Relationship between physical factors and subjective image quality of cone-beam computed tomography images according to diagnostic task

Jin-Woo Choi, DDS, MS, a Sam-Sun Lee, DDS, MSD, PhD, b Soon-Chul Choi, DDS, MSD, PhD, b Min-Suk Heo, DDS, MSD, PhD, b Kyung-Heo Huh, DDS, MSD, PhD, b Won-Jin Yi, PhD, b Se-Ryong Kang, BS, c Dong-Hun Han, DDS, MSD, PhD, d and Eun-Kyung Kim, DDS, MSD, PhD a

Objective. This study was designed to investigate the relationship between physical factors and the subjective quality of cone beam computed tomography (CBCT) images used for different diagnostic tasks.

Study Design. CBCT images of a real skull phantom and a SedentexCT IQ phantom were acquired under different exposure conditions (one Dinnova3 CBCT scanner, 60-110 kV and 4-10 mA). Radiologists evaluated subjective image quality of real skull phantom images for each diagnostic task. On the basis of the evaluation results, the images were classified into two groups: acceptable and unacceptable. The modulation transfer function (MTF), contrast-to-noise ratio (CNR), and image uniformity were measured using the SedentexCT IQ phantom images. The differences in physical factors were evaluated.

Results. MTF and CNR values showed statistical differences in image quality in two groups with regard to all diagnostic tasks. In the maxilla, MTF and CNR values showed no significant differences between periapical diagnosis and implant planning in the acceptable groups. Higher MTF and CNR values were required in the periapical diagnosis compared with the implant planning of the mandible.

Conclusions. This study proved that MTF and CNR values have a significant association with subjective image quality. The diagnostic task should be considered in evaluation of CBCT image quality. (Oral Surg Oral Med Oral Pathol Oral Radiol 2015;119:357-365)

Cone beam computed tomography (CBCT), which was introduced in 1998, involves lower radiation exposure and monetary costs compared with multi-detector computed tomography (MDCT). 1,2 Although the radiation doses of CBCT are known to be lower than those of MDCT, they are still higher than those of conventional dental radiography. Therefore, the optimization of CBCT is essential for protection against radiation. 2 Optimization can be achieved by reducing exposure while maintaining the image quality at an acceptable clinical level. 2 However, many types of CBCT devices are currently available on the market, and the equipment features, radiation doses, and image quality of these units vary. 1-3 For the optimization of these different devices, it is essential to assess the image quality by using a standardized method. 4

Image quality can be assessed by using a subjective evaluation method or quantitative measurement of physical factors using test phantoms. CBCT images are used for many clinical purposes, such as implant planning, impacted teeth, trauma, and periapical diagnosis. Therefore, the evaluation of image quality should be based on the accomplishment of these various diagnostic tasks. 5-7 Subjective evaluation is being used as the gold standard to assess the image quality in a task-based approach, but the possibility of standardizing this method is limiting due to its inherent subjectivity. 3,5,6,8 Therefore, it is important to investigate the relationship between subjective evaluation and the physical qualities of CBCT images for the development of a standardized method. There is a general consensus that the physical factors of spatial resolution, contrast resolution, and image noise are related to clinical image quality and are used for quality control programs in multi-detector CT devices. 9,10 Standardized phantoms are required to measure the physical factors of CBCT images.
image factors, and recently, the quality control phantom for CBCT was developed as part of the SedentexCT project.4

In the process of optimization, the acceptable clinical level of the image quality and the radiation dose may differ according to diagnostic tasks.2 Previous studies have reported that a reduction of the radiation dose can be achieved for implant planning with CBCT images.5,11,12 However, these studies did not measure the physical factors of an image, and research assessing the relationship between subjective evaluation and the physical image quality in CBCT devices is lacking. This study is designed to investigate the relationship between physical factors and the subjective evaluation of the quality of CBCT images in different diagnostic tasks.

MATERIAL AND METHODS

CBCT system

CBCT images were obtained by using a Dinnova3 CBCT scanner (HDXwill Inc., Seoul, Korea). Dinnova3 has an amorphous silicon flat panel detector. Further, the voxel size of 0.3 mm × 0.3 mm × 0.3 mm was used. A pulsed x-ray beam was rotated 360 degrees around the phantom, and the exposure time was 12 seconds (scan time: 24 seconds). A total filtration of 2.8-mm aluminum was used. The computed tomography dose index value was 3.183 mGy (120 kV, 120 mAs, field of view [FoV]: 200 mm × 190 mm). The FoV of 200 mm × 190 mm was used to obtain the complete image of the SedentexCT IQ phantom (SedentexCT IQ, Leeds Test Objects Ltd., Boroughbridge, UK). The phantom consists of a head-sized cylindrical phantom housing (diameter: 160 mm; height: 162 mm) made of polymethyl-methacrylate (PMMA) and 42 cylindrical inserts (diameter: 35 mm; height: 20 mm) (Figure 1). To obtain CBCT images with different image qualities, 24 combinations of six different tube voltages and four different tube currents were used for a skull phantom and a SedentexCT IQ phantom (60, 70, 80, 90, 100, and 110 kV; and 4, 6, 8, and 10 mA).

Measurement of physical image quality

To measure the physical factors of the images, CBCT images of a SedentexCT IQ phantom were acquired under 24 exposure combinations. The images were saved in the digital imaging and communications in medicine (DICOM) format, and the MATLAB program (R2010 b, Mathworks Inc., Natick, MA) was used to analyze the modulation transfer function (MTF), contrast-to-noise ratio (CNR), and image uniformity.

Under each exposure condition, a total of 640 slices of an axial image were loaded in the MATLAB program, and suitable inserts for each physical measurement were selected manually. All image factors were measured repeatedly for reliability: 10 times in the MTF measurement and three times in the other measurements.

Modulation transfer function

The point spread function (PSF) insert in row 4 of the phantom was used to calculate the MTF. The PSF insert is made of a stainless steel wire (diameter: 0.25 mm) suspended in air. MTF 50 and MTF 10 values were calculated using the fast Fourier transform method from the one-dimensional PSF. The wire area was magnified, and the image contrast was increased to find the exact center of the wire. When the window level and width were adjusted until the center of the wire was seen within 5 pixels and the brightest pixel was set as the center, the MTF values of four directions (vertical, horizontal, and two diagonal) were calculated automatically (Figure 2). In all, 10 slices of PSF insert images were used to obtain a reliable result.

Contrast-to-noise ratio

The pixel intensity value inserts in row 3 of the phantom were used to calculate the CNR. The inserts were made of five material disks (diameter: 25 mm) suspended in PMMA. Among them, aluminum (AL) and polyoxymethylene (POM) inserts were used to obtain image factors representing high- and low-contrast resolutions. The density values of AL, POM, and PMMA were 2.70, 1.42, and 1.20 g/cm³, respectively.
The CNR was calculated from the mean difference in the pixel value between the contrast material and the background (PMMA) divided by the average standard deviation within the material and the background, by using formula (1).

\[
\text{CNR} = \frac{\mu_{\text{obj}} - \mu_{\text{bg}}}{\frac{\sigma_{\text{obj}} + \sigma_{\text{bg}}}{2}} \quad (1)
\]

where \(\mu_{\text{obj}}\) and \(\mu_{\text{bg}}\) denote the average attenuation coefficients of the contrast material and the background, and \(\sigma_{\text{obj}}\) and \(\sigma_{\text{bg}}\) represent the standard deviation of the contrast material and the background. The region of interest was selected from the central area of each insert, and the marginal area was not included.

**Image uniformity**

The uniform PMMA layer in the base part of the phantom was used to calculate image uniformity. Image uniformity was calculated from the average pixel value difference between the center and the four peripheral regions of interest (upper, lower, right, and left regions) (Figure 3).

**Subjective evaluation**

CBCT images of a real skull phantom with a soft-tissue replica (X-ray phantom, head; product number 7280, Erler Zimmer Co., Lauf, Germany) were acquired under
the same 24 combinations of tube voltages and currents. Images were saved in the DICOM format. All 24 sets of images were reconstructed into three planes (axial, coronal, and sagittal) with a slice thickness of 0.3 mm. All sectional images were presented to five radiologists for a subjective evaluation of image quality. In all, three 20.8-inch monochrome monitors (ME315 L, Totoku Electric Co., Tokyo, Japan) with a resolution of 2048 × 1536 pixels were used, and images of each plane were displayed on a different monitor (Figure 4).

All observers had a trial session before the evaluation, and the evaluation was performed individually in a random, irreversible order. The observers were blinded to the exposure conditions, and they were allowed to adjust the brightness and the contrast of the images. Each observer evaluated the left maxillary first molar area first and then the right mandibular first molar area.

To ensure reliability of the evaluation, observers were asked about the visibility of three anatomic structures before evaluating the image quality for the diagnostic tasks of periapical diagnosis and implant planning (Table I). The following six-grade scale was used to answer five statements: strongly agree (6), agree (5), slightly agree (4), slightly disagree (3), disagree (2), and strongly disagree (1). The evaluation was repeated after an interval of 2 weeks to calculate intraobserver reliability.

We classified image quality into two groups—acceptable and unacceptable—by using the two agreement criteria of the average score and the consensus. The average score criterion used an average observer score of 3.5 as the threshold for both the acceptable and unacceptable groups. In the case of the consensus criterion, only images that obtained a score of more than 4 from all the observers were defined as the acceptable group.

Table 1. Evaluation statement of subjective image quality

<table>
<thead>
<tr>
<th>Maxilla: #26 Mesiobuccal root area</th>
<th>Mandible: #47 Mesial root area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clear inferior border of maxillary sinus</td>
<td>1. Clear border of mandibular canal</td>
</tr>
<tr>
<td>2. Clear lamina dura and periodontal ligament space</td>
<td>2. Clear lamina dura and periodontal ligament space</td>
</tr>
<tr>
<td>3. Clear trabecular bone pattern</td>
<td>3. Clear trabecular bone pattern</td>
</tr>
<tr>
<td>5. Image quality sufficient for implant planning</td>
<td>5. Image quality sufficient for implant planning</td>
</tr>
</tbody>
</table>

Statistical analysis

Intra- and interobserver reliabilities of the subjective evaluations were calculated as weighted κ-value by using an online calculator (http://vassarstats.net/index.html).

On the basis of the distribution of the samples, an independent t test or a Mann—Whitney U test was used to evaluate the differences in the physical image factors (MTF 50, MTF 10, CNR AL, CNR POM, and uniformity) between the acceptable and the unacceptable image quality groups by using IBM SPSS Statistics version 21 (IBM Corp., Armonk, NY). In the case of physical image factors with significant differences, the receiver operating characteristic (ROC) curve was used to determine the cutoff value by using IBM SPSS Statistics version 21.

An independent t test or a Mann—Whitney U test was used to investigate the differences in physical factors according to the type of jaw or diagnostic task. A statistical significance level of $P < .05$ was used.
RESULTS

In the subjective evaluation, the average weighted \( k \)-value was 0.59 (0.48-0.78) for intraobserver reliability and 0.41 (0.29-0.59) for interobserver reliability, corresponding to moderate agreement. Figure 4 shows examples of CBCT images from the acceptable and unacceptable groups.

The physical measurement results for the different diagnostic tasks are shown in Figures 5, 6, and 7. There were no statistical differences in image uniformity between the acceptable group and the unacceptable group in most diagnostic tasks (see Figure 7). Image uniformity did not seem to have a considerable influence in the subjective evaluation. The MTF 50, MTF 10, CNR AL, and CNR POM showed statistical differences between the acceptable group and the unacceptable group in all diagnostic tasks irrespective of the criteria (\( P < .05 \)) (see Figures 5 and 6).

An ROC curve was used to determine the cutoff value of MTFs and CNRs between the acceptable group and the unacceptable group. The area under the curve (AUC) and the cutoff values of MTFs and CNRs are given in Table II. In most cases, AUC was higher than 0.9 and the significance level was calculated as 0.000, which implies that the cutoff value is useful in the distinction between the acceptable and the unacceptable groups.

The differences in MTF and CNR values of the acceptable group according to the type of jaw are presented in Table III. There were no significant differences in the MTF and CNR values between the maxilla and the mandible for the same tasks. The required physical image quality difference depending on the type of jaw was considered unnecessary. In the case of the maxilla, there were no statistical differences in the MTF and CNR values between the
acceptable image groups of periapical diagnosis and implant planning. However, higher MTF and CNR values were required in the periapical diagnosis of the mandible than in the implant planning of the mandible (Table IV).

**DISCUSSION**

This study investigated the relationship between physical factors and subjective image quality and attempted to assess the differences in the required physical image quality according to diagnostic tasks. The evaluation
was performed using a single CBCT device, SedentexCT IQ phantom, and real skull phantom. There were significant differences in the MTF values between clinically acceptable and unacceptable images in all tasks. However, the usefulness of the MTF is not clearly proven, since the cutoff value of MTF 10 by ROC was low (0.67-0.76), and we are not sure whether the accomplishment of a diagnostic task depends on these small differences of spatial resolution or these differences in MTF values were just the result of the image quality classification by other factors. In addition, the spatial resolution can be influenced by various factors, such as voxel size, current, noise, FoV, and geometric accuracy.

Spatial resolution can be measured by using a line pair chart or an MTF calculation. However, the use of a line pair insert of the SedentexCT IQ phantom will not yield a detailed result. In previous research, MTF values were calculated by measuring the PSF of one or two directions in CBCT images with the use of a wire-suspended phantom. In this study, we measured the MTF in four directions because in the pilot measurement, the MTF values of the diagonal directions differed considerably from the vertical and horizontal values. This corresponds to the result obtained by Ozaki et al, which shows a difference between the MTF of the radial direction and that of the azimuthal direction. Further studies are required to develop a standardized MTF measurement of different CBCT images.

Limited literature is available with regard to the MTF and diagnostic acceptability, but there are earlier studies on the voxel sizes and the diagnostic accuracy of CBCT images. Özer reported that there were no significant differences in accuracy in the diagnosis of a root fracture among CBCT images of different voxel sizes. In contrast, Wenzel et al reported higher sensitivity in high-resolution CBCT images than in low-resolution images. Liedke et al reported that the voxel size did not influence the diagnostic accuracy of external root resorption, whereas another study reported that high-resolution CBCT images performed similarly or better than low-resolution images in the detection of simulated internal resorption.

### Table II. Cutoff value of MTF and CNR in the acceptable and unacceptable groups according to diagnostic task

<table>
<thead>
<tr>
<th>Task</th>
<th>By average score criteria</th>
<th>By consensus criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MTF 50</td>
<td>MTF 10</td>
</tr>
<tr>
<td>PD of Mx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff value</td>
<td>0.37</td>
<td>0.67</td>
</tr>
<tr>
<td>AUC</td>
<td>0.980</td>
<td>0.967</td>
</tr>
<tr>
<td>P value</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>IP of Mx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff value</td>
<td>0.36</td>
<td>0.70</td>
</tr>
<tr>
<td>AUC</td>
<td>0.982</td>
<td>0.935</td>
</tr>
<tr>
<td>P value</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>PD of Mn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff value</td>
<td>0.39</td>
<td>0.70</td>
</tr>
<tr>
<td>AUC</td>
<td>0.988</td>
<td>0.984</td>
</tr>
<tr>
<td>P value</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>IP of Mn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff value</td>
<td>0.36</td>
<td>0.70</td>
</tr>
<tr>
<td>AUC</td>
<td>0.982</td>
<td>0.935</td>
</tr>
<tr>
<td>P value</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

MTF, modulation transfer function; CNR, contrast-to-noise ratio; AL, aluminum; POM, polyoxymethylene; PD, periapical diagnosis; IP, implant planning; Mx, maxilla; Mn, mandible; AUC, area under receiver operating characteristic (ROC) curve.

### Table III. Differences in physical factors of acceptable groups according to the type of jaw

<table>
<thead>
<tr>
<th>Task</th>
<th>By average score criteria</th>
<th>By consensus criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MTF 50</td>
<td>MTF 10</td>
</tr>
<tr>
<td>PD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mx</td>
<td>0.592</td>
<td>0.592</td>
</tr>
<tr>
<td>Mn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mx</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Mn</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MTF, modulation transfer function; CNR, contrast-to-noise ratio; AL, aluminum; POM, polyoxymethylene; PD, periapical diagnosis; IP, implant planning; Mx, maxilla; Mn, mandible.

*Obtained using Mann–Whitney U test.
†Obtained using independent t test.
been disagreement over the voxel sizes and the diagnostic accuracy of CBCT images, spatial resolution is considered an important image quality factor in dental radiography, and further studies on the spatial resolution of CBCT images are required.2,18-21

It is unquestionable that the CNR can be a useful standard factor for evaluation of image quality, and previous studies have measured the CNR to evaluate the technical image quality of CBCT images.4,22-24 In this study, AL and POM inserts were used to obtain CNR values representing high- and low-contrast resolutions. Publication 93 of the International Commission on Radiologic Protection states that low-contrast resolution has a closer relationship to clinical image quality compared with high contrast resolution because high contrast resolution is not considerably influenced by an increase in noise level.7 Although CBCT in dentistry is mainly used for high-contrast anatomic structures, such as tooth and bone lesions, the periodontal ligament space and lamina dura, which have a low contrast compared with trabecular bone, are important anatomic structures and should be clearly depicted. Earlier studies have already revealed that the periodontal ligament space and lamina dura are less visible compared with other anatomic structures in several protocols with different CBCT devices.3,24 Therefore, low-contrast resolution can be considered an important image factor in the observation of the periodontal ligament space and lamina dura in CBCT devices. Bamba et al reported that a low-contrast-resolution phantom made of low-density polyethylene, POM, and polytetrafluoroethylene is useful in distinguishing between the performances of the devices’ contrast resolutions.14

Average CNR values of POM in the acceptable group ranged from 7.55 to 9.27, and the cutoff values were 3.92 to 7.35. Information from further research that uses different CBCT devices will be helpful in identifying the relationship between CNR values and the subjective image quality. However, image noise is influenced by the voxel size, and the CNR values from different devices should be compared with great care. Pauwels et al reported that CNR values are highly device dependent due to the differences in hardware and software.4,24 The relationship between CNR values and the visibility of various-sized rod inserts are worth investigating. Because the measurement of CNR values using Image J or MATLAB software is difficult and time consuming, evaluating the image contrast by using the visibility of the rod inserts would be more convenient in the use of CBCT devices. Of interest, accreditation programs of MDCT devices now include a rod visibility test.10

There was no distinction in the type of jaw for the assessment of image quality on the same task, and there were no statistical differences in MTF and CNR values between the periapical diagnosis and the implant planning of the maxilla (see Table IV). These results do not correspond to a previous task-based study, which showed different decision levels between the upper and lower jaw and between implant planning and periapical diagnosis.5 It is speculated that the reason for these differences in our results is the use of different criteria in each study. Although the same six-grade scale was used, Lofthag-Hansen et al used separation criteria in each study. Although the same six-grade scale was used, Lofthag-Hansen et al used separation criteria between “totally agree” and “agree,” whereas we used the criteria “acceptable” and “unacceptable” images.

In the case of the mandible, higher MTF and CNR values were required to perform the task of periapical diagnosis compared with implant planning (see Table IV). This result corresponds to the result obtained by Lofthag-Hansen et al.5

The use of a single CBCT device with a large FoV is the limitation of this study. Therefore, this result or the cutoff values cannot be compared directly with those obtained in the case of other CBCT devices. Future work will focus on other CBCT devices and different exposure conditions.

**CONCLUSIONS**

This study proves that the MTF and the CNR, among physical image factors, have a significant association with subjective image quality. In the case of the maxilla, the diagnostic task may be of little importance for the subjective evaluation of diagnostic acceptability.

---

**Table IV. Physical factors of acceptable groups according to diagnostic task**

<table>
<thead>
<tr>
<th>Jaw</th>
<th>Task</th>
<th>By average score criteria</th>
<th>By consensus criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MTF 50</td>
<td>MTF 10</td>
</tr>
<tr>
<td>Mx</td>
<td>PD</td>
<td>0.97</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>IP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn</td>
<td>PD</td>
<td>0.034†</td>
<td>0.037†</td>
</tr>
<tr>
<td></td>
<td>IP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MTF, modulation transfer function; CNR, contrast-to-noise ratio; AL, aluminum; POM, polyoxymethylene; PD, periapical diagnosis; IP, implant planning; Mx, maxilla; Mn, mandible.

*Obtained using Mann–Whitney U test.

†Obtained using independent t-test.
However, in the case of the mandible, the periapical diagnosis required better physical image quality compared with implant planning.

This study was supported by grant no 04-2014-0076 from the SNUDH Research Fund.

REFERENCES


Reprint requests:
Sam-Sun Lee
Department of Oral and Maxillofacial Radiology
School of Dentistry, Seoul National University
Chang-kyeong-gung-ro 62-1
Jongno-gu
Seoul 110-768
Korea
raylee@snu.ac.kr