Intrarater Reliability in the Ultrasound Diagnosis of Medial and Lateral Orbital Wall Fractures With a Curved Array Transducer

Siegfried Jank, MD, DMD, Martina Deibl, MS,† Heinrich Strobl, MD, DMD,‡ Andreas Oberrauch, MD, DMD,§ Alessandro Nicasi, DMD,‖ Martin Missmann, MD, PhD,¶ and Gerd Bodner, MD, DMD#*

Purpose: The aims of the study were to document the effectiveness of ultrasound (US) in diagnosing orbital wall fractures when compared with computed tomography (CT) and to measure the intraobserver reliability of US using a curved array transducer.

Materials and Methods: From December 2003 to March 2004, 13 patients with the clinical diagnosis of an orbital trauma were investigated prospectively by CT (reference) and 2 US investigators. Both orbits were investigated. Sensitivity, specificity, accuracy, and positive and negative predictive value were calculated. The statistical difference between the 2 US investigators was calculated by a chi-square test. The interrater reliability was calculated using the λ coefficient. Values below 0.4 represent poor reliability, between 0.4 and 0.75 represent fair to good reliability, and a score > 0.75 is graded as excellent reliability.

Results: The comparison of the results of the 2 US investigators by the chi-square test showed *P* values of .385 for the medial orbital wall and .638 for the lateral orbital wall, which shows no significant difference. The λ-value for the investigation of the medial orbital wall reached 0.429, 0.714, and 0.750. The λ-value for the investigation of the lateral orbital wall yielded 0.647, 0.750, and 0.882. These values show a good and excellent inter-rater reliability.

Conclusion: The US investigation does not yet reach the diagnostic quality of CT. US could be a helpful diagnostic imaging tool in cases with clear clinical symptoms. The results of the current study and the previously published results imply that US has the potential to reach the same diagnostic quality as CT in the future, but further studies must be performed to improve the diagnostic quality of the method.

J Oral Maxillofac Surg 64:68-73, 2006

Computed tomography (CT) and conventional radiography are described to be the most used diagnostic imaging methods for orbital fractures. Conventional radiography has several disadvantages like overlying structures[1,2] which makes the interpretation of the images difficult. Therefore, conventional radiography no longer plays an important imaging role for orbital fractures in the 21st century; while CT is described in the literature to be the gold standard in the imaging of orbital fractures.[3-11] The disadvantages of CT are the cost and radiation exposure to the globe.[9,12] The sonographic investigation (ultrasound [US]) of orbital fractures seemed to be difficult because of the unsatisfying fitting of the transducer to the orbital walls.[8,13-23] Cost effectiveness is getting more and more important in the medical field all over the world. This leads to the question if cheaper alternatives to CT exist. US is a cost-effective and widely available diagnostic imaging method available in many maxillofacial departments. If US would show a satisfying diagnostic quality, the maxillofacial surgeon

Received from the Medical University of Innsbruck, Innsbruck, Germany.

*Associate Professor, Department of Oral and Maxillofacial Surgery.

†Research Assistant, Department of Biostatistics.

‡Associate Professor, Department of Oral and Maxillofacial Surgery.

§Resident, Department of Oral and Maxillofacial Surgery.

‖Resident, Department of Oral and Maxillofacial Surgery.

¶Consultant, Department of Oral and Maxillofacial Surgery.

#Associate Professor, Department of Radiology.

Address correspondence and reprint requests to Dr Jank: University of Innsbruck, Dept of Oral and Maxillofacial Surgery, Maximilianstr 10, A-6020 Innsbruck, Germany; e-mail: siegfried.jank@gmx.de

0278-2391/06/6401-0011$32.00/0

doi:10.1016/j.joms.2005.09.014

would get a quick diagnosis within several minutes in patients with suspected orbital fractures. Recently, the use of a curved array transducer was suggested to be an alternative method in the imaging of orbital fractures regarding sensitivity and specificity.[24-28] Especially in US imaging techniques, it is important to regard not only 1 investigator but to compare 2 investigators to get information about the intrarater reliability of the imaging technique. Recently published results showed good and excellent intrarater reliability in the diagnosis of orbital fractures regarding the infraorbital rim and the orbital floor by using a curved array transducer.[29] We hypothesize that US could be an alternative diagnostic imaging method in medial and lateral orbital wall fractures. The aims of the study were to document the effectiveness of US in diagnosing orbital wall fractures when compared with CT and to measure the intraobserver reliability of US using a curved array transducer.

Materials and Methods

From December 2003 to March 2004, 13 patients with the clinical diagnosis of an orbital fracture were investigated. Patients were recruited randomly from the Department of Oral and Maxillofacial Surgery of the Medical University of Innsbruck (Germany). The study design was prospective using the following inclusion criteria: 1) hematoma; 2) diplopia; 3) reduction of globe motility; and 4) exophthalmos/enophthalmos. No control cases were added to the study.

The predictor variables were CT and US. The outcome variables were the presence or absence of a fracture of the medial and/or lateral orbital wall. CT was used as the gold standard imaging diagnosis in the study. Every patient underwent a CT investigation using an axial thin slice CT with a coronal reconstruction (GE multislice; GE Healthcare, Chalfont St Giles, UK). CT investigation was performed by an experienced radiologist. The following parameters were used: 1.5 mm slices with 1.2 mm coronal reconstruction (pitch, 3.75, high quality, standard, or bone). The US investigation was performed by 2 experienced maxillofacial surgeons. For the US investigation, a 7.5 MHz curved array transducer was used (Picker CS9300; Hitachi Medical Systems Europe, Zug, Switzerland).

CT and US investigation were performed with the following dual criteria: fracture of the medial and/or the lateral orbital wall in both orbits. The medial wall was defined as the orbital plate of the ethmoid bone and the lacrimal bone; the lateral orbital wall was defined as the zygomatic bone and the greater wing of the sphenoid. The radiologist performing the CT investigation was blinded to the results of the clinical and US examination. The oral and maxillofacial surgeon performing the US investigation was blinded to the results of the CT investigation. Additionally, age and gender of the patients were collected.

Data analysis was computed using SPSS (SPSS, Chicago, IL). Positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, accuracy, significance, and intrarater reliability were determined. Statistical significance was calculated by a chi-square test. A result was considered to be significant if P value reached the common value of $< .05$. Sensitivity is an index of the performance of a diagnostic test, calculated as the percentage of individuals with a disease who are correctly classified as having the disease (ie, the conditional probability of having a positive test result given having the disease). A test is sensitive to the disease if it is positive for the most individuals having the disease.[30] Specificity is an index of the performance of a diagnostic test, calculated as the percentage of individuals without the disease who are classified as not having the disease (ie, the conditional probability of a negative test result given that the disease is absent). A test is specific if it is positive for only a small percentage of those without the disease.[30] Sensitivity is calculated by including the results into the following formula: TP/TP + FN. Specificity is calculated as follows: TN/TN+FP (TP, true-positive results; TN, true-negative results; FP, false-positive results; FN, false-negative results).

Intrarater reliability was determined by calculating the λ coefficient for nominal data. Scores for statistical measurements with the λ coefficient range from 0 to 1, where the former shows no reliability and the latter perfect reliability.[31] Values below 0.4 represent poor reliability, between 0.4 and 0.75 represent fair to good reliability, and a score > 0.75 is graded as excellent reliability. Intrarater reliability is a measure of association that reflects the proportional reduction in error when values of the independent variable are used to predict values of the dependent variable. λ measures are suitable for general use. The idea underlying λ measures is that it may be possible to predict the category of 1 variable from knowledge of the category of another variable.[32]

Results

The sample was composed of 13 subjects (9 male; 69%) and had a mean age of 38.9 years (range, 11 to 85 years). A total of 26 orbits were evaluated clinically and radiographically for evidence of orbital wall fractures. Based on CT imaging, 6 subjects had a fracture of the medial orbital wall and 8 subjects had a lateral orbital wall fracture. Based on US, 6 subjects had a medial orbital wall fracture and 8 subjects had a lateral orbital wall fracture.

Table 1. RESULTS OF THE CT AND US INVESTIGATION (INVESTIGATOR 1)

	No. of Fractures CT Investigation	No. of Fractures US Investigation	Sensitivity US	Specificity US	Accuracy US	PPV US	NPV US
Medial orbital wall							
Fracture	6	8	83	85	85	63	94
No fracture	20	18					
Lateral orbital wall							
Fracture	8	9	88	89	88	78	94
No fracture	18	17					

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

MEDIAL ORBITAL WALL

Data analysis showed the following results: the US investigation of investigator 1 was false-positive in 3 patients and false-negative in 1 patient. This leads to a sensitivity of 83% and a specificity and accuracy of 85%. PPV and NPV yielded 63% and 94%, respectively (Table 1). The US investigation of investigator 2 was false-positive in 2 patients, while no false-negative result was found. The sensitivity and specificity of investigator 2 were calculated to be 100% and 90%, while accuracy reached 92%. PPV and NPV yielded 75% and 100%, respectively (Table 2).

The comparison between investigator 1 and investigator 2 by the chi-square test regarding the medial orbital wall showed no significant difference between the investigators (P = .385). The intra-rater reliability of the medial orbital wall showed λ values of 0.429 and 0.714 for the comparison between the CT results and the US investigators, which means a good and excellent reliability. The comparison between the 2 US investigators showed a λ value of 0.750, which also means an excellent reliability.

LATERAL ORBITAL WALL

Regarding the lateral orbital wall, CT investigation showed a fracture in 8 orbits (30.8%), while in 18 patients (69.2%) no fracture was diagnosed.

Data analysis showed the following results: the US investigation of investigator 1 was false-positive in 2 patients and false-negative in 1 patient. Therefore, sensitivity and specificity yielded 88% and 89%, respectively, while accuracy was calculated to be 88% (Table 1). PPV and NPV reached 63% and 94%, respectively (Table 1). The US investigation of investigator 2 was false-positive and false-negative in 1 patient each (Table 2), which leads to a sensitivity and specificity of 88% and 94%, respectively (Table 2), while accuracy reached 92%. PPV and NPV of investigator 2 yielded 88% and 94%, respectively (Table 2).

The comparison between the US results of investigator 1 and 2 by the chi-square test showed no significant difference.

The intrarater reliability for the lateral orbital wall was calculated as follows: The λ value for the comparison between the CT results and the US investigators yielded 0.647 and 0.750, which indicates a good and excellent reliability, while λ for the comparison between investigator 1 and 2 yielded 0.882, which indicates excellent reliability as well.

Discussion

Ultrasound is an important diagnostic tool in the maxillofacial surgery in the soft tissue diagnosis. The use of ultrasound for the investigation of bony lesion-like fractures in the midfacial region is not yet reported frequently in the current literature. US is an imaging technique that could be performed by the maxillofacial surgeon by himself, because (especially in Europe) most of the departments have their own ultrasound machine. Therefore, it would be a diagnostic imaging technique that leads to a quick diagnosis

Table 2. RESULTS OF THE CT AND US INVESTIGATION (INVESTIGATOR 2)

	No. of Fractures CT Investigation	No. of Fractures US Investigation	Sensitivity US	Specificity US	Accuracy US	PPV US	NPV US
Medial orbital wall							
Fracture	6	8	100	90	92	75	100
No fracture	20	18					
Lateral orbital wall							
Fracture	8	8	88	94	92	88	94
No fracture	18	18					

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

FIGURE 1. Position of the US transducer for the investigation of the medial orbital wall overlying the medial orbital rim.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

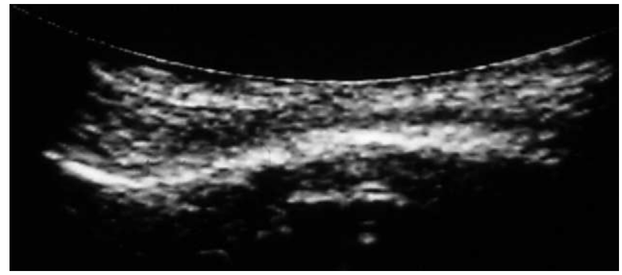

FIGURE 3. Ultrasound image of an unfractured medial orbital wall. The medial orbital wall shows an s-shaped structure.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

within a few minutes. Recently, the use of a curved array transducer for the investigation of orbital fractures was described.[24-28] However, these studies showed only 1 US investigator. To get information about the reproducibility of a method, intrarater and inter-rater reliability must be investigated. Intrarater reliability compares the results of 2 independent investigators, while inter-rater reliability compares the results of 2 independent re-investigators diagnosing already existing images. As a result of the convex shape of the transducer (Figs 1, 2) the imaging diagnosis of the bony orbital structure is much easier than with the use of a linear transducer. In former studies, sensitivity, specificity, and accuracy of the US investigation of the medial (Figs 3, 4) and lateral (Figs 5, 6) orbital wall reached values of more than 85% for both investigators. These values are a sign for the diagnostic quality of the method. The high rate of false-positive diagnoses in the current investigation has to be discussed critically, because this could lead to an unnecessary operative intervention. One reason for a false-positive diagnosis could be that the investigator probably "expects" a fracture if there are clinical signs like hyposphagma or hematoma. Therefore, the aim of further studies should lie in the reduction of false-positive diagnoses. The US investigators have to discuss their false diagnoses to improve their diagnostic quality and further retrospective studies computing an error pattern analysis must be performed to reduce the rate of false-positive and false-negative diagnoses. Of course, the technical features of the US machines should be improved as well to get a better resolution. Finally, the experience of the investigator grows with the number of investigated patients. This leads to the personal conclusion of the authors of the current study, to perform an US investigation even if a CT scan already exists (just to improve the personal diagnostic quality of US). The US investigation is not

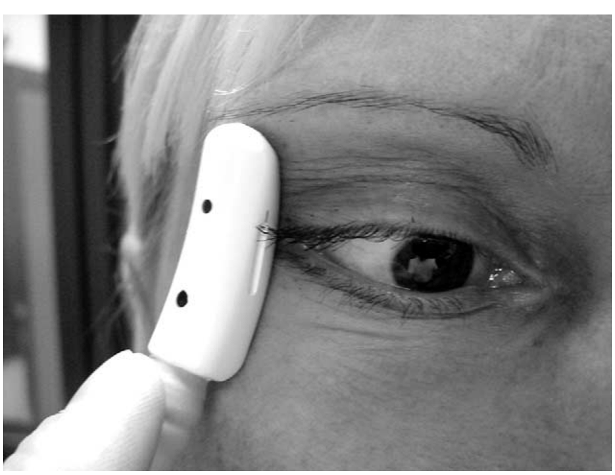

FIGURE 2. Position of the US transducer for the investigation of the lateral orbital wall overlying the lateral orbital rim.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

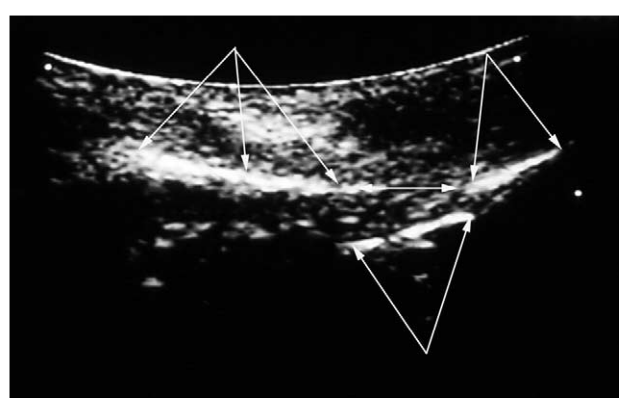

FIGURE 4. Ultrasound image of a fracture of the medial orbital rim. The *double arrow* shows the dislocation of the fracture. The *lower arrow* shows the dislocated fragment. The *upper arrow* show the unfractured area of the medial orbital wall.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

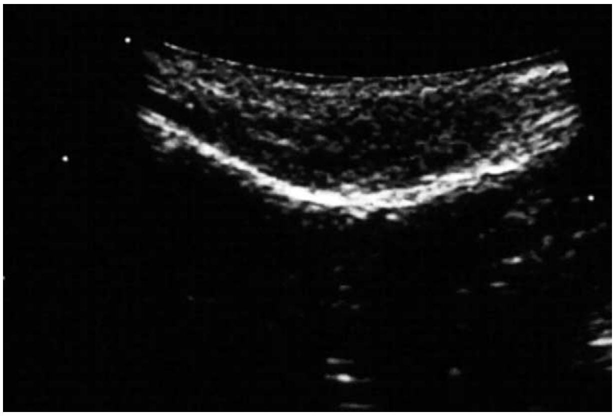

FIGURE 5. Ultrasound image of an unfractured lateral orbital wall. The lateral orbital wall shows a continuous convex structure.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

painful for the patient and does not expose the patient to any radiation, so US should be performed if the patient agrees.

The comparison between the 2 investigators by the chi-square test showed no significant difference. However, the chi-square test does not give any information if the investigators made the mistakes in the same patients or if the false diagnoses depended on the investigators themselves. The intrarater reliability was calculated to be good and excellent. These results could be interpreted that false diagnoses appeared mainly in the same patients. One could state that the calibration of the method is good and that the US investigation of the medial and lateral orbital wall is a reproducible method with a high diagnostic value. If these results are compared with recently published results[29] regarding the infraorbital rim and the orbital floor in the same cohort of patients, it is seen that the results concerning the infraorbital margin and the orbital floor are comparable to the current study. In the literature, the inter-rater reliability of the US investigation of the infraorbital-rim and the orbital floor is also described to be good and excellent.[29]

The results of the current investigation regarding the *intra*rater reliability could also be compared with the recently published results concerning *inter*-rater reliability.[33,34] These recently published results[33,34] also refer to the same cohort of patients as in the current study, so one could state that the method is reproducible. Inter-rater reliability[33] for the investigation of medial orbital wall fractures showed a λ coefficient of 0.571 and 0.538, which represents fair to good reliability while the λ coefficient for the lateral orbital wall yielded 0.765 and 0.631,[33] which represents fair to good and excellent reliability. Compared with the current investigation, intrarater reliability of the US investigation of medial and lateral orbital wall fractures shows better results than inter-rater reliability. These results show that it is difficult to re-interpret existing US images, because a print of an US image is only a static view of the situation. If the investigator has a real-time view on the screen, he can move the transducer to get a better overview, which seems to allow a better diagnostic quality. One way to solve this problem could be to save a video of the whole US investigation.

The question is if the US investigation of the medial and lateral orbital wall could replace CT, which is reported to be the current accepted gold standard imaging diagnosis of orbital trauma.[3-11] The results of this study imply that US does not yield the same diagnostic quality as CT, yet further studies must be performed to reduce the high rate of false-positive rates. On the one hand US is a cheap and widely available method, but on the other hand medial and lateral orbital wall traumas are often associated with severe midfacial and/or intracranial trauma, and of course, such injuries require a CT investigation. In the authors' opinion, US is an alternative method in the diagnosis of medial and lateral orbital wall fractures if there is a clear clinical diagnosis with the suspicion of a single medial or lateral orbital trauma. US could also be an alternative method to exclude a fracture in patients with only light clinical symptoms. US of the medial and lateral orbital wall is a method that requires investigators with a high level of experience. The decision for or against a surgical intervention should lie completely in the surgeon's hand. The maxillofacial surgeon should decide by himself if he would perform a surgical intervention on the basis of an US investigation or if a preoperative CT is required.

Therefore, US could be a helpful diagnostic imaging tool in cases with clear clinical symptoms. The results of the current study and the previously published

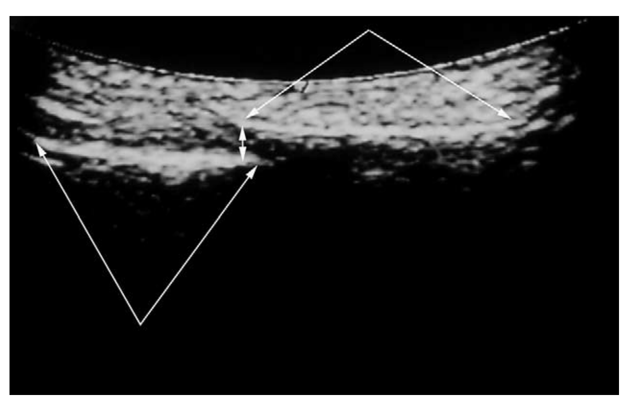

FIGURE 6. Ultrasound image of a fracture of the lateral orbital wall. The *double arrow* shows the position of the dislocation. The *upper and lower arrows* show the lateral orbital wall.

Jank et al. Intrarater Reliability in Ultrasound Diagnosis. J Oral Maxillofac Surg 2006.

results imply that US has the potential to reach the same diagnostic quality as CT in the future, but further studies must be performed to improve the diagnostic quality of the method.

References

1. Iinuma T, Ishio K, Yoshinami H, et al: Orbital wall fractures: A comparison of computed tomography and conventional views. Nippon Jibiinkoka Kaiho 96:175, 1993
2. Iinuma T, Hirota Y, Ishio K: Orbital wall fractures. Conventional views and CT. Rhinology 32:82, 1994
3. Anda S, Elsas T, Harstad HK: The missing rectus: A CT observation from blow-out fracture of the orbital floor. J Comput Assist Tomogr 11:895, 1987
4. Ball JB Jr: Direct oblique sagittal CT of orbital wall fractures. AJR Am J Roentgenol 148:601, 1987
5. Benateau H, Chevalier E, Hamon M, et al: The three dimensional spiral scanner and volume rendering technique: Importance in craniofacial traumatology and reconstructive surgery. Rev Stomatol Chir Maxillofac 103:233, 2002
6. Berado N, Leban SG, Williams FA: A comparison of radiographic treatment for evaluation of the orbit. J Oral Maxillofac Surg 46:844, 1988
7. Daly BD, Russell JL, Davidson MJ, et al: Thin section computed tomography in the evaluation of naso-ethmoidal trauma. Clin Radiol 41:272, 1990
8. Forrest CR, Lata AC, Marcuzzi DW, et al: The role of orbital ultrasound in the diagnosis of orbital fractures. Plast Reconstr Surg 92:28, 1993
9. Grove AS Jr, Tadmor R, New PF, et al: Orbital fracture evaluation by coronal computed tomography. Am J Opthalmol 85: 679, 1978
10. Imhof H, Czerny C, Dirisamer A: Head and neck imaging with MDCT. Eur J Radiol 45:23, 2003
11. Ng P, Chu C, Young N, et al: Imaging of orbital floor fractures. Austr Radiol 40:264, 1996
12. Dorbisz H, Voegeli E, Hardt N: Conventional radiology and computed tomography in facial fractures. Rontgenblatter 36: 428, 1983
13. Bleier R, Rochels R: Ultrasonic diagnosis of fractures of the bony orbit. Fortschr Kiefer Gesichtschir 32:144, 1987
14. Jenkins CM, Thuau H: Ultrasound imaging assessment of fractures of the orbital floor. Clin Radiol 52:708, 1997
15. Klinger M, Danter J, Siegert R: Ultrasound diagnosis of orbital floor fractures: An alternative to computerized tomography. Laringorhinootologie 75:242, 1996
16. Lata AC, Marcuzzi DW, Forrest CR: Comparison of real-time ultrasonography and coronal computed tomography in the diagnosis of orbital fractures. Can Assoc Radiol J 44:371, 1993
17. Levine RA: Orbital ultrasonography. Radiol Clin North Am 25:447, 1987
18. McCann PJ, Brocklebank LM, Ayoub AF: Assessment of zygomatico-orbital complex fractures using ultrasonography. Br J Oral Maxillofac Surg 38:525, 2000
19. Rochels R, Scherer U, Geyer G, et al: Echographic diagnosis of orbital floor fractures. Laringol Rhinol Otol 63:494, 1984
20. Rochels R: Ultrasonic diagnosis of fractures of the bony orbit. Fortschr KieferGesichtschir 32:144, 1987
21. Rowe LD, Miller E, Brandt-Zawadzki M: Computed tomography in the maxillofacial trauma. Laryngoscope 91:745, 1981
22. Rowe-Jones JM, Adam EJ, Moore-Gillon V: Subtle diagnostic markers of orbital floor blow-out fracture on coronal CT scan. J Laryngol Otol 107:161, 1993
23. Takashi H, Manders E, Nagamoto K, et al: Ultrasonic observation of facial bone fractures. J Oral Maxillofac Surg 54:776, 1996
24. Jank S, Strobl H, Emshoff R, et al: Bildgebende Diagnostik von medialen und lateralen Orbitawandfrakturen. Ultraschall versus CT. Mund Kiefer Gesichtschir 4:208, 2003
25. Jank S, Pomaroli A, Strobl H, et al: Sonographische anatomie bei orbitabodenfrakturen – Eine pilotuntersuchung am leichenpräparat. Z Stomatol 8:181, 2003
26. Jank S, Emshoff R, Etzelsdorfer M, et al: The diagnostic value of ultrasonography in the detection of orbital floor fractures with a curved array transducer. Int J Oral Maxillofac Surg 33:13, 2004
27. Jank S, Emshoff R, Etzelsdorfer M, et al: Ultrasound versus computed tomography in the imaging of orbital floor fractures. J Oral Maxillofac Surg 62:150, 2004
28. Jank S, Emshoff R, Etzelsdorfer M, et al: Effectiveness of ultrasonography in determining medial and lateral orbital wall fractures with a curved array transducer. J Oral Maxillofac Surg 62:451, 2004
29. Jank S, Deibl M, Strobl H, et al: Intrarater reliability in the diagnosis of orbital floor fractures with a curved array transducer. J Oral Maxillofac Surg (in press)
30. Everitt B: The Cambridge Dictionary of Statistics. Cambridge, MA, Cambridge University Press, 1998, pp 300-301,314
31. Fleiss J: Reliability of measurements, *in* The Design and Analysis of Clinical Experiments. New York, NY, John Wiley & Sons, 1986, pp 2-32
32. Upton G: Goodman-Kruskal measures of association, *in* Armitage P, Colton T, (eds): Encyclopaedia of Biostatistics. Chichester, John Wiley & Sons, 1998, pp 1721-1722
33. Jank S, Deibl M, Strobl H, et al: Interrater reliability of sonographic examinations of orbital fractures. Eur J Radiol 55:298, 2005
34. Jank S, Deibl M, Strobl H, et al: Interobserver variation of the sonographic diagnosis of orbital floor fractures and fractures of the infraorbital margin. Mund Kiefer Gesichtschir 8:337, 2004