CASE 75

Use of Mahalanobis Distance in Medical Diagnosis

Abstract: This case deals with use of the Mahalanobis–Taguchi system (MTS) and multivariate analysis together with a special medical examination focusing on liver disease. More specifically, as a medical application of statistics, 16 blood biochemical data, ages, and genders were analyzed to evaluate examinees' health conditions. This is the first research case study of the MTS method used in an application of Mahalanobis distance (D^2) to a multivariate analysis.

1. Introduction

Generally speaking, when analyzing multiple variables, we should create a good database above all. A good database is generally homogeneous. In the case of biochemical data for human beings, while data for healthy people have little variability and distribute homogeneously, those for patients suffering diseases vary widely and cannot be viewed as homogeneous.

For these reasons it is desirable to prepare a database by using examination results of healthy people (normal persons). However, when attempting to diagnose a specific disease, unless necessary examination items to diagnose the disease are included, we cannot make use of the database. In a periodic medical checkup, due to constraints of budget, personnel, or time, few examination items are selected for diagnosis. Even in a complete physical examination, only common items are checked. Therefore, if the data for a regular medical checkup or complete physical examination are used as a normal contrast, we sometimes lack examination items. Then it is quite difficult to obtain enough data on normal persons. However, when we determine a standard value in certain examination facilities, we can obtain healthy people's data. Yet it is not indisputable whether judgment of health conditions is made accurately enough. In this research, as our database, we capitalized on good biochemical data for normal persons, which we obtained by chance.

2. Database and the Base Space

For the people whose examination data can be obtained, after examining their disease histories in medical checkups at Tokyo Teishin Hospital over the last year, we selected 200 persons regarded to have no disease. Although we initially attempted to obtain data for 1000 persons, due to our limited computer capacity in the 1980s, we decided to perform an analysis using the data for only 200 persons. If a larger-scale computer were available and more people could be involved, it would be better to analyze a greater number.

Based on the performance of the automated analyzer used, we determined 18 examination items, including the following 16 biochemical data plus age and gender: total protein (TP), albumin (Alb), A/G (albumin/globulin) ratio, cholinesterase (ChE), glutamate oxaloacetate transaminase (GOT), glutamate pyruvate transaminase (GPT), lactate dehydrogenase (LDH), alkaline phosphatase (ALP), γ-glutamyltranspeptidase (γ-GTP), leucine aminopeptidase (LAP), total cholesterol (TCh), triglyceride (TG), phospholipases (PL), creatinine

Table 1
Age, gender, and biochemical data of normal persons

No.	Gender	Age	TP	Alb	A/G	ChE	GOT	GPT	LDH	γ-GTP	LAP	TCh	TG	PL	Cr	BUN	UA
1	Male	59	6.5	3.7	1.32	0.70	12	6	190	12	275	235	140	238	0.8	13	3.8
2	Female	51	7.0	4.4	1.69	0.98	21	18	247	23	340	225	87	227	1.1	15	4.5
3	Female	53	6.7	4.2	1.68	0.83	18	8	220	16	278	230	67	232	0.8	17	3.2
4	Female	52	6.6	4.3	1.87	0.90	12	6	244	15	289	171	59	175	0.8	13	3.8
5	Female	52	6.7	3.9	1.39	0.97	13	7	198	13	312	192	51	203	1.0	14	4.1
6	Female	56	7.2	4.0	1.25	0.93	22	16	188	16	304	216	86	213	1.0	13	4.2
7	Female	51	7.1	4.0	1.29	0.88	16	9	187	13	272	235	96	251	1.0	14	3.5
8	Female	50	6.6	3.6	1.20	0.71	14	5	190	12	270	149	57	165	0.9	14	3.4
9	Female	41	6.9	4.3	1.65	0.81	16	12	195	13	319	160	55	175	1.0	12	3.0
10	Female	48	7.0	4.2	1.50	0.93	21	19	230	35	411	197	110	213	1.0	8	5.2

Table 2
Normalized data of normal persons

No.	Age	Gender	TP	Alb	A/G	ChE	GOT	GPT	LDH	ALP	γ-GTP	LAP	TCh	TG	PL	Cr	BUN	UA
191	−1.42	+0.73	−0.34	−0.26	−0.08	−1.18	−0.44	−0.65	−0.06	−0.21	−0.57	+0.59	−1.24	−0.38	−0.72	+0.20	−0.39	−0.24
192	−1.32	+0.73	−0.34	+0.51	+0.94	+0.69	+0.90	−0.65	−0.92	−0.21	−0.57	+0.24	−1.36	−0.38	−1.44	−0.33	+0.23	+0.40
193	−1.32	+0.73	+1.48	+1.27	+0.04	+1.21	+1.16	+1.29	−0.41	+0.25	−0.26	+0.03	−1.09	+2.15	−1.15	+0.20	+1.16	−1.10
194	−1.90	+0.73	+1.48	+1.65	+0.49	−1.21	−0.67	−1.09	−0.68	+0.71	−0.88	+1.04	−1.75	−0.55	−1.80	+0.73	−1.32	+0.96
195	−1.23	+0.73	−0.65	+0.51	+1.32	+0.38	−1.36	−0.87	−0.51	−0.34	−0.72	−0.93	−1.24	−1.22	−1.11	+0.20	−1.63	+0.08
196	−1.03	−1.36	−0.65	+0.13	+0.75	−0.87	−1.13	−1.30	−1.87	+0.01	−0.41	+0.56	−1.30	−1.36	−1.11	+0.73	−0.39	+0.48
197	+0.80	+0.73	+0.04	−0.26	−0.36	+0.48	+1.39	+1.51	+1.57	+0.86	+2.81	+1.99	+0.25	−0.14	+0.35	+0.20	−0.39	+1.12
198	+0.45	+0.73	−0.34	−2.16	−2.05	−1.91	−0.67	−0.44	−2.18	+0.45	−0.26	+0.21	−1.45	−0.74	−0.33	+0.73	−0.39	+1.12
199	−0.45	+0.73	+0.26	+0.89	+0.78	−0.24	+0.47	+0.43	−0.24	−0.93	+1.04	+0.47	+0.31	+1.85	+0.38	+1.25	+1.18	+0.56
200	+1.29	−1.36	−0.65	−1.40	−1.11	+0.04	+0.93	+0.93	+0.96	+1.83	−0.34	−0.02	+1.79	+1.02	+1.39	−0.85	+0.54	−0.55
Total	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Ave.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 3
D^2/f and 10 log(D^2/f) of normal persons

No.	D^2/f	10 log(D^2/f)	No.	D^2/f)	10 log(D^2/f)	No.	D^2/f	10 log(D^2/f)
1	1.989	+2.987	36	0.513	−2.897	71	0.722	−1.413
2	0.609	−2.153	37	0.530	−2.761	72	1.165	+0.66
3	1.104	+0.428	38	0.567	−2.465	73	0.927	−0.328
4	0.747	−1.267	39	0.675	−1.704	74	0.946	−0.242
5	0.571	−2.432	40	0.746	−1.273	75	1.186	+0.741
6	0.893	−0.491	41	1.587	+2.006	76	0.557	−2.545
7	0.779	−1.087	42	0.859	−0.662	77	1.47	+1.674
8	0.872	−0.593	43	0.391	−4.053	78	1.244	+0.947
9	0.587	−2.316	44	0.616	−2.102	79	1.719	+2.352
10	1.274	+1.053	45	0.724	−1.400	80	0.440	−3.567
11	0.698	−1.564	46	0.813	−0.901	81	0.999	−0.005
12	0.596	−2.251	47	1.027	+0.117	82	0.707	−1.505
13	0.623	−2.053	48	1.107	+0.443	83	1.128	+0.524
14	1.249	+0.965	49	0.907	−0.422	84	0.713	−1.468
15	0.673	−1.722	50	1.773	+2.486	85	0.755	−1.219
16	0.482	−3.174	51	1.129	+0.527	86	1.086	+0.357
17	1.144	+0.586	52	0.634	−1.979	87	0.842	−0.749
18	0.529	−2.766	53	0.786	−1.047	88	0.972	−0.124
19	1.131	+0.535	54	1.368	+1.360	89	0.555	−2.560
20	1.299	+1.135	55	1.184	+0.733	90	0.761	−1.183
21	0.341368	−4.668	56	0.324	−4.900	91	1.90	+2.801
22	1.028	+0.118	57	1.412	+1.499	92	0.576	−2.395
23	0.588	−2.33	58	1.039	+0.165	93	0.775	−1.106
24	1.776	+2.495	59	0.782	+1.068	94	0.417	−3.803
25	0.657	−1.825	60	1.140	+0.570	95	1.682	+2.746
26	0.600	−2.222	61	1.213	+0.839	96	0.571	−2.433
27	0.720	−1.427	62	0.629	−2.015	97	2.251	+3.523
28	1.084	+0.349	63	1.135	+0.548	98	1.195	+0.775
29	3.442	+5.368	64	0.652	−1.858	99	0.678	−1.690
30	1.082	+0.343	65	0.811	−0.910	100	1.286	+1.093
31	0.960	−0.177	66	1.609	+2.06			
32	0.521	−2.831	67	0.905	−0.436			
33	1.076	+0.319	68	0.999	−0.003			
34	1.629	+2.120	69	2.264	+3.549			
35	0.896	−0.475	70	0.978	−0.098			

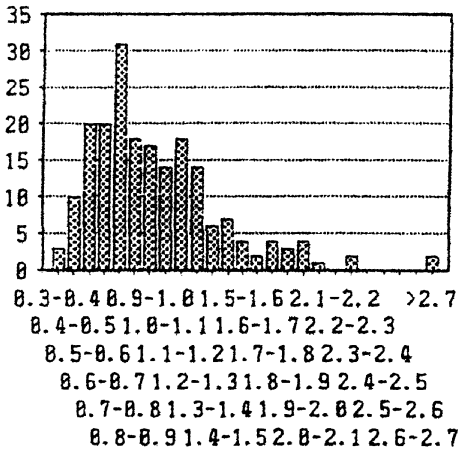

Figure 1
D^2/f of normal persons

(Cr), blood urea nitrogen (BUN), and uric acid (UA).

Since age and gender do not require extra cost and money for investigation and the biochemical data are associated with them, we considered it better to study the biochemical data along with age and gender. As the n value for age, considering that each datum is logarithmized, we set male as 10 and female as 1. Table 1 shows a part of the data.

Strictly speaking, although these data approximately follow the normal distribution, the normalization process facilitates the succeeding step. Setting the mean of the ith person's jth examination item to m_j, the standard deviation to σ_j, and the data from the ith person's jth examination item to X_{ij}, we calculated the normalized data:

$$Y_{ij} = \frac{X_{ij} - m_j}{\sigma_j} \tag{1}$$

Table 2 shows a part of the normalized data. The total sum and mean for each item result in zero, as shown in the table.

Although we can proceed with the analysis with D^2, now we use the value of D^2 divided by a degree of freedom, f, and its logarithm multiplied by 10:

$$Z = 10 \log \frac{D^2}{f} \tag{2}$$

Table 3 shows a part of D^2/f' and $10 \log(D^2/f')$. In addition, Figures 1 and 2 illustrate their distributions. Although neither distribution is completely homogeneous because the number of data are relatively small, it is considered homogeneous enough to analyze.

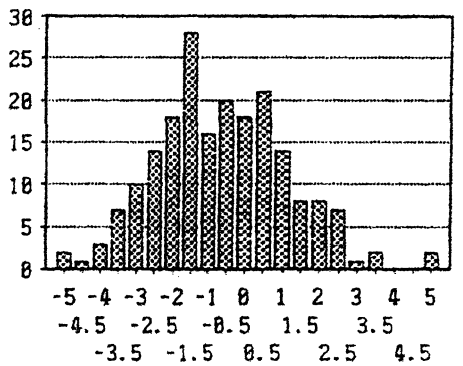

Figure 2
10 $\log(D^2/f)$ of normal persons

3. Data for Examinees Taking Special Medical Checkups

The examinees for a special medical checkup are basically clerical workers 35 years of age or older, 45 males and 50 females.

Collection of Data

The examinees were tested in the afternoon without diet restrictions. If even a single piece of data is beyond a target limit, the examinee takes the second test on another day with no breakfast.

Diagnosis of Examinees

1. If all data were within the target limit in the first test, also taking into account other data through a medical examination by interview or other checkups, we diagnosed the examinee as being within a normal limit (WNL).
2. If there were no data beyond the target limit in the second test, although a certain number of data were out of limit in the first test, we judged the examinee as WNL if the data beyond the limit were caused by the intake of

Table 4
Data of special medical examination

No.	Age	Gender	TP	Alb	A/G	ChE	GOT	GPT	LDH	ALP	γ-GTP	LAP	TCh	TG	PL	Cr	BUN	UA
1	35	Male	7.0	4.0	1.33	0.67	17	9	144	4.1	19	265	151	53	183	1.5	1.4	5.7
2	42	Female	7.5	4.2	1.27	0.67	15	10	187	5.4	14	315	202	148	218	1.3	14	5.6
3	41	Female	7.1	4.1	1.37	0.64	19	15	295	11.4	13	230	191	384	270	1.3	16	3.5
4	52	Female	6.6	3.9	1.44	0.88	14	10	243	5.8	10	289	185	96	218	1.3	13	4.7
5	53	Female	7.4	4.2	1.31	0.77	16	15	204	6.1	44	408	283	344	300	1.4	14	5.3
6	45	Female	7.3	4.3	1.43	0.81	15	14	196	5.4	15	312	169	70	181	0.9	11	3.3
7	41	Female	6.7	4.0	1.48	0.74	9	6	178	4.5	7	273	135	98	160	1.0	15	3.1
8	41	Female	6.7	3.7	1.23	0.82	16	12	180	6.3	13	269	214	56	218	1.1	14	4.4
9	48	Female	6.7	3.8	1.31	0.73	26	5	119	6.3	5	251	176	69	190	0.9	14	3.6
10	51	Female	7.1	4.3	1.54	0.99	12	9	220	6.1	7	305	305	114	203	1.2	15	4.7

Table 5
D^2/f and 10 log(D^2/f) of special medical examination

No.	D^2/f	10 log(D^2/f)	No.	D^2/f	10 log(D^2/f)	No.	D^2/f	10 log(D^2/f)
1	1.894	+4.933	21	3.114	+4.933	41	1.594	+2.025
2	1.320	+1.207	22	3.280	+5.159	42	1.872	+2.723
3	18.591	+12.683	23	3.797	+5.794	43	2.366	+3.741
4	2.160	+3.345	24	5.305	+7.670	44	0.959	−0.181
5	5.648	+7.519	25	1.537	+1.867	45	4.609	+8.813
6	0.630	−2.006	26	4.611	+6.638	46	4.274	+6.308
7	1.075	+0.314	27	2.856	+4.557	47	5.847	+7.670
8	0.846	−0.725	28	1.706	+2.320	48	1.653	+2.184
9	2.956	+4.70	29	4.848	+6.855	49	7.325	+8.648
10	0.815	−0.888	30	5.498	+7.402	50	2.076	+3.172

alcohol the previous night or meals taken before the test, and there was no necessity of considering liver disease, the examinee was diagnosed as normal.

3. If there were data out of the limit even in the second test, the examinee takes a precise examination later.
4. Among the people who took precise examinations, those who were judged to have certain signs that indicated the potential for a specific liver ailment, and thus to whom a warning should be given for future health care, were diagnosed as slightly abnormal (SAB). On the other hand, those who did not need to be warned were diagnosed as WNL.
5. People who evidenced a fatty liver due to obesity, excessive intake of sugar, a fatty liver due to diabetes, slight liver disease due to a gallstone or other abnormal organs, or to be an asymptomatic carrier of hepatitis B were judged to be in the category of slight liver disease (SLI).
6. Those with a chronic liver disease due to hepatitis B or C virus, cirrhosis, or alcoholic liver disease were diagnosed as having a liver disease (LVD).

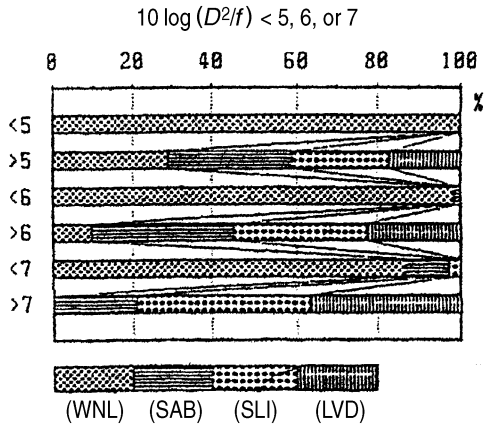

Figure 3
Final diagnosis distribution change under different thresholds

Data for Examinees

Table 4 shows a part of data for examinees who took a special medical examination. Normalizing these data using equation (1), we calculated D^2 by using the inverse matrix. Table 5 shows a part of D^2/f and 10 log(D^2/f)

4. Analysis and Results

Now we detail the analysis of the D values calculated. We studied what value should be selected as a

Table 6
Threshold for 10 log(D^2/f) and final diagnosis (estimation and 95% confidence interval)

Threshold	95% Confidence Interval	10 log(D^2/f) Less Than Threshold				10 log(D^2/f) Greater Than Threshold			
		WNL	SAB	SLI	LVD	WNL	SAB	SLI	LVD
5	Lower limit	98.19	$-\infty$	$-\infty$	$-\infty$	19.31	20.07	15.72	10.64
	Estimation	100.00	0.00	0.00	0.00	29.27	30.27	24.39	17.07
	Upper limit	$+\infty$	1.81	1.81	1.81	41.71	42.87	35.86	26.25
6	Lower limit	97.43	0.94	$-\infty$	$-\infty$	6.05	24.86	22.27	14.93
	Estimation	98.44	1.56	0.00	0.00	9.68	35.48	32.26	22.58
	Upper limit	99.05	2.57	1.74	1.74	0.1512	47.76	44.19	32.65
7	Lower limit	79.54	6.48	1.57	$-\infty$	$-\infty$	13.58	30.00	25.58
	Estimation	86.84	10.53	2.63	0.00	0.00	21.05	42.11	36.84
	Upper limit	91.65	16.05	4.38	1.77	1.77	31.16	55.25	49.75

threshold for 10 log(D^2/f) in order to obtain a good result.

Figure 3 and Table 6 go along with the following analyses.

1. *Threshold* = 5. If 5 is selected as the threshold, all cases with a value less than or equal to 5 result in WNL with no abnormal cases. However, since 29.27% of the cases with a value greater than or equal to 5 (or from 19.31 to 41.71% in a 95% confidence interval) belong to WNL, the indication is that unnecessary retesting or a precise examination was conducted.
2. *Threshold* = 6. Compared to the case of threshold = 5, the proportion of WNL among cases diagnosed as abnormal decreased to 9.68% (6.05 to 15.12%). Yet among cases with a value below 6 diagnosed as normal, 1.56% (0.94 to 2.57%) of abnormal cases were included. However, these abnormal cases were not considered to be a serious problem because all of them had only a slight abnormality.
3. *Threshold* = 7. If the threshold is set to 7, the possibility of mistakenly diagnosing a normal person for an abnormal one is eliminated. However, since 10.53% of slight abnormality and 2.63% (1.57 to 4.38%) of slight liver disease are mingled, this threshold cannot be regarded as correct.

Biochemical Data within and beyond Target Limit and Final Diagnosis

If the judgment based on biochemical data (within the target limit or not) affects our final diagnosis of slight liver disease, it can be seen from Table 7 and Figure 4 that 6.67% (3.44 to 12.54%) are included in the cases judged as normal. In addition, the fact that 65% (48.05 to 78.85%) of the cases diagnosed as abnormal are WNL, in fact, demonstrates that the

Table 7
Biochemical data and final diagnosis

95% Confidence Interval	Within Target Limits				Beyond Target Limits			
	WNL	SAB	SLI	LVD	WNL	SAB	SLI	LVD
Lower limit	87.46	$-\infty$	3.44	$-\infty$	48.05	8.08	5.94	4.56
Estimation	93.33	0.00	6.67	0.00	65.00	15.00	11.25	8.75
Upper limit	96.56	2.09	12.54	2.09	78.85	26.16	20.29	16.14

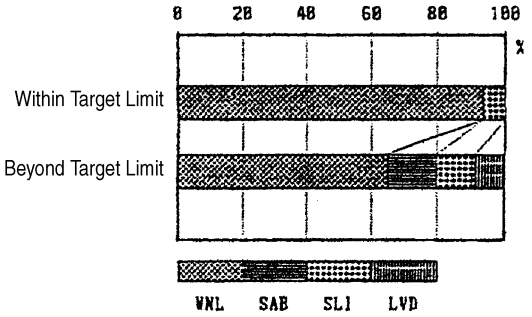

Figure 4
Within and beyond target limits and final diagnosis

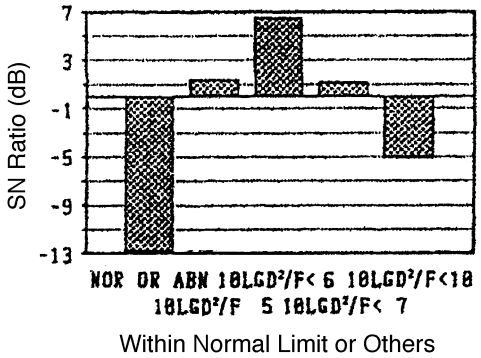

Figure 6
Within normal limits or others

method based on biochemical data cannot be used for diagnosis. Since those doctors who can judge cases with several data outside a target limit as normal are, in most cases, quite experienced, it is extremely difficult to make a similar judgment.

This is one of the reasons that we attempted to use a somewhat bothersome calculation such as D^2. It is quite easy to imagine the feeling of persons who are asked to take a retest or precise examination in a medical checkup. It is possible that some of them could have high blood pressure or a stomach ulcer due to mental stress or have an attack of angina. Thus, this study was not necessarily a waste of time, labor, and budget.

Next we studied how a limited number of examination items influences the result. Figure 5 shows the result obtained by only the five items of GOT, GPT, γ-GTP, TCh, and TG prescribed by the Industrial Safety and Health Law. Of the cases diagnosed as normal, 5.55% (2.71 to 10.94%) and 2.68% (1.23 to 5.35%) were actually SAB and SLI, respectively. Additionally, 55.93% (37.57 to 72.80 percent) of the cases judged as abnormal were categorized as WNL. Considering these results, we cannot say that constraint in the number of examination items leads to better judgment.

Determination of Data within and beyond the Target Limit

Since we were dealing with a medical checkup, we studied the relationships among the final diagnosis (of whether data were within or beyond a normal

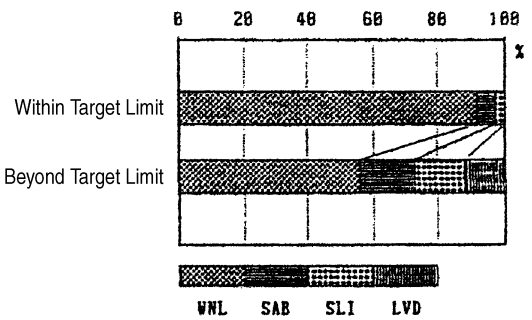

Figure 5
Within and beyond target limits using five test items and final diagnosis

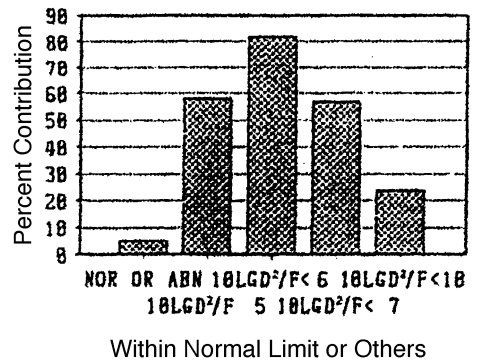

Figure 7
Percent contribution for biochemical data limits and 10 $\log(D^2/f)$ threshold

Table 8
Data within or beyond target limit versus threshold for 10 log(D^2/f) versus final diagnosis (estimation and 95% confidence interval)

Final Diagnosis	95% Confidence Interval	Within or Beyond Target Limits		10 log(D^2/f) = 5		10 log(D^2/f) = 6		10 log(D^2/f) = 7		10 log(D^2/f) = 10	
		Within Target Limit	Beyond Target Limit	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold
Within normal limit	Lower limit	87.28	47.65	98.32	20.47	97.87	7.26	80.32	$-\infty$	63.54	$-\infty$
	Estimation	93.33	65.00	1.00	29.27	98.44	9.68	86.84	0.00	76.74	0.00
	Upper limit	96.62	79.12	$+\infty$	39.95	98.85	12.79	91.43	1.69	88.20	1.97
Beyond normal limit	Lower limit	3.38	20.88	$-\infty$	60.05	1.15	87.21	8.67	98.31	13.80	98.03
	Estimation	6.67	35.00	0.00	70.73	1.56	90.32	13.16	100.00	23.26	100.00
	Upper limit	12.72	52.35	1.68	79.53	2.13	92.74	19.68	$+\infty$	36.46	$+\infty$
Contribution		5.0331		57.87		81.66		56.90		23.82	
SN ratio		-12.76		1.38		6.49		1.21		-5.05	

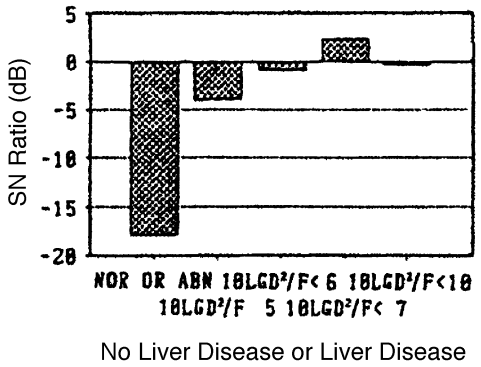

Figure 8
SN ratio for liver disease, biochemical data limits, and 10 $\log(D^2/f)$ threshold

limit), biochemical data within or beyond a target limit, and a threshold of 10 $\log(D^2/f)$, as shown in Figures 6 and 7 and Table 8.

1. *Biochemical data within or beyond the target limit.* While 93.33% of the cases within the target limit reasonably belong to WNL, 6.67% (3.38 to 12.72%) of them are categorized as abnormal. On the other hand, 65% (47.65 to 79.12%) of the cases judged as abnormal are WNL. The percent contribution computed results in 5.03, and the SN ratio was 12.76. Therefore, we do not regard this as a recommendable method.

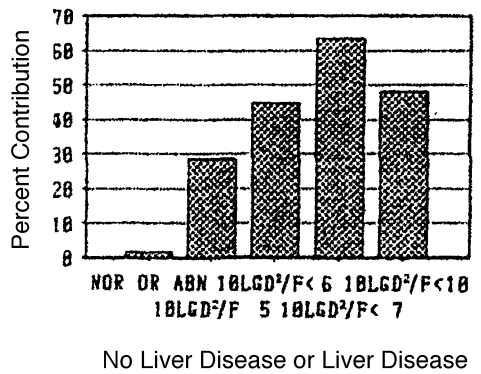

Figure 9
Percent contribution for liver disease, biochemical data limits, and 10(D^2/f) threshold

2. *Threshold for* 10 $\log(D^2/f)$ *set to* 5. All of the cases are WNL (i.e., no abnormality is mingled with the cases with a value less than or equal to the threshold). In contrast, the fact that 29.27% (20.47 to 39.95%) of the cases with a value greater than the threshold are diagnosed as WNL was regarded as a problem. A higher percent contribution and SN ratio, 57.87% and 1.38, respectively, were obtained.
3. *Threshold for* 10 $\log(D^2/f)$ *set to* 6. While 98.44% of the cases with a value below the threshold are WNL, 1.56% (1.15 to 2.13%) of slight abnormality cases were mixed with them. On the other hand, only 9.68% of the cases with a value greater than or equal to the threshold were WNL, which may be regarded as tolerable. The resulting contribution of 81.67% and SN ratio of 6.49 are a good value. In actuality, we decided that we should use this method.
4. *Threshold for* 10 $\log(D^2/f)$ *set to* 7. Although no normal case was mingled with the cases with a value greater than the threshold, 23.26% (13.80 to 36.46%) of the abnormality was mixed with the cases with a value less than the threshold. We consider this impractical. In fact, the contribution and SN ratio are 23.82% and −5.05, respectively, both of which are poor.

Relationship with Liver Disease

Despite slight digression from our main focus, we looked at the relationship among the final diagnosis of liver disease, biochemical data within or beyond a target limit, and threshold for 10 $\log(D^2/f)$. Because we dealt with an issue with a different objective, the resulting contribution and SN ratio are somewhat poor (Figures 8 and 9 and Table 9).

1. *Biochemical data within or beyond the target limit.* While 6.67% (2.66 to 15.76%) of the cases categorized as normal based on biochemical data are mixed up with the cases of liver disease, 80% (60.44 to 91.62%) of the cases categorized as abnormal do not have a liver disease based on the final diagnosis. The contribution is 1.61%, and the SN ratio turns out to be 17.87. Both values indicate a poor classification.

Table 9
Data within or beyond target limit versus threshold for 10 log(D^2/f) versus final diagnosis of liver disease

Final Diagnosis	95% Confidence Interval	Within or Beyond Target Limits		10 log(D^2/f) = 5		10 log(D^2/f) = 6		10 log(D^2/f) = 10		10 log(D^2/f) = 10	
		Within Target Limit	Beyond Target Limit	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold	Less Than Threshold	Greater Than Threshold
With liver disease	Lower limit	84.24	60.44	97.65	38.44	97.86	28.61	95.37	12.92	83.01	−∞
	Estimation	93.33	80.00	100.00	58.54	100.00	45.16	97.37	21.05	90.70	0.00
	Upper limit	97.34	91.62	+∞	76.15	+∞	62.85	98.52	32.40	95.11	2.08
With no liver disease	Lower limit	2.66	8.38	−∞	23.85	−∞	37.15	1.48	67.60	4.89	97.92
	Estimation	6.67	20.00	0.00	41.46	0.00	54.84	2.63	78.95	9.30	100.00
	Upper limit	15.76	39.56	2.35	61.56	2.14	71.39	4.63	87.08	16.99	+∞
Contribution		0.0161		28.71		45.10		63.42		48.02	
SN ratio		−17.86		−3.95		−8.89		2.39		−0.34	

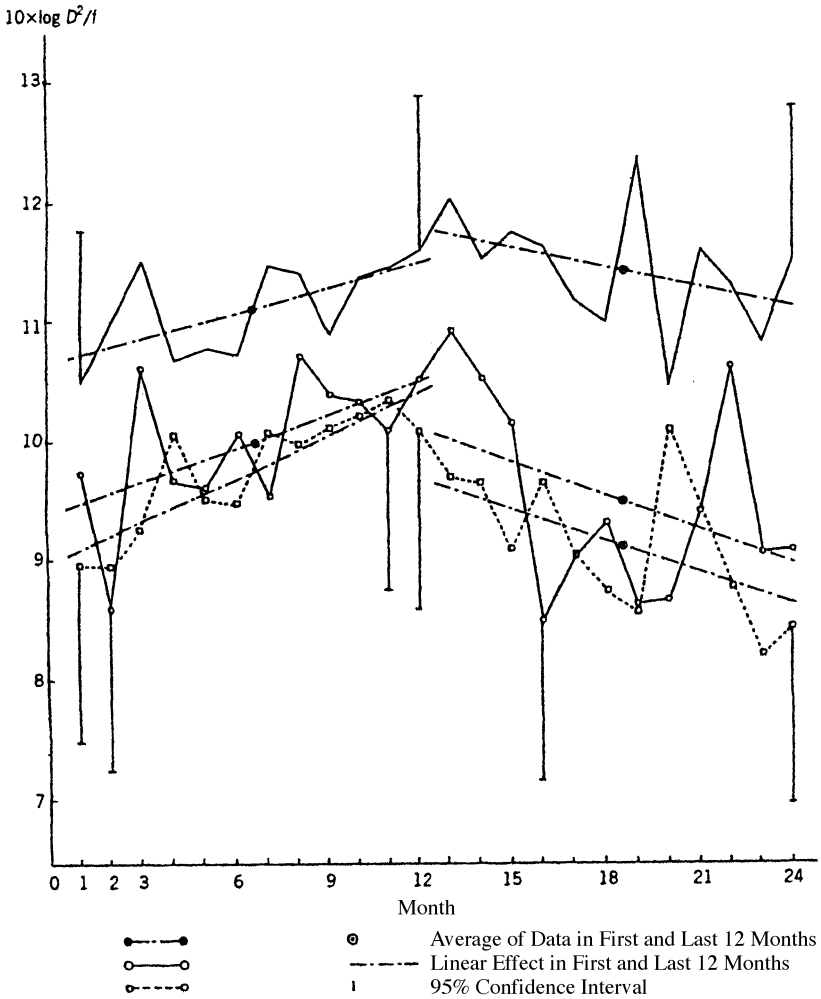

Figure 10
Transition of chronic active hepatitis

2. *Threshold for* 10 $\log(D^2/f)$ *set to* 5. No case of liver disease was mingled with cases with a value below threshold. In contrast, 58.54% (38.44 to 76.15%) of the cases with a value greater than the threshold actually have no liver disease. The contribution and SN ratio result in 28.71% and −3.96, respectively, both of which are not good enough.
3. *Threshold for* 10 $\log(D^2/f)$ *set to* 6. No case of liver disease was mixed up with cases with a value below threshold. However, 45.16% (28.61 to 62.85%) of the cases with a value greater than threshold had no liver disease. The resulting contribution and SN ratio were computed as 45.10% and −0.89, respectively.
4. *Threshold for* 10 $\log(D^2/f)$ *set to* 7. Of the cases with a value below threshold, 2.63% (1.48 to 4.63%) had liver disease. In addition, 21.05% (12.92 to 32.40%) of the cases with a value above threshold had no liver disease. We ob-

tained the best contribution of 63.42% and SN ratio of 2.39 among all for this threshold. Since our main objective was judgment of health conditions rather than diagnosis of whether or not there was liver disease, such moderate results seem reasonable enough.

5. *Threshold for* 10 $\log(D^2/f)$ *set to* 10. While 9.30% (4.89 to 16.99%) of the cases with a value below threshold had liver disease, the cases with a value greater than threshold had no case of liver disease. The resulting contribution and SN ratio were 48.02% and −0.34.

As can be seen, when we diagnose liver disease, the correctness of the judgment deteriorates slightly. To improve our judgment, we need to use a different concept in building when using the same database.

5. Another Application of Mahalanobis Distance

To date, we have attempted to apply the D^2 value to the medical field by judging whether or not certain data belong to a certain group, as Mahalanobis originally intended in his research. Here, our discussion is based on the idea that D^2 is regarded as the distance from the center of gravity of the group forming a database.

Let's take the case of people with a certain disease: for example, patients with chronic active hepatitis. Defining their distance from normal people as D^2, we can see that their distance from the gravity center of the normal group is equivalent to how serious their disease is. Since a decreasing D^2 value indicates proximity to a normal group when we keep track of the condition of patients suffering chronic active hepatitis, their health is expected to improve. On the contrary, when D^2 increases gradually, their disease is considered to be aggravated because the distance from the normal group changes by increments.

Figure 10 shows the 18-item data measured for 24 months for 21 patients with chronic active hepatitis: eight for group 1, seven for group 2, and six for group 3. Each of the three groups received a different treatment in the last 12 months.

Looking at D^2 and the linear effect line for each group, we can see that whereas health degraded during the first 12 months, for the last 12 months, the D^2 value decreased, due to the active therapy (i.e., health was improved). This implies that a time-based analysis of D^2 enables us to judge the transition of data. That is, in the medical science field, we can make a judgment on medical efficacy.

Although the Mahalanobis distance was complicated and impractical as developed by Mahalanobis, today, when computer technology is widely used, even a microcomputer can easily calculate the distance. As a result, it is regarded as one of the most broadly applicable techniques for multivariate analysis.

References

Tatsuji Kanetaka, 1987. An application of Mahalanobis distance: an example of chronic liver disease during active period. *Standardization and Quality Control,* Vol. 40, No. 11, pp. 46–54.

Tatsuji Kanetaka, 1997. An application of Mahalanobis distance to diagnosis of medical examination. *Quality Engineering,* Vol. 5, No. 2, pp. 35–44.

This case study is contributed by Tatsuji Kanetaka.