Use of the Objective Grading System of the American Board of Orthodontics to Evaluate Treatment Outcomes at the Orthodontic Graduate Program Clinic, University of Puerto Rico, 2007-2008

José J. Santiago, DMD; Carlos J. Martínez, DMD

Objective: The Objective Grading System (OGS) was introduced in 1999 by the American Board of Orthodontics (ABO) as an instrument to reduce subjectivity when evaluating cases submitted to the ABO for examination. The objectives of this study were 1) to employ the OGS to determine the percentage of treated and completed cases from the University of Puerto Rico's (UPR) Orthodontic Graduate Program Clinic that would have earned a passing OGS score (according to the ABO standards) and 2) to assess the contribution of various patient characteristics and factors to this score.

Methods: A total of 64 cases completed during 2007 and 2008 met the inclusion criteria and were evaluated by a calibrated examiner using the OGS. Logistic regression and multiple regression analyses were performed to assess the association between the explanatory variables and the overall OGS score.

Results: Of the cases evaluated, 18.8% received a passing OGS score of <20, 34.4% received a borderline score of 20-30, and 46.9% received a failing score of >30. The mean OGS score was 32.17 ± 13.03 points, similar to the mean OGS score of 34.36 ± 10.39 reported in 2004 by a similar study.

Conclusion: This study demonstrate that 53% of the completed cases at the university clinic obtained a potential passing score as per the ABO (OGS <30 points). Multiple and logistic regression analyses could neither explain the relationship between the explanatory variables and the OGS scores nor predict the probability of a particular case's passing when all variables were considered. [*P R Health Sci J 2012;31:29-34*]

Key words: Orthodontic treatment outcomes, Objective Grading System, American Board of Orthodontics, Graduate program, Dental casts

The assessment of orthodontic treatment outcomes has traditionally been accomplished by evaluating patients' post-treatment records using the subjective opinions and experience of clinicians (1). Until 1998, a standardized method of assessing post-treatment records was not available to objectively evaluate treatment outcomes. A valid and reliable measurement standard was certainly needed to facilitate the evaluation of treatment results. Several indexes and rating systems have been proposed for evaluating orthodontic treatment outcomes (2-5). These indexes compare pretreatment and post-treatment records to determine the quality of the outcome.

The American Board of Orthodontics (ABO) developed the Discrepancy Index (DI) to provide an objective evaluation of case complexity, which evaluation could lead to a better understanding of case difficulty before treatment is started (6). The DI considers common elements of the orthodontic diagnosis with respect to the dental cast's occlusal relationship and cephalometric values as an objective guide to categorize case difficulty. The DI can be used to evaluate the following clinical features of a patient's condition overjet, overbite, anterior open bite, lateral open bite, crowding, occlusion, lingual posterior cross bite, buccal posterior cross bite, ANB angle, IMPA angle, and SN-GoGn angle.

In 1994, an ABO committee was formed to begin the field testing of more precise methods to objectively evaluate orthodontic post-treatment dental casts and panoramic radiographs. In 1999, the ABO introduced the objective grading system (OGS) (7). The purpose of the OGS was to provide a standardized method for evaluating the dental casts of finished cases to determine whether those finished cases met the ABO's

Orthodontics Graduate Program, University of Puerto Rico School of Dental Medicine, San Juan, Puerto Rico

The authors have no conflicts of interest to disclose.

<u>Author correspondence to</u>: José J. Santiago, DMD, University of Puerto Rico School of Dental Medicine, PO Box 365067, San Juan, Puerto Rico 00936-5067. Email: jose.santiago12@upr.edu

standards for tooth alignment. The ABO encourages clinicians to use the OGS "at anytime [sic] in their orthodontic career[s] to determine if they are producing 'Board quality' results" (7).

The OGS detects subtle inadequacies in occlusion with a high degree of precision. It assesses the final occlusion in first, second, and third orders according to the following 8 different occlusal categories: alignment, marginal ridges, buccolingual inclination, occlusal contacts, occlusal relationships, overjet, interproximal contacts, and root angulation. From the starting score of zero, points are deducted for any discrepancy from the ideal occlusion as described by the ABO, and the sum of these deducted points gives us the score for the first phase of the ABO examination; a score of less than 20 points is considered to be a "passing" score; a score greater than 30 points is considered to be a "failing" score; a score between 20 and 30 is considered borderline for passing. This first phase score represents only part of the overall score of the ABO clinical examination for each case report. The ABO also evaluates and carefully scrutinizes the quality of the records, the appropriateness of the treatment plan, and the objectives for the positioning of the maxilla, mandible, maxillary and mandibular dentition and for the facial profile. This supplementary evaluation can add from 0 to several points to the initial phase score. After the supplementary evaluation points are added, all cases with a combined OGS score lower than 30 points will pass the ABO examination while those with scores of more than 30 will fail the examination. With this effective grading system, examinees may grade their own results before the clinical examination in order to determine whether their results will pass ABO standards.

The ABO OGS applies only to the treatment outcome and does not consider either the severity of the malocclusion or the difficulty of the treatment rendered (8).

The OGS has been widely accepted since its implementation in 1999. In the ABO clinical examination of February 2002, the fourth year in which the OGS was used, 89% of the participants passed this examination, the highest passing rate in recent years. Of the 479 cases presented for examination, only 9 (1.9%) were found to be unacceptable because of occlusion; this was the lowest post-treatment case failure rate ever (9).

The ABO's OGS has also been used in several studies to compare orthodontic treatment outcomes achieved by different treatment modalities (1, 5, 8, 10-13).

In 2004, Pinskaya et al. assessed the treatment outcomes achieved in completed cases dating from 1998 to 2000 at the Indiana University's graduate orthodontic clinic.12 In this retrospective evaluation of 521 records, only 39.7% of the finished cases had passing scores for the ABO clinical examination (defined as an OGS score lower than 30 points). Posterior to this study, several changes were implemented at the graduate orthodontic clinic to improve the quality of treatment outcomes (13). In 2006, Knierim et al. conducted a follow-up study at Indiana University that evaluated completed cases (coming from the years 2001 to 2003) and found that of 437 finished cases, 76.6% had a passing score for the ABO clinical examination, an improvement of almost 37% (13).

The objectives of the study described herein were 1) to determine the percentage of treated and completed cases from the University of Puerto Rico's Orthodontic Graduate Program Clinic that would have earned a passing score according to ABO clinical examination using the OGS and 2) to assess the contribution of the following factors to the overall OGS score: age treatment started, year treatment ended, treatment time, gender, attending faculty, patient compliance, and discrepancy index.

Methods

After obtaining the approval from the Institutional Review Board of the Medical Sciences Campus of the University of Puerto Rico, all treated and completed cases from 2007 and 2008 were identified in the archives of the UPR's Orthodontic Graduate Program Clinic.

Patients were included in the study if 1) they began and completed treatment at the UPR's orthodontic graduate clinic, 2) their post-treatment panoramic x-rays and study casts were made no later than 1 year after treatment was completed, 3) their cases conformed to the ABO's clinical examination specifications, which specifications pertain to 10 categories defined by the pretreatment malocclusion or by the Discrepancy Index scores (14).

Patients were excluded from the study if 1) their treatment records, post-treatment study cast, and/or panoramic radiograph were not available; 2) their records were not of sufficient quality for diagnostic purposes; or 3) they had craniofacial anomalies.

The ABO's OGS was used to evaluate and analyze the dental casts and panoramic radiographs of participating patients using the ABO Measuring Gauge (Figure 1). The following criteria were measured: alignment, marginal ridges, buccolingual inclination, occlusal contacts, occlusal relationships, overjet, interproximal contacts, and root angulation. For each criteria, points were deducted using the OGS guidelines and according to the degree to which each criteria deviated from the ideal (7). Each case started with an OGS score of 0, to which the sum of deducted points from the evaluation of casts and panoramic radiograph was added. Scores of 20 points or less are considered to be passing, and scores of 30 or greater fail the clinical examination of the American Board of Orthodontics (1, 7). A score between 20 and 30 was considered borderline for passing, pending the second-phase evaluation (15);these scores were not used in our study or in similar studies reported in the literature (1, 5, 8, 10, 12, 13).

An OGS score intra-examiner calibration (making use of 10 sets of stone models and panoramic radiographs) was performed

prior to data collection. Two weeks after the first measurements were taken, the examiner evaluated the same sets of models and radiographs. A kappa statistic was calculated for intra-examiner reliability and the resulting kappa statistics (k = 0.98) indicated excellent intra-examiner repeatability. The measurements of all dental casts and panoramic radiographs were performed by the principal investigator, a senior resident of the UPR's Orthodontic Graduate Program.
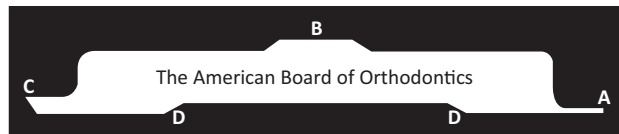

Figure 1. The American Board of Orthodontics Measuring Gauge. A) This portion of the gauge is 1 mm in width and is used to measure discrepancies in alignment, overjet, occlusal contact, interproximal contact, and occlusal relationships. B) This portion of the gauge has steps measuring 1 mm in height and is used to determine discrepancies in mandibular posterior buccolingual inclination. C) This portion of the gauge has steps measuring 1 mm in height and is used to determine discrepancies in marginal ridges. D) This portion of the gauge has steps measuring 1 mm in height and is used to determine discrepancies in maxillary posterior buccolingual inclination.

Patient record numbers were used as the study IDs for the cases and data were categorized according to gender, age, treatment time, residents in charge, faculty in charge, number of missed appointments, Discrepancy Index, and OGS score. Age was recorded for each case but was grouped (for statistical purposes) according to the eruption events in the normal development of a child. The groups were as follows: Group A included those children who were 12 years old or younger and whose second molars had not yet erupted; Group B included those children 13 to 16 years old in whom the second molars had erupted but the third molars had not; Group C included 17- to 23-year-olds in whom third molars had erupted; and Group D, included those who were 24 years old or older in whom dental development had completed. Patient compliance was defined by the principal investigator as the following: Patients with 5 or fewer missed appointments were considered compliant with treatment, while patients with 6 or more missed appointments were considered non-compliant.

Multiple linear regression analysis was employed to predict the OGS score. Logistic regression analysis was then applied to reveal the configuration of the variables for passing or failing the ABO clinical examination.

Results

A total of 128 cases were completed during the observation period of 2007 through 2008, and 64 cases met the inclusion criteria. Of the 64 cases included, 28 cases were completed in 2007 and 36 cases in 2008, which total cases comprised 27 males (42.2%) and 37 females (57.8%).

The minimum age at the beginning of treatment was 3 years and the maximum age was 44 years, with a mean age of 14.06 years (SD of 6.92). Most patients were in group A: 12 years old or younger (n = 34). For the multiple and logistic regression analyses, cases in group C, 17 to 23 y/o (n = 4), and D, 24 and older (n = 6), were grouped together because of the small number of observations for each group. Treatment duration was registered in months, with a minimum duration of 7 months, a maximum of 120 months, and a mean of 51.30 months (SD 25.05) (Table 1).

Table 1. Frequency, percentage, Objective Grading System (OGS) mean, and standard deviation (SD) for selected criteria (n = 64).

	# of patients	% of patients	Mean OGS	SD
Gender				
male	27	42.2	36.85	13.25
female	37	57.8	28.76	11.91
Treatment completed				
2007	28	43.8	31.36	12.90
2008	36	56.4	32.81	13.28
Group ages				
12 or younger	34	53.1	33.53	n.a.
13 to 16 y/o	20	31.2	31.45	n.a.
17 to 23 y/o	4	6.2	26.25	n.a.
24 or older	6	9.4	30.83	n.a.
Patient's compliance				
compliant	48	75.0	30.77	12.23
non-compliant	16	25.0	36.38	14.81
Faculty				
1	21	32.8	29.67	11.36
2	10	15.6	35.20	13.04
3	7	10.9	39.14	14.61
4	13	20.3	34.23	13.45
5	5	7.8	24.00	6.86

n.a.: not available

There were a total of 9 faculty members supervising the residents. Most cases (n = 56/64) were distributed among those faculty members coded from 1 to 5. Cases from faculty members coded 6 through 9 were not used for the multiple regression or logistic regression analyses due to the small number of cases assigned.

There were 48 compliant and 16 non-compliant patients. Of the 48 compliant patients, 21 (33%) were male and 27 (42%) were female; of the 16 non-compliant patients, 6 (9%) were male and 10 (16%) were female.

The Discrepancy Index (DI) was used as a means of establishing a generalized idea of the severity of each case at the beginning of treatment. It was also included as an explanatory variable in the logistic and multiple regression analyses. The minimum DI score obtained was 3 points, while the maximum was 57 points; the mean was 19.42 points (SD 10.74) (Table 1).

The mean OGS score in this study was 32.17 points ± 13.03 and covered a range of scores progressing from a minimum of 14

points to a maximum of 68 points. There were 12 cases (18.8%) that obtained an ABO OGS passing score of less than 20; 22 cases (34.4%) had a borderline score of between 20 to 30 points; and 30 cases (46.9%) obtained a failing score of more than 30 points (Table 2). Overall, 53.2% of the cases have scores that will be accepted by the ABO if all other supplemental aspects of the evaluation of the exam are satisfactorily met.

Table 2. Minimum, maximum, mean, and standard deviation for Objective Grading System (OGS) score, discrepancy index, patient age, and treatment time.

	Minimum	Maximum	Mean	Standard deviation
OGS score	14	68	32.17	13.03
Discrepancy index	3	57	19.42	10.74
Age at treatment start, years	3	44	14.06	6.92
Treatment time, months	7	120	51.30	25.03

Box plots were drawn to evaluate the behavior of the outcome variables (OGS score) with respect to the explanatory variables, these being gender (Figure 2a), age groups (Figure 2b), attending faculty (Figure 2c), patient compliance (Figure 2d), and the year the case was completed (Figure 3). The box plot

of OGS by gender demonstrates that males had a higher mean OGS score (36.85 ± 13.25) than did females (28.76 ± 11.91). OGS by age group showed that patients in the age group of 17 to 23 years had the lowest mean OGS score (26.25). Those 12 years old or younger had a mean OGS score of 33.53; those in the age group of 13 to 16 years had a mean OGS score of 31.45; and those 24 or older had a mean OGS score of 30.83. OGS score by patient compliance demonstrates that patients who were compliant with treatment had a lower mean OGS score (30.77 ± 12.23) than did non-compliant patients (36.38 ± 14.81). OGS by supervising faculty demonstrates that cases supervised by faculty 1 and faculty 5 had the lowest mean OGS scores (29.67 ± 11.36 and 24.0 ± 6.86, respectively), while faculty 2, 3, and 4 had mean OGS scores of 35.20 ± 13.04, 39.14 ± 14.61, and 34.23 ± 13.45, respectively. Finally, we looked at the year in which treatment was completed and observed that the group of patients completed in 2007 had a slightly lower mean OGS score (31.36 points) than did that group of patients completed in 2008 (32.80 points), a difference not considered clinically or statistically significant (p>0.05).

Correlations between the OGS score and DI ($r^2 = 0.032$) and between OGS score and treatment time ($r^2 = 0.034$) were examined, but no associations were observed.
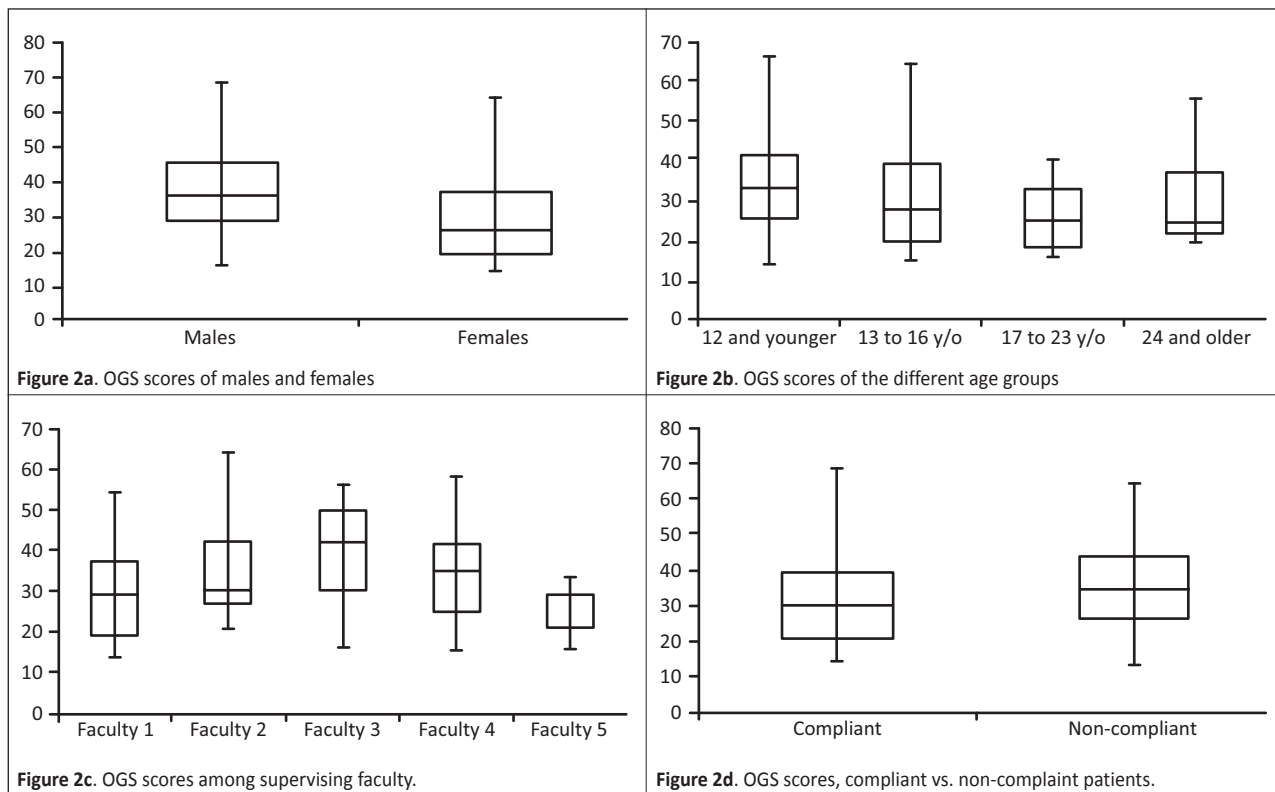

Figure 2a. OGS scores of males and females

Figure 2b. OGS scores of the different age groups

Figure 2c. OGS scores among supervising faculty.

Figure 2d. OGS scores, compliant vs. non-complaint patients.

Figure 2a-2d. Box plots comparing Objective Grading System (OGS) scores in selected factors. The line dividing each box identifies the median OGS score for that set of data. **2a**. Box plot comparing OGS scores of males and females. **2b**. Box plot comparing OGS scores between age groups: 12 years or younger; 13 to 16 years old; 17 to 23 years old and 24 years or older. **2c**. Box plot comparing OGS scores among supervising faculty. **2d**. Box plot comparing OGS scores of compliant vs. non-compliant patients.
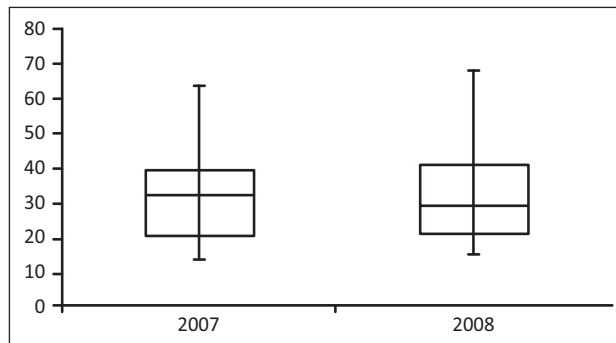
Figure 3. Box plot comparing the Objective Grading System (OGS) scores of the cases completed in 2007 with those of the cases completed in 2008. The line dividing each box identifies the median OGS score for that set of data.

Multiple linear regression analysis was employed to predict the OGS score and Logistic Regression analysis to reveal the configuration of the variables for passing or failing the ABO examination. Multiple linear regression analysis revealed that the p values for each of the parameters were greater than 0.05 and were, therefore, not significant and that the model does not explain the outcome (OGS score). For the logistic regression, the confidence intervals for all variables included the number 1; therefore, none of the factors or patient characteristics under study explained the probability of passing or failing the ABO examination given the other predictors in the model (Pass = OGS score ≤ 30).

Discussion

This observational retrospective study used the ABO's OGS to evaluate the outcome of all patients that completed orthodontic treatment during the period of 2007 to 2008 at the UPR's orthodontic graduate clinic. In the present study, we found that the percentage of patients with a potential passing score (defined as an OGS score of ≤ 30) was 53%. This percentage is higher than the 39.7% reported by the Indiana University for cases treated during the years of 1998 to 2000, but lower than the 76.6% reported in their follow-up study for cases treated from 2001 to 2003 (12-13).

No significant differences were observed in the multiple regression or logistic regression analyses. Nonetheless, some variables, such as age group, patient compliance, gender, and supervising faculty, when analyzed individually, affected the final OGS score to a greater degree. The year of case completion did not seem to affect the OGS mean score. For cases completed in 2007, the mean OGS score was 31.36 points, while those completed in 2008 had a mean OGS score of 32.80 points, suggesting that cases were treated consistently during these years. The lowest OGS scores in our study were attained by patients in the 17 to 23 years age group, , suggesting that this age group may be more receptive and responsible with regard

to their treatment. According to patient records, this age group consists largely of females in or recently graduated college and who either are undergoing treatment for the first time or are receiving r minor relapse corrections on previous treatment . The mean pretreatment age in our study group was 14.06 years, greater than the mean of 12.9 years for groups as reported by the University of Tennessee, Louisiana State University, and the University of Alabama, but similar to that of groups evaluated at the University of Illinois in Chicago (mean, 14.3 years) (1). Compared to the mean age of the participants of the study at Indiana University, the mean age of our participants was less than the mean age of 15.8 years for the period of 1998 to 2000 and 17.8 years for the period of 2001to 2003 (12-13). Although lower, our mean age still confirms the trend in aging patient population seeking orthodontic treatment as reported in recent literature.

Differences were also found between genders, with females receiving a mean OGS score of 28.76, which score is much lower than the males' mean OGS score of 36.85. In addition, patients who were compliant with treatment received a mean OGS score of 30.77 points, much lower than the mean OGS score of 36.38 received by patients who were non-compliant. These findings correlate well with the findings of a similar study reporting that a decrease in compliance resulted in higher OGS scores (13). This same study found an association between OGS scores and treatment time, but we did not observe a relationship between the same variables. This difference may be attributable to the large difference in sample sizes between our study (n = 64) and that reported in the literature (n = 521).

The attending faculty assigned to the different cases also appears to have affected the OGS scores. The cases attended by faculty 1 and 5 obtained a mean OGS score of less than 30 points, while the cases attended by faculty 2, 3, and 4 had mean OGS scores greater than 30 points. These differences may be attributable to the number of cases each faculty contributed to the total group of 64 cases. The number of cases per faculty ranged from 5 to 21. There were no previous studies available in the literature that looked into this association.

In conclusion, of the 64 cases that met the inclusion criteria in the present study, 19% received a passing score of less than 20 points, 34% received a borderline score of between 20 to 30 points, and 47% had a failing score of more than 30 points. Without considering the supplemental ABO assessment, the total possible passing rate was 53% of the cases that had been completed at the UPR's orthodontic graduate clinic. According to the multiple and logistic regression analyses, none of the explanatory variables explained the outcome or the odds ratio when all of the variables were considered.

We recommend the following: initiate follow-up studies in order to identify areas that would contribute to the attainment of a higher OGS score for rectification prior to case completion, establish stricter guidelines for clinical record keeping, and

redefine the compliance variable so that it includes the number of total visits required to finish treatment and the number of missed appointments.

Resumen

Objetivos: El "Objective Grading System" (OGS se introdujo en 1999 por el "American Board of Orthodontics" (ABO) como un instrumento para reducir la subjetividad en la evaluación de los casos sometidos para la evaluación del "ABO". Los objetivos de este estudio fueron: 1) utilizar el "OGS' para determinar el porcentaje de casos tratados que obtendrían una puntaje para aprobar el examen clínico del "ABO" en la clínica del programa graduado de ortodoncia de la Universidad de Puerto Rico y 2) determinar cómo varios factores y características del paciente contribuyen a este puntaje. Métodos: Un total de 64 casos completados durante los años 2007 y 2008 cumplieron con los criterios de inclusión y fueron evaluados utilizando el "OGS" por un examinador calibrado. Se utilizaron análisis de regresión múltiple y logística para determinar la asociación entre las variables explanadoras y el puntaje final del "OGS". Resultados: El 18.8% de los casos obtuvo una puntuación de <20, aprobando el examen; 34.4% obtuvo una puntuación entre 20-30 y 46.9% obtuvo una puntuación de >30, fracasando el examen. La puntuación promedio fue 32.17 ± 13.03 puntos. Conclusión: Este estudio demostró que el 53% de los casos completados obtuvieron una puntuación que potencialmente aprobaría el examen clínico del "ABO" (OGS <30). El análisis de regresión múltiple y la regresión logística demostraron que ninguna de las variables explicaba el resultado (OGS) o la probabilidad de éxito o fracaso cuando todas las variables se tomaron en consideración.

Acknowledgments

References

1. Cook DR, Harris EF, Vaden JL. Comparison of university and private-practice orthodontic treatment outcomes with the American Board of Orthodontics objective grading system. Am J Orthod Dentofacial Orthop 2005;127:707-12.
2. Eisman D. A method of evaluating efficiency of orthodontic treatment. Trans Eur Orthod Soc 1974;20:223-32.
3. Gottlieb E. Grading your orthodontic treatment results. J Clin Orthod 1975;9:155-61.
4. Berg R. Post-retention analysis of treatment problems and failures in 264 consecutively treated cases. Eur J Orthod 1979;1:55-68.
5. Abei Y, Nelson S, Amberman BD, Hans MG. Comparing orthodontic treatment outcome between orthodontists and general dentists with the ABO index. Am J Orthod Dentofacial Orthop 2004;126:544-8.
6. Cangialosi TJ, Riolo ML, Owens SE Jr, Dykhouse VJ, Moffitt AH, Grubb JE, et al. The ABO discrepancy index: A measure of case complexity. Am J Orthod Dentofacial Orthop 2004;125:270-8.
7. Casco JS, Vaden JL, Kokich VG, Damone J, James RD, Cangialosi TJ, et al. Objective grading system for dental casts and panoramic radiographs. Am J Orthod Dentofacial Orthop 1998;114:589-99.
8. Yang-Powers LC, Sadowsky C, Rosenstein S, Be Gole EA. Treatment outcome in a graduate orthodontic clinic using the American Board of Orthodontics grading system. Am J Orthod Dentofacial Orthop 2002;122:451-5.
9. James RD. Objective cast and panoramic radiograph grading system. Am J Orthod Dentofacial Orthop 2002;122:450.
10. Djeu G, Shelton C, Maganzini A. Outcome assessment of Invisalign and traditional orthodontic treatment compared with the American Board of Orthodontics objective grading system. Am J Orthod Dentofacial Orthop 2005;128:292-8.
11. Nett BC, Huang GJ. Long-term posttreatment changes measured by the American Board of Orthodontics objective grading system. Am J Orthod Dentofacial Orthop 2005;127:444-50.
12. Pinskaya YB, Hsieh TJ, Roberts WE, Hartsfield JK. Comprehensive clinical evaluation as an outcome assessment for a graduate orthodontics program. Am J Orthod Dentofacial Orthop 2004;126:533-43.
13. Knierim K, Roberts WE, Hartsfield J. Assessing treatment outcomes for a graduate orthodontics program: Follow-up study for the classes of 2001-2003. Am J Orthod Dentofacial Orthop 2006;130:648-55.
14. The American Board of Orthodontics. Clinical Examination Guide. [ABO Website]. June 2008 revision. Available at: http://www.americanboardortho.com/professionals/initialcertificationexhibit.aspx. Accessed September 2008.
15. The American Board of Orthodontics. Grading system for dental casts and panoramic radiographs. [ABO Website]. June 2010 revision. Available at: http://www.americanboardortho.com/professionals/downloads/Grading System Casts-Radiographs.pdf. Accessed December 2010.