

Journal section: Oral Surgery
Publication Types: Research

doi:10.4317/medoral.16.e68
<http://dx.doi.org/doi:10.4317/medoral.16.e68>

Validation of a computer-assisted system on classifying lower third molars

Isidoro Cortell-Ballester ¹, Nieves Almendros-Marqués ², Leonardo Berini-Aytés ³, Cosme Gay-Escoda ⁴

¹ DDS. Resident of the Master of Oral Surgery and Implantology. University of Barcelona Dental School

² DDS. Associate Professor of Oral Surgery. Professor of the Master of Oral Surgery and Implantology. University of Barcelona Dental School. UB-IDIBELL Research group

³ DDS, MD, PhD. Assistant Professor of Oral Surgery. Professor of the Master of Oral Surgery and Implantology. Dean of the University of Barcelona Dental School. UB-IDIBELL Research group

⁴ MD, DDS, PhD. Chairman of Oral and Maxillofacial Surgery. Director of the Master of Oral Surgery and Implantology. University of Barcelona Dental School. Oral and maxillofacial surgeon of the Teknon Medical Center, Barcelona (Spain). UB-IDIBELL Research group

Correspondence:
Centro Médico Teknon,
C/ Vilana 12,
08022 Barcelona (Spain)
cgay@ub.edu

Cortell-Ballester I, Almendros-Marqués N, Berini-Aytés L, Gay-Escoda C. Validation of a computer-assisted system on classifying lower third molars. Med Oral Patol Oral Cir Bucal. 2011 Jan 1;16 (1):e68-73.
<http://www.medicinaoral.com/medoralfree01/v16i1/medoralv16i1p68.pdf>

Received: 11-02-2010
Accepted: 31-05-2010

Article Number: 16830 <http://www.medicinaoral.com/>
© Medicina Oral S. L. C.I.F. B 96689336 - pISSN 1698-4447 - eISSN: 1698-6946
eMail: medicina@medicinaoral.com

Indexed in:

Science Citation Index Expanded
Journal Citation Reports
Index Medicus, MEDLINE, PubMed
Scopus, Embase and Emcare
Indice Médico Español

Abstract

Objectives: The present study evaluates the reliability of the Radio Memory[®] software (Radio Memory; Belo Horizonte, Brasil.) on classifying lower third molars, analyzing intra- and interexaminer agreement of the results.

Study Design: An observational, descriptive study of 280 lower third molars was made. The corresponding orthopantomographs were analyzed by two examiners using the Radio Memory[®] software. The exam was repeated 30 days after the first observation by each examiner. Both intra- and interexaminer agreement were determined using the SPSS v 12.0 software package for Windows (SPSS; Chicago, USA).

Results: Intra- and interexaminer agreement was shown for both the Pell & Gregory and the Winter classifications, $p < 0.01$, with 99% significant correlation between variables in all the cases.

Conclusions: The use of Radio Memory[®] software for the classification of lower third molars is shown to be a valid alternative to the conventional method (direct evaluation on the orthopantomograph), for both clinical and investigational applications.

Key words: Third molar, classification, software validation.

Introduction

In both Medicine and Dentistry, history and tradition have contributed to the use of numerous diagnostic methods or criteria without prior validation. New diagnostic methods, scales or classifications should be taken with caution, and not accepted until their validity, reliability and reproducibility is demonstrated. The validity of a diagnostic test is determined by its sensitivity and specificity, that is to say, its capacity to detect cases that present the defining features, as well as discarding those that do not. Reliability refers to the degree to which similar results are obtained when measurements are repeated. In current practice, knowledge about the quality of the diagnostic methods is essential.

Classically, lower third molars have been classified according to the Pell & Gregory and Winter (1) criteria, that distribute them into different groups according to the relative depth respect to the occlusal plane, the position in relation to the mandibular ramus and the inclination of the longitudinal axis. These three radiologically-determined anatomical parameters have provided the clinical information regarding position and inclination of the third molars, allowing greater precision in treatment planning. Beginning with these classifications, Pederson (2) designed a difficulty index for the surgical extraction of lower third molars that has been broadly accepted and used for the planning of this surgical procedure. In recent years the utility of the Pell & Gregory and the Winter (1) classifications as well as of the Pederson difficulty index has been questioned by diverse authors (2-5) who consider them inadequate for classifying the position and inclination of third molars, as well as inaccurate in predicting the difficulty of surgery.

These studies have evidenced the need to develop new methods to increase precision in determining the position and inclination of lower third molars.

For this reason, we decided to study the reliability and reproducibility of the Radio Memory® software (Radio Memory; Belo Horizonte, Brazil. License to the University of Barcelona) when applied to the classification of lower third molars. This software consists of several programs that allow different individualized tracings to be made and measured on the radiographs, applying them to diagnosis in different odontological disciplines such as Orthodontics, Implantology and Oral Surgery.

Material and Methods

A descriptive, observational study of agreement was made, in which 280 lower third molars of consecutive patients undergoing exodontia were selected. Patients with third molar germs below Nolla stage 8 were excluded from the study, as were those in which the first and/or second lower molars were absent; and those with very damaged third molars or those reduced to radicular remains.

The corresponding orthopantomographs were scanned at a resolution of 350 dpi and analyzed separately by 2 examiners using the Radio Memory® software (Radio Memory; Belo Horizonte, Brazil. License to the University of Barcelona) to record the position and inclination of the lower third molars according to the criteria formulated by Pell & Gregory and Winter (1). The procedure was repeated 30 days later by each examiner to establish the level of intra- and interexaminer agreement between the two observations.

Data collection

Four lines were traced on the scanned orthopantomographs (Fig. 1) as in previous studies (5,6):

- Line of the occlusal plane, established by the occlusal surface of the lower first and second molars.
- Cervical line, established by the cervical bone limit of the lower second molar.

These lines allow the molars to be classified according to their depth (Pell & Gregory position A, B and C).

- Line of the anterior margin of the ascending mandibular ramus that indicates the position of the molar with respect to the ramus (Pell & Gregory class I, II and III).

- Line of the longitudinal axis of the third molar, which forms an angle with the occlusal plane delimiting its inverted, horizontal, mesioangular, vertical or distoangular inclination (Winter classification).

Obtainment of the angle between the occlusal plane or line parallel to it and the longitudinal axis of the impacted third molar, in turn, allowed objective classification of the third molars within the Winter subclasses, as follows:

- Third molars with a negative angle ($< 0^\circ$) were considered to be inverted.
- Third molars with an angle between $0-30^\circ$ were considered to be horizontal.
- Third molars with an angle between $31-60^\circ$ were considered to be mesioangular.
- Third molars with an angle between $61-90^\circ$ were considered to be vertical.
- Third molars with an angle $> 90^\circ$ were considered distoangular.

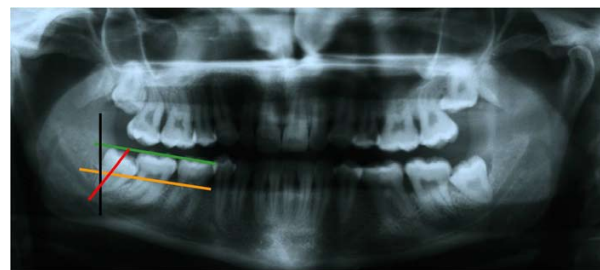


Fig. 1. Four lines were traced on the scanned orthopantomographs to assess the Winter and the Pell & Gregory classifications.

Statistical analysis

Data were analyzed using the Statistical Package for Social Science version 12.0 (SPSS; Chicago, USA. License to the University of Barcelona) applying the following statistical tests:

- Bivariate correlations (Pearson, Kendall’s Tau, Spearman rho) to evaluate intraexaminer agreement with respect to the Winter (taken as a categorical variable) and Pell & Gregory classifications.
- Bivariate correlations (Pearson) to evaluate intraexaminer agreement with respect to the Winter classification (taken as a numerical variable).
- Bivariate correlations (Pearson, Kendall’s Tau, Spearman rho, Kappa) to evaluate the interexaminer agreement in the Winter and Pell & Gregory classifications.
- Calibration by linear regression to evaluate the intra- and interexaminer agreement for both classifications.

Results

Intraexaminer agreement

The data analysis evidenced a high intraexaminer agreement for both the Pell & Gregory and the Winter classifications ($p < 0.01$) (considered as numerical and categorical variables), with 99% significant correlation between variables in all cases.

Nevertheless, for the Pell & Gregory classification the degree of intraexaminer agreement was slightly less for one of the two examiners. Even so, a high level of homogeneity was indicated.

Table 1 shows the intraexaminer agreement results obtained by applying the Pearson, Kendall’s Tau and Spearman rho correlation coefficients. Table 2 shows the intraexaminer agreement obtained by linear regression.

Table 1. Intra- and interexaminer agreement for the Winter classification (taken as categorical and numerical variables) and Pell & Gregory classification.

Winter (categorical)				
	Pearson	Kendall’s Tau	Spearman rho	Kappa
Examiner 1	0.959	0.926	0.942	-
Examiner 2	0.973	0.950	0.962	-
First Observation	0.947	0.910	0.932	0.845
Second Observation	0.935	0.887	0.914	0.803
Winter (numerical)				
Examiner 1	0.995	-	-	-
Examiner 2	0.992	-	-	-
First Observation	0.878	-	-	-
Second Observation	0.890	-	-	-
Pell & Gregory				
Examiner 1	0.866	0.840	0.877	-
Examiner 2	0.985	0.973	0.982	-
First Observation	0.716	0.651	-	-
Second Observation	0.724	0.662	-	-

The values vary between 0 and 1, where 0 = minimum and 1 = maximum agreement.

Table 2. Intra- and interexaminer agreement by linear regression for the Winter classification (taken as categorical and numerical variables) and for the Pell & Gregory classification.

	Examiner	Regression	R ²	Observation	Regression	R ²
Winter (categorical)	1	Obs1-Obs2*	0.921	1	Exam1-Exam2 [†]	0.897
		Obs2-Obs1**			Exam2-Exam1 [‡]	
	2	Obs1-Obs2*	0.947	2	Exam1-Exam2 [†]	0.875
		Obs2-Obs1**			Exam2-Exam1 [‡]	
Winter (numerical)	1	Obs1-Obs2*	0.990	1	Exam1-Exam2 [†]	0.772
		Obs2-Obs1**			Exam2-Exam1 [‡]	
	2	Obs1-Obs2*	0.984	2	Exam1-Exam2 [†]	0.792
		Obs2-Obs1**			Exam2-Exam1 [‡]	
Pell & Gregory	1	Obs1-Obs2*	0.750	1	Exam1-Exam2 [†]	0.513
		Obs2-Obs1**			Exam2-Exam1 [‡]	
	2	Obs1-Obs2*	0.971	2	Exam1-Exam2 [†]	0.524
		Obs2-Obs1**			Exam2-Exam1 [‡]	

* Dependent variable: Observation 1.
 ** Dependent variable: Observation 2.
 † Dependent variable: Examiner 1.
 ‡ Dependent variable: Examiner 2.

Interexaminer agreement

For the interexaminer agreement the results showed a very high agreement for both the Pell & Gregory and the Winter classifications (p<0.01) (considered as numerical and categorical variables), with 99% significant correlation between variables in all cases.

For the Winter classification (numerical variable) the Kappa value was also calculated, this was higher than 0.80 in all the observations, demonstrating a very high level of homogeneity between examiners.

The level of interexaminer agreement was lower for the Pell & Gregory classification, although, as with the intraexaminer agreement. Even so, the results indicated a high level of homogeneity.

Table 1 shows the interexaminer agreement values obtained by the Pearson, Kendall’s Tau and Spearman rho correlation coefficients and Kappa value. Table 2) shows the interexaminer agreement values obtained by linear regression.

Discussion

The results of this study suggest that the Radio Memory[®] software is a reliable method for use in the classification of lower third molars. The intra- and interexaminer agreement was significant for both classifications (p<0.01) although the results showed a higher agreement for the Winter classification. In this classification the values are distributed numerically among the different categories, thus greatly reducing examiner subjectivity. However, the Pell & Gregory classification contains more categories, resulting from the combination of two consecutive position parameters which may overlap, therefore producing a higher number of combinations. In addition, the categories are determined by lines traced directly by the examiner, and may therefore contain a certain degree of bias. Nevertheless, a high level of homogeneity was obtained, thus the line tracing may be interpreted as providing sufficient objectivity to classify the third molars correctly.

When analyzing interexaminer agreement by linear regression for the Winter classification (according to

the numerical angle value) we observe that more disparate values are obtained and the confidence interval increases, showing a lower level of agreement (Tables 1 and 2). This result could be interpreted as a deficit in the validity of the software. However, these values are clinically insignificant, since in the practice the Winter classification (divided into five different categories) is used to determine the inclination of the third molars and the agreement for this classification demonstrated a very high level of homogeneity. In contrast, a poor result for the Winter classification (by categories) would have had much greater importance, since the method would not have been validated for determining the lower third molar inclination according to the Winter criteria.

Technological developments over recent years have brought new diagnostic methods and new possibilities in all fields of investigation. However, any new diagnostic method, scale or classification should undergo validation testing to determine the reliability and reproducibility. In Dentistry, diverse studies (7-11) have been published with the aim of assessing the reliability and validity of different diagnostic methods. Thus, Schmitter et al. (7) assessed the reliability and validity of orthopantomography in determining condylar morphology, studying the sensitivity and specificity of this method, and taking magnetic resonance imaging (MRI) and the clinical findings as the gold standard. Their study, being a validation test, differs from ours in that it used other reference methods for the evaluation (internal validity). In our case, we assess the external validity of the Radio Memory[®] software, since we determine the homogeneity (interexaminer agreement) and the reproducibility (intraexaminer agreement) of the results.

On the other hand, Araujo et al. (8) assess the reproducibility of probing depth measurement when using a constant-force electronic probe analyzing intra- and interexaminer variability of results. Therefore, these authors studied the external validity of the test, presenting some similar methodological features as the present study.

We find only one study regarding lower third molar position (9). This article evaluated the reliability of panoramic radiography in determining the relationship of third molars with the inferior alveolar nerve, contrasting the results with images obtained by computed tomography. Although the authors analyzed interexaminer agreement, no intraexaminer agreement test was made to determine the reproducibility of the results after repeated evaluations.

Regarding classification of lower third molars position and inclination, there are no previous studies assessing the validity of new methods designed for this purpose. It seems that using a computer-assisted system in the classification of lower third molars reduces examiner interpretation bias and is sufficiently objective for both

correct surgical planning, and clinical investigation where patient selection with a given third molar position is required.

The method itself is constrained by the time required to carry out the relevant measurements to determine the position and inclination of the lower third molars, as well as by the technical equipment needed to carry them out.

In conclusion, use of Radio Memory[®] software for the classification of lower third molars is demonstrated to be a valid alternative to the conventional method (direct measurement on the orthopantomograph) for both clinical and investigational applications.

In spite of the good results obtained in the present study further studies with a larger number of examiners are necessary to confirm these findings.

References

References with links to Crossref - DOI

1. Diniz-Freitas M, Lago-Méndez L, Gude-Sampedro F, Somoza-Martin JM, Gándara-Rey JM, García-García A. Pederson scale fails to predict how difficult it will be to extract lower third molars. *Br J Oral Maxillofac Surg.* 2007;45:23-6.
2. García AG, Sampedro FG, Rey JG, Vila PG, Martin MS. Pell-Gregory classification is unreliable as a predictor of difficulty in extracting impacted lower third molars. *Br J Oral Maxillofac Surg.* 2000;38:585-587.
3. Yuasa H, Kawai T, Sugiura M. Classification of surgical difficulty in extracting impacted third molars. *Br J Oral Maxillofac Surg.* 2002;40:26-31.
4. Susarla SM, Dodson TB. Risk factors for third molar extraction difficulty. *J Oral Maxillofac Surg.* 2004;62:1363-71.
5. Almendros-Marqués N, Berini-Aytés L, Gay-Escoda C. Evaluation of intraexaminer and interexaminer agreement on classifying lower third molars according to the systems of Pell and Gregory and of Winter. *J Oral Maxillofac Surg.* 2008;66:893-9.
6. Almendros-Marqués N, Berini-Aytés L, Gay-Escoda C. Influence of lower third molar position on the incidence of preoperative complications. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2006;102:725-32.
7. Schmitter M, Gabbert O, Ohlmann B, Hassel A, Wolff D, Rammelsberg P, et al. Assessment of the reliability and validity of panoramic imaging for assessment of mandibular condyle morphology using both MRI and clinical examination as the gold standard. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2006;102:220-4.
8. Araujo MW, Hovey KM, Benedek JR, Grossi SG, Dorn J, Wactawski-Wende J, et al. Reproducibility of probing depth measurement using a constant-force electronic probe: analysis of inter- and intraexaminer variability. *J Periodontol.* 2003;74:1736-40.
9. Nakagawa Y, Ishii H, Nomura Y, Watanabe NY, Hoshiba D, Kobayashi K, et al. Third molar position: reliability of panoramic radiography. *J Oral Maxillofac Surg.* 2007;65:1303-8.
10. Raitz R, Correa L, Curi M, Dib L, Fenyo-Pereira M. Conventional and indirect digital radiographic interpretation of oral unilocular radiolucent lesions. *Dentomaxillofac Radiol.* 2006;35:165-9.
11. Pecoraro M, Azadivatan-le N, Janal M, Khocht A. Comparison of observer reliability in assessing alveolar bone height on direct digital and conventional radiographs. *Dentomaxillofac Radiol.* 2005;34:279-84.

Acknowledgments

This study has been performed by the research group “Odontological and Maxillofacial Pathology and Therapeutics” of the UB-IDIBELL Institute, with the economic support of the Oral Surgery teaching-attendance agreement between the University of Barcelona, the Integrated Health-Care Consortium and the Catalan Health Service of the Generalitat of Catalunya (Spain).

The authors thank César Pérez-López, Associate Professor of the Department of Statistics and Operative Research III, Statistics School, Madrid Complutense University (Spain), for his contribution to the statistical analysis of the study.