European Journal of Orthodontics 17 (1995) 231–240

Cluster analysis application to Class I malocclusion

Ignacio G. Espona*, Joaquin Travesi Gomez*, and
Jorge Bolanos Carmona**
Departments of *Orthodontics and **Computation Sciences, University of Granada, Spain

SUMMARY The purpose of this study was to obtain Class I malocclusion statistical subtypes by applying cluster analysis techniques, to assess clinical and cephalometric characteristics from different clusters, and to analyse sex and age effects on grouping patterns. Four-hundred-and-sixteen Spanish patients (243 females, 173 males) with Class I osseous and dental malocclusion (4 degrees >ANB>0 degrees) between 8 and 16 years of age, with no previous orthodontic treatment, were analysed. Cluster analysis was applied to the information provided by 20 variables both clinical and cephalometric (Ricketts' analysis) per patient. The grouped individuals were defined in a statistically significant manner by a greater lower incisive proclination, greater lower labial protrusion, less dental crowding, and less pogonion–NB distance. So, only the protrusive traits were statistically expressed in the cluster analysis. The grouping pattern in Class I malocclusion was shown in a more defined form at younger age levels and disappeared with age. The clustering pattern was very similar in Class I malocclusion males and females.

Introduction

Evaluating and classifying dentofacial anomalies is one of the essential objectives in orthodontics. Traditionally, this process has been carried out by methods founded on clinical (subjective) (Angle, 1899; Anderson, 1963) and cephalometric (Downs, 1948; Wylie, 1952; Riedel, 1952; Steiner, 1955; Ricketts, 1960; McNamara, 1984) criteria, and therefore, has been unable to deal with a whole set of variables in an efficient, quantifiable, and reproducible manner (Hirschfeld *et al.*, 1973).

In recent years, sophisticated statistical multivariate methods (cluster analysis, stepwise discriminant analysis, canonical analysis of discriminance, etc.) (Johnson and Wichern, 1988) have been developed and applied in taxonomy in an effort to improve the evaluation of an individual as a whole, and to make classification schemes more objective (Moyers *et al.*, 1980). Its application in the redefinition of malocclusive syndromes (Hirschfeld *et al.*, 1973; Lavelle, 1977a,b, 1989; Bhatia *et al.*, 1979; Moyers *et al.*, 1980; Solano Reina, 1982; Petrovic *et al.*, 1986; Moyers, 1988; Finkelstein *et al.*, 1989; Keeling *et al.*, 1989; Fine and Lavelle, 1992) has been mainly centred on the Class II malocclusive syndrome (Solano Reina, 1982; Moyers, 1987; Johnson and Wichern, 1988). To date few studies that mathematically define categories within Class I and Class III malocclusions (Finkelstein *et al.*, 1989) are available.

The purpose of this study was:

Table 1 Cluster analysis variables.

Clinical	
1.	Overjet
2.	Overbite
3.	Lower intercanine distance
4.	Lower intermolar distance
5.	Lower discrepancy
6.	Upper discrepancy
Cephalometric (Ricketts' analysis):	
7.	Molar relationship
8.	Facial convexity
9.	Lower face height
10.	Mandibular incisor protrusion
11.	Maxillary incisor protrusion
12.	Mandibular incisor inclination
13.	Maxillary incisor inclination
14.	Lower lip protrusion
15.	Facial depth
16.	Facial axis
17.	Mandibular plane angle
18.	Maxillary height
19.	Mandibular arc
20.	Mandibular corpus length

(1) to obtain Class I malocclusion statistical subtypes by applying cluster analysis techniques;
(2) to assess clinical and cephalometric characteristics from different clusters;
(3) to analyse sex and age effects on grouping patterns.

Cluster analysis

A cluster analysis is a multivariate procedure used to detect natural data groupings on the basis of the similarity of data (Tsuchiya *et al.*, 1992). A classification procedure is based on generalization and results in a loss of information. In a sample of N individuals, the best classification with no loss of information is of N categories with one individual in each category. As the number of categories is reduced, and more and more individuals are grouped together, the loss of information increases (Bhatia *et al.*, 1979). At any one stage in the analysis the loss of information could be measured by the ESS (error of the sum of squares) or similarity coefficient, obtained by the total sum of squared deviations of every point from the mean of the cluster to which it belongs (Ward, 1963; Finkelstein *et al.*, 1989).

Cluster analysis involves a series of steps: in the first cycle every possible pairing of the individuals is considered, and the two most similar individuals resulting in the least ESS are

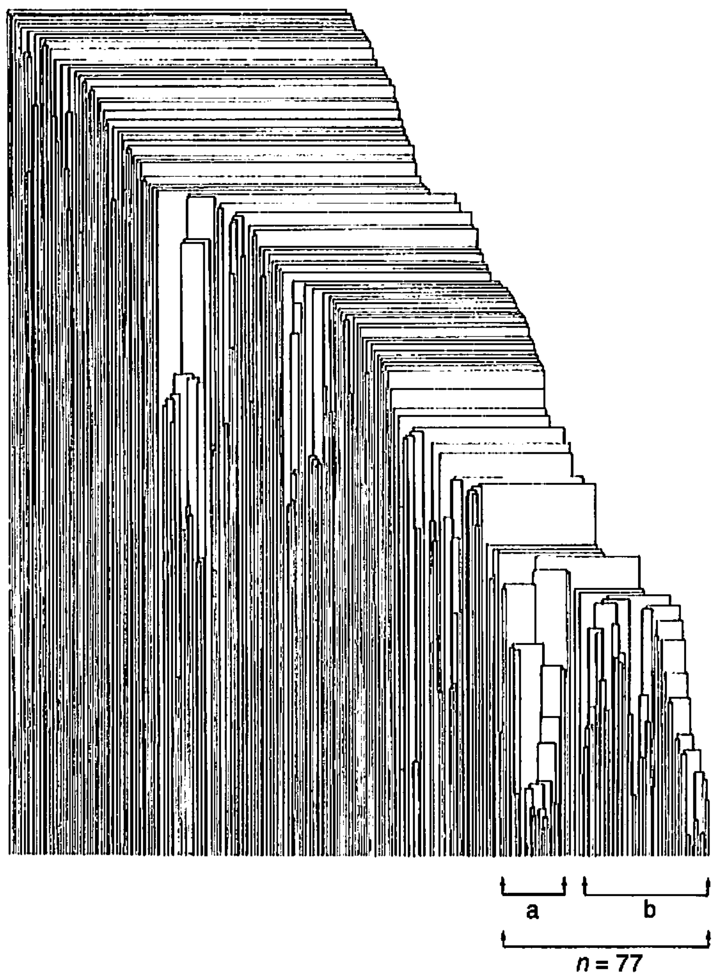

Figure 1 Dendogram for Class I malocclusive females (8–16 years of age), showing a nuclear structure ($n=77$) with two subgroups (a + b).

Table 2 Statistical comparison between nuclear ($n=77$) and non-nuclear ($n=166$) groups from dendogram corresponding to Class I malocclusive females 8–16 years of age.

	Nuclear group		Non-nuclear group		Level of significance
	Mean	SD	Mean	SD	
Clinical variables					
Age	11.00	1.98	11.29	1.96	—
Overjet	2.84	1.72	3.46	2.84	—
Overbite	2.21	1.67	2.42	2.47	—
Midline deviation	1.33	1.03	1.26	1.08	—
Lower intercanine distance	24.86	1.21	24.62	2.25	—
Lower intermolar distance	39.70	1.93	39.56	2.99	—
Curve of Spee	1.34	0.84	1.81	1.04	—
Upper discrepancy	−2.46	2.96	−3.20	4.41	—
Lower discrepancy	0.02	2.88	−1.41	4.36	$P<0.01$
Cephalometric variables (Steiner analysis)					
SNA (degrees)	79.37	2.80	78.89	3.40	—
SNB (degrees)	76.36	2.74	76.04	3.41	—
SND (degrees)	73.44	2.68	73.41	3.36	—
ANB (degrees)	3.05	1.36	2.85	1.91	—
SL distance (mm)	43.99	5.84	44.34	7.70	—
SE distance (mm)	20.63	2.98	20.71	3.26	—
Mandibular to SN (degrees)	37.13	4.19	37.39	6.26	—
Occlusal to SN (degrees)	21.09	3.53	21.15	4.92	—
1 to NA (mm)	3.62	1.87	3.99	2.94	—
I to NA (degrees)	22.25	5.61	22.62	7.78	—
I to NB (mm)	4.39	1.64	3.53	2.37	$P<0.001$
1 to NB (degrees)	24.26	4.59	21.70	6.16	$P<0.001$
Upper lip to LE (mm)	0.24	2.02	−0.21	2.19	—
Lower lip to LE (mm)	1.29	2.06	0.46	2.61	$P<0.01$
Pogonion to NB (mm)	1.41	1.17	2.02	1.48	$P<0.01$
Cephalometric variables (Ricketts' analysis)					
Molar relationship	−0.83	1.78	−0.58	2.42	—
Lower incisor extrusion	1.81	1.42	2.14	1.87	—
Interincisal angle	130.42	8.55	132.89	11.28	—
Facial convexity	2.29	1.53	1.79	2.13	—
Lower face height	45.60	3.03	46.02	4.93	—
Upper molar position	13.26	3.39	13.46	3.87	—
Lower incisor protrusion	1.96	1.78	0.98	2.50	$P<0.001$
Upper incisor protrusion	5.48	1.93	5.45	3.11	—
Lower incisor inclination	22.47	4.17	20.81	5.36	$P<0.05$
Upper incisor inclination	27.10	5.67	26.29	8.32	—
Occlusal plane to ramus	−1.04	4.18	−1.81	3.84	—
Occlusal plane inclination	22.24	4.31	22.69	4.34	—
Lip protrusion	−0.48	2.24	−1.42	2.86	$P<0.01$
Upper lip length	25.26	2.03	25.64	2.69	—
Lip embrasure to occlusal plane	−4.18	1.87	−4.40	2.62	—
Facial depth	87.47	2.25	87.07	3.06	—
Facial axis	87.02	2.88	86.80	3.99	—
Facial taper	65.80	3.36	65.49	4.77	—
Mandibular plane angle	26.71	3.94	27.42	5.64	—
Maxillary depth	89.78	2.46	88.86	3.02	$P<0.05$
Maxillary height	57.82	2.22	57.66	3.31	—
Palatal plane	1.38	3.44	1.33	3.05	—
Cranial deflection	28.54	2.09	28.69	2.62	—
Cranial length anterior	55.64	3.01	56.07	2.97	—
Posterior facial height	57.42	5.08	56.97	5.52	—
Ramus position	74.74	3.57	75.14	4.71	—
Porion location	−40.02	3.54	−39.97	3.61	—
Mandibular arc	26.53	6.21	24.62	7.62	$P<0.05$
Mandibular corpus length	67.09	4.37	67.26	4.98	—

fused (Bhatia *et al.*, 1979). In each new cycle a new individual is included in one of the subgroups (clusters) already formed, with the condition of adding minimal variability to the group. In addition, the union of every possible cluster pair is considered at each step in the analysis, with a cluster pair being combined only where fusion resulted in the minimum increase in the error sum of squares. Once the process is finished, the number of clusters from the original sample can be decided based on the assessment of sudden increase in the 'loss of information', which indicates an excessive subgrouping (Finkelstein *et al.*, 1989).

In the cluster represented in vertical form (dendogram), the horizontal lines serve to group cases that are similar. Objects in a given cluster tend to be similar to each other in some sense and objects in different clusters do not tend to show this similarity (Tsuchiya *et al.*, 1992). The proximity to the base line indicates the order in which the groups are formed.

Subjects and methods

Sample

A sample of 416 Spanish patients (243 females and 173 males) with Class I osseous and dental malocclusion (4 degrees > ANB > 0 degree and Class I Angle molar relation) between 8 and 16 years of age, with no previous orthodontic treatment, was analysed. For each patient, 108

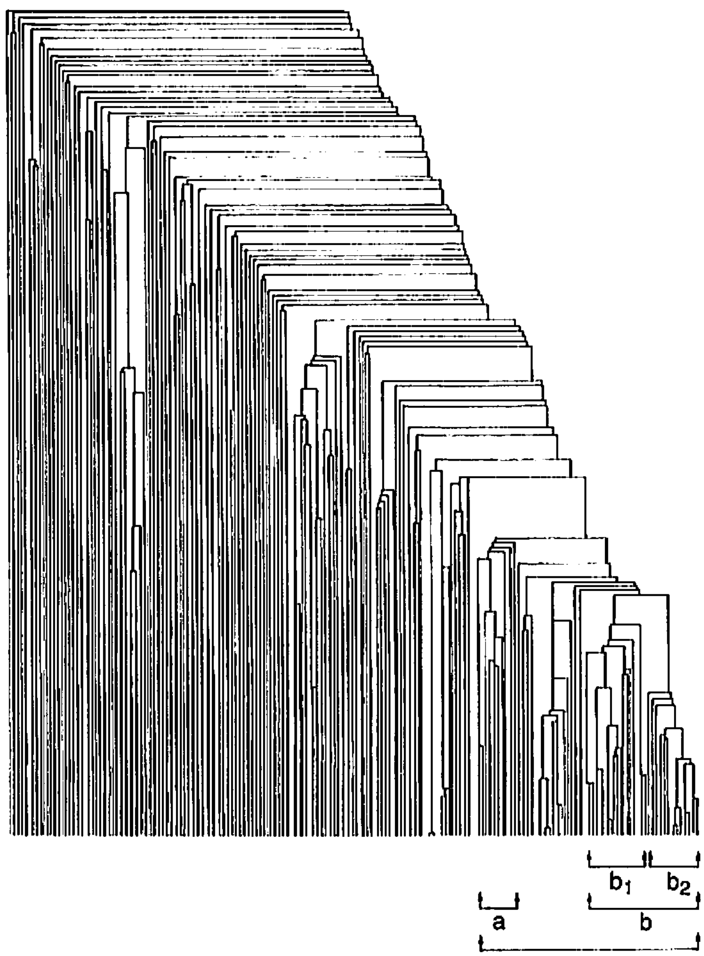

Figure 2 Dendogram for Class I malocclusive females (8–12 years of age), showing a complex nuclear structure with two different sized subgroups (a and $b = b_1 + b_2$) and some intermediate individuals.

clinical and cephalometric variables were recorded (Steiner and Ricketts' analyses).

Statistical analysis

Cluster analysis. A statistical cluster analysis of average linkage was applied to the information provided by 20 variables per patient, both clinical (six variables) and cephalometric (14 variables from Ricketts' analysis) (Table 1). Average linkage cluster analysis was chosen due to its greater capacity with respect to others (single and complete linkage cluster analysis) to define more meaningful and numerous subgroups. The 20 variables selected permitted a representation of the basic characteristics of each patient and without repetitions between the same.

In the first stage of the research cluster analysis was applied to the complete sample. In the second stage the sample was studied separately at two intervals of 4 years (8–12 and 12–16 years of age). Females and males were analysed separately.

The reproducibility of the cluster analysis for both sexes was checked using two subsamples obtained at random and also an additional application of other cluster analysis (single and complete linkage).

Comparisons of means between clusters were completed by the Wilcoxon non-parametric test for independent samples. Statistical assessments were completed at the 95 per cent level of significance. Moreover, a discriminant analysis was carried out by adjusting a logistic regression model step by step with the BMDP software program in order to determine which variables best defined the nucleus. All the variables were used for this, except those from Steiner's cephalometric analysis. Variable input level of $P<0.10$ was accepted.

Results

Females

The dendogram obtained for Class I malocclusive females 8–16 years old (Fig. 1) shows a nuclear structure ($n=77$) in which two subnuclei (a and b) are distinguished. The subnucleus b, in turn, is composed of two distinct populations. Table 2 shows the descriptive statistic of this cluster's nuclear individuals compared with that of the remaining individuals. Females in nuclear structure showed, in a statistically significant manner, greater lower incisive protrusion and proclination, greater lower lip protrusion, greater maxillary depth, greater mandibular arc, less lower dental crowding and less pogonion–NB distance.

The logistic regression analysis showed the lower incisor protrusion, upper lip length, mandibular arc, and maxillary depth variables to be discriminant thus confirming the results of the multivariate analysis.

When the female population was split into two age intervals, two clearly distinct dendograms were obtained. The first (females 8–12 years old: Fig. 2) showed a complex nuclear structure in which two subnuclei (a and $b=b_1+b_2$) of different sizes were distinguished with some individuals in between. The second dendogram (females 12–16 years old: Fig. 3) revealed on the other hand a scarce grouping pattern.

Males The dendogram obtained for Class I malocclusive males 8–16 years old (Fig. 4) shows a complex nuclear structure ($n=70$) in which two subnuclei (a and b) are distinguished.

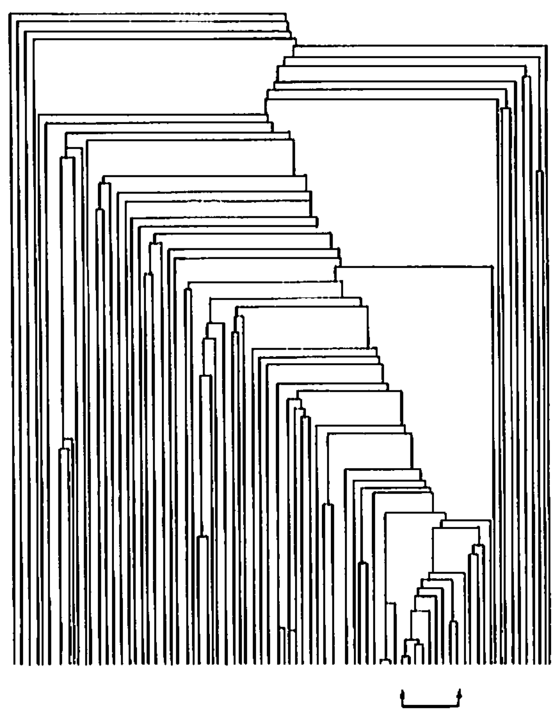

Figure 3 Dendogram for Class I malocclusive females (12–16 years of age). A short grouping pattern is observed.

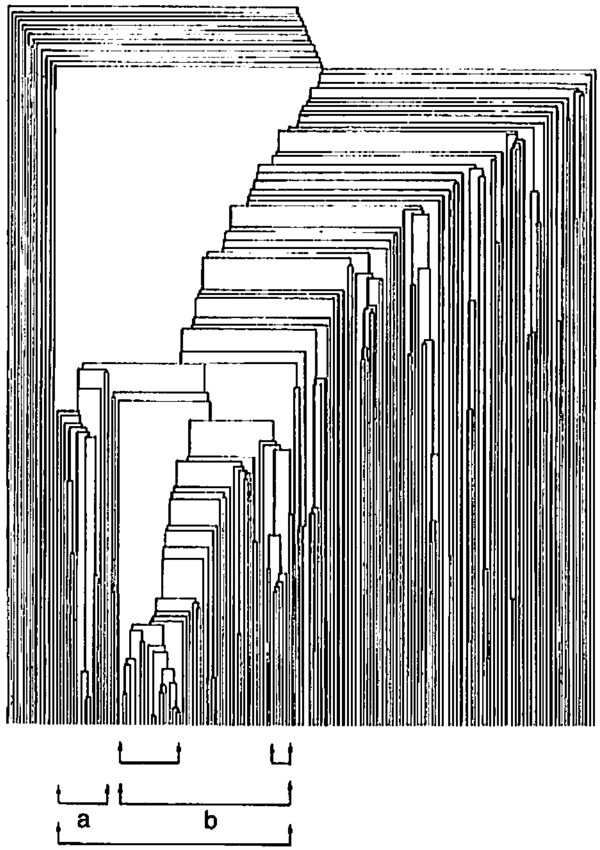

Figure 4 Dendogram for Class I malocclusive males (8–16 years of age), showing a complex nuclear structure ($n = 70$) with two different-sized subgroups (a + b).

The right-hand subnucleus (b), in turn, is composed of several populations. Table 3 shows the descriptive statistic of this cluster's nuclear individuals compared with that of the remaining individuals. Males in nuclear structure showed, in a statistically significant manner, greater lower incisive proclination, greater lower lip protrusion, more mesial molar relationship, less upper dental crowding, less pogonion–NB distance, greater mandibular arc, and an occlusal plane closer to the centre of the branch.

The logistic regression analysis showed the mandibular arc, lip protrusion, molar relationship, and facial taper as discriminant thus confirming the results of the multivariate analysis.

When the male population was split into two age intervals, two clearly distinct dendograms were obtained. The first (males 8–12 years old: Fig. 5) again showed a nuclear structure with two subnuclei (a and b) of different sizes, while the second (males 12–16 years old: Fig. 6) revealed a stepped order of the individuals, without defined groupings.

Discussion

Methodologically, our study differs from previous ones (Hirschfeld *et al.*, 1973; Lavelle, 1977a,b, 1989; Bhatia *et al.*, 1979; Moyers *et al.*, 1980; Solano Reina, 1982; Moyers, 1988; Keeling *et al.*, 1989; Finkelstein *et al.*, 1989; Fine and Lavelle, 1992) in the application of a high number (20) of clinical as well as cephalometrical variables and, in turn, linear as well as angular. A greater number of variables was not considered so as not to multiply the importance of certain characteristics (Everitt, 1983). The information provided for each patient allowed

Table 3 Statistical comparison between nuclear ($n=70$) and non-nuclear ($n=103$) groups from denodgram corresponding to Class I malocclusive males 8–16 years of age.

	Nuclear group		Non-nuclear group		Level of significance
	Mean	SD	Mean	SD	
Clinical variables					
Age	10.52	1.71	11.22	2.02	—
Overjet	2.84	1.98	3.59	3.95	—
Overbite	2.78	1.60	2.70	3.00	—
Midline deviation	1.17	1.15	1.13	1.13	—
Lower intercanine distance	25.39	1.44	25.71	2.29	—
Lower intermolar distance	40.53	1.94	40.99	2.71	—
Curve of Spee	1.44	0.79	1.59	1.17	—
Upper discrepancy	−1.53	3.64	−3.19	4.62	$P<0.05$
Lower discrepancy	0.27	3.63	−0.87	4.48	—
Cephalometric variables (Steiner analysis)					
SNA (degrees)	80.29	2.89	79.63	3.74	—
SNB (degrees)	77.04	0.33	76.62	3.58	—
SND (degrees)	74.11	2.73	74.02	3.43	—
ANB (degrees)	3.24	1.29	3.01	1.77	—
SL distance (mm)	47.20	6.05	47.25	8.38	—
SE distance (mm)	20.48	2.88	21.15	3.42	—
Mandibular to SN (degrees)	35.64	4.34	36.54	6.30	—
Occlusal to SN (degrees)	19.65	3.56	20.01	5.46	—
1 to NA (mm)	3.65	2.09	3.88	3.60	—
I to NA (degrees)	21.53	5.47	21.94	9.28	—
I to NB (mm)	4.53	1.79	3.91	2.64	—
1 to NB (degrees)	24.89	4.64	22.90	6.85	$P<0.05$
Upper lip to LE (mm)	1.42	1.76	0.80	2.28	—
Lower lip to LE (mm)	2.34	1.97	1.22	2.83	$P<0.01$
Pogonion to NB (mm)	1.54	1.23	2.24	1.94	$P<0.01$
Cephalometric variables (Ricketts' analysis)					
Molar relationship	−1.14	1.87	−0.25	2.76	$P<0.05$
Lower incisor extrusion	1.88	1.53	2.27	2.76	—
Interincisal angle	130.32	8.27	132.14	12.36	—
Facial convexity	2.44	1.43	1.80	2.20	—
Lower face height	45.73	3.73	45.99	4.68	—
Upper molar position	13.34	3.44	13.80	3.98	—
Lower incisor protrusion	1.89	1.88	1.14	2.85	—
Upper incisor protrusion	5.69	2.06	5.40	3.88	—
Lower incisor inclination	23.04	4.38	22.04	5.87	—
Upper incisor inclination	26.63	5.51	25.81	9.95	—
Occlusal plane to ramus	−0.05	3.81	−1.75	4.01	$P<0.01$
Occlusal plane inclination	21.93	4.29	22.78	4.45	—
Lip protrusion	0.57	2.11	−0.62	3.01	$P<0.01$
Upper lip length	26.19	2.31	26.70	2.72	—
Lip embrasure to occlusal plane	−4.04	2.57	−4.42	2.84	—
Facial depth	87.52	2.36	87.36	3.20	—
Facial axis	87.17	3.27	87.04	4.41	—
Facial taper	66.51	3.61	65.67	4.51	—
Mandibular plane angle	25.95	4.05	29.26	5.57	—
Maxillary depth	89.97	2.44	89.21	3.36	—
Maxillary height	56.95	2.80	57.29	3.54	—
Palatal plane	1.70	2.73	1.78	3.34	—
Cranial deflection	28.92	2.32	28.69	3.20	—
Cranial length anterior	57.25	3.37	57.66	3.81	—
Posterior facial height	59.19	5.38	58.36	5.11	—
Ramus position	74.88	3.53	74.36	4.06	—
Porion location	−41.19	3.52	−41.10	3.63	—
Mandibular arc	28.13	6.88	24.58	8.14	$P<0.01$
Mandibular corpus length	68.36	3.99	69.34	4.68	—

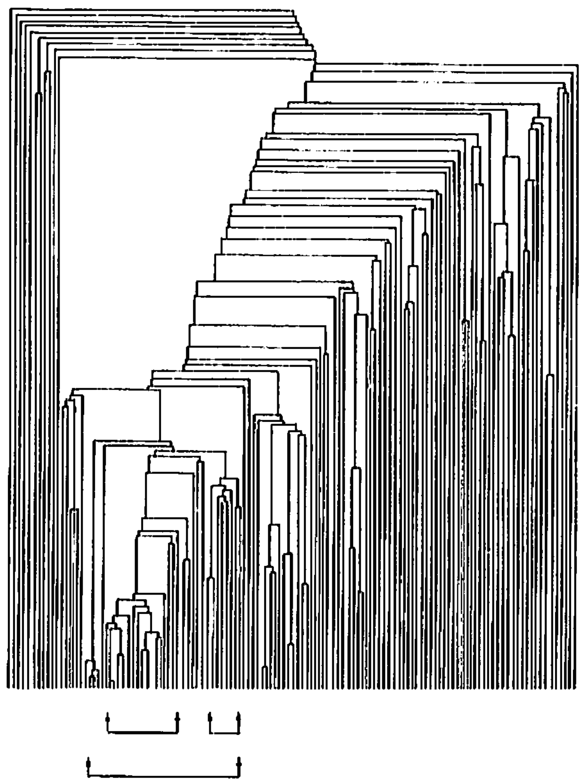

Figure 5 Dendogram for Class I malocclusive males (8–12 years of age), showing a nuclear structure with two different sized subgroups (a + b).

a three-dimensional definition and considered additionally the degree of dental crowding of both arches. The repetition obtained in the grouping pattern can be explained by the levelling of sex and age characteristics of the sample individuals, as well as by their high number.

Our results for the Class I malocclusive population by means of the cluster analysis technique show a high percentage of individuals, females as well as males, that systematically remain outside defined groupings. This finding allows us to emphasize the high heterogeneity of the Class I malocclusion from a numerical point of view.

The individuals that are arranged in groups constituting a nucleus are mainly characterized by protrusion of the incisors and of the lower lip, and slight crowding in the lower arch. The subpopulations of the nuclei do not represent clinically distinguishable entities. These findings show a notable agreement with those of Bhatia *et al.* (1979), who, from five different clusters associated to skeletal Class I, identified four with no clinical rendering and one characterized by its biprotrusive traits.

The discriminant analysis confirms that the differences are observable when based on a reduced number of variables among which the lower incisor protrusion, upper lip length, lower lip protrusion, mandibular arc, maxillary depth, molar relationship, and facial taper, stand out.

In our numeric taxonomy about Class I malocclusive population we have not found a reflection (Solano Reina, 1982) of the multiple clinical categories indicated previously by different authors (Langlade, 1981; Houston and Tulley, 1984). Traits such as osseous dental crowding, posterior cross-bites, open bites, etc., would not therefore constitute specific headings within the Class I malocclusive syndrome from a purely statistical point of view.

Grouping patterns were very similar in Class I malocclusion for females and males. This clustering pattern was shown in a more defined

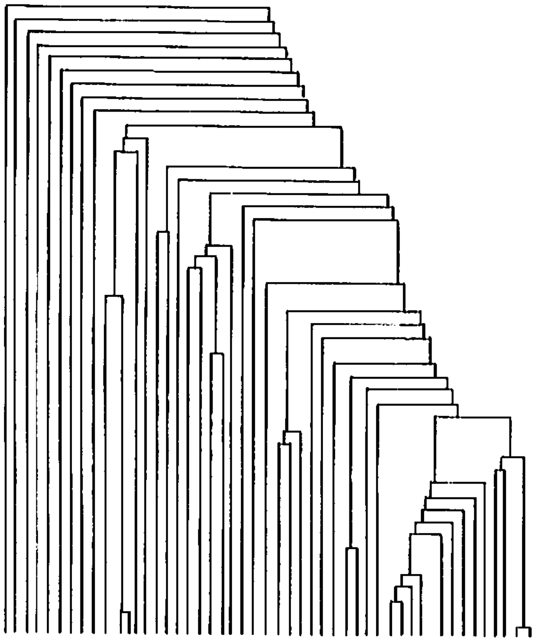

Figure 6 Dendogram for Class I malocclusive males (12–16 years of age), showing absence of defined groupings and a stepped order of the individuals.

form at younger age levels and disappears with age, especially in males. This appears to suggest an increase in the diversification of the sample individuals as their age increases, a hypothesis suggested previously by Bhatia *et al.* (1979), for whom the differences between individuals would not be expressed at early ages. We cannot rule out, however, the fact that the older groups constituted by smaller number of individuals plays a certain role.

We consider that, although the great clinical validity of the concept of Class I malocclusion cannot be denied, at least statistically this class appears to be a very heterogeneous entity. Further studies will have to clarify and define in a more specific way what differential qualities characterize the Class I malocclusive syndrome and, where appropriate, separate from that syndrome some entities currently included in the same that possess their own entity.

Conclusions

1. Of the multiple clinical entities constituting the Class I malocclusive syndrome, only the protrusive traits were statistically expressed in the cluster analysis.
2. The grouping pattern in Class I malocclusion is shown in a more defined form at younger age levels and disappears with age.
3. The clustering pattern is very similar in Class I malocclusion males and females.
4. Since cluster analysis is sensitive to method, sample, and variable selection these findings would need to be confirmed in further studies.

Address for correspondence

Dr Ignacio G. Espona
Camino de Ronda, 74, 4°, 1
18004-Granada
Spain

Acknowledgements

This study was supported in part by research grant BOJA 38/1989 of the Faculty of Dentistry, University of Granada, Spain, and formed part of a thesis for the PhD Degree in the University of Granada. We would like to express our thanks to Angela Barnie for reviewing the English version of the text.

References

Angle E H 1899 Classification of malocclusion. Dental Cosmos 41: 248–357

Anderson G M 1963 Ortodoncia práctica. Editorial Mundi, Buenos Aires

Bhatia S N, Wright G W, Leighton B C 1979 A proposed multivariate model for prediction of facial growth. American Journal of Orthodontics 75: 264–281

Downs W B 1948 Variations in facial relationship: their significance in treatment and prognosis. American Journal of Orthodontics 34: 812–840

Everitt B 1983 Cluster analysis. Heinemann Educational Books, London

Fine M B, Lavelle C L B 1992 Diagnosis of skeletal form on the lateral cephalogram with a finite element-based expert system. American Journal of Orthodontics and Dentofacial Orthopedics 101: 318–329

Finkelstein M, Lavelle C L B, Hassard T 1989 The role of cluster analysis on traditional cephalometric dimensions. Angle Orthodontist 59: 97–106

Hirschfeld W J, Moyers R, Enlow D H 1973 A method of deriving subgroups of a population: a study of craniofacial taxonomy. American Journal of Physical Anthropology 39: 279–290

Houston W J B, Tulley W J 1986 Class I malocclusions. In: Textbook of orthodontics. Wright, Bristol, pp. 154–163

Johnson R A, Wichern D W 1988 Applied multivariate statistical analysis. Prentice-Hall, New Jersey

Keeling S D, Riolo M L, Martin R E, Ten Have T R 1989 A multivariate approach to analyzing the relation between occlusion and craniofacial morphology. American Journal of Orthodontics and Dentofacial Orthopedics 95: 297–305

Langlade M 1981 Diagnostic orthodontique. Maloine Editeur, Paris

Lavelle C L B 1977a An analysis of the craniofacial complex in different occlusal categories. American Journal of Orthodontics 71: 574–583

Lavelle C L B 1977b A cephalometric study. Angle Orthodontist 47: 111–117

Lavelle C L B 1989 A retrospective cephalometric study of Class I patients. British Journal of Orthodontics 16: 17–24

McNamara J A 1984 A method of cephalometric evaluation. American Journal of Orthodontics 86: 449–469

Moyers R E 1987 Handbook of orthodontics. Year Book Medical Publishers, Chicago

Moyers R E, Riolo M, Guire K, Wainwright R, Bookstein F 1980 Differential diagnosis of Class II malocclusions. Part 1. Facial types associated with Class II malocclusions. American Journal of Orthodontics 78: 477–494

Petrovic A G, Lavergne J M, Stutzmann J J 1986 Tissue-level growth and responsiveness potential: growth rotation and treatment decision. In: Vig, P, Ribbens, K A (eds) Monograph No. 19, Craniofacial Growth Series. Center for Human Growth and Development, University of Michigan, Ann Arbor

Ricketts R M 1960 Cephalometric synthesis. American Journal of Orthodontics 46: 647–673

Riedel R A 1952 The relation of maxillary structures to cranium in malocclusion and normal occlusion. Angle Orthodontist 22: 142–145

Solano Reina E 1982 Estudio craneométrico y electromiográfico del síndrome de clase II. MSc Complutense University, Madrid

Steiner C C 1955 Cephalometrics for you and me. American Journal of Orthodontics 39: 729–765

Tsuchiya M, Lowe A A, Pae E K, Fleetham J A 1992 Obstructive sleep apnea subtypes by cluster analysis. American Journal of Orthodontics and Dentofacial Orthopedics 101: 533–542

Ward J H 1963 Hierarchical grouping to optimize an objective function. American Journal of Statistical Association 19: 236–244

Wylie L 1952 Assesment of antero-posterior dysplasia. Angle Orthodontist 22: 38–40