33

Epidemiology in Gynecological Diseases

Sisty J. Moshi

INTRODUCTION

This chapter will help you to plan preventive measures to control gynecological diseases in your area. The preceding chapters have talked about specific preventive measures and risk factors for each disease. This chapter will explain how to assess the need for such measures in your area and how to collect data on gynecological diseases in order to set priorities and discuss these priorities with policy makers. Furthermore this chapter will help you to better understand scientific literature on gynecological diseases by explaining statistical terms that are most commonly used.

BASIC EPIDEMIOLOGY

Epidemiology is the study of the distribution and determinants of health-related states or events, in a specified population and the application of this study in the control of a health problem[1].

History traces epidemiology back to 500 BC when Hippocrates wrote his book named '*On Airs, Waters and Places*'. In this book there was a chapter on environmental influences on health. This chapter mentioned that weather (hot, cold and winds), water quality (contamination), drinking and eating habits (behavior), indolence, exercise and labor have influence on disease occurrence. Thus Hippocrates concluded that disease does not just happen. There were several assumptions but all meant that diseases or other health events do not occur at random; instead they occur in specific populations exposed in a particular way (at risk). Also he assumed that diseases have causal relationships with some risk factors and so preventive measures can be identified. Therefore researchers investigate these factors and apply the knowledge to control disease.

Diseases or other health-related events do not occur at random. They occur in specific groups of people, exposed in a particular way (at risk) to a condition/environment that causes the disease in question. All diseases have cause(s). Also all diseases have preventive factors that can be identified. The problem is usually how to link the diseases to their causes so that prevention can be planned. Researchers try to connect diseases to possible causes in order to identify the definite cause. The knowledge from research is used to plan (or research) for effective control/prevention measures together with treatment if possible.

Gynecological diseases like other diseases have a causal relationship with some factors in the population/environment. These factors may be physical or/and social. Someone may suffer from gynecological diseases either due to her physical condition/exposure (e.g. nutritional status, environment, exposure to bacteria or viruses etc.) or social conditions (education, income level, culture etc.). So while dealing with gynecological disease clinically, it is recommended to look at its epidemiology. This is important not only for gynecological diseases because in many regions of the world large proportions of women are at more risk for diseases and complications compared with men. Many women are often less educated, have less power in decision-making, have lower economic status etc., and as a consequence are more prone to diseases and complications through late presentation for treatment.

There are several ways to analyse disease epidemiologically. Some of the common ways are mentioned below. The more advanced ones are not in the scope of this book.

Models of causation

Although people get diseases in various ways, all mechanisms are grouped into two: communicable and non-communicable. This has resulted in classifying diseases as either communicable or non-communicable:

1. *Communicable diseases* are the diseases that can be transmitted from one person to another with or without vector. Communicable diseases are commonly caused by micro-organisms. For example, sexually transmitted diseases (gonorrhea, syphilis, chlamydia, human papillomavirus etc.).
2. *Non-communicable diseases* are the diseases that cannot be transmitted from one person to another. Examples of non-communicable diseases include trauma/injuries, tumors (benign or malignant), congenital malformations etc.

This is one way of classifying diseases. There are several other ways of classifying diseases. Some of them include classifications based on the level of disease in the community (sporadic, endemic, epidemic or pandemic), or acute versus chronic diseases etc.

Epidemiological triad

For diseases to occur there must be an interaction between the host (human being), an agent and the environment. There are some factors among the three that favor the occurrence of a disease. For a disease to occur there must be an agent that usually causes that disease (e.g. bacteria, fungi, virus, mechanical force etc.). Also there should be a favorable environment for an agent to be effective in causing that disease. Lastly, there should be a person with some characteristics that cause her to get that disease. Interaction of these three factors is mandatory for disease occurrence. Interference with any of these factors will prevent disease from occurring, and can be considered as a preventive measure.

Host factors

These are conditions/characteristics that exist in the human being which influence the occurrence of disease. This includes things like behavior (promiscuity, lack of condom/contraceptive use, smoking etc.) age, socioeconomic status, nutrition, underlying diseases (e.g. diabetes mellitus, HIV/AIDS) parity, culture (e.g. female genital mutilation, vaginal cleansing, early marriages, polygamy), genetics etc.

Agent factors

These are conditions/characteristics that favor the ability of the causative agent to cause the disease. These include infectivity (the ability of a disease agent to enter, survive and multiply in the host), pathogenicity (ability to cause disease) or virulence (the ability of the agent to cause death).

Environmental factors

These are conditions/characteristics existing in the environment that favor the ability of an agent to cause disease or favor the susceptibility of the host to disease. Examples are availability of disease vectors, population density, cleanliness, air quality, weather, special environments (hospitals, day-care centers, refugee camps etc.)

For example, if a woman has a promiscuous behavior, and on top, she does do not use condoms (host factors) she may get pregnant unintended. She might decide for an induced abortion that may be illegal in her country, so she will have a criminal abortion in a place where sterility is poor (environmental factors). Staphylococcus aureus (very infective and pathogenic – agent factors) may be introduced into the uterus. The woman may get a severe infection, and later suffer from infertility as she could not afford proper treatment for her infection due to poverty (host factor).

Measures of disease frequency

They serve to measure 'how much disease is out there' at a given point in time. You can already see that by measuring disease frequency you can control whether your preventive activities were successful by comparing measures of frequency for the same disease at different moments, e.g. before and after your preventive activities.

This measurement can be done in various ways but the aim is always to obtain information about the pattern of infection, disease, mortality or any other event related to health in human populations. For example, how many people are living with cancer of the cervix in a population, or what fraction of a population has cancer of the cervix, or how quickly/slowly does a population produce

new patients (cases) of cancer of the cervix, how many deaths are caused by cancer of the cervix etc.

For measuring disease frequency there are several terms that are always referred to. They will be defined here:

- *A case* is the person in a population who has a particular disease or to whom the event of interest (e.g. sickness, death etc.) occurs. It is a countable instance in the population or study group with a particular disease. For example if dealing with cancer of the cervix, then a patient with cervical cancer may be referred to as a case.
- *A population at risk* is a group of people who share characteristic(s) that makes each member susceptible to a particular health event. The members share the same characteristics (e.g. host factors, environmental exposure) as cases except that they don't have the disease in question. This is usually the population from where the cases are originating. It is also called the vulnerable population. For example, if dealing with cancer of the cervix, the population at risk is all women in that population who have no cancer of the cervix but are exposed to human papillomavirus.

There are several ways to measure disease frequency. They can be used individually or in combination depending on what is to be achieved. These include count, ratio, proportion, rate, prevalence, incidence, and attack rate.

- *Count* is the simplest measure of disease frequency. It means simply to count the number of cases, i.e. how many cases there are. Although it is very basic in epidemiology and has many limitations, it is very useful. It is the starting point to get to other measures.
- *Proportion* is a quotient of two numbers of which the numerator is part of the denominator. It gives the information about what fraction is affected (or not affected) out of the total population. As numerator and denominator have the same dimension, dimensional contents cancel out, and so a proportion is commonly a dimensionless quantity. It is calculated by dividing cases by the population at risk:

$$\text{Proportion} = \frac{\text{Number of cases}}{\text{Population at risk (including cases)}}$$

It can be expressed as a fraction, decimal or percentage. For example if 7 women get cervical cancer in a population of 2500 women, then the proportion of women with cervical cancer is 7/2500 which is 0.0028 or 0.28%.

- *Ratio* is also a quotient of two numbers of which the numerator is not necessarily a part of the denominator. The numerator and denominator may have the same or completely different dimensions. It may be beds per doctor, e.g. 850 beds/10 doctors (85 beds per doctor) or 100 females/20 male (5:1) etc.
- *Rate* is also a quotient of two numbers. It tells about the speed or frequency of occurrence of an event in a specified population at risk or over a certain period of time. It gives the information of the number of new disease cases in a specified population or in a specified period of time. Components needed to calculate a rate are the numerator (number of events/patients in that time), denominator (population at risk in that time) and time in which the events occur.

$$\text{Rate} = \frac{\text{Number of cases in a year}}{\text{Total population in that year}}$$

Example: if in a period of 1 year, 14 women get cervical cancer in a population of 5000 women, then the rate is 14/5000 = 0.0028 per year or 0.28%.

Rates can be used to compare events in a certain specified number of populations like number of deaths per every 1000 or 100,000 people. For example maternal mortality rate is the number of deaths per every 100,000 live births, or under-5 mortality rate is the number of under-5 deaths per every 1000 live births. It can differ per region, country etc., per time, e.g. each year there may be new rates. The use of rates rather than raw numbers is essential for comparison of experience between populations at different times, different places or among different classes of persons. For example in the current year, under-5 mortality rate may be 217 deaths per every 1000 live births. If parents are educated and health services improved, the following year the rate may change to 150 deaths per every 1000 live births.

- *Prevalence* is the number of existing disease cases in a given population at risk, commonly at a specific point in time. This is point prevalence. Prevalence may be calculated for a period of time (period prevalence). It gives the information

about average number of people who have the disease of interest in a certain population at a time. It may be constant or it may change with time. For example if in January of the current year there are 70 patients with pelvic inflammatory disease (PID) in a population of 20,000 people, then the prevalence of PID in that community is 70/20,000 which is 0.0035 or 0.35%. If control measures are taken, and in January of the following year the number of cases drops to 20, then the prevalence changes to 20/20,000 which is 0.001 or 0.1%.

- *Incidence* is the number of new cases that occur in a given population at risk, during a defined period or the amount of people who develop the disease of interest in that specified period. It is used to measure the risk of developing a disease of interest within a specified period of time. Although sometimes loosely expressed simply as the number of new cases during a time period, it is better expressed as a proportion or a rate. It can be expressed as incidence rate or incidence density. It is calculated by using the following formula:

$$\text{Incidence} = \frac{\text{Number of new events in specified period}}{\text{Average number of persons exposed to risk during this period}} \times 10^n$$

The use of 10^n is to make it a whole number for easy interpretation. For example if in a year the number of people who acquire PID is 20 in a population of 2500 women, then the incidence is 20/2500 or 0.008 women per year. This becomes more meaningful if multiplied by a factor 10^n. If expressed per every 1000 women (i.e. $n = 3$) then incidence is $20/2500 \times 1000$ or 8 cases per every 1000 women per year. Incidence is also a rate. In some circumstances it is expressed in different ways depending on the health event in question. Some examples include morbidity rates, mortality rates, case fatality rates, attack rates etc.

Difference between incidence and prevalence

Assume the incidence is the amount of water being brought to a dam over a specific period of time (e.g. from 1 January to 31 December) and prevalence is the amount of water in a dam at a certain point in time (e.g. 1 August). If there is no leakage in the dam, the amount of water in the dam (prevalence) will depend on the amount of water being brought in by the river (incidence). If there are leakages in the dam (recovery, outmigration, deaths), prevalence will also be influenced by these leakages. So prevalence depends on incidence rate and duration of illness[1] (Table 1).

Table 1 Incidence and prevalence

Incidence	*Prevalence*
New cases or events over period of time	All cases at point/period of time
Useful for studying factors causing disease, 'risk'	Useful for measuring size of problem and planning
Numerator of incidence includes only new cases that occurred during a given time period	Numerator of prevalence includes all cases present at a point in time regardless of when the illness began

THE USE OF EPIDEMIOLOGY IN DISEASE MANAGEMENT

Clinicians aim to cure patients of their diseases. But epidemiologists, in addition to curing the patients, aim to control or eliminate their diseases in the community. So epidemiologists look at the disease/health event not only in its clinical perspective, but in its totality. Totality means to look on its distribution, determinants, affected population and how to control it.

Distribution

This looks at the distribution of the cases/patients. It serves to find out *who* is affected. Is it children, old people, prisoners, students, farmers, female, etc? *When* does the condition occur? Is it during teen age, rain or dry season, etc? *Where* does it occur? Is it in Africa, rural area, in the tropics, etc? *How much*, i.e. what is the number of cases? All these factors can assist in establishing the frequency of the disease and its pattern. Knowing all these factors will help to plan cost-effectively on how to manage the disease in question.

Determinants

This looks at *what* are the causes or contributing factors to the occurrence of disease, or *why* an individual gets that 'disease'. For example what are the

determinants of infertility, meaning why does someone get infertility? This means what are factors that contribute for someone to become infertile [prevalence of sexually transmitted disease (STI), lack of condom use, poverty, number of partners].

It also looks at *how* the 'disease' occurs. For example how does someone become infertile? What influences the occurrence of the infertility? Is a contact with the causative agent enough for the infertility to occur or is there something else (influencing factors, e.g. severity and duration of 'disease', obesity, hormonal imbalance, age) needed? Careful analysis of determinants helps to find and strengthen the evidence of the link between the disease and its causes so as to come up with effective control measures.

Specified population

This means the group of people that share a certain disease (health problem). For example hyperemesis occurs commonly in primigravid women. So 'primigravid women' is the specified population for hyperemesis gravidarum. The same applies to the population that gets cervical cancer, or ectopic pregnancy etc. If the affected population is known, management efforts will be focused on that small group. The success rate will be higher compared with dealing with the whole population randomly.

Prevention

Prevention is defined as actions aimed at eradicating, eliminating or minimizing the occurrence of disease, or retarding the progress of the disease and its impact. It is best defined in levels. There are four levels of prevention – primordial, primary, secondary and tertiary.

1. *Primordial prevention* is the combination of actions and measures that reduce or eliminate the emergence of risk factors in the population. Here efforts are directed toward discouraging those who are not yet at risk (e.g. children) from developing risk behaviors. For example early sexual debut and multiple sexual partners are risk factors for the development of cancer of the cervix or STI. At primordial level of prevention, children will be sensitized to delay their sexual debut and to avoid multiple sexual partners. Commonly primordial prevention is achieved through individual and mass education.
2. *Primary prevention* can be defined as the action taken prior to the onset of disease, which removes the possibility that the disease will ever occur. It is done in those people who have risk factors but the disease process has not yet started (i.e. pre-pathogenesis phase of a disease). It is done through health promotion (health education, environmental control, nutritional interventions, lifestyle/behavior changes etc.) and specific protection (chemoprophylaxis, immunization, nutritional supplementation, etc.)[2]. For example primary prevention of cervical cancer may include the use of human papillomavirus vaccine.
3. *Secondary prevention* is defined as efforts that aim to stop the progression of the disease at its early stage and prevent complications and death. This includes specific interventions like early diagnosis (e.g. screening tests, and case-finding programs etc.) and adequate treatment. It aims to arrest the disease progress by diagnosing it before it reaches an irreversible stage. For example, visual inspection with acetic acid (VIA), visual inspection with Lugol's iodine (VIL) and Papanicolaou (Pap) smear aim to discover cancer of the cervix before it is obvious and cryotherapy is used to remove precancerous cells in the cervix and this will be sufficient to arrest the progression to cancer of the cervix.
4. *Tertiary prevention* consists of measures aimed at limiting the impact of long-term disease and disability by eliminating or reducing impairment, disability and handicap that may be caused by it. It aims to minimize suffering and maximize the quality of life in the remaining time of living. It is mainly rehabilitation. For example tertiary prevention of cervical cancer involves the diagnosis and treatment of confirmed cases of cancer. Treatment is through surgery, radiotherapy and sometimes chemotherapy.

Some authors talk about *quaternary prevention*. This consists of actions that identify patients at risk of over-diagnosis or over-medication and that protect them from excessive medical intervention which may result in iatrogenesis, (inadvertent adverse effects or complications resulting from medical treatments). For example over-diagnosis of obstructed labor in a hospital may result in a lot of women

who have undergone cesarean section in the surrounding community. This may increase the rate of ruptured uterus in subsequent pregnancies if emergency services are not very good. Training on the proper management of labor may be considered as quaternary prevention

Successful prevention depends on knowledge of causation, dynamics of transmission, identification of risk factors and risk groups, availability of prophylactic or early detection and treatment measures, an organization for applying these measures to appropriate persons or groups, and continuous evaluation of and development of procedures applied. Without these, the likelihood of successfully preventing any disease in the community is very low.

Outcome of preventive measures

These are usually categorized into control, elimination, eradication and extinction depending on the achievement:

- *Control* The reduction of disease incidence, prevalence, morbidity or mortality to a locally acceptable level as a result of deliberate efforts; continued intervention measures are required to maintain that reduction. For example, if prevalence of STIs in a community is low (0.1%), and is kept around that value, then STIs in that community are said to be under control.
- *Elimination* Reduction to zero of the incidence of a specified disease in a defined geographical area as a result of deliberate efforts. Here continued intervention measures are also needed to maintain that situation; otherwise the disease may come back from other regions. For example, obstetric vesico-vaginal fistula has been eliminated in industrialized countries but is still endemic in third-world countries.
- *Eradication* Permanent reduction to zero of the disease. The agent of the disease may be available in laboratories somewhere in the world. Here intervention measures are no longer needed except to make sure the agent does not leak either accidentally or deliberately (e.g. in terrorism) from the laboratory. An example is smallpox.
- *Extinction* The specific infectious agent no longer exists in nature or in the laboratory. No disease has ever reached this stage.

ESTIMATION OF THE BURDEN DISEASE

Disease burden is the effect/impact of a disease in a community measured by financial cost, mortality, morbidity or other indicators[3]. It is often quantified in terms of quality-adjusted life years (QALYs), disability-adjusted life years (DALYs) or years lost due to disability (YLD) which combines the burden due to both death and morbidity into one index. This allows for the comparison of the disease burden for varying risk factors or diseases. It also makes it possible to predict the possible impact of health interventions in a community.

The number of cases seen in the health facility is the fraction of actual disease burden in the community. It is like the tip of an iceberg: what is seen appearing above the water is a very small fraction of what is underneath. The difference between what is seen in the health facility and what is actually present in the community depends on the health-seeking behavior of the respective community. If most of the sick people in the community seek medical treatment in health facilities (good health-seeking behavior), the difference is small and vice versa.

- A burden of illness study estimates the true burden of a disease in the community. (In-depth description can be found in epidemiology books.) This study uses the data available in the health system registers (out-patient, laboratory, in-patients etc.) together with surveillance data and household surveys. The aim of this study is to obtain estimates of the burden of gynecological diseases in the community.
- To gain a better understanding of how your surveillance system and laboratories are working.
- Effective planning for health services.
- To promote cooperation and collaboration among various government sectors.
- To advocate for gaining money for future studies and affect policy change.
- To justify initiation of preventive measures.
- It also helps to confront data deficiencies in measuring population health to guide investment in the health system.

Quality-adjusted life years

QALYs take into account both quantity and the quality of life generated by healthcare interventions. It is the measure of the life-expectancy corrected

for loss of quality of that life caused by diseases and disabilities. Some health interventions do not prolong life but do significantly improve the quality of life, e.g. palliative care for terminally ill patients with cervical cancer. A year of normal health is given a QALY of 1 while death has a QALY of 0.

Disability-adjusted life years

DALYs are a measure of the burden of disease and reflect the potential years of life lost due to premature death (PYLL) and equivalent years of 'healthy' life lost by virtue of being in states of poor health or disability. These disabilities can be physical or mental. One DALY can be thought of as one lost year of 'healthy' life.

All these measures of disease burden are used to measure the impact of disease burden in the community in order to compare different diseases for a prioritization in planning for health services.

MEASURES OF DISEASE ASSOCIATION AND IMPACT

To measure a disease association means to quantify the relationship between exposure and disease among two groups. (Exposure is the process by which a disease-causing agent comes into contact with a person in such a way that the person may develop the relevant outcome, such as a disease.) Exposure is used loosely to mean not only exposure to foods, mosquitoes, a partner with a sexually transmissible disease, or a toxic waste dump, but also inherent characteristics of persons (e.g. age, race, sex), biologic characteristics (immune status), acquired characteristics (marital status), activities (occupation, leisure activities), or conditions under which they live (socioeconomic status or access to medical care). The measures of association described in the following section compare disease occurrence among one group with disease occurrence in another group. Examples of measures of association include risk ratio (relative risk), odds ratio and risk difference. In this chapter only relative risk and odds ratio will be discussed.

Risk

This is the probability that an event will occur, e.g. that an individual will become ill or die within a stated period of time or by a certain age. It is the probability of an individual developing a change in health status over a fixed time interval.

$$\text{Risk} = \frac{\text{Number of new cases during a specified period}}{\text{Size of population at start of period}}$$

Relative risk

Relative risk (RR, also referred as rate ratio or risk ratio) compares the risk of developing a disease (any other health event, e.g. death) between two groups of which one group is exposed to a risk factor and another group is not exposed. It is done by dividing the risk in the exposed group by the risk in the unexposed group The risk may be in the form of incidence, attack rate, etc. The two groups are typically differentiated by demographic factors such as gender (e.g. males versus females) or by exposure to a suspected risk factor (e.g. did or did not use condoms).

$$\text{Relative risk} = \frac{\text{Risk of disease in exposed group}}{\text{Risk of disease in comparison group}}$$

For the calculation of measures of disease association, two-by-two tables are very useful. If you are interested in using them, please refer to books on epidemiology, e.g. those cited in the references.

Odds ratio

Odds ratio is the ratio of the odds of a disease among the exposed group compared with the odds of the disease in the unexposed group. (Odds ratio is the ratio of two probabilities. One probability is that the event of interest will occur, the other probability is that it will not occur.)

We use the odds ratio to determine whether exposure is associated with disease. This is because in case–control studies, incidence of disease among the exposed group is not calculated directly, through the probability of developing disease among the exposed group. A case–control study estimates the proportion of cases with exposure and the proportion of controls with exposure in order to see if a certain exposure leads to a certain disease.

EFFECT OF A DISEASE ON AN INDIVIDUAL AND THE SOCIETY

A measure of impact is the reflection of the burden that an exposure contributes to the frequency of

disease in the population. A measure of public health impact is used to place the association between an exposure and an outcome into a meaningful public health context. Whereas a measure of association quantifies the relationship between exposure and disease, the measure of public health impact quantifies the extent to which an exposure contributes to the occurrence of the disease.

Economic impact

The economic impact of diseases is very profound. A sick person/community is unproductive and consumes more resources to survive. This is double impact on the economy (not producing and at the same time the need for more to survive). Diseases in general, particularly chronic diseases deprive individuals of their health and productive potential. The burden of chronic diseases may invariably challenge individual or household income and savings, and compete with investment activities; instead of saving or investing, the money is spent on medical care. From a country's perspective, chronic diseases reduce life-expectancy and ultimately economic productivity, thus depleting the quality and quantity of a country's labor force. This may result in lower national output and national income (GDP and GNI). There has been some description in the literature of how diseases reduce intergenerational skills and wealth transfer[4]. Schooling of the children is affected, propagating the spiral of ill health and poverty. The burden in the healthcare system is increased. A lot of resources including drugs, manpower, time etc. are required, and all these cost a lot of money. If a sick person dies, the funeral will also consume some resources.

Social impact

The social impact arising from diseases is of significant importance. It is hard to over-emphasize the trauma and hardship that family members are forced to bear when one of them is sick. Assume a mother has cancer of the cervix. Provision of meals at home and other household activities done by the mother cease (mother cannot look after her children), so children have to take care of themselves and their cleanliness, cook food, prepare each other for school etc. Girls (sometimes boys) drop out of school to help their mother and take care of their siblings ruining their prospects for education, a well-paid work and thus their future lives. A husband may be needed to take over some responsibilities that were initially been done by his wife. He will be overburdened, unable to concentrate on his bread-winning activities or to participate in his usual social activities. He may abandon the family, creating more problems to the children. The patient – a mother – will be unable to fulfill her roles in the family. This will be psychological torture for herself. Because of smell and increasing pain, she will be unable to mix with other people so she will be unable to attend her social activities in the community. People will avoid her. Her death will result into longstanding grief for the family and produce orphans in the family/community. Some of the gynecological diseases do not cause death (e.g. fistulas) but have almost the same social consequences except death. It is very important to look at a patient as an entity and not as an individual because most of the time what is behind her is not known. Mismanagement of one patient may mean mismanagement of a certain group of people, e.g. a family, and vice versa.

STATISTICS IN GYNECOLOGY

While managing gynecological diseases, it is recommended to keep a record of the patients. Several parameters can be recorded in registers for analysis at the end of certain periods, preferably a year. This will help to track the characteristic of diseases in the surrounding community and nationwide to plan for management and prioritize financial means.

Recording

It is very important to keep record of all gynecology in- and out-patients who are given medical services in a health facility. This is the very basic step in epidemiology (see above). It is very basic because for the information to be analysed it must have been recorded previously. If no information has been recorded, there will be nothing to count, nothing to divide and nothing to interpret. So it is important to record all (gynecological) patients given health services in a health facility. Information to collect may include sex, age, marital status, ethnicity, residency, occupation, initial diagnosis, final diagnosis, investigation results, treatment given, result of treatments and recurrence. Additional information may be collected depending on the guideline from the public health authority in the area. It is important to collect meaningful data

that enable your staff to do their clinical work AND collect data. This means that you should aim at collecting a realistic amount of data as otherwise the providers recording them will be overburdened and you will end up with no or incomplete datasets. So you should discuss the items of interest with your colleagues and make sure data are not recorded for the recording only. It is very important however to bear in mind that your patients at the health facility are only the tip of the iceberg. In order to have a realistic picture of frequency and burden of disease you would need data from the community as well. How to obtain these?

Some services should be provided in the community as well, e.g. home-based care. The providers should gather data as well. The other possibility is to conduct household surveys. This is expensive as it needs a lot of trained interviewers but is often done nationally, e.g. for HIV and STI or within a research set-up. If you don't perform a household survey in your region yourself try to gather evidence from the literature on your region or similar areas to estimate the rest of the iceberg of burden of gynecological disease.

Counting and compilation

At various time-frames depending on the prevailing situation, counting is done to get the number of gynecology patients treated. Counting will be done in each individual register. Then numbers from various registers are compiled to get the figures for the whole health facility. The number of cases in various health facilities is compiled to get the number of cases in administrative areas (district, region/province, country, continent and then global).

At the level of health facilities, just counting may give the clue to the underlying problem. Take an example of septic abortion. Bearing in mind that septic abortions are difficult and expensive to manage, and may cause death, the number of cases in one health facility might be low. Only by counting and compiling information from several healthcare facilities, a monthly report may show that there is a sharp increase in the number of patients who were treated for septic abortions in an area. If the place of residency was recorded, it may be known where most of these patients come from. Health workers may go to the area and investigate what is the cause of the sudden increase of these cases finding that there is an inexperienced private provider that just started to operate in the area. Training of this provider will cut down the problem, cut down unnecessary expenditures, admissions, deaths and the long-term complications.

Analysis and interpretation

The information collected is analysed at certain intervals and on various levels of administrative areas. This can be done by calculating proportions, rates, frequencies and all other measures of disease frequency mentioned above. Complex analysis of health-related information becomes possible and meaningful as the number of cases increases. Analysis at facility levels may be simple counting and dividing. At national, continental and global levels, information is compiled on a large scale. Number of cases and their related information is from a huge area. In epidemiology, evidence of linkage between diseases and the causes (or any other related factor) becomes stronger as the number of events increases. The information that may not give useful meaning if analysed using single facility records, may give very useful information if analysed using records compiled from various health facilities as explained above. This useful information is brought back to individual health facilities as evidence-based preventive strategy. If implemented, the burden of diseases around the facility goes down, and the scarcity of resources is lessened.

In summary, in most cases, information collected by individual health workers is the one used to design preventive measures, treatment or research.

REFERENCES

1. Bailey L, Vardulaki K, Langham J, Chandramohan D. Chapter 1: Basic concepts and applications of epidemiology. In: Bailey L, Vardulaki K, Langham J, Chandramohan D, eds. *Introduction to Epidemiology*. Berkshire, UK: Open University Press, 2006;3–18
2. Tewari KS, DiSaia PJ. Primary prevention of uterine cervix cancer: focus on vaccine history and current strategy. *Obstet Gynecol Clin North Am* 2002;29:843–68
3. Bradshaw D, Groenwald P, Bourne D. Measuring burden of disease – an essential foundation to improve health. Cape Town: MRC Burden of Disease Research Unit and UCT Department of Public Health and Family Medicine, 2007
4. Dele A, Anderson S. An estimation of the economic impact of chronic noncommunicable diseases in selected countries. World Health Organization, Department of Chronic Diseases and Health Promotion (CHP), 2006

Further reading

AbouZahr C, Vaughan JP. Assessing the burden of reproductive ill-health: questions regarding the use of disability-adjusted life years. *Bull World Health Organ* 2000;78:655–6 [An interesting read about the consequences of morbidity and its abscense in priority setting for health care]

Bonita R, Beaglehole R, Kjellström T. *Basic Epidemiology*, 2nd edn. Geneva: World Health Organization, 2006 (ISBN 92 4 154707 3). [A very good book on epidemiology which can be downloaded from the internet for free]

Centers for Disease Control and Prevention (CDC). Principles of Epidemiology in Public Health Practice. An Introduction to Applied Epidemiology and Biostatistics, 3rd edn

Atlanta, GA: CDC. http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf

Miquel P, Sander G, Last J. *Dictionary of Epidemiology*, 5th edn. New York: Oxford University Press, 2008

Paneth N. Measures of disease association, 2005. http://www.pitt.edu/~super1/lecture/lec8061/001.htm [A nice web-based presentation on measures of association and, for those who want to go deeper, of impact]

Rasha S. Concepts of prevention and control. PhD Community Medicine Suez Canal, University Egypt, 2005. http://www.pitt.edu/~super7/32011–33001/32311.ppt

Sadana R. Measuring reproductive health: review of community-based approaches to assessing morbidity. *Bull World Health Organ* 2000;78:640–54