ELSEVIER

Computer Methods and Programs in Biomedicine 58 (1999) 69–81

Computer Methods and Programs in Biomedicine

Classification of urinary stones by cluster analysis of ionic composition data

R.E. Abdel-Halim [a], R.E. Abdel-Aal [b,*]

[a] *Division of Urology, Department of Surgery, King Khalid University Hospital, P.O. Box 7805, Riyadh 11472, Saudi Arabia*
[b] *Energy Research Laboratory, Research Institute, P.O. Box 1759, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*

Received 10 September 1997; received in revised form 18 May 1998; accepted 27 May 1998

Abstract

The cluster analysis technique is considered for classifying kidney stones based on data for nine chemical analysis parameters. A set of 214 stones is used, which has been previously classified using empirical classification rules into three stone types using the percentage concentrations of the urate, oxalate, and phosphate radicals. We investigate whether cluster analysis utilising data on all parameters leads to different classifications and explore the possibility of other effective classifiers. We also compare the performance of various clustering techniques, distance and similarity measures and data standardisation methods. Results indicate that inclusion of the additional six parameters does not improve the classification accuracy. Best matching with the empirical classification (6% error) is achieved using the average linkage (between groups) clustering method and the squared Eculidean distance measure without data standardisation. Excluding these three main radicals causes a 63% matching error. Cluster analysis results suggest that carbon ions alone provide a single classifier for the three stone types, giving a matching error of $\approx 10\%$ with the empirical classification.

Keywords: Urinary stones; Types of stones; Non-infection stones; Classification; Stone analysis; Elemental microanalysis; Cluster analysis; Ionic composition; Wet chemistry analysis

1. Introduction

Classification of urinary stones according to their chemical composition is an important requirement for proper management of the disease. Quantitative chemical analysis of the urinary stones has been preferred by many investigators, e.g. [1–5]. Of particular importance are microanalytical techniques used for stones below 5 mg in weight [5]. Almost all these investigators apply calculations, based on certain assumptions, to the

* Corresponding author. Tel.: +966 3 8604320; fax: +966 3 8604281; e-mail: radwan@kfupm.edu.sa

0169-2607/99/$ - see front matter
PII S0169-2607(98)00075-3

findings of chemical analysis to arrive at estimates for the composition of compounds. Accordingly, several classifications of urinary stones have been suggested [3,6–9], but it is difficult to compare them since chemical analysis results are presented differently [10].

More recently, a different empirical approach for the classification of urinary stones was proposed [11–14]. The scheme is based only on the percentage composition of detected ions, without the need for such cumbersome calculations which introduce an element of uncertainty [10]. It takes into account both trace and minimum amounts of ions present. Based on the percentage concentration of the urate, oxalate, and phosphate content of the stone, a collection of 214 non-infection (category I) stones was classified into three main types: urate (U), oxalate (X) and phosphate (P), with memberships of 43, 144, and 27 stones, respectively. With this empirical classification, a stone was classified as U (type 1) if the percentage concentration of the urate radicals with respect to the total radical content was $>20\%$. Similarly, a stone was classified as X (type 2) if the percentage concentration of the oxalate radicals was $>40\%$. A stone was classified as P (type 3) if it did not satisfy either of the above conditions and had the percentage concentration of the phosphate radicals $>10\%$. Percentage concentrations of the urate, oxalate, and phosphate radicals in the stone will be referred to as URATE, OXAL and PHOS, respectively. The same three variables were used to further classify each main type into a number of subgroups [12–14]. For example, OXAL was used to separate type 1 stones (U) into four subgroups and type 3 stones (P) into two subgroups. Cluster analysis results reported here consider only classification into the three main stone types.

The empirical classification scheme described above provides a simple method for classifying the stones into three major groups, which requires the measurement of only three parameters. However, it does not utilise much of the chemical analysis data available [11–14] (to be described in Section 2). Moreover, while the empirical classifier was quite successful in separating the urate stones (type 1) from the rest of the population, separation between the oxalates and phosphates (types 2 and 3, respectively) was marginal. This was particularly the case for some stones belonging to one subgroup of the phosphate stones (subgroup 1 with $10 \leq \mathrm{OXAL} < 40$, which represents the mixed type of these stones [11–14]). In this subgroup, boundary cases existed which had OXAL as high as 38.9% and PHOS as low as 10.8%. Inclusion of this subgroup of mixed stones in the phosphate type was in good agreement with the results of the X-ray diffraction technique [15].

Cluster analysis techniques provide means for classifying a given population into groups, based on similarity or closeness measures, through unsupervised learning. It would be interesting to see if cluster analysis based on all the chemical analysis data available yields a similar classification into three clusters as that obtained through the simple empirical scheme above. With agglomerative hierarchical clustering the user can specify the number of clusters required for the solution, at which clustering stops. The objective of this study is to compare the results of cluster analysis with those of the empirical classification described above, with the empirical classification being used as a reference in the absence of a 'gold standard'. In particular, the study attempts to establish whether clustering with all the nine variables available produces significantly different classifications. Obtaining similar classifications with both techniques would indicate that the empirical method is an efficient approach and that the most relevant descriptors for the stone type in the given setting are the percentages of urate, oxalate, and phosphate radicals. Clustering performance by subsets of the data variables is considered in search of other effective classifiers. The study also investigates the effect of using various clustering methods, distance and similarity measures, and standardisation techniques which are applied to the data prior to clustering.

2. The stones data set

The stones data set used for this study consists of 214 urinary stones spontaneously passed or removed in toto (operatively or endoscopically).

Table 1
Averages and standard deviations (in parenthesis) of the nine variables for each of the three stone types and for the total population of 214 stones

Stone type	Variable								
	CT_R	CA	URATE	OXAL	PHOS	C	H	N	MG
(a) Without data standardisation									
Type 1 (U)	513.33	2.85	79.22	5.30	0.80	32.19	2.69	28.73	0.06
	(197.01)	(3.68)	(18.73)	(8.20)	(1.92)	(3.81)	(0.37)	(5.6)	(0.11)
Type 2 (X)	1684.67	23.57	1.00	49.79	3.70	16.78	1.86	1.00	0.08
	(290.53)	(4.11)	(3.25)	(4.62)	(3.57)	(1.83)	(0.26)	(1.09)	(0.13)
Type 3 (P)	1660.70	25.46	0.27	22.59	27.76	10.29	1.63	1.00	0.81
	(292.02)	(3.03)	(0.28)	(14.94)	(12.69)	(3.86)	(0.35)	(0.60)	(0.87)
Overall	1446.28	19.64	16.62	37.42	6.15	19.06	2.00	6.50	0.17
	(542.86)	(9.31)	(32.65)	(19.86)	(9.90)	(7.42)	(0.46)	(11.30)	(0.41)
(b) With data standardised by variable as Z-scores									
Type 1 (U)	−1.72	−1.80	1.92	−1.62	−0.54	1.77	1.50	1.93	−0.27
	(0.36)	(0.40)	(0.57)	(0.41)	(0.19)	(0.51)	(0.80)	(0.49)	(0.28)
Type 2 (X)	0.44	0.42	−0.48	0.62	−0.25	−0.31	−0.30	−0.49	−0.21
	(0.54)	(0.44)	(0.10)	(0.23)	(0.36)	(0.25)	(0.56)	(0.10)	(0.31)
Type 3 (P)	0 39	0.62	−0.50	−0.75	2.18	−1.18	0.79	−0.49	1.57
	(0.54)	(0.32)	(0.01)	(0.75)	(1.28)	(0.52)	(0.77)	(0.05)	(2.14)

For case (b), overall variable averages are all zeros and overall standard deviations are all unity (not shown).

Before samples were taken for the routine methods of quantitative chemical and elemental microanalysis [12–14], the whole stone was subjected to ex-vivo computerised axial tomographic scanning (CT) for the determination of stone density using a modification (unpublished data) based on the work of Drach et. al. [16]. The CT reading determined (CT_R; in Hounsfield units) was one of the nine parameters studied for each stone. The other eight parameters are the percentage concentration of eight ions and radicals: calcium (CA), urate, oxalate, phosphate, carbon (C), hydrogen (H), nitrogen (N) and magnesium (MG). Importance of the percentage concentration of CA, urate, oxalate and phosphate in the classification of urinary stones is well documented in many studies, although data were usually interpreted as compounds and not as ions or radicals [3,6–9]. Recent studies confirmed the reliability of microanalytical determination of C, H and N for the quick identification of stones, particularly those that weigh < 1 mg, which are too small for using wet chemical analysis [11–14]. It was also shown that the percentage concentration of MG differentiates between infection and non-infection stones [11–14]. All samples in this study belong to the non-infection stones category (MG $< 3\%$).

Table 1 lists the average and standard deviation for each of the nine variables. Values are given separately for each of the three main stone types in the empirical classification [12–14], together with overall values for all the stone cases in the data set. These results are shown both without standardisation, Table 1 (a), and with data standardised by variable as Z-scores having overall zero means and unity standard deviations for each variable over the whole data set, Table 1(b). Table 1(a) indicates that the choice for the three main parameters, URATE, OXAL, and PHOS, as basis for the empirical classification of the three stone types is justified. Disparity among type averages for these three variables is greater than that for any other variable. For example, the type averages for URATE have their ratios, relative to type 3 average, as 293.4:3.7:1. This shows that URATE should adequately separate type 1 from both types 2 and 3, but will not be as effective in separating type 2 from type 3. For the variable H,

this ratio is 1.65:1.14:1, indicating that H would not be a good discriminator for the three types combined. Variable C has the better ratio of 3.13:1.63:1 and, therefore, may perform adequately in classifying the three types. Variables such as CT_R, CA, N may discriminate type 1 against types 2 and 3 combined. Among the three main radicals, URATE has the largest average to standard deviation ratio for stones of type U, while for types X and P the ratio is largest for OXAL and PHOS, respectively. This supports the empirical choice of these variables to classify the three stone types, as the three sub-populations adequately cluster with minimum spread about the mean values of the respective variables.

3. Cluster analysis techniques

Cluster analysis is widely used in many areas of applied sciences to place cases of a given population into groups or clusters suggested by a data set of key variables; such that cases in a given cluster tend to be 'similar' or 'close' according to some similarity or distance measure [17] Choice of variables used for clustering is a crucial step which should be carefully considered by the domain expert, because excluding important variables may lead to misleading grouping. When variables have different scales, distance measures will be biased towards reflecting contributions from variables expressed in larger numbers, and therefore, data should be re-scaled or standardised prior to clustering. Commonly used standardisation methods include transformations into Z-scores (zero mean and unity standard deviation) or dividing by the mean, the standard deviation, the range, or the maximum value [18].

In agglomerative hierarchical clustering, each case is considered as a separate cluster initially, and groups are formed by merging cases into bigger and bigger clusters, eventually ending up with a single group if the analysis is allowed to continue. Usually, the user specifies a solution containing a given number of clusters. Cases are grouped based on their 'nearness' according to a given distance or similarity measure [18]. Merging cases/clusters into bigger clusters depends on the clustering method used. Methods differ in their use of the distance measure selected in deciding which cases or groups to be combined at each step of the clustering procedure and the final clustering solution depends on the combination of the clustering method used and distance measure adopted. Commonly used clustering methods fall into three main categories: linkage, centroid, and variance. Linkage methods define nearness of two classes in terms of the distances between member cases of the two clusters. The single linkage (nearest neighbour) method uses the distance between the closest two cases in the two candidate clusters as an indication of their nearness, while the complete linkage (furthest neighbour) method uses the distance between the furthest two points. Other linkage methods include the average linkage (between groups) method which define the distance between two clusters as the average of the distances between all possible pairs of cases made up of one case from each cluster. On the other hand, the average linkage (within groups) method computes a similar average distance but between all pairs of a resulting cluster. At each agglomeration step, the two clusters chosen for merging would be those that lead to the minimum such distance in the resulting bigger cluster.

Both the centroid and variance methods use the notion of the cluster centre, defined as the mean vector of the variables for all cases within the cluster. The centroid method considers the distance between two clusters as the distance between their centres. Here, the centroid of a merged cluster is a weighted combination of the centroids of the two individual clusters, with the weights being proportional to the sizes of the two clusters. The median method is similar to the centroid method, but the centroids are weighted equally to avoid de-emphasising the influence of smaller clusters. The Ward's minimum variance method is an example of the variance or error-sum-of-squares category of clustering methods. The method computes the sum of the squared Euclidean distance between each case and the cluster centre and considers it as a dispersion 'error' which should be minimised. The two clusters to be joined are selected among all possible candidate cluster pairs such that this dispersion variance is minimised in the resulting cluster.

The most commonly used distance measure is the squared Euclidean distance. Between two cases, this is defined as sum of the squared differences between the values of all variables for the two cases. It should be used with the centroid, median, and Ward's clustering methods. The Euclidean distance is the square root of that distance. The cosine distance is the cosine of the two vectors of the variables for the two cases. Similarly, the Pearson correlation distance is the Pearson correlation coefficient between these two vectors. The city block distance is the sum of the absolute values for the differences between the values for each variable in the two vectors. The Chebychev distance is the maximum absolute value of these differences. The Minkowski distance is obtained by raising the absolute difference for each variable to the nth power and taking the nth root of the sum for all the variables.

Different clustering methods produce cluster solutions having different characteristics related to size, shape, and dispersion [19]. For example, average linkage methods tend to form clusters of equal variance while the Ward's method tends to form clusters of equal sizes. Performance of the various methods depends heavily on the nature of the problem; e.g. how well the clusters are separated, how widely each cluster is dispersed, and whether or not they have equal membership. If the clusters are widely separated, almost any clustering method would be able to successfully categorise them, while data representing poorly separated clusters present a more difficult problem. In a series of simulation studies [20] on artificial data sets obtained using pseudo-random numbers, it was found that the best overall performance was obtained with either average linkage methods or the Ward's minimum variance method while the single linkage method almost always gave the poorest overall performance [19].

4. Cluster analysis results

A number of cluster analysis methods were used for the automatic classification of the stones data described in Section 2 using the cluster module of the Windows version of the SPSS package [18]. Performance was evaluated using the empirical classification [12–14] as a reference. The following clustering methods were evaluated: (1) average linkage between eight groups; (2) average linkage within groups; (3) single linkage (nearest neighbour); (4) complete linkage (furthest neighbour); (5) centroid method; (6) median method; (7) Ward's method. Various distance measures were investigated, including: (a) Euclidean; (b) squared Euclidean; (c) cosine; (d) Pearson correlation; (e) Chebychev; (f) block; (g) Minkowski, with $n = 3$. A number of techniques were explored for data standardisation prior to clustering, including: (i) no standardisation; (ii) Z-scores; (iii) range -1 to $+1$; (iv) range 0 to $+1$; (v) maximum magnitude = 1; (vi) S.D. = 1; (vii) mean = 1.

4.1. Effect of the clustering method

Table 2 shows the classification performance of the seven clustering methods listed above on the stones data set with data for all the nine variables used by the clustering procedure (later we investigate the effect of selecting variables that take part in clustering). The default squared Euclidean distance measure was used throughout, with data always standardised as Z-scores. The empirical scheme [12–14] provided the reference classification into the three stone types: U, X, and P, having memberships of 43, 144, and 27 samples, respectively, out of the total number of 214 stones considered. Therefore, the table gives the number of samples in each of the these three types which are classified into each of three corresponding stone types by cluster analysis. Diagonal elements in the 3×3 matrices in the table indicate correctly classified samples, while off-diagonal elements in a row represent the number of misclassifications into the other two types. Shown also is the total percentage classification error for each method. The table indicates that, with the exception of the single linkage (nearest neighbour) method (labelled 3 in the table), all clustering methods give adequate classification results, which closely match those of the reference empirical classification [12–14] with an overall error of ≈ 8–11%. Best performance is obtained with methods la-

Table 2
Classification performance of the various clustering methods, with the empirical classification as a reference

Clustering method		Classification performance				
		Empirical type	Classified as type:			Total classification error (%)
			1	2	3	
1	Average linkage between groups	1	43	0	0	
5	Centroid	2	0	144	0	7.9
7	Ward	3	0	17	10	
6	Median	1	43	0	0	
		2	0	144	0	8.9
		3	0	19	8	
4	Complete linkage (furthest neighbour)	1	43	0	0	
		2	0	144	0	10.3
		3	0	22	5	
2	Average linkage within groups	1	42	1	0	
		2	0	144	0	11.1
		3	0	23	4	
3	Single linkage (nearest neighbour)	1	43	0	0	
		2	141	2	1	79
		3	27	0	0	

Clustering with all nine variables.
Distance measure: squared Euclidean.
Data standardised by variable as Z-scores.

belled 1, 5, and 7, while performance of the single linkage method (3) is quite poor with the error approaching 80%. This agrees with general observations on the performance of the average linkage, Ward's, and single linkage methods given in Section 2 above. Type 1 (U) is classified almost error-free by all methods, including method 3. Type 2 (X) is also correctly classified by all but the single linkage (3) method, where almost all cases are incorrectly classified as type 1. It is clear that this method fails to classify the cases into three reasonably-sized groups, nearly mistaking them all for a single type (type 1). The single linkage method has been known to sacrifice performance in the recovery of compact clusters in return for the ability to detect elongated and irregular clusters [19]. Type 3 (P) classification is the main source of error with all the other clustering methods. For example, 63% of the population of this type is misclassified as type 2 (X) by methods 1, 5, and 7.

4.2. Effect of distance/similarity measures

Table 3 shows the effect of using different distance measures on the performance of the average linkage (between groups) method. Best performance is obtained with the squared Euclidean and the city block measures, while the Chebychev measure gives the worst performance. The results indicate that measures based on differences between variables perform better than those adopting other criteria, such as the cosine measure, or those ignoring much of the available difference data, such as the Chebychev measure which uses only the maximum value of such data.

4.3. Effect of data standardisation prior to clustering

The effect of data standardisation before clustering is shown in Table 4 for the average linkage (between groups) clustering method and the squared Euclidean distance measure. With all nine

Table 3
Classification performance of the various distance measures, with the empirical classification as a reference

Distance/similarity measurement		Classification performance				
		Empirical type	Classified as type			Total classification error(%)
			1	2	3	
b	Squared Euclidean	1	43	0	0	
f	City block	2	0	144	0	7.9
		3	0	17	10	
g	Minkowiski, $n=3$	1	42	1	0	
		2	0	144	0	8.4
		3	0	17	10	
a	Euclidean	1	41	2	0	
		2	0	144	0	8.9
		3	0	17	10	
c	Cosine	1	43	0	0	
d	Pearson correlation	2	1	139	0	14.5
		3	0	27	0	
e	Chebychev	1	43	0	0	
		2	144	0	0	78
		3	16	7	4	

Clustering with all nine variables.
Clustering method: average linkage between groups.
Data standardised by variable as Z-scores.

variables used, it was found advantageous to standardise the data using any standardisation method, except that of standardising to mean = 1, which proved inferior to the case of no standardisation. Where used, all standardisation was carried out by variable, since standardisation by case produced inferior results to the case of no standardisation, with errors as high as 86% using Z-scores. Referring to Table 1 (a), since all variables other than CT_R are percentage concentrations with a maximum value of 100, the major requirement for standardisation arises in connection with the CT_R variable which can be as high as 2250 in the raw data set. It is noted that the empirical classification [12–14] was performed without standardisation and using only the three main variables URATE, OXAL, and PHOS. Cluster analysis results under the same conditions give the best matching with the empirical classification [12–14], with a total classification error of $\approx 6\%$. Clustering with all three main variables excluded produces errors as large as 63%, indicating the importance of these radicals for classifying the stones (see Table 4).

4.4. Effect of the choice of variables used in clustering

The effect of the choice of variables used for clustering on the classification accuracy was investigated. As shown in Table 5, excluding the CT_R variable, the H variable, or the MG variable alone, and performing the cluster analysis with the remaining eight variables in each case, causes only a minor increase in the classification error (an extra case or two being misclassified) as compared to when all the nine variables were utilised. Excluding these three variables simultaneously also has the same effect. When the six variables: CT_R, H, MG, CA, C, and N are excluded simultaneously, leaving only the main three variables used empirically [12–14], we get the same classification accuracy obtained with all the nine variables utilised. This shows that these three

Table 4
Effect of data standardisation by variable prior to cluster analysis on the classification performance, with the empirical classification as a reference

Standardisation method (standardisation by variable)		Classification performance				
		Empirical type	Classified as type			Total classification error (%)
			1	2	3	
ii	Z scores	1	43	0	0	
iii	Range −1 to +1	2	0	144	0	7.9
iv	Range 0 to +1	3	0	17	10	
v	Maximum magnitude = 1					
vi	S.D. = 1					
vii	Mean = 1	1	43	0	0	
		2	144	0	0	73
		3	9	4	4	
i	No standardisation	1	43	3	0	
		2	0	61	83	46
		3	0	13	14	
	No standardisation, excluding the CT_R variable	1	43	0	0	
		2	0	144	0	7.9
		3	0	17	10	
	No standardisation, excluding all three main variables: (URATE, OXAL, PHOS)	1	40	3	0	
		2	0	17	127	63
		3	0	5	22	
	No Standardisation, using only the three main variables: (URATE, OXAL, PHOS)	1	43	0	0	
		2	0	144	0	6.1
		3	0	13	14	

Clustering with all nine variables unless otherwise mentioned.
Clustering method: average linkage between groups.
Distance measure: squared Euclidean.

Table 5
Effect of the choice of variables used in the cluster analysis on the classification performance, with the empirical classification as a reference

Step	Variables used in clustering	Classification performance				
		Empirical type	Classified as type:			Total classification error (%)
			1	2	3	
1	All nine variables	1	43	0	0	
		2	0	144	0	7.9
		3	0	17	10	
2	Excluding CT_R alone or excluding H alone	1	41	2	0	
		2	0	144	0	8.9
		3	0	17	10	
3	Excluding MG alone	1	43	2	0	
		2	0	143	0	8.4
		3	0	16	11	
4	Excluding CT_R, H, MG (Clustering using main six variables only: CA, URATE, OXAL, PHOS, C, N)	1	43	1	0	
		2	1	143	0	8.9
		3	0	16	11	
5	Excluding CT_R, H, MG, CA, C, N (Clustering using main three variables only: URATE, OXAL, PHOS)	1	43	0	0	
		2	0	144	0	7.9
		3	0	17	10	
6	Excluding the main variables: URATE, OXAL, PHOS	1	42	1	0	
		2	0	144	0	11.2
		3	0	23	4	

Clustering method: average linkage between groups.
Distance measure: squared Euclidean.
Data standardised by variable as Z-scores.

variables are key parameters for the classification of the stones data being considered and that their choice as a minimum set of classifiers [12–14] is quite appropriate. Adding more variables does not significantly improve the accuracy of classification by cluster analysis. The last entry in Table 5 shows the results obtained by clustering with the main three variables (ORATE, OXAL, PHOS) excluded, giving the poorest total classification error of 11.2%. Deterioration in performance is negligible on types 1 and 2, but is quite significant for type 3 where the within-type error rate increases from 63 to 85%.

Limiting the number of variables to only three (ORATE, OXAL and PHOS) simplifies further analysis involving the average squared Euclidean distance to the group centres. Using data on variable averages without standardisation (Table 1 (a)), the average distance matrix to group centres is given by:

$$\mathbf{D} = \begin{vmatrix} 412.0 & 8517.3 & 7670.3 \\ 8150.7 & 44.3 & 1363.5 \\ 7629.1 & 1689.4 & 370.1 \end{vmatrix} \quad (1)$$

The element in row i and column j, dij; $i, j = 1, 2, 3$ is the average distance within group i to the centre of group j. Diagonal elements of the square matrix reflect within-group spread, while other elements indicate between-group separation. For example, d_{22} measures the spread within type 2 stones, while d_{23} measures how much type 2 stones are further removed from the centre of type 3 and, therefore, how unlikely it would be to misclassify them as type 3 by a clustering algorithm. The smallest spread is seen to be within type 2, while type 1 exhibits the widest within-type spread. Accurate discrimination between the various types requires small within-type spread and large inter-type separation. Normalising the elements of each row to its diagonal element gives:

$$\mathbf{D}' = \begin{vmatrix} 1 & 20.8 & 18.6 \\ 184 & 1 & 30.8 \\ 20.6 & 4.6 & 1 \end{vmatrix} \quad (2)$$

High values for the normalised (non-diagonal) elements in a row indicate potential for good discrimination for the stone type corresponding to that row. The normalised matrix above suggests that best performance would be expected for type 2. It appears highly unlikely that type 2 stones are classified as type 1 ($d'_{21} = 184$). Worst performance is expected to manifest itself in misclassifying type 3 as type 2 ($d'_{32} = 4.6$). Cluster analysis results given in the last entry of Table 4 confirms these expectations.

Standardising the data for the three main variables as Z-scores by variable gives $\mathbf{D}'$ as

$$\mathrm{D}'_{(\mathrm{Z-scores})} = \begin{vmatrix} 1 & 21.8 & 27.9 \\ 58.2 & 1 & 41.9 \\ 7.6 & 4.6 & 1 \end{vmatrix} \quad (3)$$

Comparison between Eqs. (2) and (3) indicates that, using only the main three variables (ORATE, OXAL, PHOS), clustering without standardisation promises better overall results. In this case the average linkage (between groups) method gave the best classification performance, as the last entry of Table 4 shows. Comparison with entry 5 in Table 5 for clustering with the main three variables using Z-scores standardisation shows that classification error for type 3 is reduced from 63% with standardisation to 48% without standardisation.

4.5. Clustering with only one variable

In the empirical classification [12–14], each of the three stone types was determined by a simple test on each of the three main variables, as given by the classification rules in Section 1. It is interesting to see how cluster analysis performs when operating on the chemical analysis variables individually. Using only one variable at a time, standardisation has no effect. Clustering with the URATE variable alone into two classes only [type 1 (U) and non-type 1 (X and P combined)] matches the empirical classification [12–14] to 1% as shown in Table 6. The table also shows the results of using each of OXAL and PHOS to separate type 2 (X) and type 3 (P), respectively in 2-cluster solutions. Shown also are results obtained when using the remaining six parameters individually for cluster analysis in light of expec-

Table 6
Classification performance using only one variable at a time, with the empirical classification as a reference

Single variable used in clustering		Performance				
		Classification of empirical stone types				Total classification error (%)
		Type	Classified as type			
			1	2 and 3		
1	URATE	1	41	2		0.9
		2 and 3	0	171		
			2	1 and 3		
2	OXAL	2	144	0		11.2
		1 and 3	24	46		
			3	1 and 2		
3	PHOS	3	10	17		7.9
		1 and 2	0	187		
			1	2 and 3		
4	CT_R	1	40	3		1.4
		2 and 3	0	171		
5	CA	1	43	0		0.0
		2 and 3	0	171		
6	H	1	38	5		2.3
		2 and 3	0	171		
7	N	1	39	4		1.9
		2 and 3	0	171		
			2	1 and 3		
8	C	2	144	0		14.5
		1 and 3	31	39		
			3	1 and 2		
9	MG	3	4	23		10.7
		1 and 2	0	187		
			1	2	3	
10	C	1	39	4	0	
		2	0	144	0	9.8
		3	0	17	10	

Clustering method: average linkage between groups.
Distance measure: squared Euclidean.
All are 2-cluster solutions, except for the last entry where a 3-cluster solution is used

tations for each variable mentioned in Section 2, based on the disparity between type averages for each variable as listed in Table 1 (a). Results indicate that CA perfectly separates type 1 stones from the rest of the population with no errors, outperforming even URATE for this purpose. Each of CT_R, H, and N adequately performs this task with a slightly poorer accuracy. C ions separate type 2 from the rest of the population with a modest accuracy of 14.5%, but can serve as a single classifier for all three stone types with an overall accuracy of 9.8%. The percentage C ions are maximum for type 1 stones (20.05–35.81) and minimum for type 3 stones (3.0–16.86). For type 3 the values fall in the middle (11.53–23.53). Overlap between the three ranges accounts for the difficulty in attaining complete separation between the three stone types and the largest overlap between types 2 and 3 leads to the largest error. The results confirm the usefulness of ele-

mental microanalysis parameters in the quick identification of urinary stones [12–14].

5. Conclusions

Cluster analysis techniques were used for the classification of non-infection urinary stones into three stone types, based on chemical analysis data for nine variables. Results are generally supportive of earlier findings by a simple empirical classification scheme [12–14]. Best agreement between the empirical and cluster analysis results (total error = 6%) is obtained by clustering with the main three variables (URATE, OXAL, PHOS) (without standardisation) using the average linkage within groups method and the squared Euclidean distance measure. Using all the nine variables in clustering does not improve performance. Excluding the three main variables increases the error considerably (to 63%), indicating their importance as classifiers. Standardisation of the data by variable has no significant effect unless CT_R is included, due to its significantly larger values. Therefore, the empirical rules of Section 1 provide a simple and effective means of classifying the stones. The study confirmed the effectiveness of other single parameters, such as CT_R, CA, H, N in separating type 1 stones from the rest of the population. C alone can classify all three stone types with an overall accuracy of 9.8%. Performance with other clustering methods and similarity and distance measures generally agrees with properties documented in the literature.

Acknowledgements

Support by the Research Institute of King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia and by the Division of Urology, Department of Surgery, Faculty of Medicine, King Saud University, Riyadh, Saudi Arabia is greatly appreciated.

References

[1] A. Hodgkinson, A combined qualitative and quantitative procedure for the chemical analysis of urinary calculi, J. Clin. Path. 24 (1971) 147–151.
[2] A. Hodgkinson, A. Williams, An improved calorimetric procedure for urinary oxalate, Clin. Chem. Acta 36 (1972) 127.
[3] M. Fuss, J. Simon, D. Verbeelen, M. Weiser, P. Van Landuyt, J Elskens, Chemical analysis of renal stones from 377 Belgian patients by using qualitative or quantitative methods, Eur. Urol. 4 (1978) 90–93.
[4] C.J. Farrington, M.L. Liddy, A.C. Chalmers, A simplified sensitive method for the analysis of renal calculi, Am. J. Clin. Path. 73 (1980) 96–99.
[5] L. Larsson, B. Sorbo, H.G. Tiselius, S. Ohman, A method for quantitative wet chemical analysis of urinary calculi, Clin. Chem. Acta 140 (1984) 9–20.
[6] E.J. Westbury, Some observations on the quantitative analysis of over 1000 calculi, Br. J. Urol. 46 (1974) 215.
[7] D.J. Sutor, D.J. Wooley, J.J. Illingworth, Some aspects of the adult urinary stone problem in Great Britain and Northern Ireland, Br. J. Urol. 46 (1974) 275–288.
[8] D.J. Sutor, D.J. Wooley, J.J. Illingworth, A geographical and historical survey of the composition of urinary stones, Br. J. Urol. 46 (1974) 393–407.
[9] E. Takasaki, An observation on the composition and recurrence of urinary calculi, Urol. Int. 30 (1975) 228.
[10] B. Otnes, Urinary stone analysis, methods, materials and value, Scand. J. Urol. Nephrol. (Suppl. 71) (1983) 1–109.
[11] R.E. Abdel-Halim, A.O. Baghlaf, A.B. Farag, Clinical chemical study of urinary stones in Saudi Arabia: l-uric acid stones, in: P.O. Schwille, L.H. Smith, W.G. Robertson, W. Vahlensieck (Eds.), Urolithiasis and Related Clinical Research, Plenum Press, New York, 1985, pp. 715–718.
[12] R.E. Abdel-Halim, A. Al-Sibaai, A.O. Baghlaf, Ionic associations within 460 non-infection urinary stones: A quantitative chemical analytical study applying a new classification, Scand. J. Urol. Nephrol. 27 (1993) 155–162.
[13] R.E. Abdel-Halim, A. Al-Sibaai, A.O. Baghlaf, The structure of large lamellar urinary stones: a quantitative chemical analytic study applying a new classification scheme, Scand. J. Urol. Nephrol. 27 (1993) 337–341.
[14] A.O. Baghlaf, A.A. Al-Sibaai, R.E. Abdel-Halim, M.S. Al-Hadramy, S. Abdul-Wahab, A.W. Noorwali, A. Al-Waseef, Chemical composition of urinary stones in Jeddah: a study of 331 stones and a proposed new classification, in: V.R. Walker, et al. (Eds.), Urolithiasis, Plenum Press, London, New York, 1989, pp. 713–714.
[15] T.A.K. Ghori, K.C. Nagpal, S.Z. Ali, X-ray diffraction study of urinary stones from the Meerut region, Indian J. Med. Res. 75 (1982) 894–898.
[16] G.W. Drach, B.J. Hillman, P. Tracey, J.A. Gaines, Computed tomographic analysis of the composition of renal calculi, in: P.O. Schwille, L.H. Smith, W.G. Robertson,

W. Vahlensieck (Eds), Urolithiasis and Related Clinical Research, Plenum Press, London, New York, 1985, pp. 743–746.

[17] H.C. Romesburg, Cluster Analysis for Researchers, Life-time Learning Publications, Belmont, California, 1984.

[18] SPSS for Windows: SPSS Professional Statistics, Release 5, SPSS, Chicago, 1992.

[19] SAS/STAT User's Guide, version 6, 4th ed., SAS Institute, Cary, NC, 1989.

[20] G.W. Milligan, A review of Monte Carlo tests of cluster analysis, Multivar. Behav. Res. 16 (1981) 379-407.