

# Experts' clinical Diagnosis Test as a Gold Standard for Cephalometric Evaluation of Vertical Facial Excess

## Diagnóstico clínico de expertos como patrón de oro para evaluación cefalométrica del exceso vertical facial

Salomón Yezioro Rubinsky and Javier Eslava-Schmalbach

Facultad de Odontología y Facultad de Medicina. Universidad Nacional de Colombia.  
syezioror@unal.edu.co, jheslavas@unal.edu.co

Received 13<sup>th</sup> October 2006/Sent for Modification 6<sup>th</sup> Jun 2007/Accepted 10<sup>th</sup> August 200

### ABSTRACT

**Objective** Cephalometric measures are used to evaluate vertical facial excess (VFE), however anyone of them have been validated against a gold standard to this purpose. Also, there are differences between cephalometric results and clinical evaluation. This study pretends to validate experts' clinical diagnosis test (ECDT) as gold standard for severe VFE, with the purpose of validating further against it the cephalometric measures results.

**Methods** A consensus (Delphi method) was done to determine if ECDT could be used as gold standard for evident VFE (n=12 experts). A scale of 9 items was initially built from literature. Validity: A convenience sample was used (n=24), which were tested by 3 experts twice. Principal factor analysis was made. Internal consistency was evaluated with Cronbach's alpha coefficient. Inter-observer and intra-observer agreement was measured using Kendall concordance coefficient. ECDT's medians were compared between groups with VFE, using Kruskal Wallis test.

**Results** Eleven of the twelve experts agreed that clinical diagnosis can be used as a gold standard for VFE. After Principal factor analysis a 6 items' scale was made. Internal consistency was high (Cronbach's alpha= 0.8051). ECDT's medians were different in groups with different qualitative appreciation of VFE by experts, in first ( $p<0.0001$ ) and second evaluation ( $p<0.0001$ ). A significant interobserver agreement was found (Kendall tau-b,  $p<0,01$ ), and a significant intra-observer agreement too (Kendall tau-b,  $p<0.0005$ ).

**Conclusions** Experts' clinical diagnosis test (ECDT) could be used as gold standard for VFE. Later on, Cephalometric measures should be evaluated using the experts' clinical diagnosis as gold standard.

**Key Words:** Cephalometry, vertical dimension, physical examination (source: MeSH, NLM).

## RESUMEN

**Objetivo** Aunque las medidas cefalométricas se utilizan para evaluar el exceso vertical facial (VFE), ninguna de ellas ha sido validada contra un patrón de oro. Además, hay diferencias entre los resultados cefalométricos y la evaluación clínica. Determinar si el diagnóstico clínico de expertos (ECDT) puede ser usado como prueba de oro en casos severos de exceso vertical facial (VFE)

**Materiales y Métodos** Consenso de expertos (Método Delphi,  $n=12$ ) para determinar si ECDT podría ser usado como prueba de oro en casos severos de VFE. Para ello se construyó escala con 9 características clínicas reportadas en la literatura. Validación: Muestra por conveniencia de 24 pacientes, evaluada por 3 expertos 2 veces. Se realizó análisis de Factores Principales. La consistencia interna fue evaluada con el Coeficiente alfa de Cronbach. La concordancia inter e intra-observador se evaluó con el coeficiente de concordancia Kendall. Las medianas de ECDT fueron comparados entre los grupos con y sin VFE usando la prueba de Kruskal Wallis.

**Resultados** Once de doce expertos consensuaron que el diagnóstico clínico podría usarse como prueba de oro para VFE. Se elaboró una escala con 6 ítems, luego del análisis de factores principales. La consistencia interna fue alta (Cronbach's = 0.8051). Las medianas de ECDT fueron diferentes en los grupos con y sin VFE en la primera ( $p<0.0001$ ) y segunda evaluación ( $p<0.0001$ ). Hubo una alta concordancia interobservador (Kendall  $p<0,01$ ), e intraobservador (Kendall  $p<0.0005$ ).

**Conclusiones** ECDT puede ser utilizado como prueba de oro para diagnosticar el exceso vertical facial severo.

**Palabras Clave:** Cefalometría, técnicas y procedimientos diagnósticos, dimensión vertical, examen físico (*fuentes: DeCS, BIREME*).

Cephalometry is a diagnostic tool used in orthodontics that is questioned for two reasons: a. Most of the average measures were derived from biometric averages obtained from small samples considered aesthetically ideal and b. most of the population does not fit into those measures due to the known variability found in all the biological and ethnic characteristics analysed (1). Hence, it's considered that such measures are limited to individual diagnosis (2-3). On the other hand, clinicians prefer to use quantitative diagnostic methods that ease generalization and not considering each individual independent and unique characteristics. Therefore, it's necessary to evaluate the amount of information obtained from cephalometric analysis. The Receptor Operator Characteristic curve would allow to evaluate the diagnostic accuracy of cephalometric measures, but its use is not possible because is necessary to have a gold standard to compare them (4-6). Many investigators have reported studies that allow to evaluate sensitivity and specificity of the vertical and

sagittal cephalometric measurements using as gold standard: overbite, open anterior bite and interoclusal relations in first molars measures (7-9). None of them used facial clinical evaluation as a gold standard. The aim of this study was to validate the experts' clinical diagnosis test (ECDT) as a way of measuring the facial vertical excess (VFE).

## MATERIALS AND METHODS

### Design

The study was developed in two phases: First phase consisted in an experts' consensus, selecting 12 from a list of 40 dentists from the most important schools of dentistry of the Country that fulfilled the expert's definition criterion (convenience sampling). Each one of the clinicians was evaluated with a self responded standardized survey to avoid the interviewer's bias. The Delphi method was used to obtain the general consensus which facilitates the study conduction and avoids the introduction of a leadership bias likely to appear in a group consensus methodology (10-11). On the second phase, 3 experts were selected. They evaluated independently and blindly the esqueletical characteristics of the 24 patients. The sample size was determined assuming a probability of disagreement of 5 % with a maximum disagreement probability of 15 %. Considering that for this study it was more relevant to obtain an agreement level in the diagnosis of the facial vertical excess, and not in its absence, 3 subjects with apparent vertical excess were included for each subject without apparent vertical excess for the evaluation (3:1 ratio). All patients that arrived consecutively to the Postgraduate Program Clinics in Dentistry of the National University of Colombia and Private Clinics were included during the period from February 2003 until January 2004.

### Procedure

The clinical evaluation on the patients was made from sagittal and frontal clinical photographs. To evaluate intra-observer agreement, 6 to 8 days were awaited between the first and second evaluation.

### Definition of terms

**Expert:** Nationally recognized Orthodontists or Maxillofacial Surgeons graduated from several national or international dentistry schools, with more than 10 years of clinical experience and an academic bond to a recognized postgraduate program.

**Experts' Clinical Diagnosis Test:** Clinical evaluation of vertical facial excess by the selected experts using a questionnaire. Initially were included nine clinical

variables of routine use in the diagnosis and evaluation of the extra oral facial characteristics of an orthodontic patient, reported in the literature (12-14). These variables were: Facial profile, Facial type, Anterior lower face height, Size of the body and mandibular ramus, mandibular plane inclination, gonial angle, menton position and anterior lower face height proportion. After principal factor analysis, ECDT included, finally, six variables: Facial profile, Anterior lower face height, Size of the body and mandibular ramus, gonial angle, and menton position. Each one of the variables were re-categorized as classifier of VFE or not.

Qualitative appreciation of facial vertical excess (QA-VFE): Subjective global evaluation by the expert considering if the patient had or hadn't a diagnosis of facial vertical excess. If the appreciation agreed in both evaluations (positive or negative) individuals were classified as positive or negative. In case of disagreement between both evaluations, individuals were classified as undetermined. This last subjects were classified as without VFE when Receiver Operative Characteristics of ECDT was made.

Total score: Total punctuation of ECDT.

Intra-observer agreement: Level of agreement between the results of the first and the second evaluation by each one of the experts, for each one of the study variables.

Inter-observer agreement: Level of agreement between experts for every one of the variables used in clinical diagnosis, during both evaluations.

### Analysis

Principal factor analysis (varimax rotation) was used to build ECDT. Internal consistency was evaluated with Cronbach's alpha coefficient. Inter-observer and intra-observer agreement was measured using Kendall tau-b concordance coefficient. ECDT's results were compared among three groups of VFE (with, without and undetermined, from QA-VFE), using Kruskal Wallis test. Shapiro Wilks test was used to establish if the variables had a normal distribution. Receiver Operative Characteristics of ECDT was made with the second evaluation of subjects. The database was typed in an Excel® sheet, and the statistical analysis was made with STATA v8.1.

## RESULTS

Eleven of the twelve experts agreed that clinical diagnosis could be used as a gold standard to identify evident facial vertical excess. Therefore, each of them was interviewed again independently, letting them know the results of the initial consensus and giving them the option to reassure or change their previous concept; all of them confirmed their initial opinion.

Principal factor analysis showed a 99,6 % of cumulated variance in two factors (Table 1), represented by six of nine included variables (varimax rotation) (Table 2). The final instrument included: Facial profile, Anterior lower face height, Size of the mandibular body, Mandibular plane inclination, Gonial angle and Menton position. Internal consistency of variables was high (Cronbach's  $\alpha = 0.80591$ ). ECDT was clinically and statistically different in patients qualified through QA-VFE in both, first and second evaluation (Table 3). Inter-observer agreement was high in both, first and second evaluation (Kendall tau-b,  $p < 0.01$ ). Intra-observer agreement was high too (Kendall tau-b,  $p < 0.0005$ ). ECDT showed a sensitivity of 86,3 %, a specificity of 75 %, an accuracy of 81,9 % and area under ROC curve of 0.8620 (Figure 1).

**Table 1.** Eigenvalues and cumulated variance. Principal factor analysis

Factor	Eigenvalue	Difference	Proportion	Cumulative
1	3.26	2.14	0.74	0.74
2	1.11	0.77	0.25	0.99
3	0.34	0.15	0.07	1.07
4	0.18	0.10	0.04	1.11
5	0.08	0.13	0.01	1.13
6	-0.04	0.05	-0.01	1.12
7	-0.10	0.06	-0.02	1.10
8	-0.16	0.12	-0.03	1.06
9	-0.28		-0.06	1.00

**Table 2.** Principal factor analysis. Varimax rotated factors loading

Variable	1	2	3	4	Uniqueness
Facial profile	0.101	<b>0.753</b>	0.003	-0.074	0.416
Facial type	0.368	0.038	0.276	-0.433	<b>0.598</b>
Anterior lower face height	<b>0.631</b>	0.110	0.054	-0.388	0.435
Size of the mandibular body	<b>0.557</b>	0.162	0.398	-0.059	0.479
Size of mandibular ramus	0.013	0.390	0.058	-0.026	<b>0.777</b>
Mandibular plane inclination	<b>0.933</b>	0.131	0.046	-0.093	0.099
Gonial angle	<b>0.909</b>	0.085	0.138	-0.015	0.144
Menton position	0.275	<b>0.696</b>	-0.009	0.041	0.437
Anterior lower face height proportion	0.329	-0.117	0.491	-0.143	<b>0.610</b>

**Table 3.** Median and 95 % Confidence interval of ECDT score in patients qualified "with", "without" or "undetermined" of vertical facial excess

Variable	n	Median	[95 % Conf. Interval]	
Patients without VFE at first and second QA_VFE				
ECDT score1	22	1	0	3
ECDT score2	22	0,5	0	1
Patients without agreement about VFE at first and second QA_VFE				
ECDT score1	6	4	3	5.9
ECDT score2	6	2	0.1	5.8
Patients with VFE at first and second QA_VFE				
ECDT score1	44	5	4	5
ECDT score2	44	4	3	4

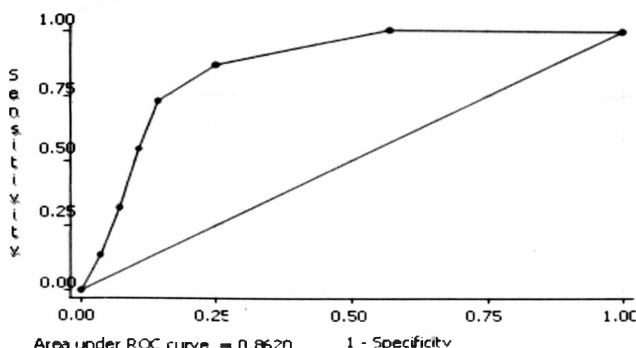
QA\_VFE: Qualitative appreciation of vertical facial excess

Score 1=first evaluation; Score 2=second evaluation

Kwallis,  $p=0.0001$  in both, first evaluation and second evaluation

## DISCUSSION

The ECDT is represented by six of nine clinical variables used to facial evaluation, and has the ability to discriminate patients with severe VFE. Also, inter and intra-observer agreement shows a high reproducibility of results. Clinical diagnosis should be considered as a gold standard of facial vertical excess, in severe cases, especially those related with the lower third of the face where cephalometric measures are usually used as a part of the diagnosis (15-17). The search for a gold standard based on clinical diagnosis is a remarkable fact of this study. Until now, other published studies have used controvertible radiographic or intra-oral clinical characteristics as gold standard.

**Figure 1** Roc curve of ECDT (second evaluation of subjects)

These studies use the presence of anterior open bite or the measures of vertical overbite as a gold standard. The presence of an anterior open bite does not mean necessarily a skeletal alteration as could happen in patients with

thumb sucking and tongue-thrusting. In other words, a patient with skeletal vertical characteristics could or could not have an open bite. Therefore it is relevant to propose a gold standard from the clinical evaluation, obtained from the experts, although limited to severe cases that allows in further studies to establish the level of association among the clinical aspects considered as gold standard and cephalometric aspects of vertical evaluation. Additionally, it will allow the use of the ECDT as a gold standard in sagittal malocclusions, comparing to previous studies or evaluating the need and reproducibility of new cephalometric analysis (18). The questionnaire made to evaluate the facial vertical characteristics of the sample, included all the variables mentioned in the literature associated with vertical facial excess. It's very unusual to find all of them in one subject, due to the individual variability. Principal factor analysis let us to reduce to six the variables in the test. It is really difficult to get the nine variables included initially, in a patient. Values obtained with the ECDT in the VFE subgroups patients were definitively different between them. ECDT showed higher sensibility than specificity in the study patients when QA-VFE made by experts was used as gold standard. This could be related with the higher proportion of subjects with apparent VFE in the sample. Operative characteristics of ECDT were very good. Clinical diagnosis made by experts is not available in the daily practice. Because of that, an instrument that let the clinician to know the presence or not of VFE in a patient is really useful. However, further comparisons will have to be made with cephalometric tests, because they have been used routinely without having had a validation process against clinical evaluation. This study allows rationalizing the use of radiographs to obtain a skeletal diagnosis of vertical abnormalities. It also provides a way to objectively identify which one of the cephalometric variables is more reliable.

The difficulty to change a test that is being used as gold standard is evident when this has not been validated as gold standard and when there's not a recognized better gold standard, as it occurred in this study. Within this study, the experts considered that the ECDT could be used as a gold standard for vertical facial excess. The high inter and intra-observer agreement level and the ability to discriminate VFE, suggest the ECDT as a valid and reliable test for making the diagnosis of VFE. The limitation of not having a better gold standard to validate new diagnostic tests has been shown in other clinical scenarios, in which clinical evaluation is pretended to be replaced by laboratory test (19). The use of ECDT, as a gold standard for vertical facial excess, remains as an option for the academic community. The ECDT can be used as a way to measure the severe vertical facial excess. Considering the fact that there's not a way to use experts to diagnose vertical excess in our daily practice,

it is necessary to compare different diagnostic cephalometric tests' results with ECDT, using ECDT as gold standard. This will be the subject of the next phase of this investigation. With this study the patients' clinical evaluation is made again relevant, and it is proposed a more rational use of other invasive and expensive diagnostic tests ♦

**Acknowledgements.** To the DIB for sponsoring this research work and to Doctors Cesar Beltrán and Pilar Idrobo for their help in collecting the samples.

## REFERENCES

1. Jacobson A, Caufield PW. Introduction to Radiographic cephalometry. Philadelphia, PA: Lea & Febiger; 1985. p. 1-13.
2. Hixon EH. The norm concept and cephalometrics. *Am. J. Orthod* 1956; 42: 898-906.
3. Horowitz SL, Hixon EH. The nature of orthodontic diagnosis. Saint Louis, MO: The C.V. Mosby Company; 1966. p. 299-302.
4. Ardila E. Estrategias Diagnósticas en epidemiología Clínica. *Revista Facultad de Medicina Universidad Nacional de Colombia* 1994; 42: 157-165.
5. Browner WS NT, Cummings S. Designing a new study. In: Helley S, Cummings S, Browner W, Grady D. Chapter 9. N.H, editors. *Designing Clinical Research*. Philadelphia, PA, Lippincott Williams & Wilkins; 2000.
6. Riegelman RK, Hirsch R. Cómo estudiar un estudio y probar una prueba: lectura crítica de la literatura médica. 3rd ed. Organización Panamericana de la Salud. 1998. p. 110-122.
7. Hildebolt CF, Wardlaw DW, Smith RJ, Hertweck DW. Cephalometrics of anterior open bite: A receiver operating characteristic (ROC) analysis. *Am J Orthod Dentofac Orthop* 1992; 101: 234-242.
8. Freudenthaler JW, Celar AG, Schneider B. Overbite depth and anteroposterior dysplasia indicators: the relationship between occlusal and patterns using the receiver operating characteristic (ROC) analysis. *European Journal of Orthodontics* 2000; 22(1): 75-83.
9. Unae KH, Young HK. Determination of Class II and Class III skeletal patterns: Receiver operating characteristic (ROC) analysis on various cephalometric measurements. *Am J Ortho Dentofacial Orthop* 1998; 113(5): 538-45.
10. Maercke AM, Eika B. What are the clinical skills levels of newly graduated physicians? Self - assessment study of an intended curriculum identified by a Delphi process. *Medical Education* 2002; 36: 472-478.
11. Steward J. Is the Delphi technique a qualitative method? *Medical Education* 2001; 35: 922-923.



12. Bishara ES. Textbook of Orthodontics. 1ra Edición, Philadelphia,PA: W.B Saunders Company; 2001.p.418.
13. Vaden JL, Pearson LE. Diagnosis of the vertical dimension. *Seminars in Orthodontics* 2002; 8: 120-129.
14. Subtelny JD. Early Orthodontic Treatment. 1ra Edition. Carol Stream, Illinois: Quintessence Publishing Co, Inc; 2000. p. 183-205.
15. Steiner C. Cephalometrics for you and me. *Am. J. Orthod* 1953; 39: 729 -755.
16. Sassouni V. Analysis of dentofacial vertical proportions. *Am. J. Orthod* 1964; 50: 801-823.
17. McNamara JA. A method of cephalometric evaluation. *Am. J. Orthod* 1984; 86: 449 - 469.
18. Baik CY, Ververidou M. A new approach of assessing sagittal discrepancies: the Beta angle. *Am J Ortho Dentofacial Orthop* 2004; 126 (1):100-5.
19. Bradshaw CS, Morton AN, Garland SM, Horvath LB, Kuzevska I, Fairley CK. Evaluation of a point-of-care test, BVBlue, and clinical and laboratory criteria for diagnosis of bacterial vaginosis. *J Clin Microbiol* 2005; 43 (3):1304-8.