CHAPTER 4

ANALYSIS OF DATA

The purposes of this chapter are (1) to present the data from the primary and secondary sources, (2) to strengthen the analytical and critical ability of the writer by assessing the data which are available and (3) to explain to the reader how the costs, determined in subsequent chapters, were obtained given the limited quantity and quality of data available.

Section 3.3.3 (Table 3.1) lists the nature of data required for the cost model and the actual sources available. This chapter examines the quantity and quality of data available, explores the feasibility of using secondary data, the reliability of the small sample of primary data and the possibility of using derived data where secondary and sampled data were not sufficient for the cost model.

4.1 Behavior of Patients in Receiving Leprosy Care

Application of the principles of decision tree yields a complex pattern of decisions and alternative actions available to patients after an initial decisions as to which of the service points will first be visited. The decisions and alternatives are presented in Figures 4.1 to 4.5

Figure 4.1 Decisions and Possible Actions on Where to Seek Care

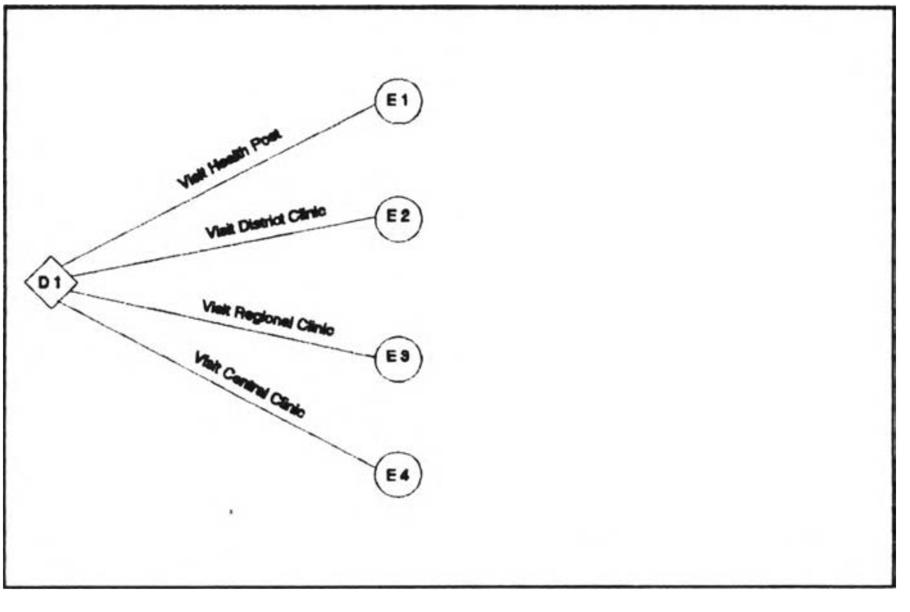


Figure 4.2 Decisions and Possible Actions on Attending a Health Post

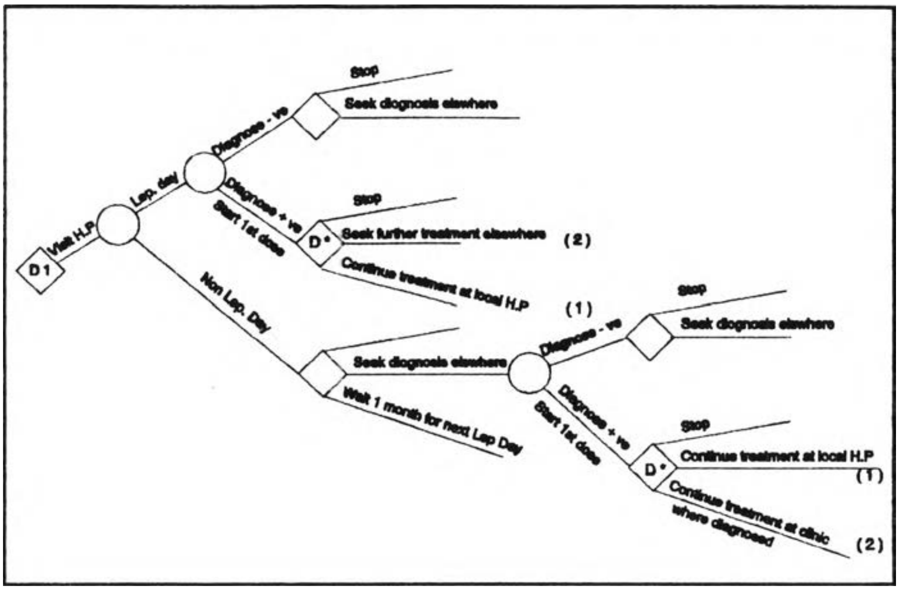


Figure 4.3 Decisions and Possible Actions on Attending a District Clinic

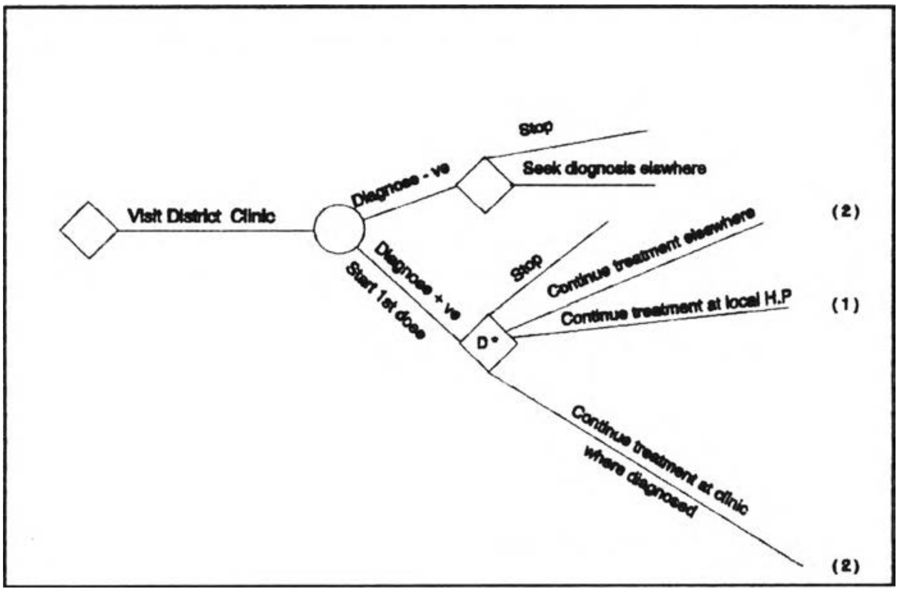


Figure 4.4 Decisions and Possible Actions on Attending a Regional Clinic

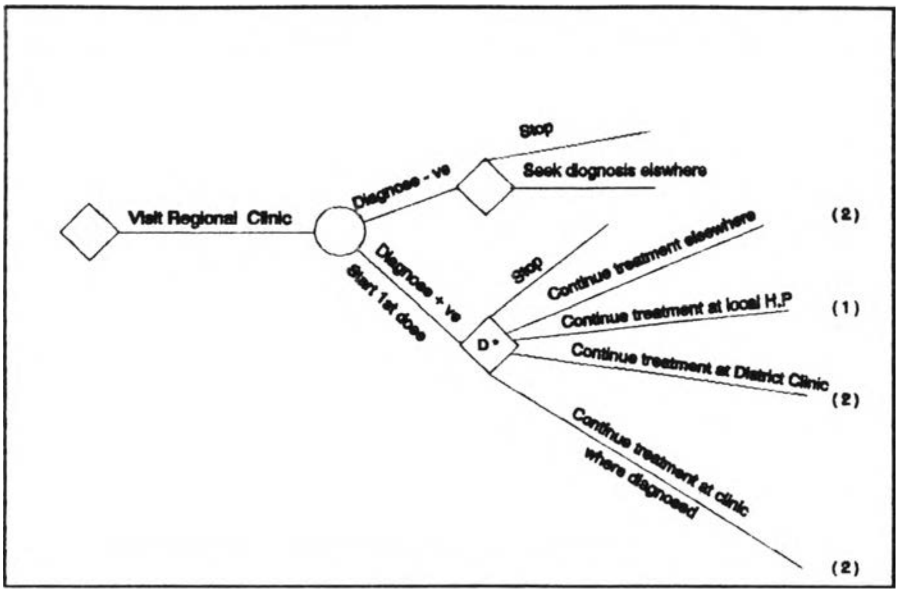
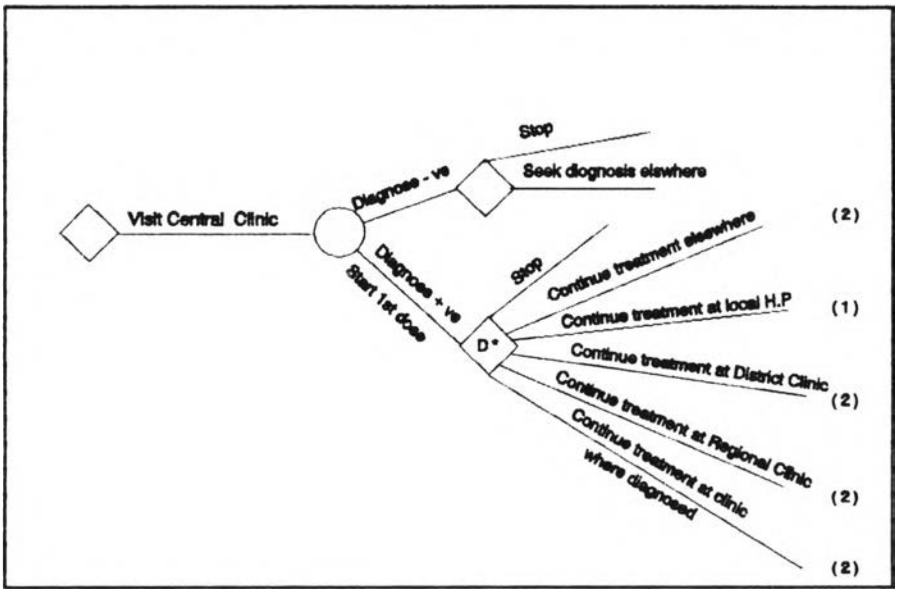


Figure 4.5 Decisions and Possible Actions on Attending the Central Clinic



Having identified the decisions and possible actions one must ask what are the determinants of the decisions? The thesis which underpins this research is that patients travel to outstation clinics to avoid the embarrassment of their condition being made known to local people. But that aspiration, no matter how strong, is presumably mediated by opportunities and constraints. Therefore a demand for leprosy services at a clinic can be constructed in a form of general function to illustrate relationships of factors which may affect demand as follows:

X_{tj} = total number of patients attending the clinics in district j

X_j = the demand for leprosy services at a particular clinic outstation clinic in district j

n_{ij} = number of existing cases in district i which is feasible to attend clinic at district j

ct_{ij} = total costs to be incurred by patients from district i in receiving care at clinic in district j

r_{ij} = distance to travel from district i to given clinic in district j

q_{ij} = geographic feasibility of travel from (i) to given j

d = state of disease of patient

e = fear of exposure (stigma)

a_t = attractiveness of other treatment center

l = confidence in local center

t = time feasibility of patients

$$X_j = f(n_{ij}, ct_{ij}, r_{ij}, q_{ij}, d, e, a_t, l, t)$$

The significance of each variable remains in doubt and requires separate study but it was expected that distance would show a significant relationship to demand since many constraints will be largely related to distance. This is examined in Section 4.2.1.

4.2 Analysis of Secondary Data

4.2.1 Number of Outstation and Local Patients

Most of the data from secondary sources are in aggregate form simply detailing the number of patients registered in a district with no breakdown into local patients and outstation patients from other districts. The exception is the five regional clinics (j = 10, 30, 39, 51 and 65) and the central clinic (j = 28). For these six clinics data are available on the districts (i) from which they drew outstation patients, distances from districts (i) to each clinic (j) (r_{ij}), the number of patients from outstation districts attending the clinics (XD_{ij}, XR_{ij}, XC_{ij} and the number of local patients attending these particular clinics (XD_l, XR_l, XC_l) (Appendix 2).

The fundamental data required for the cost models are XD_l, XR_l, XC_l and XD_{ij}, XR_{ij}, XC_{ij} for every clinic in which the costing is to be implemented. Since the data are only available for six clinics, to complete the cost models, these data have to be estimated or derived based upon available data and assumptions.

Distance r_{ij} is expected to be a determinant of travel costs, time costs and related food and accommodation costs. As such an inverse relationship may be expected between the number of outstation patients at the clinic in district j (X_j), and the distance that patients travel from each district i (r_{ij}) to attend the clinic at district j. A direct relationship is also expected between (X_j) and the total number of patients attending the clinic in district j (X_{tj}).

$$X_j = f(r_{ij}, X_{tj})$$

The validity of this expectation was tested by using least squares regression of the secondary data available for the six sample clinics, with r_{ij} and X_{tj} as the independent variable and X_j as the dependent variable (j = 10, 28, 30, 39, 51 and 65). No such inverse relationship between X_j and r_{ij} was observed for each of the six clinics or for the six clinics as a whole. This implies that distance is not a primary determinant of demand for services at a given clinic for those patients who would travel to another district for treatment. However, it must be recognized that the reverse may be the case for the majority who do not travel.

4.2.2 Distance Between Districts

As explained in section 3.2.4 the distance between districts (r_{ij}) which is required for deriving data was only available for the j = 10, 28, 30, 39, 51 and 65, the five regional clinics and the central clinic. Data from a topographic map showed where travel between the 73 districts in the five regions was feasible, not beyond the central clinic, and provided data on travel distance (+/- 10km). The data show that movement between districts is actually very limited (Table 4.1)

Table 4.1 Feasible Movements Between Districts

Potential movements (73 districts) Assumes movement between each i,j	5256
Feasible movements Assumes two way movement between i,j combinations which are actually feasible	468 8.8% of potential movements
Known movements When j = 10 from 9 districts j = 28 from 39 districts j = 30 from 37 districts j = 39 from 8 districts j = 51 from 7 districts j = 65 from 8 districts	108 23% of feasible movements

The 360 feasible movements between i,j, where the number of outstation patients is not known, have an average travel distance of 123 km.

4.3 Analysis of Primary Data

The procedure for sample surveys was presented in Section 3.2.3. The data gathered from the sampled patients are presented and analyzed in this section.

4.3.1 The Sampled Clinics and Patients

Patients were interviewed at three clinics; the regional clinic in the Eastern Region ($j = 10$), the regional clinic in the Central Region ($j = 30$) and the Central Clinic in the Central Region ($j = 28$). The regional clinic in the Central Region ($j = 30$) is atypical since it is a mission clinic which, by virtue of its central location and quality of service, attracts more outstation patients than other regional clinics.

The number of outstation patients attending these sampled clinics in 1993, from each district (i), is shown in Appendix 2. In each of the sample clinics 30 patients were selected at random and interviewed.

4.3.2 Cost Data for Local Patients

The fundamental data required for cost models are T_1, D_1, a_1, b_1, c_1, where $i=j$ and T_{ij}, D_{ij}, a_{ij}, b_{ij}, c_{ij}, where $i \neq j$, for every value of i and j.

It is assumed in the model that O_1 , T_1 , D_1, a_1 ,b_1 ,c_1 are constant and can be obtained from the mean value of the sampled patients attending local clinics. The means and S.D of the data for the three clinics are presented in (Table 4.2).

Table 4.2 Summarized Data for Local Patients at Three Clinics

j	n	Data for	Mean	S.D.	Max.	Min.
10	14	T_1	2.0	1.4	5.0	0.16
28	11		0.5	0.3	1.0	0.16
30	12		3.3	1.4	5.0	1.0
Total	37		2.0	1.6	5.0	0.16
10	14	D_1	4.6	2.7	12.0	1.0
28	11		3.1	3.7	14.0	1.0
30	12		3.7	1.4	6.0	2.0
Total	37		3.9	2.7	14.0	1.0
10	14	a_1	7.2	9.2	20.0	0
28	11		10.2	14.2	50.0	0
30	12		5.3	5.9	14.0	0
Total	37		7.4	10.1	50.0	0
10	14	b_1	9.1	3.6	20.0	5.0
28	11		12.4	4.5	25.0	10.0
30	12		11.7	3.5	20.0	8.0
Total	37		10.9	4.0	25.0	5.0
10	14	c_1	0	0	0	0
28	11		0	0	0	0
30	12		0	0	0	0
Total	37					0

Clinics j = 28 and j = 30 are both in the Central Region and only 6 km apart. If the data from these clinics show a significant and systematic difference in means to clinic j = 10 then there may be some justification for using the means of the latter clinics for other Regions. The only significant difference is in the time delay T_1 between the onset of symptoms and initially seeking care. The sample of local patients at the central clinic reported an average of 0.5 years compared with 2.0 and 3.3 years for clinics j = 10 and j = 30.

In the absence of convincing justification, the overall mean will be used in the cost model.

4.3.3 Cost Data for Outstation Patients

The fundamental cost required for cost models are T_{ij}, D_{ij}, a_{ij}, b_{ij}, c_{ij}, where i ≠ j, for every value of i and j. It was assumed (Section 3.2.4) that there are significant relationships between distances from i to j (r_{ij}) and D_{ij}, a_{ij} ,b_{ij}, c_{ij}. There is perhaps less reason for T_{ij}, time delay between onset of symptoms and seeking care, to be related to the distance i, j.

The validity of this assumptions was tested by fitting least squares regression to the data available from the surveys at the three sample clinics, with r_{ij} as the independent variable and T_{ij}, D_{ij}, a_{ij}, b_{ij} and c_{ij} as the dependent variables (j = 10, 28 and 30).

Only travel cost, a_{ij}, showed a significant relationship with distance with R^2 of 0.9 for the combined data and a coefficient, passing through the origin of 0.92. No significant relationship was observed between distance r_{ij} for T_{ij}, D_{ij}, b_{ij}, or c_{ij}. It follows that a_{ij} could be estimated from the distance (r_{ij}) but distance can not be used to estimate the value of other variables.

An important question is how representative the small samples were of the i districts providing outstation patients to the j clinics. The data presented in Table 4.3 shows at least one patient from only about 30% of feeder districts.

Table 4.3 Residential Districts of Sampled Patients

j = 10		j = 28		j = 30	
i	n	i	n	i	n
4	1	4	1	3	1
7	1	10	1	4	1
9	9	21	1	10	2
10	14	22	1	17	1
14	2	23	1	20	1
15	3	25	2	21	1
		26	1	22	1
		27	1	26	1
		28	11	27	1
		31	3	30	12
		32	2	31	4
		33	3	33	1
		34	2	34	2
				73	1
Local	n = 14 = 47%	Local	n = 11 = 37%	Local	n = 12 = 40%
Outstation	n = 16 = 53%	Outstation	n = 19 = 63%	Outstation	n = 18 = 60%
% Of possible districts	= 44%	% Of possible districts	= 28%	% Of possible districts	= 35%

Patients in the sample were questioned about whether they were accompanied by a relative. Twelve of the local patients (32%) and 31 of the outstation patients (59%) were accompanied by a relative.

4.4 Implications of Data Analysis

Two fundamental assumptions in the models from which derived data was to be obtained can not be used, based upon the available data.

1. A relationship was NOT found between the number of outstation patients attending a particular clinic (X_j), distance from each district providing outstation patients to j (r_{ij}) and the total number of patients attending clinic j (X_{tj})

2. There are NOT significant relationships between the distance from i to j (r_{ij}) and T_{ij}, D_{ij}, b_{ij}, and c_{ij}. A relationship was found for a_{ij}.

The original objectives of deciding how to determine potential cost savings for leprosy patients, if attending at local clinics, and the magnitude of that cost saving remain. Under these circumstances an alternative approach must be taken which is applied in Chapter 5.

The cost model can only be used where X_{ij} is known, that is at the six clinics providing secondary data on the source and number of outstation patients. In addition a number of assumptions must be made.

Assumptions

1. O_1 , T_1 , D_1, a_1 ,b_1 ,c_1 are constant and can be obtained from the mean value of the sampled patients attending local clinics.

2. a_{ij} can be derived from r_{ij}.

3. The arithmetic mean of the costs incurred by the sample of outstation patients, from each of the three sample clinics, T_{ij}, D_{ij}, b_{ij}, and c_{ij} is the same for all outstation patients attending each of these three clinics.

The summarized statistics of the outstation patients from the three sample clinics (Table 4.4) show that there is not a significant difference between the means.

Table 4.4 Summarized Data for Outstation Patients at the Three Clinics

j	n	Variable	Mean	S.D.	Max.	Min.
10	16	T_{ij}	2.8	2.0	8.0	0.5
28	19		2.9	2.3	7.0	0.41
30	18		2.2	1.1	5.0	1.0
Total	53		2.6	1.9	8.0	0.41
10	16	D_{ij}	12.4	9.9	46.0	1.0
28	19		17.9	9.1	41.0	6.0
30	18		24.3	37.0	37.0	14.0
Total	53		18.0	10.1	46.0	1.0
10	16	a_{ij}	60.1	39.4	140	10.0
28	19		172.6	162.7	600	20.0
30	18		256.6	210.9	600	46.0
Total	53		162.1	172.3	600	10.0
10	16	b_{ij}	10.3	2.9	15.0	5.0
28	19		24.2	7.5	35.0	10.0
30	18		24.9	6.6	35.0	12.0
Total	53		20.2	8.9	35.0	5.0
10	16	c_{ij}	19.3	17.7	60.0	0
28	19		27.9	15.7	50.0	0
30	18		29.2	12.0	50.0	10.0
Total	53		25.7	15.5	60.0	0
						0