Psychophysical evaluation of calibration curve for diagnostic LCD monitor

Radiat Med (2006) 24:■■–■■
DOI 10.1007/s11604-006-0085-3

ORIGINAL ARTICLE

Psychophysical evaluation of calibration curve for diagnostic LCD monitor

Masanobu Uemura · Yoshiyuki Asai
Michihiro Yamaguchi · Hideki Fujita · Yuuko Shintani
Shigeru Sanada

Received: March 8, 2006 / Accepted: July 30, 2006

Abstract
Purpose. In 1998, Digital Imaging Communications in Medicine (DICOM) proposed a calibration tool, the grayscale standard display function (GSDF), to obtain output consistency of radiographs. To our knowledge, there have been no previous reports of investigating the relation between perceptual linearity and detectability on a calibration curve.
Materials and methods. To determine a suitable calibration curve for diagnostic liquid crystal display (LCD) monitors, the GSDF and Commission Internationale de l'Eclairage (CIE) curves were compared using psychophysical gradient δ and receiver operating characteristic (ROC) analysis for clinical images.
Results. We succeeded in expressing visually recognized contrast directly using δ instead of the just noticeable difference (JND) index of the DICOM standard. As a result, we found that the visually recognized contrast at low luminance areas on the LCD monitor calibrated by the CIE curve is higher than that calibrated by the GSDF curve. On the ROC analysis, there was no significant difference in tumor detectability between GSDF and CIE curves for clinical thoracic images. However, the area parameter Az of the CIE curve is superior to that of the GSDF curve. The detectability of tumor shadows in the thoracic region on clinical images using the CIE curve was superior to that using the GSDF curve owing to the high absolute value of δ in the low luminance range.
Conclusion. We conclude that the CIE curve is the most suitable tool for calibrating diagnostic LCD monitors, rather than the GSDF curve.

Key words Grayscale standard display function · Psychophysical gradient δ · Visually recognized contrast · Diagnostic LCD monitors · ROC analysis

M. Uemura (✉) · Y. Asai · Y. Shintani
Department of Central Radiology, Kinki University Hospital, 377-2 Ohnohigashi, Osaka-Sayama, Osaka 589-8511, Japan
Tel. +81-72-366-0221; Fax +81-72-366-0206
e-mail: uemura@radt.med.kindai.ac.jp

M. Yamaguchi · H. Fujita
Department of Radiology, Osaka Prefectural Habikino Hospital, Habikino, Osaka, Japan

S. Sanada
Department of Radiological Technology, Graduate School of Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

Introduction

With the recent introduction of Digital Imaging Communication in Medicine (DICOM) into medical image networking, multivendor system construction has become common. Thus, the consistency of imaging data has become important with the increasing complexity of these systems.

In 1998, DICOM proposed a calibration tool, the grayscale standard display function (GSDF),[1] to obtain output consistency of radiographs. This ideal specificity is called *perceptual linearization*. In Europe, the values used by the International Illumination Commission (Commission Internationale de l'Eclairage, or CIE) are L*, a*, and b*; and the color measurement method is called CIELAB. The CIELAB function[2] suggested by the CIE has been used worldwide. Conventionally, clinical images are mostly presented by application-specific display functions that assign contrast nonuniformly

according to clinical need. However, it is probably an elusive goal to linearize all types of medical images perceptually under various viewing conditions by one mathematical function (GSDF).

On the other hand, tumor detectability, as determined by receiver operating characteristic (ROC) analysis, is one of the most important factors for clinical diagnosis. Our literature search revealed no studies showing differences in the clinical usefulness of the GSDF and CIE curves. In addition, there are currently no guidelines available to perform these functions in a clinical setting. During the 1960s, Kanamori[3,4] proposed the psychophysical gradient δ, corresponding to visually recognized contrast, which can be used for inspection of perceptual linearization. In previous articles,[5–7] we confirmed the usefulness of applying δ to liquid crystal display (LCD) monitors. The objectives of the present work are (1) to compare, using ROC analysis, the suitability of the GSDF and CIE curves as calibration tools for clinical images and (2) to explain the relation between tumor detectability and visually recognized contrast using psychophysical analysis.

Materials and methods

Calculation of GSDF and CIE curves

Figure 1 shows the new grayscale test pattern (NGTP)[8] used in the present study to confirm the calibration accuracy of the LCD monitor. The NGTP consists of 256 segments, made up of 16 × 16 squares without gaps. The

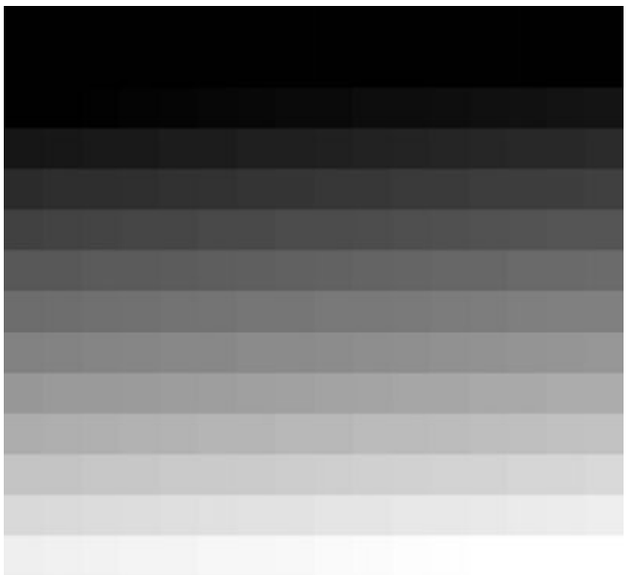


Fig. 1. New grayscale test pattern (NGTP) used in this study.

performance of the LCD monitor used in the present study was as follows: Maximum luminance was 450 cd/m^2; spatial resolution for horizontal and vertical directions was 1240 pixels and 1048 pixels, respectively; and the resolution for the digital input signal and the output luminance was 8 bits and 11 bits, respectively. In this system, the minimum difference in luminance between two segments is 0.2 cd/m^2.

DICOM standard proposed the following two equations[1] for calculating the GSDF curve

$$\log_{10} L(j) = \frac{a + c(\ln(j)) + e(\ln(j))^2 + g(\ln(j))^3 + m(\ln(j))^4}{1 + b(\ln(j)) + d(\ln(j))^2 + f(\ln(j))^3 + h(\ln(j))^4 + k(\ln(j))^5}$$

$$j = 1 \sim 1023; \quad a = -1.30; \quad b = -2.58 \times 10^{-2}; \quad c = 8.02 \times 10^{-2}$$
$$d = -1.03 \times 10^{-1}; \quad e = 1.36 \times 10^{-1}; \quad f = 2.88 \times 10^{-2};$$
$$g = -2.55 \times 10^{-2}; \quad h = -3.20 \times 10^{-3}; \quad k = 1.30 \times 10^{-4};$$
$$\textit{and} \quad m = 1.36 \times 10^{-3} \tag{1}$$

and

$$j(L) = A + B(\log_{10}(L)) + C(\log_{10}(L))^2 + D(\log_{10}(L))^3 + E(\log_{10}(L))^4 + F(\log_{10}(L))^5 + G(\log_{10}(L))^6 + H(\log_{10}(L))^7 + I(\log_{10}(L))^8$$
$$A = 71.4; \quad B = 94.6; \quad C = 41.9; \quad D = 9.82; \quad E = 0.282;$$
$$F = -1.19; \quad G = -0.180; \quad H = 0.147; \quad I = -0.0170 \tag{2}$$

where j denotes the just noticeable difference (JND) index. Equation (1) shows conversion of the JND index into the luminance, and Eq. (2) shows the reverse conversion. The minimum and maximum JND values agree with the minimum and maximum luminance on the LCD monitor, respectively, which are used to determine the JND index as a function of the digital driving level (DDL) shown in Eq. (3).[1]

$$J(DDL) = J_{\min} + \frac{DDL}{2^n - 1}(J_{\max} - J_{\min})$$
$$DDL = 0, 1, 2, 3, \ldots 2^{n-1} \tag{3}$$

where n denotes the number of bits showing the resolution of input signals to the LCD monitor. The CIELAB

proposed a modified cube root between the luminance $L'(p)$ and a perceived brightness variable, L^*, as shown in Eq. (4).[2]

$$L^* = 116 \times \sqrt[3]{\frac{L'(p)}{L'_{\max}}} - 16 \quad for \quad \frac{L'(p)}{L'_{(\max)}} > 0.00886$$
$$L^* = 903.3 \times \frac{L'(p)}{L'_{\max}} \quad for \quad \frac{L'(p)}{L'_{(\max)}} \le 0.00886 \tag{4}$$

In this scale, L^* varies between 0 and 100. A perceptually linear display curve, $L'(p)$, is obtained if the above function is inverted and L^* is identified with the DDL p values, where p is the presentation value (e.g., $p_{\max} = 255$).

$$L'(p) = \left[\frac{\left(\frac{100p}{p_{\max}} + 16\right)}{116}\right] \times L'_{\max} \quad for \quad \frac{p}{p_{\max}} > 0.08$$
$$L'(p) = \frac{1}{903.3} \times \left(\frac{100p}{p_{\max}}\right) \times L'_{\max} \quad for \quad \frac{p}{p_{\max}} \le 0.08 \tag{5}$$

To confirm the accuracy of the calibration of the LCD monitor used in this study, the NGTP was displayed on the LCD monitor, which was calibrated to the GSDF or CIE curve. The luminance of each segment was then measured using a luminance meter (LS-100; Konica Minolta, Tokyo, Japan) to compare them with the calculated curves made using Eqs. (1)–(5).

Psychophysical analysis

Figure 2 shows the relation between the luminance and contrast in a general sinusoidal pattern. The contrast is expressed by

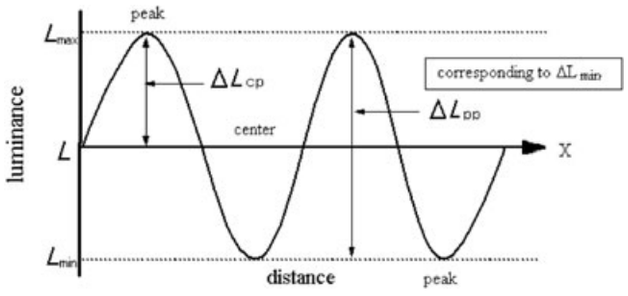


Fig. 2. Relation between ΔL_{cp} and ΔL_{pp} in the sinusoidal pattern used for definition of physical contrast and visual perception of objects, respectively. When the amplitude is just at the threshold level of visual perception, ΔL_{pp} is called $\Delta L_{\min}$

$$C_S = \frac{(L_{\max} - L_{\min})}{(L_{\max} + L_{\min})} = \frac{\Delta L_{cp}}{L} \tag{6}$$

where L and ΔL_{cp} are the luminance and the luminance difference between the center and peak levels in the sinusoidal pattern, respectively. In Fig. 2, when the amplitude of the sinusoidal pattern is just at the threshold level of visual perception, the luminance difference between the peak-to-peak levels is called the minimum perceptible luminance difference $\Delta L_{\min}$. This is because we perceive objects, including the sinusoidal pattern, by always using the difference between the maximum and minimum luminance. When the relation between ΔL_{cp} and $\Delta L_{\min}$ is considered as shown in Fig. 2, the threshold contrast C_t is expressed as

$$C_t = \frac{\Delta L_{\min}}{2L} \tag{7}$$

In psychophysical analysis, the psychophysical contrast is obtained by multiplying the physical contrast by a modulation transfer function (MTF) of the human eye. The MTF(Te) of the human eye (Te is defined as the inverse of C_t, expressed using the peak-to-peak of the sinusoidal pattern[9]) is given by

$$T_e = C_t^{-1} = \frac{2L}{\Delta L_{\min}} \tag{8}$$

According to Eqs. (6) and (7), the physical contrast C_p of the consecutive two segments is expressed as

$$C_P = \frac{\Delta L_{PP}}{2L} \tag{9}$$

where L denotes the mean luminance of the sinusoidal pattern, and ΔL_{pp} is the luminance difference corresponding to the peak-to-peak level, as shown in Fig. 2. Therefore the visually recognized contrast (S) is expressed as

$$S = \frac{\Delta L_{PP}}{2L} \times \frac{2L}{\Delta L_{\min}} = \frac{\Delta L_{PP}}{\Delta L_{\min}} \tag{10}$$

When the difference in the luminance between two segments is very small, ΔL_{pp} is proportional to the physical gradient G of the calibration curve. Thus, the psychophysical gradient δ is given by

$$\delta = \frac{G}{\Delta L_{\min}} \tag{11}$$

To compare the δs of the GSDF and CIE curves, the *G*s of those curves were substituted into Eq. (11), and the value of $\Delta L_{\min}$ obtained by the Barten model[10,11]

$$\frac{\Delta L_{\min 5}}{L} = \frac{2}{P_1}\exp\left(P_2 d^6\right)\sqrt{\frac{P_3}{d^2 L} + P_4}$$
$$P_1 = 1.147681\times10^{-1}, \quad P_2 = 1.579137\times10^{-6},$$
$$P_3 = 3.962775\times10^{-5}, \quad P_4 = 1.356244\times10^{-7}, \qquad (12)$$
$$d = 4.6 - 2.8\tan h\left\{0.4\cdot\log_{10}\left(0.625L\right)\right\}$$

was used for the denominator, similar to the DICOM standard.

ROC analysis of the tumor shadow

To evaluate the differences in detectability between the GSDF and CIE calibration curves, ROC analysis of clinical images was performed using the standard digital image database "Chest Lung Nodules and Nonnodules" provided by the Japanese Radiation Technology Society.[12] The database consists of 154 tumor shadow images and 93 nontumor shadow images, which are used for ROC analysis as positive and negative images, respectively. Altogether, 50 tumor shadow thoracic images and 50 nontumor shadow thoracic images corresponding to the positive and negative samples, respectively, were extracted randomly from the database and used for ROC analysis. The method for discrete confidence rating test results (five-category rating method) was carried out by six radiologists with 3 years or more of diagnostic experience. The observation conditions were as follows: room illuminance, 180 lx; maximum luminance of the LCD monitor, 450 cd/m^2; distance and time of observation were free. Light sources in the vicinity of the samples were covered with black paper to avoid the effects of glare. Observations were performed under steady-state adaptation to the room illuminance. Analysis was performed using the curve-fitting program ROCKIT for ROC analysis provided by Metz's ROC Software Users Group.[13,14] The results of ROC analysis were averaged for each calibration curve. Finally, the differences between the GSDF and CIE calibration curves were examined by the two-tailed *t*-test using the area below the ROC curve (Az),[15,16] with $P < 0.05$ indicating significance.

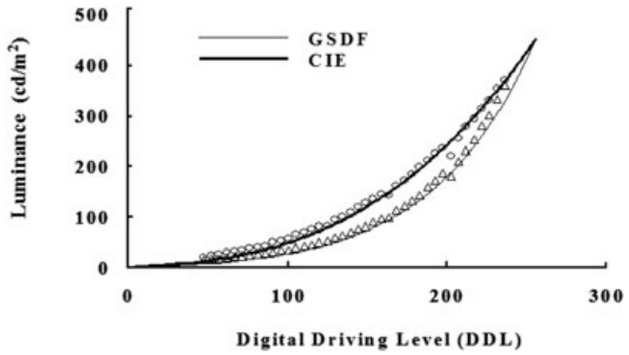


Fig. 3. Grayscale standard display function (*GSDF*) values (*triangles*) and Commission Internationale de l'Eclairage (*CIE*) values (*circles*) as a function of the digital driving level (DDL), which are measured using the NGTP and calculated by the Digital Imaging Communication in Medicine (DICOM) formula (*thin line*) and AAMP (*thick line*)

Results

Comparison of GSDF and CIE curves

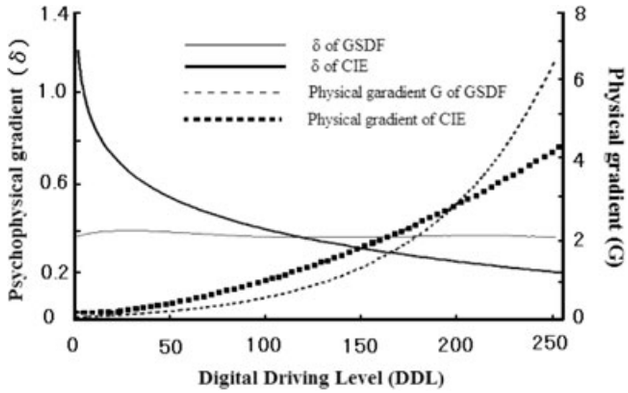
The open triangles and open circles in Fig. 3 express the measured GSDF and CIE calibration curves, respectively, as a function of DDL. Two solid lines in Fig. 3 express the calculated GSDF curve (thin line) and CIE curve (thick line), respectively, obtained using Eqs. (3)–(7). In both curves, there is good agreement between the measured value and the calculated value. The changes in luminance found by calibration using the GSDF curve are smaller up to the intermediate region on the abscissa than those found by calibration using the CIE curve; the differences are markedly larger above this region.

Comparison by physical gradient *G* and psychophysical gradient δ

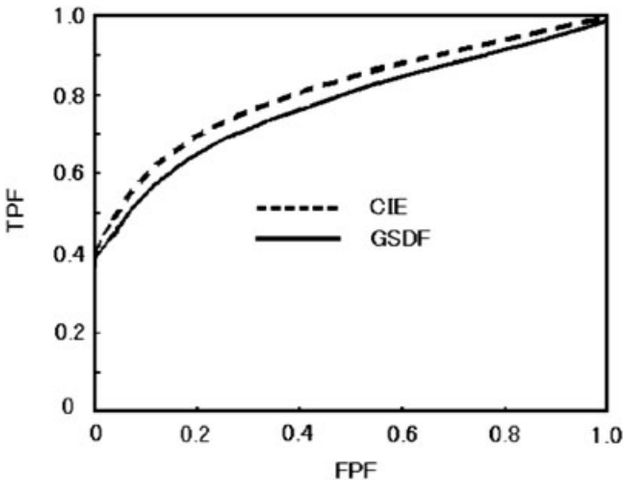
The broken lines and solid lines in Fig. 4 show the physical gradient *G* and the psychophysical gradient δ curves, respectively. In both broken and solid lines, it is also shown that the thin and thick lines are the results for the GSDF and CIE curves, respectively. On the δ, the GSDF curve is almost constant during the changes in DDL, whereas the CIE curve decreases markedly with an increase of DDL. Two curves of the GSDF and the CIE intersect in the immediate neighborhood of 200 DDL in the *G*, and around 120 DDL in the δ, respectively.

ROC analysis of tumor shadows

Figure 5 shows the results of ROC analysis of the thoracic tumor shadow images after calibration of the



3 **Fig. 4.** Psychophysical gradients (δ) and physical gradients (G) as a function of the DDL



4 **Fig. 5.** Average receiver operating characteristic (ROC) curve showing chest lung nodule detection performance. TPF, ■■; FPF, ■■

GSDF or CIE curve in the LCD monitor at a maximum luminance of 450 cd/m^2. The area parameter Az was 0.8331 with the CIE curve and 0.774 with the GSDF curve, respectively. This fact indicates that the CIE curve has slightly higher detectability than the GSDF curve. However, the paired t-test yielded a value of $p = 0.666$, which means that the difference is not significant.

Discussion

As can be seen from Fig. 3, the LCD monitor used in this study is accurately calibrated to the GSDF or CIE curve. Accordingly, a reliable comparison of these curves can be performed. The GSDF curve has been used for a wide range of image-output devices. This is because DICOM has emphasized the importance of making perceptual linearization on the calibration curve. Moreover, in clinical diagnosis, high detectability of a tumor shadow under actual conditions should be considered in preference to any other factor. To our knowledge, there have been no previous reports of investigating the relation between perceptual linearity and tumor detectability on calibration curves. Therefore, the discussion focuses on that relation.

We succeeded in expressing visually recognized contrast directly using δ instead of the JND index of the DICOM standard. As can be seen from Fig. 4, the δ value of the GSDF curve becomes almost constant for abscissa DDL (i.e., the perceptual linearization is valid), whereas that of the CIE curve decreases markedly with an increase in DDL. However, the absolute value of δ of the CIE curve is markedly larger than that of the GSDF curve in the low DDL range. Thus, the CIE curve can provide more visually recognized contrast to observers than the GSDF curve in low luminance areas on the LCD monitor.

On the other hand, as shown in Fig. 5, there is no significant difference regarding tumor detectability between the GSDF and CIE curves for clinical thoracic images. With ROC analysis, however, the area parameter Az of the CIE curve is superior to that of the GSDF curve. The lung field that occupies most of the thoracic image is in the low luminance range (e.g., approximately 20–40 cd/m^2 for maximum luminance of 450 cd/m^2). A comparison of Figs. 4 and 5 indicates that the improvement in tumor detectability on clinical thoracic images can be related to the absolute value of the psychophysical gradient δ rather than the shape of the δ curve, such as perceptual linearization. Moreover, it has become feasible to explain the above relation only by using the psychophysical gradient δ, not the physical gradient G, because the intersected point between the GSDF and CIE curves in G is inconsistent with the ROC analysis results.

From the viewpoints of tumor detectability and the absolute value of δ, we do not support use of the GSDF curve proposed by the DICOM standard for consistency of output gradation characteristic of clinical images. It is not suitable as a calibration tool for the diagnostic LCD monitor.

Conclusion

To determine a suitable calibration curve for diagnostic LCD monitors, GSDF and CIE curves were compared using the psychophysical gradient δ and ROC analysis for clinical images. The δ of the GSDF curve became almost constant for DDL changes (i.e., the perceptual

G

linearization was valid), whereas that of the CIE curve decreased markedly with an increase in DDL. The absolute value of δ was markedly larger in the CIE curve than in the GSDF curve at low luminance areas. Tumor detectability of the CIE curve for clinical images was superior to that of the GSDF curve owing to the high absolute value of δ in the low luminance range. By synthetically considering these factors, we conclude that the CIE curve, not the GSDF curve, is the suitable tool for calibrating diagnostic LCD monitors.

References

1. Digital Imaging and Communication in Medicine (DICOM). Part 14: Grayscale Standard Display Function. Rosslyn, VA, USA: National Electrical Manufacturers Association; 1999. p. 1–14.
2. Assessment of display performance for medical imaging systems. American Association of Physicists in Medicine Task Group 18 Imaging Informatics Subcommittee. 2005.
3. Kanamori H. The determination of the optimum density and the density-range of radiographs from visual effects. Jpn J Appl Phys 1964;3:286–94.
4. Kanamori H. Determination of optimum film density range for roentgenograms from visual effects. Acta Radiol Diagn 1966;4:463–76.
5. Asai Y, Shintani Y, Yamaguchi M, Uemura M, Matsumoto M, Kanamori H. Evaluation of grey-scale standard display function as a calibration tool for diagnostic liquid crystal display monitors using psychophysical analysis. Med Biol Eng Comput 2005;43:319–24.
6. Yamaguchi M, Fujita H, Asai Y, Uemura M, Ookura Y, Matsumoto M, et al. Psychophysical analysis of monitor display function affecting observer diagnostic performance of CT image on liquid crystal display monitors. Eur Radiol 2005;15:2487–96.
7. Uemura M, Asai Y,Yamaguchi M, Shintani Y, Sanada S. Using psychophysical analysis to compare gradation characteristics among three typical medical imaging display devices. Jpn J Radiol Technol 2005;61:1587–91.
8. Yamaguchi M, Fujita H, Uemura M, Asai Y, Wakae H, Ishifuro M. Development and evaluation of a new gray-scale test pattern to adjust gradients of thoracic CT imaging. Eur Radiol 2004;14:2357–61.
9. Patel AS. Spatial resolution by the human visual system: the effect of mean retinal illuminance. J Opt Soc Am 1966;56:689–94.
10. Barten PGJ. Physical model for the contrast sensitivity of the human eye. Proc SPIE 1992;1666:57–72.
11. Barten PGJ. Spatio-temporal model for the contrast sensitivity of the human eye and its temporal aspects. Proc SPIE 1993;1913:2–14.
12. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. AJR Am J Roentgenol 2000;174:71–4.
13. Metz CE. ROC methodology in radiologic imaging. Invest Radiol 1986;21:720–33.
14. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol 1989;24:234–45.
15. Shiraishi J. The evaluation of the diagnosability—the experimental method of the ROC analysis. Jpn J Radiol Technol 1999;55:362–8.
16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.

AUTHOR QUERY FORM

Dear Author

During the preparation of your manuscript, the questions listed below have arisen. Please answer **all** the queries (marking any other corrections on the proof enclosed) and return this form with your proofs.

Query no.	Querys
1.	AU: Is your meaning conveyed? If not, please reword to clarify.
2.	AU: Please define "AAMP."
3.	AU : The deleted material is explained within the figure.
4.	AU: Please define TPF and FPF.