Biological Forum – An International Journal **7**(2): 1082-1086(2015)

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Evaluation the accuracy and repeatability of radiological diagnosis of well-defined unilocular radiolucent lesions by oral and maxillofacial radiologists

H. Bashizadeh Mokhtar*, M. Sahebjamee** and P. Bolandian***
**Department of Oral and Maxi Logacial Radiology, Faculty of Dentistry,*
Tehran University of Medical Sciences, Tehran. Iran
***Department of Oral Medicine Faculty of Dentistry, Tehran University of Medical Sciences, Tehran, Iran*
***Department of Perosthodantics, Faculty of Dentistry, Qazvin University of Medical Sciences, Tehran, Iran*

(Corresponding author: P. Bolandian)
(Received 28 August, 2015, Accepted 29 November, 2015)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Diagnosis of unilocular radiolucent lesions with respect to their similar appearance on radiographs has always been problematic. The present study aimed to determine the accuracy and repeatability of radiological diagnosis of unilocular well-defined radiolucent lesions that defined by Oral and Maxillofacial radiologists. In a diagnostic study, radiographs of 48 unilocular well-defined radiolucent lesions (in a ratio of 4 periapical, pericoronal, interradicular and unrelated with teeth) were chosen from patients files of the Department of the Faculty of Dentistry of Tehran University and were detected differentiallyby3radiologists.When one of 3 diagnoses was compatible with histopathologic diagnosis (golden standard) the diagnosis was correct, other wise an incorrect diagnosis was considered. One month later 2 observer assessed those images and intraobserver and interobserver repeatability assessed with kappa test.
Accuracy of observers distinguish was averaged 8/68% (7/66%, 9/72% and 7/66% for three radiologists). Also the interobserver repeatability of each two observers was 2/81%, 3/83% and4/85% in comparison (with an average of 3/83%). Kappa values of these evaluationswere557/0, 625/0and656/0(mean613/0) respectively. Repeatability of two intraobserver were 8/95% and81/3% (mean 6/88%), respectively (895/0 and542/0 for kappa values).
In general, oral and maxillaofacial radiologists had accurately detection for unilocular radiolucent lesions. The mean repeatability of the diagnoses was at the appropriate range. However, due to some differences, it seems definite diagnosis of lesions based on radiographs cannot bean appropriate criterion in the treatment plan.

Keywords: unilocular radiolucent lesions, diagnostic accuracy, repeatability

I. INTRODUCTION

Radiographs are important in treatment planning for surgical removal. So they can evaluate encroachment on vital structures, size of the lesion, extent into soft tissue, and requirements for reconstruction. Therefore radiography allows for creation of a radiologic differential diagnosis (Apostole *et al.*, 2003). Diagnosis of bone lesions has always been problematic, especially unilocular radiolucent lesions because of their similar radiography (O'Reilly *et al.*, 2000). Maxillofacial bone lesions have different radiographic views including radiolucent, radiopaque or mixedseen. Radiolucent lesions are most common type of maxillofacial lesions and are classified in periapical, pericoronal, interradicular and unrelated with teeth lesions (Wood and Goaz 1997). A well-defined unilocular radiolucency has distinct radiopaque borders that show discrete radiopaque foci (Curran *et al.*, 1997, Philipsen *et al.*, 2002).

Causes of radiolucent lesions are different so they should be distinguished from normal structures and maxillofacial anatomy. Due to the radiolucent nature of most aggressive, malignant tumors and maxillofacial lesions and so since the dentist may be the first person who observes them in radiography understanding and recognition of this lesion is so important. Radiographic appearances of most unilocular radiolucent lesions are well-defined and it seems impossible to differential diagnosis just by radiography (Myoung *et al.*, 2001). Today, by development of digital radiography techniques use of those techniques in medicine and dentistry is considered (Wenzel and Hintze 1993). Despite the availability of complex techniques such as tomography, in most cases, if unilocular radiolucent lesions are small in size and have cystic appearance, panoramic radiography is only available diagnostic tool in many health centers.

This study was performed to determine the accuracy and repeatability of radiological diagnosis of well-defined unilocular radiolucent lesions by oral and maxillofacial radiologists in Tehran University of medical sciences, faculty of dentistry in 2008-2009.

MATERIALS AND METHODS

A diagnostic study was performed on 48 radiographic images of unilocular radiolucent lesions by participating three oral radiologists. In the Faculty of Dentistry, Tehran University of medical sciences. Initially, 48 radiographs of well-defined unilocular radiolucent lesions were selected by following conditions:

1. High quality of the stereotype for expected diagnosis.
2. Completeness of patient medical records, including history, clinical findings and pathologicfinal diagnosis.

Lesions consisted of 12 cases periapical radiolucency, 12 cases pericronal radiolucency, 12 cases interradicular radiolucency and 12 cases of single cyst like radiolucency unrelated to teeth. Clinical signs and symptoms in patients with slides from lesions radiography were given to 3 expert of oral radiology that they had relatively equal educational background, later asked them to propose their 3 differential diagnoses. There was no time limiting for observer and they did their observations in standard and constant condition (equal light). Diagnosis of radiology and pathology (golden standard) were compared and if one of three diagnoses was conformed the point was 1 otherwise point of zero was considered. Golden standard of all the radiographic diagnoses of radiolucent lesion were assessed by microscopic examination and recorded in their files. To evaluate inter observer repeatability between the 3 observers, diagnoses results which were provided by them were compared. Also to assessment repeatability intraobserver, a month later, 2 of radiologists participating in the evaluation repeated all the diagnosis stages for unilocular radiolucent lesions. At this stage, if the one of the provided diagnoses was confirmed with golden standard, the diagnosis considered true. Results of two times observation were considered for basis of statistical assessment. Detection accuracy of well-defined unilocular radiolucent lesions was determined based on three observer correct diagnoses of total lesions by percentage. To evaluate the intra observer repeatability of diagnoses with obtained points in two diagnostic stages and comparison of inter observer repeatability between two observers Kappa test was used. The result of kappa test were classified include poor agreement, slight agreement, fair agreement, moderate agreement, substantial and almost perfect agreement for under zero, 0-0/2, 0/21-0/4, 0/41-0/6, 0/61-0/8 and 0/81-1/0, respectively.

RESULTS

Each group of periapical, pericoronal, interradicular and unrelated with teeth lesions radiolucency was 12 cases of the 48 unilocular radiolucent lesions, 25% for each group. First radiologists diagnosed 32 lesions correctly (66/7%). However, 16 diagnoses were wrong (33/3%). The second radiologist diagnosed 35 lesions correctly (9/72%). The results were repeated at next observation in one month later (72/1% for correct diagnoses and 21/9% for wrong diagnoses). The third observer diagnosed 32 lesions correctly (66/7%), also he diagnosed 37 lesions correctly (77/1%) one month later. Evaluation of the first two observer interobserver repeatability was performed with kappa test and result showed the agreement rate was 81/2% (0/557 for kappa test) between two observers, so that agreement classified in moderate group (Table 1). Then evaluation of the first and third observer interobserver repeatability was performed with kappa test and result showed the agreement rate was 83/3% (0/625 for kappa test), so that agreement classified in substantial group (Table 2). As well as evaluation of the second and third observer interobserver repeatability was performed with kappa test and result showed the agreement rate was 85/4% (0/656 for kappa test), so that agreement classified in substantial group (Table 3). The second and third observer intraobserver repeatability was evaluated in two stages, at the first time and one month later.

Table1: Comparison of the diagnostic results of the first and second observer for well-defined unilocular radiolucent lesions radiolucent.

First observer / Second observer	Incorrect diagnosis	Correct diagnosis	Total
Incorrect diagnosis	10 (20/8%)	3 (6/3%)	13 (27/1%)
Correct diagnosis	6 (12/5%)	29 (60/4%)	35 (6/3%)
total	16 (33/3%)	32 (66/7%)	48 (100%)

(kappa=0/557, p<0/0001)

Table 2: Comparison of the diagnostic results of the first and third observer for well-defined unilocular radiolucent lesions radiolucent.

First observer / Third observer	Incorrect diagnosis	Correct diagnosis	Total
Incorrect diagnosis	12 (25/0%)	4 (8/3%)	16 (33/3%)
Correct diagnosis	4 (12/5%)	28 (58/3%)	32 (66/7%)
total	16 (33/3%)	32 (66/7%)	48 (100%)

(kappa=0/625, p<0/0001)

Comparison of the second observers repeatability was performed by using kappa test at two times and showed the agreement was 95/8% (0/895 for kappa test). This rate was classified in almost perfect group (Table 4). As well as the third observer intraobserver repeatability evaluation was performed for two stages of diagnosis by kappa test and results showed 81/3% agreement (0/542 for kappa test). This rate was classified in moderate group.

Table 3: Comparison of the diagnostic results of the second and third observer for well-defined unilocular radiolucent lesions.

Second observer / Third observer	Incorrect diagnosis	Correct diagnosis	Total
Incorrect diagnosis	11 (22/9%)	5 (10/4%)	16 (33/3%)
Correct diagnosis	2 (4/2%)	30 (62/5%)	32 (66/7%)
total	13 (27/1%)	35 (72/9%)	48 (100%)

(kappa=0/656, p<0/0001)

Table 4: The results of two time (beginning of study and one month later) diagnosis of the second observer in compare.

First observation / One month later observation	Incorrect diagnosis	Correct diagnosis	Total
Incorrect diagnosis	12 (25/0%)	1 (2/1%)	13 (27/1%)
Correct diagnosis	1 (2/1%)	34 (70/8%)	35 (72/9%)
total	13 (27/1%)	35 (72/9%)	48 (100%)

(kappa=0/557, p<0/0001)

Table 5: The results of two time (beginning of study and one month later) diagnosis of the third observer in compare.

First observation / One month later observation	Incorrect diagnosis	Correct diagnosis	Total
Incorrect diagnosis	9 (18/8%)	7 (14/6%)	16 (33/3%)
Correct diagnosis	2 (4/2%)	30 (62/5%)	32 (66/7%)
total	11 (22/9%)	37 (77/1%)	48 (100%)

(kappa=0/542, p<0/0001)

DISCUSSION

The results showed successful diagnosis for repeatability and reliability of these lesions by radiologists. As well as accuracy of the diagnoses was appropriate and acceptable. So that 3 radiologists differentiated 68/8% of lesions properly. The accuracy of observers diagnosis were 66/7%, 72/9% and 66/7% respectively. Despite this issue above results indicate in 32% of cases it may be impossible to differentiate lesions from normal anatomical sites, although oral and maxillofacial radiologists are able to diagnosis most of well-defined unilocular radiolucent lesions. However, this probability for failure to diagnosis had been in research literature, so these failures should be decreased. Raitz *et al* (2006) evaluated diagnostic accuracy of conventional and digital radiography techniques in the diagnosis of radiolucent lesions. The probability of correct diagnosis was similar for these lesions in both techniques and was about 56/6% and in compare with current study results were slightly lesser. As well as Bohany *et al* (2000) studied radiographic diagnostic indicators in inflammatory diseases. The results showed mean diagnostic accuracy about 70/2% that was similar to this studies result. Results of Bashizadefakhar *et al* (2003) showed diagnostic accuracy rate about 83/75% for radiolucent lesions of posterior teeth roots, so their results are higher than our results and so on very significant. This may be because observers in mentioned study were dentistry student and oral and dental radiologist, so they had furthermore capabilities to diagnosis. Although type of lesions were different in two studies, but the results are not justifiable because the difference is about 15%. With further research, focus on the type of lesions and use of different observers it may be possible to identify various reasons which are related to it. The diagnostic accuracy of observers based on ratio of correct diagnoses to total number of diagnoses express in percentage, including 2 diagnostic parameters, specificity and sensitivity, which are used with other diagnostic indicators to indicate performance of radiography methods and observers. However, computing the percent of observers correct diagnoses are possible, but using this index cannot evaluate performance of diagnostic tests accurately.

In this study because of particular circumstances and relying on differentiate diagnosis, correct diagnosis percentage index for radiolucent lesions was used; so that 3 priority which provided by different observers was compatible with golden standard that was obtained on histopathological evaluations basis, point 1 was given for correct diagnosis and zero for wrongs.

In addition to the diagnostic index of accuracy, repeatability of radiographic diagnosis is important in providing an accurate diagnosis and appropriate treatment. Repeatability is a quantitative diagnostic target in which observers can provide similar diagnosis. Repeatability is expressed as interobserver (agreement rate of observers) and intraobserver (agreement rate of two stages diagnosis for one observer).The results showed observers had reliable and repeatable diagnoses for well-defined unilocular radiolucent lesions; such that study of agreement and repeatability for the first and second radiologists showed they had similar diagnoses for 81/2% of cases. The first and third radiologists had also 83/3% agreement for cases of study and so on 85/4% of agreement was for diagnoses of second and third radiologists. The rate of evaluated number of kappa test 0/557, 0/625 and 0/656, respectively were for these 3 comparisons (0/613 in average), and classified in moderate and substantial groups. Despite this issue it might be considered that diagnosis were incorrect in 14/6%-18/8% of cases. In this study evaluation of intraobserver repeatability was similar in 95/8% and 81/3% (in average 88/6) of cases for second and third observer in twice observation. Evaluated kappa rate for two stages of observations were 0/895 and 0/542, so classified in almost perfect and moderate groups. According to these results, it seems that the observers diagnoses are based on constant and determined parameters so high repeatability of diagnoses obtained. Despite this, it should be noted, although both the radiologists and had an educational background and evaluated radiographies in constant conditions, however, in some cases have been different diagnoses in the first stage that it was considerable especially for the third observer. It is obvious that finally this level of discrepancy in diagnoses made of unilocular radiolucent lesions lead to indistinguishable from normal adjacent anatomic sites and patients will be exposed to in appropriate treatments. In terms of methodology, the rate of diagnos is agreement in order to provide an appropriate treatment should be in the range of 98% -85% to ensure their repeatability (WHO). The results of this study show that the reliability of the diagnoses made by oral and maxillofacial radiologists for unilocular radiolucent lesions are within the WHO specified range. Evaluation of observations repeatability in the present study was performed by kappa test. This testis one of the most reliable methods to determine rate of repeatability and agreement for observations an disused in several studies. In this test, the rate of agreement made is removed from the chance and consequently the use of it is appropriate criteria for determining the observers agreement with each other (Donner *et al.*, 1996). Also, the obtained kappa is classified in 6 classes, poor, slight, fair, moderate, considerable, substantial and almost perfect agreement, and so on makes it easier to compare the results of different studies (Naitoh *et al.*, 1998). The researches that have been focused on rate of radiographic diagnosis repeatability in dentistry, have reported different values: the rate of kappa in evaluations of observers agreement by using two techniques, conventional and digital radiographyin the diagnosis of radiolucent lesions of the jaws on the study of Raitz *et al* (2006) was reported in the range of 0/332-0/708. In other words, the rate of observers internal repeatability was good for two techniques utilization and observers have had same function for both techniques. As well as in their study type of radiolucent lesion had no effect on the accuracy of observers diagnosis and repeatability rate of results. As well as the result of intraexaminer reliability test that had done by Apostole *et al* (2003) showed that the value of the kappa-statistic for the permanent teeth was 0.87 with a confidence interval ranging between 0.78 and 0.96, while for the primary teeth the corresponding value was 0.86 with an interval between 0.82 and 0.89. This results were in group of perfect agreement.

The rate of interobserver repeatability in study of Bashyzadhfakhar *et al* (2003) on radiolucent lesions diagnosis of tooth root end was 59% (Bashizadeh *et al.*, 2002-03), so was estimated lesser than results of present study (83/3). Low agreement obtained in their study may due to type of observers who were employed by them; because in their study diagnoses were accomplished by dentistry student, while in present study oral and maxillofacial radiologists were employed, undoubtedly in compare to students they have higher abilities to diagnosis different lesions. The mean intraobserver repeatability for study of Bashyzadhfakhar *et al* (2003) were 31/5% and 36% for the two cases, which were lesser than result of current study (88/6%). It is obvious senior students in their study in comparing to oral radiologists of this study had provided different diagnoses for twice evaluations. The rate of intra examiner repeatability in study of Molven *et al* (2002) was 83% for diagnosis radiography of teeth periapical disease and so on rate of interobserver repeatability was estimated 83% and86% that were in a good range. The study of Bohay *et al* (2000) on diagnosis of inflammatory disease in the periapical region of the posterior teeth by 6 observers showed rate of interobserver reliability about 0/54 (intraclass coefficient) and the rate of intraobserver repeatability 0/66 in average. It is obvious results of these studies are not comparable to results of present study because these have been performed on caries lesions. Also Curtis *et al* (2007) studied on dental caries prevention and the results were 0/78 and 0/84 for interobserver and intraobserver reliability, respectively.

As is clear the results are in substantial and almost perfect classes. Differences between observers in the diagnosis of various oral diseases or in their assessment of the two stages of diagnosis is seen, may be due to natural and biological differences in different lesions or differences indiagnostic procedures. When the diagnosis of a disease in a particular phase these two factors will be integrated and inseparable. However, in case of repeated measurements at more than one stage, it is not only possible to determine the differences between different observers, but also differences in the diagnosis made by a single observer can be measured (De Paola and Alman 1972).

The prevalence of disease, educational backgrounds and job experience of observers and also quality of radiographs are effective in repeatability of radiographic diagnosis (Molven *et al.*, 2002). When the prevalence is lower than normal, it seems that repeatability rate of diagnoses is higher than the rates which are obtained by chances (Bulman and Osborn 1989, Wulff and Gøtzsche 2000).

Of course using of kappa test can reduce the chances for agreement percentage and provides reliable results for this field. Considering the results of present study, in some cases diagnosis of unilocular radiolucent lesions was different between observersor in twice assessment, as well as all of the observers were professors of faculty of dentistry and had acceptable experience for different lesions, it seems this issue have had worse situation among other observers, such as dentists or students of dentistry. Therefore, it is obvious that relying on radiographic findings will not be able to provide appropriate plans to treatment patients. To prevent different diagnosis of these lesions, it suggested that clinical examination of radiolucent lesions and radiographic judgments of them be at same time, as well as further education be presented to students and dentists in continuing education programs.

CONCLUSION

Considering the accuracy of68/8% it seems the diagnostic accuracy for oral and maxillofacial radiologists has been proper to identification of well-defined unilocular radiolucent lesions. As well as, According to the mean intra observer repeatability 88/6% and mean interobserver repeatability 83/3% for observers it seems they had repeatable diagnoses for radiolucent lesions and this values are in defined range by the WHO. Despite this issue, due to some differences among observers diagnosis or during diagnostic evaluation by an observer in two stages, It seems definite diagnosis of radiolucent lesions based on radiography is not an appropriate criterion for treatment plan.

REFERENCES

Apostole P. Vanderas, Constantinos Manetas, Matina Koulatzidou, (2003). Progression of Proximal Caries in the Mixed Dentition: A 4-year Prospective Study. *Pediatric Dentistry*, **3**, 229-234.

Bulman JS, Osborn JF. (1989). Measuring diagnostic consistency. *Br Dent J*; **166**: 377-381.

Bashizadeh Fakhar Horie, Kharazifard Mohammad Javad, Gholizadeh Azam. (2002-03). Survey the susceptibility, specificity and repeatability of periapical radiographies diagnosis on radiolucent lesions of the end posterior teeth root. Thesis of dentistry doctorate. Faculty of dentistry, Tehran University. **4201**. 2002-2003.

Bohay RN. (2000). The sensitivity, specificity, and reliability of radiographic periapical diagnosis of posterior teeth. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*; **89**: 639-642.

Curran AE, Miller EJ, Murrah VA (1997). Adenomatoid odontogenictumor presenting as periapical disease. *Oral Surg Oral Med OralPathol Oral Radiol Endod* **84**: 557.

Curtis B, Evans RW, Sbaraini A, (2007). Recruitment and standardization of a group of Australiandentists for a multi practice study on dental caries prevention. *Australian Dental Journal*; **52**(2):106-111.

De Paola PF, Alman J. (1972). Assessment of the reliability of radiographic diagnosis in a clinical caries trial. *J Dent Res*; September-October **51**(5):1431-1437.

Donner A, Eliasziw M, Klar N. (1996). Testing the homogeneity of kappa statistics. *Biometrics*; **52**:176-183.

Escobar, E, Godoy, L. & Peafiel, C. (2007). Odontogenic Cysts: Analysis of 2.944 cases in Chile German Ochsenius. *Med Oral Pathol Oral Cir Bucal*, Vol.**12**, pp:E85-91, eISSN1698-6946.

Myoung H, Hong SP, Hong SD, Lee JI, Lim CY, Choung PH, (2001). Odontogenic keratocyst: review of 256 cases for recurrence and clinicopathologic parameters. *Oral Surg Oral Med Oral Pathol Oral RadiolEndod*; **91**: 328-333.

Molven O, Halse A, Fristad I. (2002). Long-term reliability and observer comparisons in the radiographic diagnosis of periapical disease. *IntEndod J.*, **35**(2): 142-147.

Naitoh M, Yuasa H, Toyama M, Shiojima M, Nakamura M, Ushida M, Iida H, Hayashi M, Ariji E. (1998). Observer agreement in the detection of proximal caries with direct digital intraoral radiography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, **85**(1): 107-112.

O'Reilly M, O'Reilly P, Todd CEC, Altman K, Schfler K. (2000). An assessment of the aggressive potential of radiolucencies related to the mandibular molar teeth. *Clin Radiol*, **55**: 292-295.

Philipsen HP, Srisuwan T, Reichart PA (2002). Adenomatoid odontogenictumor mimicking a periapical (radicular) cyst: A case report. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* **94**: 246.

Raitz R, Correa L, Curi MM, Dib LL, (2006). Fenyo-Pereira: Conventional and indirect digital radiographic interpretation of oral unilocular radiolucent lesions. *Dentomaxillofacial Radiology*; **35**:165-169.

Wulff HR, Gøtzsche PC. (2000). Rational Diagnosis and Treatment. Evidence-Based Clinical Decision-Making. London, UK: Blackwell Science, **29.**

Wood NK and Goaz PW. (1997). Differential diagnosis of oral and maxillofacial lesions. 5th ed. St Louis: The CV Mosby Co. **252**, 449.

Wenzel A and Hintze H. (1993). Perception of image quality in direct digital radiography after application of various image treatment filters for detectability of dental disease. *Dentomaxillofac Radiol*, **22**: 131-134.

World Health Organization. Oral health surveys. Basic methods. 4th ed. WHO. Geneva, Switzerland: **13-15**, 62-63.