Effect of Distance and Population Size on Patient Trips in a Prefecture of Japan: Application of a Transportation Distribution Model to the Demand for and Supply of Health Services

Masanobu KUDO, Junichi MISUMI and Akira SHIMAOKA

Department of Public Health and Hygiene, Oita Medical University, Oita

Abstract

The gravity model, a method for analyzing transportation distribution in transportation engineering, was used to explain patient trips between ten health service regions in a Japanese prefecture. The OD (Origin-Destination) tables were constructed with zoning by regions, distinguishing between out- and inpatients. The observed trips were determined from the data of a survey conducted in 1992 that reported the locations of patient residences and chosen medical facilities.

The base values used in the model calculations were the population size of each region and the road distance between the centers of regions. Problems of intrazonal trip were avoided by setting a mean intrazonal movement distance.

This model was calibrated by the linear regression method with simultaneous validation by the index of correlation coefficients. The model was found to accurately simulate the effect of distance on the choice of medical facilities and the differences between the characteristics of in- and out-patients. The population value in use showed the relation not only with demand but also the supply of clinical services.

It was suggested that the model presented here was useful in the allocation of medical resources and would help explain the relationship between suppliers and consumers of medical services.

Key words: transportation distribution, gravity model, health services, distance, population size

Introduction

The health care system in Japan covers universal demands at a reasonable cost¹⁾. However, the aging population is placing new strains on the system. The Japanese government has responded with measures to restrain "excessive" demands on national and municipal budgets caused by the welfare and National Health Insurance systems²⁻⁵⁾.

The current Japanese government policy (Iryou-Hou: Medical Service Act: first revision 1986 ; second revision 1993) requires prefectural governors to design primary health service plans which divide prefectures into zones named as secondary health service regions under a program to cover the medical demands of inpatients arising in each region. The total bed

number in the region has been capped during planning and used as a benchmark in determining an appropriate allocation of medical resources.

Although the capacities of health service regions were assumed to meet local demands, in reality there are movements of patients between regions. In addition, the total cost of health care and welfare is rising steadily despite the caps on bed numbers, which are intended to restrain the expansion of equipment and medical facilities⁶⁾.

The previous researches of the trips of patients⁷⁻¹²⁾ in medical field have not suggested useful quantitative models. The pioneer of this problem in social fields¹³⁻¹⁶⁾ employed an entropy model of A.G. Wilson^{17,18)} and succeeded to some extent, but was not popular in medical field.

This study set out to provide a quantitative explanation on how medical services are produced and distributed in a Japanese prefecture through a macro view of movements in regional populations. Transportation distribution analysis¹⁹⁾ was used for this purpose. The population of the prefecture was divided into clusters based on secondary health service regions, and the gravity model^{20,21)} of W. Isard, which was simpler than entropy models,

Received Apr. 23 1997/Accepted Dec. 4 1998

Reprint requests to: Masanobu KUDO,

Department of Public Health and Hygiene, Oita Medical University, Hazama, Oita 879-5593, Japan.

TEL. +81(97)586-5742

FAX. +81(97)586-5749

Email: xkudo@oita-med.ac.jp

was used to account for the distribution of services with regard to distance and population size. In addition, the relationships between the population and the variables assumed to reflect medical resource distribution were studied by multivariate analysis.

Methods

Data

The area examined in this study was Oita Prefecture on the east coast of Kyushu Island in Japan (Fig. 1). The population of the prefecture was about 1,233,000 (October 1992)⁶⁾ and the area 6,336 km². The sex and age distributions of the population are shown in Table 1. The ratio of people aged 65 and over was 16.8 % (the Japanese national figure; 13.1 %) at that time.

The secondary health service regional divisions were used as the geographical population clusters for zoning in this study. The features of those ten regions (R1-10) in 1992 — name, population size, numbers of clinics, hospitals, doctors and beds²²⁾ — are shown in Table 2. The populations of the regions are presented as circles in Fig. 2.

Table 1 The population of Prefecture (1992).

	Sex	Age groups (years)	Population
Total			1,232,846
	Male	Total	581,453
		-14	111,649
		15-64	387,587
		65-	82,217
	Female	Total	651,393
		-14	106,025
		15-64	420,962
		65-	124,406

A patient survey was conducted on October 13 (Tuesday) 1992 by the prefectural government with the collaboration of the prefectural Medical Association.

Questionnaires were sent to all of the clinics and hospitals in the prefecture except dental clinics and hospital dental departments. Each facility was requested to record all patients who visited or stayed on that day and had clinical records. The questionnaire requested patients' sex, birth date, residence, out- or inpatient status, disease names (two selected diagnoses, main and sub one), location of the facility and other items. No patient names were recorded. The return rates were 97.6 % from hospitals and 82.5 % from clinics.

Patients who attended and were hospitalized, or who had been hospitalized and were discharged on the day of the survey were classified as inpatients. Patients who were not hospitalized and those who received house calls were defined as outpatients. The inpatients who consulted other departments or clinics were not counted as outpatients. Outpatients who attended other departments or clinics were counted separately and this multiple attendance was included in the outpatient data. Diseases were coded based on ICD-9 with the cooperation of the prefectural office. Patient residences and places of medical facilities were also coded.

Patients residing in other prefectures were excluded from this analysis: 2.3% of the outpatient and 5.5 % of the inpatients. Patients from Oita seeking care in other prefectures could not be detected in this survey, but their numbers were judged not so large as to confuse the analysis. About 140 records of ambulatory patients were excluded from analysis because of defects in those data.

The ten regions were arranged in a ten-by-ten matrix that was called OD (Origin-Destination) table in which the patient trips or movements were classified from the survey data based on the locations of the residence and the medical facility according to the zoning by regions. Such tables were made using SPSS

Fig. 1 Location of Oita Prefecture and its ten health service regions.

Right: Japan, with Oita Prefecture indicated by an arrow. Left: Enlarged view of Oita Prefecture showing the boundaries of its ten health service regions, whose names appear in Table 2.

Table 2 The populations and other variables of 10 health service regions (1992).

Names of health service regions	Number				
	Population	Hospitals	Clinics	Beds	Doctors
Higashi-Kunisaki (R1)	40,355	2	24	491	54
Beppu-Kitsuki-Hayami (R2)	195,863	36	159	6,711	470
Oita (R3)	471,020	58	284	10,643	1,125
Usuki-Tsukumi (R4)	63,197	5	44	899	73
Saiki (R5)	89,576	9	62	1,616	135
Ohno (R6)	56,540	6	42	918	75
Taketa-Naoiri (R7)	31,241	5	26	611	50
Hita-Kusu (R8)	113,893	20	81	2,179	143
Nakatsu-Shimoge (R9)	86,759	14	71	2,218	158
Usa-Takada (R10)	84,402	13	67	1,648	104

Fig. 2 Populations of health service regions.

The area of each circle corresponds to the population size of its region. Scale is shown at left.

Table software^{23,24)} for out- and inpatients, and for other subgroups. The values of one hundred cells in each table were estimated using the model described later in this paper.

It was assumed that travel time would differ between inpatients and outpatients; the former were hospitalized and stayed at the facility on the day of survey, while the latter presumably made round trips between the facilities and their homes on the day of the survey. Because of these differences, all analyses were done separately on the out- and inpatient data.

The patient residence data were considered reliable since they were based on the administrative records of the Japanese Health Insurance System.

In full consideration for the privacy of patients and facilities, the data were compiled with the permission of the appropriate divisions of the prefectural government.

Model

The original of the gravity model^{19,20,21)} was:

$$T_{ij} = k \cdot \frac{P_i \cdot P_j}{D_{ij}^2}$$

where i and j are the zone i and zone j as the origin and the destination of the transportation each, T_{ij} is the quantity of the transportation from zone i to zone j , P_i and P_j is the population of zone i and j each, D_{ij} is the social distance between zone i and zone j as the resistance of transportation, and k is a coefficient.

With some modifications on the original model after preliminary analysis, the expression of this model used in this study was:

$$T_{ij} = \alpha \cdot \frac{P_i^\beta \cdot P_j}{D_{ij}^\gamma} \cdot C_i$$

where i represents the region of the patient's residence as origination of trip, and j that of the medical facility as destination of trip, expressed as a number from one to ten according to the map in Fig. 1. T is the trip number (patient's record number), P_i is the population of region i , P_j the population of region j , D_{ij} the road distance between regions i and j . Alpha, beta and gamma are coefficients. C_i is the ratio of elderly aged 65 and over in region i . The coefficient alpha reflects the total number of patient records.

The modifications for realistic modeling are explained below.

Firstly, the population size of trip origin, P_i , was powered by a coefficient (beta) because the both ways flow of patients between regions with large and small populations was assumed to be asymmetric. The population size of trip destination, P_j , was not powered by coefficient because one coefficient was enough to adjust the balance of P_i and P_j in computing of multiple regression model by logarithmic values.

Secondly, road distances between regions were employed for the distance in original model. In geographical terms, the center of gravity of the regional population was set at the largest population town or city in the region, and the municipality's center was settled as the starting point and as the end point in determining the road distance between regions. These municipalities were all sites of district Health Service Bureau offices. The D_{ij} values that were obtained from a published road map were expressed in a ten-by-ten matrix and ranged from 14 to 152 km.

Thirdly, when the patient's residence and the medical facility were in the same region, the distance was zero by definition and the denominator of the model expression became zero. This problem was solved by setting mean intrazonal movement distance (MIMD), the average distance that patients would travel within a single region. In computation, when $i = j$, zero values of D_{ij} were replaced by the MIMD.

Finally, the ratio of population aged 65 and over was multiplied to adjust for the regional characters of higher attendance and hospitalization rates in regions with more elderly citizens. This modification was intended not to limit the population of analysis only to the elderly, but to compensate for regional imbalance in generation of patients.

Table 3 Summary of coefficient alpha and correlation values for inpatient group.

Disease (ICD-9)	Sex	Age class (years)	Number of patient	Alpha value	Correlation coefficient (R)	95% Confidence intervals of R	R square	F value+
All			21135	0.02860	0.873	(0.816-0.913)	0.761	1,389.91
	Male	Total	9993	0.02051	0.860	(0.798-0.903)	0.739	1,246.24
		-14	407	0.00146	0.791	(0.704-0.855)	0.626	787.46
		15-64	5685	0.01331	0.847	(0.780-0.894)	0.717	1,130.03
		65-	3901	0.01528	0.895	(0.847-0.928)	0.800	1,712.39
	Female	Total	11142	0.01853	0.902	(0.857-0.933)	0.813	1,841.90
		-14	314	0.00116	0.657	(0.529-0.756)	0.432	419.58
		15-64	4234	0.00992	0.878	(0.823-0.916)	0.770	1,454.52
		65-	6594	0.02549	0.910	(0.869-0.939)	0.828	2,030.32
Diabetes (250)			637	0.00090	0.882	(0.829-0.919)	0.778	1,509.56
	Male	Total	270	0.00062	0.863	(0.803-0.906)	0.745	1,285.90
		-14	2	*				
		15-64	145	0.00043	0.818	(0.741-0.874)	0.669	927.84
		65-	123	0.00060	0.776	(0.684-0.844)	0.602	724.89
	Female	Total	367	0.00066	0.862	(0.801-0.905)	0.743	1,268.39
		-14	0	*				
		15-64	105	0.00031	0.619	(0.481-0.727)	0.383	361.12
		65-	262	0.00114	0.825	(0.750-0.879)	0.681	971.80
Cerebrovascular disease (430-438)			2183	0.00292	0.893	(0.845-0.927)	0.798	1,686.31
	Male	Total	899	0.00187	0.883	(0.831-0.920)	0.780	1,527.64
		-14	2	*				
		15-64	244	0.00063	0.840	(0.770-0.889)	0.705	1,072.83
		65-	653	0.00264	0.877	(0.822-0.915)	0.768	1,438.44
	Female	Total	1284	0.00213	0.872	(0.815-0.912)	0.760	1,383.31
		-14	3	*				
		15-64	136	0.00038	0.533	(0.376-0.660)	0.284	265.28
		65-	1145	0.00444	0.864	(0.803-0.906)	0.746	1,288.26
Heart disease (390-429)			1901	0.00251	0.909	(0.867-0.938)	0.826	2,004.82
	Male	Total	571	0.00121	0.864	(0.804-0.907)	0.747	1,294.63
		-14	3	*				
		15-64	150	0.00042	0.734	(0.628-0.813)	0.539	585.98
		65-	418	0.00173	0.856	(0.793-0.901)	0.733	1,211.47
	Female	Total	1330	0.00217	0.901	(0.857-0.933)	0.813	1,840.01
		-14	1	*				
		15-64	156	0.00043	0.634	(0.499-0.738)	0.401	382.40
		65-	1173	0.00447	0.891	(0.842-0.926)	0.794	1,654.09
Neoplasm (140-239)			1797	0.00243	0.859	(0.796-0.903)	0.737	1,237.20
	Male	Total	944	0.00196	0.844	(0.776-0.892)	0.712	1,104.45
		-14	7	*				
		15-64	380	0.00090	0.630	(0.495-0.736)	0.397	377.50
		65-	557	0.00225	0.839	(0.770-0.889)	0.704	1,069.91
	Female	Total	853	0.00145	0.848	(0.782-0.895)	0.719	1,142.02
		-14	14	*				
		15-64	384	0.00091	0.590	(0.445-0.705)	0.348	324.61
		65-	455	0.00187	0.852	(0.787-0.898)	0.726	1,174.83

*Model could not be calculated because of smallness of records concerned.

+All of the F values in this table were significant because F (98,1) for 0.1% p value was 11.5.

The estimating model was expressed by one hundred Tij arranged in a ten-by-ten matrix on a spreadsheet which corresponded to the observed OD table.

The calibration and validation were evaluated by the calculation of the correlation coefficient between those one hundred estimated Tij and those corresponding observed Tij in the OD table. In this computation the zero cells in the OD table of the observed Tij were replaced with 0.5 for the use of logarithmic values in calculation, because when the observed case was zero the expected values from the Poisson distribution, resulted in one at the maximum expectation when the population exceeded one thousand.

The computation steps of calibration and evaluation were as

follows:

- 1: Macro search for adjustment of alpha, beta and gamma using Microsoft Excel software.
- 2: Macro adjustment and optimization of the MIMD value using Microsoft Excel software.
- 3: Specific adjustment using linear regression model in REGRESSION procedure of SPSS software^{23,24)}.

The expression used in the third step was:

$$\text{Log}T_{ij} = \text{Log}(a) + \beta \cdot \text{Log}P_i + \text{Log}P_j - \gamma \cdot \text{Log}D_{ij} + \text{Log}C_i$$

The first and second steps were carried out with attention to the validation index of correlation efficiencies. The third step was

Fig. 3 Distribution of patients among health service regions.

The area of each circle corresponds to the number of patient trips originated in the region and recorded in the survey. Inpatients are represented on the left, and outpatients on the right. Scale is shown on the left.

Fig. 4 Distribution of inpatients by origin and destination.

The starting point (residence) and endpoint (facility locations) of patient trip are categorized by zoning of health service regions. R1 to R10 represent the 10 regions named in Table 2.

done in two ways: in one condition all variables had coefficients, and in the other the number of coefficients was exactly limited as in the model expression. Both computations were done using the real and the logarithmic values of patient record numbers.

After these steps two models for out- and inpatients were decided and correlation coefficients for subgroups of sex, age, and some diseases were separately calculated with fixed coefficients, except the alpha values that definitely came from the total number of patients in each subgroup. The diseases in the analysis were diabetes (ICD-9 code 250), cerebrovascular disease (code 430-438), heart disease and hypertension (code 390-429) and neoplasm (code 140-239).

The interactions of health service resources, medical demands, and population were examined in multi-plotting and cluster plotting of these variables by region using the S language²⁵⁾

Fig. 5 Plot of outpatient trips by estimated and observed values.

The estimated values were derived from the gravity model. Points shown as small circles were plotted on a logarithmic axis; each indicates the number of trips categorized by 10 zones of the location of medical facilities and 10 zones of patient residences ($N=100$), as shown in Fig. 4 for inpatients. The vertical line of points on the far left on the observed value axis represents zero values that were adjusted to 0.5.

on a UNIX system.

Results

Outlines of the Oita Prefecture patient survey of 1992

The records, including patients from other prefectures, included about 78,000, of which 22,000 were inpatients and the other 56,000 were outpatients. The rate of the inpatients in the population on the day was 1,816 per 100,000 and the rate for outpatients was 5,554 per 100,000. The distribution of out- and inpatients classified by patient residence was shown in Fig. 3 and the distribution of inpatients in an OD table indicated a

Table 4 Summary of coefficient alpha and correlation values for outpatient group.

Disease (ICD-9)	Sex	Age class (years)	Number of patient	Alpha value	Correlation coefficient (R)	95% Confidence interval of R	R square	F value+
All	Male	Total	55015	0.02992	0.914	(0.875-0.941)	0.835	2,132.00
		-14	23885	0.02086	0.922	(0.886-0.947)	0.850	2,370.92
		15-64	3250	0.00495	0.912	(0.871-0.940)	0.831	2,067.73
		65-	12217	0.01225	0.919	(0.881-0.945)	0.844	2,258.62
	Female	Total	8418	0.01383	0.926	(0.891-0.949)	0.857	2,484.75
		-14	31130	0.01947	0.925	(0.890-0.949)	0.855	2,452.75
		15-64	2674	0.00309	0.911	(0.870-0.939)	0.829	2,046.52
		65-	15449	0.01123	0.918	(0.881-0.944)	0.843	2,246.37
Diabetes (250)	Male	Total	1439	0.00085	0.900	(0.855-0.932)	0.811	1,819.59
		-14	742	0.00068	0.872	(0.816-0.912)	0.761	1,385.43
		15-64	0	*				
		65-	432	0.00047	0.869	(0.811-0.910)	0.755	1,346.09
	Female	Total	310	0.00058	0.814	(0.736-0.871)	0.663	906.90
		-14	697	0.00045	0.879	(0.825-0.917)	0.773	1,474.27
		15-64	2	*				
		65-	290	0.00024	0.850	(0.785-0.897)	0.722	1,157.46
Cerebrovascular disease (430-438)	Male	Total	405	0.00047	0.835	(0.764-0.886)	0.697	1,039.85
		-14	1867	0.00104	0.892	(0.844-0.926)	0.796	1,671.91
		15-64	892	0.00079	0.891	(0.841-0.925)	0.793	1,643.04
		65-	0	*				
	Female	Total	219	0.00025	0.862	(0.801-0.905)	0.743	1,269.29
		-14	673	0.00116	0.873	(0.816-0.913)	0.761	1,388.95
		15-64	975	0.00058	0.859	(0.797-0.903)	0.738	1,240.96
		65-	0	*				
Heart disease (390-429)	Male	Total	9015	0.00509	0.926	(0.892-0.950)	0.857	2,500.82
		-14	3341	0.00291	0.922	(0.885-0.947)	0.849	2,349.53
		15-64	7	*				
		65-	1415	0.00145	0.911	(0.871-0.940)	0.831	2,066.30
	Female	Total	1919	0.00319	0.894	(0.847-0.928)	0.800	1,706.94
		-14	5674	0.00344	0.904	(0.860-0.935)	0.817	1,894.12
		15-64	4	*				
		65-	1823	0.00132	0.890	(0.840-0.924)	0.791	1,626.85
Neoplasm (140-239)	Male	Total	3847	0.00408	0.889	(0.839-0.924)	0.791	1,619.42
		-14	854	0.00243	0.903	(0.859-0.934)	0.815	1,868.61
		15-64	366	0.00196	0.870	(0.813-0.911)	0.758	1,364.71
		65-	4	*				
	Female	Total	134	0.00090	0.576	(0.428-0.694)	0.332	308.70
		-14	228	0.00225	0.861	(0.799-0.904)	0.741	1,257.50
		15-64	488	0.00145	0.882	(0.829-0.919)	0.777	1,506.06
		65-	3	*				

*Model could not be calculated because of smallness of records concerned.

+All of the F values in this table were significant because F (98,1) for 0.1% p value was 11.5.

tendency to gather in two regions (R2,R3) with relatively large populations (shown as a 3D graphic expression in Fig. 4) .

Calibration and validation results

For outpatients, the estimated values of the coefficients were 0.03 in alpha, 0.33 in beta and 1.8 in gamma when the MIMD was set at 2.9 km, which was smaller than the shortest value of observed Dij. Analysis using the linear regression model under condition a) (all variables have the respective coefficients) resulted in the total correlation of 0.92, and under condition b) (coefficients correspond to the model expression) a correlation of 0.91 by logarithmic values.

For inpatients no changes of beta and gamma and small

change of alpha were needed; the MIMD alone was changed to 5 km. In this case the correlation under the condition a) was 0.89 and that under the condition b) was 0.87.

The correlation coefficients of the two conditions were so similar that condition b), which closely resembled the model, was selected after analysis. In the plots of estimated and observed values by out and in-patient all points gathered closely: outpatient data are plotted in Fig. 5.

Summaries of subgroups by sex, age and selected diseases are shown in Table 3 for inpatients, and Table 4 for outpatients. The alpha values were adjusted for each subgroup and other coefficients were fixed. The Ci values were adjusted for each sex group and the populations of each region were adjusted for sex

Fig. 6 Distance effects on outpatient trips destined for R3.

Each point indicated region and was plotted separately for the observed and the estimated values. The vertical axis was the outpatient trips destined for R3 and the horizontal axis was the distance from the region to R3. The estimated values' line did not smoothly curve because the P_j values of each region had effected.

*R3, distance=2.9km (MIMD: mean intrazonal movement distance)

Fig. 7 The outpatient trip destination and the regional population.

The patient flow in a region could be divided into three categories by its origin and destination: the first is moving in a same region, the second is moving out to other regions, the third is moving in from other regions. Each point of this graph indicates a region and is plotted by the population and by the values of outpatient trips according to the three categories above. The trip value was summed by the region of concern. For the third category, the population of the region and the total patient trip from other regions were matched.

and age groups. In some data sets with over 200 records, the square of the correlation efficient was over 0.7. In all categories, and even for individual diseases, the results were well consistent with the model for both out- and inpatients judging by the F-values of every category and the relatively narrow ranges of confidence limits.

The largest value in the difference between the estimated and observed values was in the outpatient trips which were originated in R3 and destined to R2 about 20 km distance from R3.

Effects of distance and population

An example of the effects of distance on patient trips in the model was shown in Fig. 6. In the figure, outpatient trips destined for R3 were plotted for the observed and estimated values. Most of the observed values were distributed near to the curve drawn from the estimated values. Since the vertical axis was with the logarithmic scale some differences of the trip values between the observed and estimated values were remarkably shown. However, those were relatively small in the total number. In other plotting of trips destined for single region and their distance, fitting results were similar.

The relationships between the trip and the population of origin and destination were shown in Fig. 7 for outpatient. The

trip of the patients who moved within a region ($P_i \Rightarrow P_i$) was highly correlated to the regional population size (P_i). There was a high correlation of the moving-in trip ($P_j \Rightarrow P_i$) and the population size of the destined region (P_i) that must be intrinsically independent each. What to be worthy of notice was a bend of the point rows of extrazonal trips destined to other regions ($P_i \Rightarrow P_j$). The point of bend was placed at about 90,000 of the regional population size (P_i).

Interactions of medical demands, resources and zone population

In Fig. 8 the variables related to health services were plotted on a logarithmic scale; each point represents a health service region. High correlation of each variable can be in the plots. The values of these correlation are shown in Table 5.

Cluster analysis indicated the Euclidean distance between variables with the clustering method of group means (see Fig. 9). In the cluster plotting those variables largely fell into the three groups. One included the variables of the number of hospitals, clinics, and doctors; the second contained the number of patients (out- and in-) and beds. The third group was the variable of population size and was placed in midway between the other two groups.

Table 5 Correlations of variables for the population and other variables (using logarithmic values).

	Hospitals	Population	Clinics	Beds	Doctors	Inpatient	Outpatient
Hospitals	1.000	0.923	0.966	0.971	0.921	0.966	0.946
Population		1.000	0.988	0.971	0.980	0.958	0.989
Clinics			1.000	0.993	0.978	0.980	0.994
Beds				1.000	0.982	0.995	0.983
Doctors					1.000	0.982	0.969
Inpatient						1.000	0.967
Outpatient							1.000

Discussion

Use of distance for modeling the population movements

Previous studies^{7-9,13-16,26-30)} have investigated the relation between geographical distance and population movements. Colman²⁶⁾ studied the relocation caused by marriage in a nationwide survey of Britain, applying a geographical theory of central place which predicted that in large populations the movement was small, and that in small populations on the other hand the movement was more frequent and often directed to larger populations. Relethford et al.²⁷⁻²⁹⁾ studied the exogamy and migration rates of some districts and found that the relationships of such movement with population size, geographic distance and temporal change persisted in multiple regression analysis. Those studies described movements with mathematical models but they could not predict actions within the population.

In social medicine, the problem of the disposition of services has been addressed in terms of geographical distances; some studies have considered the distance to hospitals as a factor in attendance choice. Using hospital records, Studnicki²⁷⁾ delineated the distribution of hospital service for deliveries in Baltimore with the optimal accessibility model, and McLafferty³⁰⁾ predicted the hospital utilization pattern in New York City with the multiplicative competitive model. In those papers, distance was recognized as a significant factor in choosing a hospital. However, Egunjobi³¹⁾ insisted that proximity, which was the leading factor of 'choosing a hospital' accounted for about 30% of the choice and the factors of 'quality of hospital' and 'presence of relatives living in hospital town' together for 40% based on interviews of about 860 patients in Nigeria. And also suggested the maze of interactions among those variables that was to be expected; such indistinct results would follow from the limitations of the interview data.

Indeed, those studies advanced no quantitative models for estimating the actual flow of patients or the distribution of health services. Taket and Mayhew¹³⁾ applied the entropy model for planning hospital services in London, in which the origin and destination zones were separately defined and then the problem of intrazonal trips was latently avoided. There were some applications of their model for other districts¹⁴⁻¹⁶⁾. The complexity of model formula and the stiffness of model settings had spoiled their usefulness. In medical fields, they seemed not to influence researchers.

There was a work with multiple regression analysis predicting the geographic distribution of chronic homebound patients³⁰⁾ in which used were three variables of distance, relationship with senile dependency and index of family socioeconomic capacity with results of the square of correlation coefficient up to 0.42. They predicted the number of ill per 10,000 inhabitants and the population size effects were not mentioned.

The assumptions of this simple model

From discussion above, fundamentally this model depended on the assumption that medical demands would occur in proportion to the population size and that the medical resources would correlate to the population size. Another assumption was that the distance would be a large factor in the choice of medical facilities as geographical, economic or social cost. These three assumptions were put to practical use in the gravity model.

In Fig. 7 the intrazonal trips of each region were in good

correlation with the regional population size and the extrazonal trips destined to the region also showed good correlation with the population size of destined region. The former suggests that the regional variance of the producing rate of the outpatient was relatively small. The latter suggests a hypothetical power of the sucking force or attraction of the population size; a fundamental assumption of the gravity model. The bend of the point rows of extrazonal trips destined to other regions was placed at about 90,000 of the regional population size and it indicates an effect of balance that region with more than 90,000 population could detain self-originated patients and with less than 90,000 could not. These findings show that the asymmetric flows of patients which we assumed in the above method section existed. In the observed imbalances of patient trip, the flow from smaller population regions tended to be larger than that in the opposite direction. Therefore, the former two assumptions were validated to be realistic.

As shown in Fig. 6 outpatient's number inversely proportional to the distance was successfully simulated by the model expression. Therefore the third assumption was also appropriate for the model.

Fig. 8 Multiplot by pairs of the clinical resource, population, and patient status variables.

The logarithmic values of the six variables were plotted by pairs. Each point shown as an asterisk in this Fig. represents a region. Both variables of patient status, out- and inpatient, were categorized by the placements of facilities.

The modifications and adjustments of the model

The discrepancies in regional attendance rates were thought to be caused mainly by differences in age distributions. A modification was made by directly multiplying of rates of the residents aged 65 and over because the disproportionately high attendance rate of elderly raised the rate of patient total remarkably. The high attendance rates of those elderly outpatients would be caused by multiple visits that could not be detected in this survey. Such multiple visits, however, would not have skewed the results since these trips also reflect the total health care needs of the district.

The road length was employed as social distance after preliminary studies with direct or other social distances. Oita Prefecture includes a broad mountainous area and a wide gulf on the east side (Fig. 1). Judging from the size of the area, walking, small vehicles, the transportation car (personal or public) and railway were thought to be the primary means of daily travel for attendance or admission to medical facilities in this prefecture. In shorter distance less than about 2 km, walking or bicycle would be the main means. Only one center of health service area was out of the rail connections. The paved national and regional roads were well maintained, but in 1992 the highway, motor vehicles only, served very limited portions of the prefecture. These conditions would contribute to the equality of distance effects. For airplanes the district is too small.

The suitability of the model

Although the gravity model estimates not the total amount of patient trips but the states of distributions and, it appears that the model may be applied universally with comparable predictive accuracy in this district.

It partially would come from the economical accessibility. The health insurance system in Japan covers most of the population. The relatively low premiums have assured broad participation and reduction of patient payments by 70 to 90%. For the elders and the patients with high cost, special discounts are provided. The easy accessibility makes the choice of facilities free from severe economic considerations. These conditions in Japan would be profitable for the assumption of cost in traveling the distance^{20,21)} too.

The different characteristics of the out- and inpatient categories were absorbed by simply using hypothetical values of the MIMD. The MIMD of inpatients was larger than that of outpatients, suggesting that the distance to facilities in the same region had less effect on the choice by serious patients who needed admission or hospital treatment than outpatients who would tend to choose facilities relatively near their residences. On the other hand, in the crossing movements over region's border, the difference between out- and inpatients would be small. As related in the results the largest differences between observed and estimated values were in the flow from Beppu-Kitsuki-Hayami (R2) to Oita (R3) for both out- and inpatients. The predicted values were larger than those observed, with differences of 996 (1.7% of total observed) for outpatients and 853 (3.8% of total observed) for inpatients. The R2 region had the second-largest population, and the centers of the two regions were very near and connected by highway and railway. This result is partially explained by the relatively high density of medical facilities in district R2. As concerns that the trip's scale was logarithmic, total fitness of the model explained the effects of distance on the patient trip effectively.

In subgroups by sex or age class the correlation coefficients were slightly lower than those of groups' total because some numbers of each group were too small for analysis. Most of the categories with sufficient cases showed high values of the square of the correlation efficiencies. If we adjusted beta and gamma for each condition of subgroups, the correlation would be improved.

Factors such as disease and age class or regional discrepancies in trips would need further investigation.

Relationships between health services' demands, resources and population

In our results, medical demands that were measured by patient trips had a good correlation with all of the used variables related to the resources of medical services. In multivariate analysis, population size fell between two interpreted clusters of medical resources and medical demands. Concerning interactions among variables, the population size would contain both of the factors for medical resources working as supplying side and the factors for medical demands reflecting patient number. These interpretations would suggest that the population size itself can be used as a representation of variables for predicting the amount of health services in supply.

Then it may be inferred that demands of medical service exist in proportion to its size, and that public and private medical facilities supply medical services not haphazardly but according to a law that population size in essence dictates the supply of medical services in terms of quality and quantity.

Although the number of beds was the nearest, namely strongly correlated, variable to the patient number in Fig. 9, it seemed that this artificially controlled variable could not reflect all the essential characteristics of demand or supply that involve quality aspects, for example levels of skill or clinical technology.

The quality of services

The problem of how population size had reflected the amount of medical resource capacities had been discussed in aspects of quantity above and the quality of medical services was

Fig. 9 Cluster plot of the clinical resource, population, and patient status variables.

The cluster tree made using group means indicates the Euclidean distances between the logarithmic values of these variables. Patient status was categorized as in Fig. 8. The scale above the tree indicates the distance values.

not directly analyzed in this study because of the data restriction. Giving consideration to this problem, the need for special resources such as advanced techniques or equipment does exist, and a large, concentrated population would generally demand such services in plenty. From the side of supplier, a typically big city would have large hospitals or many hospitals with highly skilled clinicians. In this model, the aspects of these quality of the services were included in the effects of the population size. Exceptions would be made in case of some prominent clinicians or facilities which were in population thin districts, and which could attract and gather patients from distant areas. The result from this simple model, however, may provide an explanation that even such attractive clinics or facilities would largely depend on the regional demands of their immediate areas.

Model's simplicity

Since the model complexity and data accuracy are in the relationship of trade-off⁹⁾, the more variables had added at the model example in a linear regression, the more the accuracy of data would be necessary, which means the raise of the investigation costs. The complex model would have good fitness but the universal fitness, in case of application to other districts, would be small and the interpretation of the model would become difficult.

The simplicity of this model could contribute to the economy of investigating cost and the easiness for interpretation would be reasonable for practical use.

In further studies, questions of quality aspects, for example, should be considered in detail. It must be carried out carefully to add new related variables because the result would easily become a maze of interactions among the variables just as in some previous studies⁹⁾. Some efforts would be necessary to avoid then, such as preliminary investigation of variables using multivariable analysis.

Application to other districts

This study examined regions with relatively large areas and a moderate population density (195 persons per km²). To apply this model to a large metropolitan area or a densely populated region it would be necessary to use other analogues for social distance, such as traveling time or direct distance, because those areas would be connected closely by dense transportation flows competing with each other and equalizing the moving costs. The setting of the population category such as the size of clusters

would be important too. Nukata et al.¹⁰⁾ proposed splendid suggestions in treating a high density population in which they used a meshed map of Yokohama City and proposed that the number of visits to medical facilities was dependent on the size of the population related to attractiveness of the facilities, and on the population density of the region related to the specialties of the facilities. Those hypotheses of the attractiveness and specialties of the facilities were only supported by a tendency in the data but must be of help in the further study of quality aspects in the micro and macro views.

In contrary for the population thin districts, the distance settings for various methods of transportation would be needed. If the zoning population or area was small³⁰⁾, the problem of intrazonal trip would be adequately avoided or used carefully.

Settings outside of Japan, adjustments for financial accessibility would be necessary in view of economic and services' quality concerns.

Conclusion

The findings indicate that the choices of clinical facilities were greatly influenced by social distance, interpreted as road distance in this study. They also indicate that the demand for and supply of medical services depend on the population size of the district, which in itself seems to present an economic justification and yardstick for provision of medical resources. For the simple modeling the population size value could be used to avoid the maze of interactions among variables. The applicability of this method would be high in investigating the relationships of medical resources to population.

Acknowledgments

This study was greatly supported by the Department of Health and Environment, Oita Prefectural Government in Japan. We thank all the facility staff concerned with the survey for their cooperation and the staff of the related divisions of the prefectural government for providing data and the cooperation in data processing. We are also grateful to Dr. M. Omoto of Toho University for help in searching references and thank to colleagues at our department for advice and supports.

A part of this work was presented at the 54th General Meeting of the Japanese Society of Public Health, Yamagata, Japan, 12-14 October 1999

References

- 1) Ikegami N. The economics of health care in Japan. *Science* 1992; 258: 614-8.
- 2) Babazono A, Ogawa T, Babazono T, et al. The effect of sharing medical costs for the aged 70 or more on the financial balance of the health insurance society. *Jpn J Hyg* 1991; 46: 890-7.
- 3) Babazono A, Ogawa T, Babazono T, et al. The effect of a cost sharing provision in Japan. *Family Practice* 1991; 8: 247-522.
- 4) Ikegami N. Japanese health care: low cost through regulated fees. *Health Affairs* 1991; 10: 87-109.
- 5) Ikegami N. Japan: maintaining equity through regulated fees. *Journal of Health Politics. Policy & Law* 1992; 17: 689-713.
- 6) Kousei Toukei Kyokai. Special issue; 1993 Trends in the National Health (in Japanese). *Kousei no Shihyou* 1993; 40: 242-3.
- 7) Studnicki J. The minimisation of travel effort as a delineating influence for urban hospital service areas. *Int J Health Services* 1975; 5: 679-93.
- 8) McLafferty S. Predicting the effect of hospital closure on hospital utilisation patterns. *Soc Sci & Med* 1988; 27: 255-62.
- 9) Egunjobi L. Factors influencing choice of hospitals: a case study of the northern part of Oyo State, Nigeria. *Soc Sci & Med* 1983; 17: 585-9.

- 10) Nukata A. A study of geographical placement of medical service. *Jpn J Hyg* 1985; 40 (suppl 13): 389.
- 11) Wakabayashi K, Nakamura K, Masaki M, et al. Visiting behavior to medical institutions of 26 intractable diseases patient receiving financial aid for treatment in Saitama Prefecture. *Jpn J Public Health* 1992; 39: 729-38.
- 12) Wakabayashi K, Nakamura K, Masaki M, et al. Characteristics of medical institutions visited by patients with 26 intractable diseases who are residents of Saitama prefecture. *Jpn J Hyg* 1995; 50: 839-48.
- 13) Taket A, Mayhew L. Interactions between the supply of and demand for hospital services in London. *Omega the Int J Mgmt Sci* 1981; 5: 519-26.
- 14) Mayhew LD, Gibberd RW, Hall H. Predicting patient flows and hospital case-mix. *Environ and Plann A* 1986; 18: 619-38.
- 15) Mayhew LD, Leonardi G. Equity, efficiency, and accessibility in urban and regional health-care system. *Environ and Plann A* 1982; 14: 1479-507.
- 16) Royston GHD, Hurst JW, Lister EG et al. Modelling the use of health services by populations of small areas to inform the allocation of central resources to larger regions. *Socio-Econ Plann Sci* 1992; 26: 169-80.
- 17) Wilson AG. A statistical theory of spatial distribution models. *Transpn Res* 1967; 1: 253-69.
- 18) Wilson AG. A family of spatial interaction models, and associated developments. *Environ and Plann* 1971; 3: 1-32.
- 19) Ortuzar JD, Willumsen LG. Modelling transport second edition. Chichester: Jhon Wiley & Sons Ltd., 1994.
- 20) Suzuki K. Analysis of structure of population distribution. Tokyo: Taimeido Publishing, 1985.
- 21) Dodd ST. The interactance hypothesis: A gravity model fitting physical masses and human groups. *Am Soc Rev* 1950; 15: 245-56.
- 22) Department of Health and Environment. The year book of public health policy 1992. Oita: Oita Prefectural Government, 1993.
- 23) SPSS Inc. SPSS for Windows base system user's guide release 6.0. Chicago: SPSS Inc., 1993.
- 24) SPSS Inc. SPSS for Windows advanced statistics user's guide release 6.0. Chicago: SPSS Inc., 1993.
- 25) Becker RA, Chambers JM, Wilks A.R. The new S language. Pacific Grove, Ca: Wadsworth, 1988.
- 26) Coleman DA. The geography of marriage in Britain, 1920-1960. *Ann Hum Biology* 1977; 4: 101-32.
- 27) Relethford JH. Cross-cultural analysis of migration rates: effects of geographic distance and population size. *Am J Physic Anthropol* 1992; 89: 459-66.
- 28) Relethford JH, Mielke JH. Marital exogamy in the Aland Islands, Finland, 1750-1949. *Ann Hum Biol* 1994; 21: 13-21.
- 29) Mielke JH, Relethford JH, Eriksson AW. Temporal trends in migration in the Aland Islands: effects of population size and geographic distance. *Hum Biol* 1994; 66: 399-410.
- 30) Segura JM, Sanz F, Bastida N, et al. Description and validation of a predictive model of the geographic distribution of chronic home-bound patients. *Atencion Primaria* 1996; 17: 292-4.