ELSEVIER

Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE @ DIRECT®

Information Sciences 159 (2004) 155–165

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

# Comparison of clustering methods for clinical databases

Shoji Hirano *, Xiaoguang Sun, Shusaku Tsumoto

*Department of Medical Informatics, School of Medicine, Shimane Medical University, 89-1 Enya-cho, Izumo, Shimane 693-8501, Japan*

Received 27 September 2002; received in revised form 30 December 2002; accepted 10 March 2003

**Abstract**

Clustering methods can be viewed as unsupervised learning from a given dataset. Even without domain knowledge or labels such as the names of diseases given by medical experts, these methods generate partition of datasets. In some cases, these new generated classes lead to discovery of a new disease or new concept. This paper discusses how clustering methods work on a practical medical data set. For comparison, the following four clustering methods were selected and evaluated on a dataset on meningoencephalitis: single- and complete-linkage agglomerative hierarchical clustering, Ward's method and rough clustering. For comparison, a single similarity measure, a linear combination of the Mahalanobis distance between numerical attributes and the Hamming distance between nominal attributes was given to each clustering method. Usefulness of the clustering methods was evaluated from the following viewpoints: (1) the quality of generated clusters, (2) correspondence between the attributes used to generate the high-quality clusters and clinical knowledge. The experimental results showed that the best clusters were obtained using Ward's method where the clinically reasonable attributes were selected, which also suggested that this similarity measure would be applicable to the medical data sets.

[*] Corresponding author.
*E-mail addresses:* hirano@ieee.org (S. Hirano), tsumoto@computer.org (S. Tsumoto).

0020-0255/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2003.03.011

## 1. Introduction

A database on clinical laboratory examinations (clinical labo-exam database) is a complex, high-dimensional database that contains multiple types of attributes. Generally, a labo-exam database contains two types of attributes: numerical attributes and categorical attributes. The latter can be further classified into ordered attributes and nominal attributes. For example, values such as 1.2 mg, 4.9 IU/*l* are of numerical attributes in which an origin exists and thus we can define distance of two values using subtraction of distances from the origin. Values such as *mild*, *moderate*, *severe* are of ratio attributes in which we can define their order but cannot quantify their distances. Values such as *positive*(+), *negative*(−) are of nominal attributes in which we can discriminate them on the description but cannot define their order and distance. This means that analysis of such a database requires appropriate definitions of similarity measures to characterize differences between objects. However, there has not been discussed a standard similarity measure which incorporates characteristics of such clinical databases. The aim of this work is to investigate how the conventional similarity measures work on the practical medical databases.

We examined the following widely used similarity measures, (1) Mahalanobis distance for numerical attributes, (2) Hamming distance for nominal attributes, (3) linear combination of the Mahalanobis distance and Hamming distance for mixed attributes, with four types of clustering methods. The list of clustering methods included (1) single-linkage agglomerative hierarchical clustering (AHC) [1] (2) complete-linkage AHC, (3) Ward's method. Besides, we included (4) rough clustering [2] to the list because it can deal with relative similarity. Through the clustering experiments, we compared the differences between the generated clusters and the diagnostic classes. Usefulness of the similarity measure was evaluated from the following aspects (1) quality of the generated clusters, (2) clinical reasonability of the attributes used to generate the high-quality clusters. Results on the meningoencephalitis data set containing 140 objects and 32 attributes showed that the best clusters were obtained using Ward's method with clinically reasonable combination of attributes.

This paper is organized as follows. Section 2 gives brief explanation of the meningoencephalitis data set. Section 3 shows the similarity measures and Section 4 shows the clustering methods used in the experiments. Section 5 shows experimental results on the meningoencephalitis data set. Finally, Section 6 concludes the technical results.

## 2. Database

In the subsequent sections, the clustering methods were applied to the meningoencephalitis diagnosis data set [4] containing 140 objects. Table 1

Table 1
Structure of the meningoencephalitis data set

| Obj # | Age | Sex | Fever | Focal | ... | DIAG |
|---|---|---|---|---|---|---|
| 1 | 10 | M | 10 | + | ... | BACTERIA |
| 2 | 12 | M | 5 | − | ... | BACTERIA |
| ⋮ | | | | | ⋮ | |
| 140 | 23 | F | 10 | + | ... | VIRUS |

shows structure of the data set. Each object contains 33 attributes including 1 class attribute and 32 other attributes. The class attributes and 12 of 32 attributes are nominal and others are numerical. The class attribute DIAG represents type of the meningitis, that is, bacteria or virus. In this data set, two of 32 attributes, polynuclear cell count (Cell_Poly) and mononuclear cell count (Cell_Mono) are important in class assignment. If polynuclear cells are dominant, a patient is diagnosed as bacterial meningitis. Otherwise, if mononuclear cells are dominant, a patient is diagnosed as virus meningitis.

## 3. Similarity measures

Let $U = \{x_1, x_2, \ldots, x_N\}$ be the set of objects where $N$ denotes total number of objects. Also let us assume that each object has $p = p_c + p_d$ attributes where $p_c$ is the number of numerical attributes and $p_d$ is that of nominal attributes. Then we denote an object $x_i = \{x_i^1, x_i^2, \ldots, x_i^p\}$, where $x_i^j$ denotes the $j$th attribute value of object $x_i$.

### 3.1. Similarity for numerical attributes

In order to measure similarity for numerical attributes, we adopt the Mahalanobis distance:

$$d_{\mathrm{M}}(x_i, x_j) = \left\{(x_i - x_j)^{\mathrm{T}} \Sigma^{-1} (x_i - x_j)\right\}^{1/2},$$

where $\Sigma$ denotes the variance–covariance matrix given by examples. If all of the attributes are independent, and if all of the attribute values are standardized, the Mahalanobis distance of objects exactly matches the Euclidean distance given below.

$$d_{\mathrm{E}}(x_i, x_j) = \{(x_i^1 - x_j^1)^2 + (x_i^2 - x_j^2)^2 + \cdots + (x_i^{p_c} - x_j^{p_c})\}^{1/2}.$$

### 3.2. Similarity for categorical attributes

In order to measure similarity for categorical attributes, we adopt the Hamming distance that counts the number of attributes for which two objects have different attribute values.

$$d_{\mathrm{H}}(x_i, x_j) = \frac{1}{p_d} \sum_{k=1}^{p_d} \delta(x_i^k, x_j^k),$$

$$\delta(x_i^k, x_j^k) = \begin{cases} 1 & \text{if } x_i^k = x_j^k \\ 0 & \text{otherwise.} \end{cases}$$

### 3.3. Similarity for mixture attributes

If objects have both numerical and categorical attributes, their similarity is calculated as a weighted sum of the Mahalanobis distance $d_{\mathrm{M}}(x_i, x_j)$ of numerical attributes and the Hamming distance $d_{\mathrm{H}}(x_i, x_j)$ of nominal attributes as follows:

$$d(x_i, x_j) = \frac{p_c}{p} d_{\mathrm{M}}(x_i, x_j) + \frac{p_d}{p} d_{\mathrm{H}}(x_i, x_j).$$

## 4. Clustering methods

### 4.1. Agglomerative hierarchical clustering

Hierarchical clustering (HC) has been widely applied to cluster analysis since it can visualize hierarchical structure of clusters with a dendrogram. Basically, there are two types of algorithms for HC: agglomerative HC (AHC) and divisive HC (DHC). AHC initially assigns an independent cluster to each object. Then it seeks the most similar pair of clusters and merges it into one cluster. This process is repeated until all of the initial clusters are merged into single cluster. DHC is an inverse procedure of AHC. It starts from single cluster and finally divides it into object number of clusters. Practically, in both methods, merge or spilt is terminated when the step of similarity to merge/split the next clusters is remarkably large.

AHC has several options in determining the strategy of merging clusters. Some of them are listed below.

#### 4.1.1. Single linkage

One way to select a cluster is to take the intergroup dissimilarity to be that of the closest pair:

$$d_{\mathrm{SL}}(G, H) = \min_{x_i \in G, x_{i'} \in H} d(x_i, x_{i'}),$$

where $G$ and $H$ are clusters to be merged in the next step. The clustering based on this distance is called single linkage agglomerative clustering, also called nearest-neighbor technique.

### *4.1.2. Complete linkage*

The other way to select a cluster is to take the intergroup dissimilarity to be that of the furthest pair:

$$d_{\mathrm{CL}}(G, H) = \max_{i \in G, i' \in H} d_{ii'},$$

where $G$ and $H$ are clusters to be merged in the next step. The clustering based on this distance is called complete linkage agglomerative clustering, also called furthest-neighbor technique.

### *4.1.3. Ward's method*

The Ward's method is based on the sum of squares within a cluster. Let $x_{li}^k$ and $n_l$ denote the value of $k$th attributes of $i$th examples in a cluster $l$ and the number of examples in $l$. Then, the sum of square of cluster $l$, $S_l$ is defined as:

$$S_l = \sum_{i=1}^{n_l} \sum_{k=1}^{p} \left( x_{li}^k - \bar{x}_{l\cdot}^k \right)^2,$$

where $\bar{x}_{l\cdot}^k$ is defined as:

$$\bar{x}_{l\cdot}^k = \frac{1}{n_l} \sum_{i=1}^{n_l} x_{li}^k.$$

Let us consider the integration of cluster $l$ and cluster $m$ into cluster $lm$. The sum of square of cluster $m$, $S_m$ is defined as:

$$S_m = \sum_{i=1}^{n_m} \sum_{k=1}^{p} \left( x_{mi}^k - \bar{x}_{m\cdot}^k \right)^2.$$

Then, the sum of square of cluster $lm$ is obtained as:

$$S_{lm} = S_l + S_m + \Delta S_{lm},$$

where $\Delta S_{lm}$ is given as:

$$\Delta S_{lm} = \frac{n_l n_m}{n_l + n_m} \sum_{k=1}^{p} \left( \bar{x}_{l\cdot}^k - \bar{x}_{m\cdot}^k \right)^2.$$

The two clusters will be integrated when $\Delta S_{lm}$ is minimum with respect to all the clusters.

### *4.2. Rough-sets based clustering*

Generally, if similarity of objects is represented only as a relative similarity, it is not an easy task to construct interpretable clusters because some of important measures such as inter- and intra-cluster variances are hard to be defined. The rough-set based clustering method is a clustering method that clusters objects according to the indiscernibility of objects. It represents denseness of objects according to the *indiscernibility degree*, and produces interpretable clusters even for the objects mentioned above.

The clustering method lies its basis on the indiscernibility of objects, which forms basic property of knowledge in rough sets [3]. Let us first introduce some fundamental definitions of rough sets related to our work. Let $U \neq \phi$ be a universe of discourse and $X$ be a subset of $U$. An equivalence relation, $R$, classifies $U$ into a set of subsets $U/R = \{X_1, X_2, \ldots, X_m\}$ in which following conditions are satisfied:

(1) $X_i \subseteq U, X_i \neq \phi$ for any $i$,
(2) $X_i \cap X_j = \phi$ for any $i,j$,
(3) $\cup_{i=1,2,\ldots,n} X_i = U$.

Any subset $X_i$, called a category, represents an equivalence class of $R$. A category in $R$ containing an object $x \in U$ is denoted by $[x]_R$. For a family of equivalence relations $\mathbf{P} \subseteq \mathbf{R}$, an indiscernibility relation over $\mathbf{P}$ is denoted by $\mathrm{IND}(\mathbf{P})$ and defined as follows

$$\mathrm{IND}(\mathbf{P}) = \bigcap_{R \in \mathbf{P}} \mathrm{IND}(R).$$

The clustering method consists of two steps: (1) assignment of initial equivalence relations and (2) iterative refinement of initial equivalence relations. In the first step, we assign an initial equivalence relation to every object. An initial equivalence relation classifies the objects into two sets: one is a set of objects similar to the corresponding objects and another is a set of dissimilar objects. Let $U = \{x_1, x_2, \ldots, x_n\}$ be the entire set of $n$ objects. An initial equivalence relation $R_i$ for object $x_i$ is defined as

$$R_i = \{\{P_i\}, \{U - P_i\}\},$$
$$P_i = \{x_j | s(x_i, x_j) \geqslant S_i\}, \quad \forall x_j \in U,$$

where $P_i$ denotes a set of objects similar to $x_i$. Namely, $P_i$ is a set of objects whose similarity to $x_i$, $s$, is larger than a threshold value $S_i$. The threshold value $S_i$ is determined automatically at a place where $s$ largely decreases. A set of indiscernible objects obtained using all sets of equivalence relations corresponds to a cluster. In other words, a cluster corresponds to a category $X_i$ of $U/\mathrm{IND}(\mathbf{R})$.

In the second step, we refine the initial equivalence relations according to their global relationships. First, we define an indiscernibility degree, $\gamma$, which represents how many equivalence relations commonly regards two objects as indiscernible objects, as follows:

$$\gamma(x_i,x_j) = \frac{1}{|U|}\sum_{k=1}^{|U|}\delta_k(x_i,x_j),$$

$$\delta_k(x_i,x_j) = \begin{cases} 1 & \text{if } [x_k]_{R_k} \cap ([x_i]_{R_k} \cap [x_j]_{R_k}) \neq \phi \\ 0 & \text{otherwise.} \end{cases}$$

Objects with high indiscernibility degree can be interpreted as similar objects. Therefore, they should be classified into the same cluster. Thus we modify an equivalence relation if it has ability to discern objects with high $\gamma$ as follows:

$$R'_i = \{\{P'_i\},\{U - P'_i\}\}$$

$$P'_i = \{x_j|\gamma(x_i,x_j) \geqslant T_h\}, \quad \forall x_j \in U.$$

This prevents generation of small clusters formed due to the too fine classification knowledge. $T_h$ is a threshold value that determines indiscernibility of objects. Therefore, we associate $T_h$ with roughness of knowledge and perform iterative refinement of equivalence relations by constantly decreasing $T_h$. Consequently, coarsely classified set of sequences are obtained as $U/\mathrm{IND}(\mathbf{R}')$.

## 5. Experimental results

We performed clustering experiments on the meningoencephalitis data set using the four types of clustering methods and similarity measures described in Sections 3 and 4. The purpose of this experiment was to examine whether or not the similarity measures can be used to generate good clusters in practical medical datasets. Thus we clustered the data and evaluated accuracy and coverage between the generated clusters and diagnostic classes. Besides, we evaluated whether the clusters that have high accuracy and coverage were generated with respect to clinically meaningful attributes.

In medical data sets, statistical natures of attributes are usually different. For example, some attributes may follow normal distribution and others may follow other parametric distributions. We here assumed that an attribute whose variance is very large follows log-normal distribution, and apply logarithmic transformation to such an attribute. Based on the observation, we selected the threshold of variance as 1000 and applied transformation to the following attributes: WBC, CSF_CELL, Cell_Poly, and Cell_Mono.

Clustering experiments were performed as follows.

1. Select an arbitrary set of attributes.
2. Cluster the data according to the selected attributes. The similarity measures and clustering methods in Section 4 are used for clustering.
3. Examine importance of the selected attributes. When the clustering result well represents classification of the target class $C$, the attributes are considered to be important factors. We used following measures:

$$\text{Accuracy (positive predictivity)} : \mu(C) = \frac{|C \cap D|}{|C|}$$

$$\text{Coverage (sensitivity)} : \nu(C) = \frac{|C \cap D|}{|D|}$$

$$\text{Importance} : I(C) = \min(\mu, \nu)$$

   where $C$ denotes generated clusters, $D$ denotes diagnostic classes, $\mu$, and $\nu$ denote classification accuracy and coverage of $C$ as to $D$, respectively.
4. Iterate steps 1–3 employing various sets of attributes. Finally, important sets of attributes are obtained by referring to $I(C)$.

We examined all possible combinations of 1–3 attributes (${}_3C_{32} + {}_2C_{32} + {}_1C_{32} = 5488$ pairs).

Table 2 shows the result obtained using single-linkage AHC. This table shows the top 10 sets of attributes that have large importance values $I(C)$. The notion '*' represents an arbitrary attribute. For example, {*, *, CSF_PRO} represents a set of three attributes in which two of them are arbitrary but CSF_PRO is always included. This notion was introduced because some sets of attributes produced the clusters that have exactly the same importance, accuracy and coverage values, e.g. {ONSET, KERNIG, CSF_PRO} and

Table 2
Top 10 important combinations of attributes (single-linkage AHC)

| Rank | $I(C)$ | $\mu(C)$ | $\nu(C)$ | Attrib 1 | Attrib 2 | Attrib 3 | # Comb |
|---|---|---|---|---|---|---|---|
| 1 | 0.813 | 0.814 | 0.813 | * | * | CSF_PRO | 66 |
| 2 | 0.807 | 0.807 | 0.861 | * | CSF_CELL | | 12 |
| 3 | 0.779 | 0.779 | 0.864 | Cell_Poly | Cell_Mono | CSF_PRO | 1 |
| 4 | 0.771 | 0.771 | 0.780 | STIFF | ESR | Cell_Mono | 1 |
| 5 | 0.764 | 0.764 | 0.900 | CSF_CELL | Cell_Poly | Cell_Mono | 1 |
| 6 | 0.757 | 0.757 | 0.879 | * | Cell_Mono | CSF_PRO | 12 |
| 7 | 0.750 | 0.750 | 0.953 | AGE | * | * | 66 |
| 8 | 0.750 | 0.750 | 0.952 | * | CSF_CELL | Cell_Mono | 12 |
| 9 | 0.750 | 0.750 | 0.914 | Cell_Poly | Cell_Mono | | 1 |
| 10 | 0.743 | 0.743 | 0.942 | AGE | * | CSF_CELL | 12 |

{LOC_DAT, FOCAL, CSF_PRO}. The number of such sets of attributes are shown in the column '# Comb'.

The highest importance grade in single-linkage AHC was 0.813, which was assigned to 66 sets of attributes. Most of the other grades were less than 0.8. This means that the method could not find good partition of the classes. Since attributes appeared in the table are clinically reasonable, the reason of low importance grades were due to poor clustering ability of single-linkage AHC in this data set ('chain-effect' was observed in some cases).

Table 3 shows the result obtained using complete-linkage AHC. The highest importance grade was 0.900, which was assigned to 12 sets of attributes. All of the other importance grades in the table were higher than those of single-linkage AHC, and were assigned to more specific types of attributes. At least one of the two important factors, Cell_Poly and Cell_Mono, was successfully included in each of the attribute sets.

Table 4 shows the result obtained using Ward's method. The highest importance grade was 0.950, which was the highest of all in this experiment and was assigned only to 1 set of attributes. Besides, most of the other importance grades in the table were higher than 0.9. Similarly to complete-linkage AHC, Cell_Poly and/or Cell_Mono, was successfully included in each of the attribute sets. It was very interesting that combination of these two important factors yielded the best result.

Table 5 shows the result obtained using rough clustering. The parameters for rough clustering were set as follows: $T_h = 0.5$ and # of refinement = 5. The highest importance grade was 0.886, which was assigned to 1 set of attributes. Range of grades was almost similar to that of complete-linkage AHC and most of the grades were higher than 0.83. Cell_Poly and/or Cell_Mono, were included in each of the attribute sets. Besides, CRP was appeared as an important factor.

Table 3
Top 10 important combinations of attributes (complete-linkage AHC)

| Rank | $I(C)$ | $\mu(C)$ | $v(C)$ | Attrib 1 | Attrib 2 | Attrib 3 | # Comb |
|---|---|---|---|---|---|---|---|
| 1 | 0.900 | 0.900 | 0.926 | AGE | * | Cell_Poly | 12 |
| 2 | 0.885 | 0.886 | 0.885 | * | Cell_Poly | | 12 |
| 3 | 0.864 | 0.864 | 0.926 | Cell_Poly | | | 1 |
| 4 | 0.864 | 0.864 | 0.915 | BT | CSF_CELL | Cell_Mono | 1 |
| 5 | 0.857 | 0.964 | 0.857 | * | Cell_Poly | Cell_Mono | 12 |
| 6 | 0.857 | 0.857 | 0.890 | Cell_Poly | CSF_CEL-L7 | | 1 |
| 7 | 0.857 | 0.857 | 0.860 | STIFF | Cell_Poly | | 1 |
| 8 | 0.857 | 0.857 | 0.860 | * | STIFF | Cell_Poly | 12 |
| 9 | 0.839 | 0.936 | 0.839 | Cell_Poly | Cell_Mono | | 1 |
| 10 | 0.829 | 0.829 | 0.913 | STIFF | CSF_CELL | Cell_Mono | 1 |

Table 4
Top 10 important combinations of attributes (Ward's method)

| Rank | $I(C)$ | $\mu(C)$ | $\nu(C)$ | Attrib 1 | Attrib 2 | Attrib 3 | # Comb |
|---|---|---|---|---|---|---|---|
| 1 | 0.950 | 0.950 | 0.950 | Cell_Poly | Cell_Mono | | 1 |
| 2 | 0.914 | 0.914 | 0.939 | Cell_Poly | | | 1 |
| 3 | 0.907 | 0.907 | 0.936 | * | HEADACHE | Cell_Poly | 12 |
| 4 | 0.907 | 0.907 | 0.936 | FEVER | Cell_Poly | Cell_Mono | 1 |
| 5 | 0.907 | 0.907 | 0.936 | LOC | CRP | Cell_Poly | 1 |
| 6 | 0.907 | 0.907 | 0.923 | Cell_Poly | CSF_GLU | | 1 |
| 7 | 0.906 | 0.907 | 0.906 | * | Cell_Poly | CSF_GLU | 12 |
| 8 | 0.900 | 0.900 | 0.926 | WBC | Cell_Poly | | 1 |
| 9 | 0.893 | 0.893 | 0.931 | * | NAUSEA | Cell_Poly | 12 |
| 10 | 0.893 | 0.893 | 0.923 | * | Cell_Poly | CSF_CELL7 | 12 |

Table 5
Top 10 important combinations of attributes (rough clustering)

| Rank | $I(C)$ | $\mu(C)$ | $\nu(C)$ | Attrib 1 | Attrib 2 | Attrib 3 | # Comb |
|---|---|---|---|---|---|---|---|
| 1 | 0.886 | 0.886 | 0.921 | CRP | Cell_Poly | | 1 |
| 2 | 0.886 | 0.886 | 0.921 | CRP | Cell_Poly | CSF_CELL3 | 1 |
| 3 | 0.871 | 0.871 | 0.881 | LOC | Cell_Poly | | 1 |
| 4 | 0.850 | 0.864 | 0.850 | * | BT | Cell_Poly | 12 |
| 5 | 0.845 | 0.871 | 0.845 | SEIZURE | Cell_Poly | | 1 |
| 6 | 0.843 | 0.843 | 0.846 | * | CRP | CSF_CELL7 | 12 |
| 7 | 0.836 | 0.836 | 0.920 | Cell_Poly | CSF_CELL3 | | 1 |
| 8 | 0.836 | 0.836 | 0.851 | * | CSF_CELL7 | Cell_Mono | 12 |
| 9 | 0.834 | 0.850 | 0.834 | FEVER | Cell_Poly | | 1 |
| 10 | 0.830 | 0.843 | 0.830 | COLD | Cell_Poly | | 1 |

## 6. Conclusions

In this paper, we have investigated how the conventional clustering methods works on the practical medical data set. The similarity measure used was linear combination of the Mahalanobis distance between numerical attributes and the Hamming distance between nominal attributes. The measure was used in conjunction with the four types of clustering algorithms: single- and complete-linkage AHC, Ward's method and rough clustering. The result showed that the clusters were surely generated with respect to two types of clinically reasonable attributes and that the highest accuracy and coverage were obtained with Ward's method. In our experiments, nominal attributes were not appeared in the list of important attributes. This also suggested that the similarity measure successfully decreased similarity of objects when nominal attributes were selected.

This paper was a preliminary study on empirical comparison of clustering methods for medical data sets. Further analysis including different combinations of similarity measures, clustering methods, and data sets will be subject of future research.

## Acknowledgements

This work was supported by the Grant-in-Aid for Scientific Research on Priority Areas(B) (no. 759) "Implementation of Active Mining in the Era of Information Flood" by the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

[1] B.S. Everitt, S. Laudau, M. Leese, Cluster analysis, Fourth Edition, Arnold Publishers, 2001.
[2] S. Hirano, S. Tsumoto, An indiscernibility-based clustering method with iterative refinement of equivalence relations—rough clustering, Journal of Advanced Computational Intelligence and Intelligent Informatics 7 (2) (2003) 169–177.
[3] Z. Pawlak, Rough Sets Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[4] Available from <URL: http://www.shimane_med.ac.jp/med_info/open_data/>.