CHAPTER IV
DATA EXERCISE

4.1 <u>Introduction</u>

The data exercise should include primary and secondary, and both qualitative and quantitative data. Primary data were collected in the form of interview method for PH 521 (Fundamental Science and Research Skills).

In that study, I and one of my colleagues interviewed foreign migrants in different places in Bangkok with pre-structured questionnaire and obtained data included both quantitative and qualitative aspects (refer "Health Care Seeking Pattern of Foreign Migrants in Bangkok"). Statistical inferences were also drawn by using Epi Info and SPSS soft ware.

Secondary data analysis exercise was done with <u>cause specific mortality data base exploration</u>. Some primary data were also collected to support and/or to complete data base (e.g. interview, focus group discussion and consultation to experts). These are necessary to achieve my personal learning needs.

According to lectures given during MPH Course, there are two types of secondary data. Individual or micro data and aggregate or ecological data (Adit, 1995). The individual data concern with personal matters like sex, age, occupation etc.. Therefore, researcher can trace the respondents. But in aggregate data, it is impossible.

Cause specific data base exploration is essential in finding suitable data bases for future study "prioritization of diseases by premature mortality cases in Myanmar". Thailand has more or less sufficient data bases to learn as first hand experience before actual task.

In this type of study, first, the researcher has to review the literature thoroughly according to research question in that study. Then identify data bases that might include the variable of interest. After becoming familiar with each of the data set, he/she has to discuss with experienced persons and consult to experts.

Choosing of best data base must be done and gain access to the data. Compilation and analysis of data has to be carried out later. During analysis, suitable statistical methods must be used(Adit, 1995).

In my study, individual diseases were investigated to know their burden. Cause specific mortality data according to specific age groups was the mandatory requirement to calculate the burden of diseases with indicator potential years of life lost (PYLL). Literature revealed that mortality has long been used as the main indicator to assess the progress of health care. Mortality is expressed in many forms, such as life expectancy, infant mortality, and mortality from specific diseases or in specific subgroups (Chitr, 1995).

In many instances, mortality rates do not tell the entire picture of burden of disease. Burden of illness involve not only prevalence and incidences, but also peoples'

perceptions (Chitr, 1985). Assumption in measurement of premature mortality is that deaths at younger ages (especially productive ages) are considered as a greater loss of life than a death at age 90 (Murray, et al, 1994).

Statistical equation of PYLL is as follows:(CDC, 1986)

$$PYLL = \sum_{x=0}^{x=1} d_x \, (L-x)$$

d_x = number of deaths (in specific age group)
L   = potential limit to life
x   = age of death

The above equation of PYLL was seems to be easy, but getting the number of deaths for each disease was a bit tough. Even in the some developed countries, estimates had been done for certain diseases and conditions.

Cause of death was derived from death certificates. MOPH compile data from registered deaths. Validity in certification of cause of death depends on the person who attended that occasion (Mahidol University, 1988) Usually, established market economy countries were found to be better than other regions.

I found even within the same country, better in some regions than others due to presence of established health care facilities, refresher training, accessibility of information and better economic status etc.. Cultural pattern and life styles were also supposed to have influence on the validity of cause of death.

Table 1. Deaths in Government Hospital by Region, Thailand 1993

Region >	Central	Southern	Northern	North-eastern
Deaths in Hospital	22183	4286	8751	15090
Percentage of Total Deaths	21.1	14.0	14.1	17.2

Source: Public Health Statistics, 1993. MOPH

Table 1. shows that central region had higher percentage of government hospital deaths than others (MOPH, 1995). Presence of health care institutions (both government and private), better economy, higher knowledge cf residents and easy accessibility enhanced hospitalization may be relevant factors. This made subsequent accurate certification and classification of causes of death.

But I found North-Eastern region also had a significant number of government hospital deaths. Literature review revealed this region was the least developed area in Thailand. Therefore, other local socio-cultural factors should also be taken into account and explored as qualitative background if it is feasible.

4.2 Data Collection Methods

During data search, some recorded data were found to be incomplete and estimation of the number of deaths became a necessity. Various data bases were explored according to

advice of experts. Based on the following exercises, I could write an essay on Trend and Burden of Communicable Diseases in Thailand (refer essay) and I got first hand experience in data collection and compilation for proposed study. (refer proposal)

Thailand performed regular surveys in the past (Institute for Population and Social Research, Mahidol University, (IPSR, MU, 1988) National census is one of the source for mortality and some morbidity cases (if specific questions were included). Health and welfare survey, nutritional survey and survey of population change (SPC) were found to be very useful data sources. Specific surveys like morbidity and mortality differentials (IPSR, MU, 1988) were carried out to get more information on causes of death, disease prevalence, incidence and other conditions.

Apart from a few countries which have accurate health management information systems, many are faced with deficiency in morbidity and mortality data. Therefore, estimation methods were developed to get an overall picture. (Murray, Lopez & Jamison, 1993)

In estimation of cause of death in a country, first we need to understand the local and global trends in causes of death. (IPSR, MU, 1986) Naturally, the demographic transition in the country preceded the epidemiologic transition. These two transitions can change the disease morbidity and mortality patterns.

The following data sources and methods were found to be useful in estimation of causes of death. They were:

(1) Vital registration with certification of causes of death

(2) Sample registration systems

(3) Household surveys

(4) Population surveillance systems

(5) Epidemiological estimates

(6) Causes of death models

(7) Deaths in hospitals

(8) Cancer registries

4.2.1 Household Surveys

Census and National Health and Welfare Survey (NHWS) were found to be regular surveys in Thailand. In NHWS, National Statistical Office (NSO) attempted measures to understand better the health status of the Thai population.

4.2.2 Epidemiological estimate

This is one of the useful method for estimation of mortality by age and sex from a cause.(Murray, 1993) Epidemiologically, incidence of a disease can be mapped out by the equation. Number of deaths from specific causes can be calculated if we know incidence rates. This method is especially suitable for communicable diseases. For NCDs, it is difficult to know the definite onset(incidence),

incubation period and duration of illness. Also prevalence is relatively high in compare with CDs (e.g. Hypertension = life long disease).

Equation is as follows:

$$P = I \times D$$

P = prevalence
I = incidence
D = duration of illness

Duration of a disease is more or less the same in different localities. Therefore, it can be obtained from other data sources. Prevalence can be mapped out by the population surveys. Existing population surveys data bases were explored to get prevalence data.

Incidences of some diseases or conditions can be traced by survey method even after many months due to their typical nature (e.g. measles in children, abortion in women, motor vehicle accidents etc.).(IPSR, MU, 1988)

From incidence, remission and case fatality rates, approximate mortality from specific disease in a locality could be calculated. Data bases in Thailand were found to be under reported and so it was necessary to perform these estimates for some diseases and conditions.

4.2.3 Cause of Death Models

In Thailand, similar to other Asian countries, some diseases have no available data to estimate mortality by cause. Even epidemiological estimates could not helpful.

Therefore, those were estimated by cause of death models.
These models were developed from the past patterns of the
epidemiological transition in the countries with good vital
registration over the past decades. The example of cause of
death model was described in appendix 1 .(Murray, Lopez &
Jamison, 1993)

Many developing countries may not have complete vital
registration or a good sample registration system. Some
countries experienced that the sum of estimated deaths due
to individual causes for an age group was greater than the
total mortality for that age group. This was due to doubling
of deaths and this can be adjusted by using a squeeze
algorthms.(Murray, Lopez & Jamison 1993)

Table 2 . Yearly Ill-defined Deaths in Thailand

Year	1987	1988	1989	1990	1991	1992	1993
ases	111472	106667	110112	109176	113536	117119	17807
of otal eath	47.85	46.13	44.66	43.24	42.95	42.54	1.23

Source: Public Health Statistics, 1993. MOPH

Above table shows magnitude of ill-defined cases in the
country. It was so interesting and I traced detail
information to know which region has highest ill-defined

rates. I intended to trace down to the district level to
know the degree of ill-defined cases. But due to the
inaccessibility of data for many reasons, I could reach only
the regional level.

Even in analysis of morbidity and mortality
differential project data, (project period = 4 years) the
research team claimed that they could perform only regional
and national level analysis due to limitation of data.(IPSR,
MU, 1988)

Table 3. Symptoms, Signs and Ill-defined Deaths by Region (1993)

Region >	Central	Southern	Northern	North-Eastern
Number of ill-defined death cases	14180	16154	26331	38624
Percentage of total deaths	13.52	53.0	42.13	43.17

Source: Health Policy and Planning Unit, MOPH

Table 3. shows percentage of ill-defined cause of death
in different regions during 1993. More than ten percent of
deaths were found to be ill-defined even in the most
developed central region.

4.2.4 Qualitative Data Inquiry

I performed a focus group discussion (FGD) with a group
of Thai doctors. I got opportunity to discuss with some
government medical officers attending a Ministry of Public
Health Executive Training Program in Health Systems Reform

in my college (C8 group). Some are working in the hospital wards, some are working in health offices and some have dual charges. I asked and discussed with them about the problem of ill-defined causes of death and we got following reasons as contributing factors.

<u>Patients' side</u>
1. Presence of dead on arrival cases
2. Arrived to hospital with critically ill conditions in which some clinical and laboratory diagnoses were not feasible and died within few minutes/hours
3. Lack of knowledge to give history or lack of attendant to narrate the course of illness
4. Absence of typical signs and symptoms
5. Extremes of ages of patients

<u>Doctors' side</u>
1. Lack of specific training or refresher training to write accurate diagnosis to meet the requirement of ICD coding.
2. High work load (no time to bother complete or incompleteness in diagnosis of cause of death
3. Carelessness in writing cause of death.
4. Decrease in post mortem diagnosis practices

<u>Medical Record Department</u>
1. No intra-sectoral co-ordination
2. Impatient in data entry

3. Negligence in record keeping

Then I asked them to write cause of death on a given statement individually. I got quite interesting output from them.

Situation Statement

An AIDS patient was admitted in your ward for treatment of pulmonary tuberculosis.

One evening, he developed massive haemoptysis and died.

Q: What will be the cause of death for this patient ?

Options: 1. Pulmonary tuberculosis

2. AIDS

3. Other: Specify _____

Table 4. Output From C8 group

Cause 1	Cause 2	Other	Total
6	3	6	15

Source: Self administered questionnaire, March, 1996

Above table shows variation in making diagnosis of stated problem among C8 doctors. Actually, the rules of the ICD state that the underlying cause i.e. " the disease or injury which initiated the sequence of events which led to death " should be written for that case. Therefore, some

studies should be targeted for finding influencing factors of ill-defined cases at institutional level.

Ill-defined causes of death in under five children is also common especially in developing countries. Lopez (1993) reported that majority of deaths of children in developing countries are very vague. Because they are living in complex epidemiological environment and they are faced with multiple infections aggravated by malnutrition and poverty. Therefore, to set a single underlying cause of death is very difficult in these countries if infections are still common. (Lopez, 1993)

I explored another data bases to know more about mortality data concerning with ill-defined causes of death. I found Morbidity and Mortality Differential Survey (MMD) 1985, conducted by Mahidol University.(IPSR, MU, 1988)

In MMD survey, causes of death inquires were done at the place of residence of the deceased person. Therefore, it was claimed to be more accurate. (IPSR, MU, 1988) Their findings indicated that about 30 percent were instant deaths and no treatment was required. Only 49% of total deceased persons seek physician care immediately before they die.

In causes of death entering on a death certificate, a health personnel should write all those diseases, morbid conditions or injuries which either resulted in or contributed to death and the circumstances of the accident or violence which produced any such injuries. Therefore, accuracy of causes of death depends very much on who attended the deceased person.

Table 5 .**Percentage of distribution of deceased persons by type of attendant at death, by region, 1985.**

Type of attendant	BKK	Central	North	Northeast	South
(no. of deaths)	(112)	(269)	(360)	(298)	(260)
Medical doctors	63.4	48.0	52.8	49.7	37.3
Paramedics	2.7	3.7	4.2	6.4	4.2
Unqualified per:	0.9	8.9	13.3	18.5	25.0
Instant death (noTx)	33.0	39.4	29.7	25.5	33.5

<u>Source</u>: Morbidity and Mortality Differential survey, 1985

Table 5. shows Northeast region has higher doctor attendant rate comparable to more developed and prosperous Central region. Therefore, underlying socio-economic and cultural factors should be investigated to know the situation.

Mortality data may vary markedly from group to group according to variation in diagnostic fashion and depend on who attended the deceased person. Cultural differences appear to play a significant role in the classification of an external cause of death e.g. suicide. (due tc social and religious stigma associated with the act, and in such cases the death would usually be classified to the brcader category of external causes)

Tendency to code into ill-defined category among certifiers can cause difficulty in comparison of causes to other region or countries.

Hospitals are more reliable source of causes of death than other places. If death occur outside the health

institution without qualified health personnel, it becomes
dubious.

Table 6 .**Prevalence of Illness or Injury and Hospitalization
Per 1,000 population by Region, 1985**

Region	Illness or Injury	Hospitalization
Whole Kingdom	156.4	51.3
Urban	134.2	64.1
Rural	161.5	48.3
North	154.2	65.6(42.50%)
Northeast	144.6	38.9(26.90%)
Central (excluding BM)	184.7	59.9(32.43%)
South	151.6	34.1(22.49%)
Bangkok Metropolis	139.5	65.2(46.74 %)

Source: Morbidity and Mortality Differential Survey,
1985

Table 6. shows percentage of hospitalization in
different regions by prevailing illness or injury. Northeast
region had lower percentage in hospital attendance.
Theoretically, the more hospitalization (i.e. more hospital
deaths), the more accurate in the causes of death. But here
seems to be confounded by some cultural factors.

Household members of severely ill persons may usually
send to the hospital just before they die due to lack of
knowledge, inaccessibility and inequality in distribution of
health institutions in less developed regions. Some economic

and social reasons might interfere with keeping dead body at home.

According to deaths inquiry in MMD project, death rates on ill-defined causes by region were discovered as follow:

Table 7. **Death rates per 100,000 population on cause of death by region**

	Whole	urban	rural	BMA	Central	North	N-E	South
1. (Ill-defined Causes of Deaths)	34.9	23.8	46.2	18.6	43.7	31.9	23.7	52.7
2. (Senility without Psychosis)	69.8	42.1	98.0	29.2	87.4	74.4	59.1	84.3

Source: Morbidity and Mortality Differential Survey, 1985

Above table shows that death due to Ill-defined causes were lower in Northeast (least developed region) than central and north (more developed regions). This was found to be contradict with percentage of hospitalization for illness and injuries in each region (refer table 6) and data obtained from death certificates by Ministry of Public Health.

Injury Garbage Codes

As economy grows and changes in life style in the country, more and more cases of motor-vehicle accidents, occupational injuries and other injuries prevailed. So some problems in ICD coding comes out.

In Basic Tabulation List (BTLs), E 560 (E 980-E989 in the 3 digit codes) is found to be " Injury undetermined whether accidentally or purposely inflicted ". Sometime this code contains a lot of injuries. Therefore, for political reason, these should be redistributed into specific injuries.(Murray, lopez & Jamison, 1993)

4.2.5 Allocation of Ill-defined Causes of Death

Since we need to know the age specific causes of deaths, redistribution of ill-defined causes into specific diseases is mandatory. For all deaths causing more than 0.1% of total deaths were found to be necessary for proper allocation.

Preston (1971) assumed that majority of deaths in children under five is related to communicable diseases. (Murray et al, 1994) After age five years, all are put into non-communicable disease deaths and allocated proportionately. I learned that the choice of into which group to reallocate chapter XVI deaths should be based on all information available in the country.

Depending on study and locally available information, redistribution into different categories of injuries has to be carried out to get completeness of data in burden of disease calculations.

In Mexico, the condition of ill-defined causes of death in 1992 was shown in next page.

Table 8. Death Classification in Mexico, 1992

Percentage	>0.1%	<0.1%	Total
Classified	76.6	4.8	81.4
Residual	14.3	4.3	18.6
Total	90.9	9.1	100.0

Source: Designing and Implementing a National Burden of Disease Study, 1994.

From above residual diseases, i.e. 14.3%, half (i.e.7.3%)were redistributed to specific disease categories according to locally prevailing diseases and conditions. (Murray, lopez & Jamison, 1993)

In Thailand, community level mortality inquires were carried out in 1985. Sample population revealed about 8% of ill-defined causes of death (excluding senility without psychosis which was about 16%). Allocation of ill-defined cases will not be difficult if we use this survey findings. But I believe some details are still lacking and ill-defined causes are still troublesome in Thai mortality data.

4.3 Discussion

It is difficult to say mortality data obtained by MMD survey are whether representative or not. They tried to achieve representative samples by two stage cluster sampling. But they could only investigated less than one percent of deaths occurred in the country.(MOPH, 1995) They used verbal autopsy method to obtain cause of death in interviews. Interviewers were not physicians.

Nevertheless, this survey is the latest available mortality survey conducted at community level and it revealed some determining factors of discrepancies in cause of death certification at different regions.

Ill-defined causes were found to be stable through out the years 1987 to 1993.(MOPH, 1995) Literature review revealed that early years were mainly contributed by ill-defined causes of children (due to complexity in epidemiological environment).(Lopez, 1993) In the latter years, it can be explained by Health transition.(Murray et al, 1994)

The more gain in economy and the more develop the nation, which lead to accumulation of older population in the country. These old people become contribute larger part of ill-defined cases especially when they reach beyond 70 years.(IPSR, MU, 1988)

Changes in the disease pattern from infectious to non-infectious diseases had made difficult to estimate the causes of death with epidemiological models. Majority of non-communicable diseases have insidious onset and varying length of incubation and duration of illness.

Emerging new diseases like Acquired Immune Deficiency Syndrome (AIDS) also cause various systemic diseases. An AIDS patient can die of tuberculosis secondary to immune deficiency. Complexity in manifestation of this disease also enhanced difficulty in diagnosis by health care providers.

Prevalence of non-communicable metabolic disease like diabetes mellitus (due to life style changes) can also

contribute ill-defined conditions if the clinician did not notice it.

Like Thailand, in many countries, a large proportion of deaths were found to be coded into international classification of diseases (ICD) revision nine, chapter XVI, symptoms, signs and ill-defined conditions.(Murray, Lopez & Jamison 1993) Also, there were many other so called garbage codes were found during data search, for example, codes as circulatory or pulmonary were ill-defined or very general. Many coders used different garbage codes in their countries.

4.4 Conclusion

Therefore, more extensive survey is necessary to judge the real situation of ill-defined causes of death. But for the time being, based on some available epidemiological, social and demographic data, the task of re-allocation of some ill-defined cases should be done to know the approximate burden of disease for resource allocation by PYLL indicator.