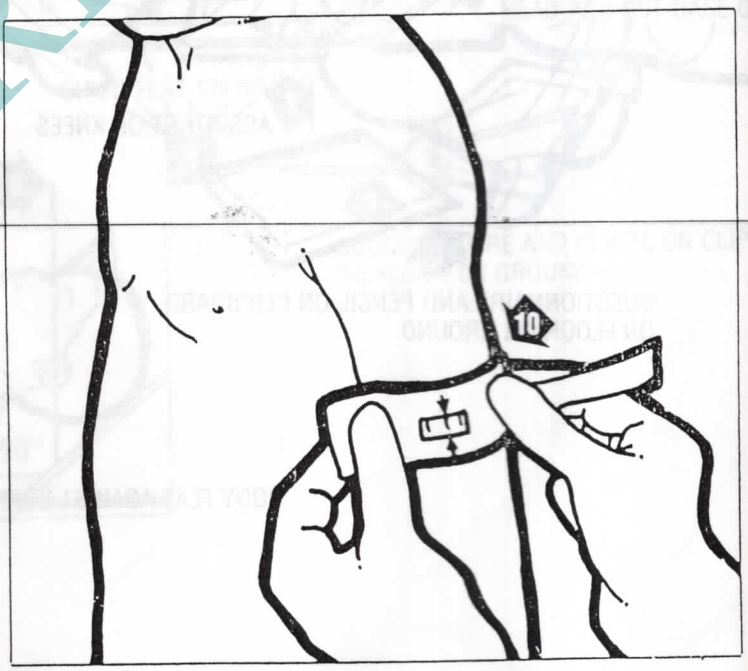



8 TAPE TOO TIGHT




9 TAPE TOO LOOSE




10 CORRECT TAPE POSITION FOR ARM CIRCUMFERENCE

ECRDA

## Appendix 4: Sample Survey Proposal Form

| | |
|---|---|
| 1. Name of Investigator | |
| 2. Name of Co-investigator(s) | |
| 3. Full Title of Project | |
| 4. Duration of Project | |
| 5. Study Area | |
| 6. Total Cost of Project | |
| 7. Address<br>Telephone<br>Mail | |

CCRDA

9. SUMMARY (Do not exceed space)

CCRDA

## III. STATEMENT OF THE PROBLEM AND BACKGROUND REVIEW

CCRDA

## 11. GENERAL AND SPECIFIC OBJECTIVES

CCRDA

12. METHODS

CCRDA

## 16. DISSIMINATION AND UTILIZATION OF RESULTS

CCRDA

| COST OF THE PROJECT. Give an itemized listing of the direct costs involved in the project. Omit cents. | | | |
|---|---|---|---|
| PERSONNEL COSTS | | | |
| Title | Rate | Duration | Total |
| | | | |
| Subtotal | | | |
| CONSUMABLES | | | |
| Description | Unit cost | Amount (#) | Total |
| | | | |
| Subtotal | | | |
| EQUIPMENT | | | |
| Type of equipment | | Estimated cost in ETH. Birr | |
| | | | |
| Subtotal | | | |
| TRANSPORTATION | | | |
| Description | Unit cost | Amount (#) | Total |
| | | | |
| Subtotal | | | |
| GRAND TOTAL COST OF THE PROJECT (Birr) : | | | |

CCRDA

## 18. ASSURANCE OF THE INVESTIGATOR

The undersigned agrees to accept responsibility for the scientific, ethical a technical conduct of the research project.

Name : _____

Signiture: _____ Date: _____

CCRDA