This is a post-peer-review, pre-copyedit version of an article published in Radiological Physics and Technology.
The final authenticated version is available online at: https://doi.org/10.1007/s12194-019-00503-z

Authors:
Takaaki Isayama[1,2], Sadamitsu Nishihara[3] and Hideki Otsuka[3]

Title:

Proposal of a new method to prove that unnecessary information is not drawn on the image using statistical analysis

Affiliations and Addresses of authors:
1): School of Health Sciences, Tokushima University, 3-18-15, Kuramoto-cho, Tokushima 770-8503, JAPAN
2): Present Address: Graduate School of Health Sciences, Tokushima University, 3-18-15, Kuramoto-cho, Tokushima 770-8503, JAPAN
3): Tokushima University Graduate School of Biomedical Sciences, 3-18-15, Kuramoto-cho, Tokushima 770-8503, JAPAN

e-mail address, telephone and fax numbers of the corresponding author:
Name: Sadamitsu Nishihara
E-mail: nishihra.sadamitsu@tokushima-u.ac.jp
Telephone: +81-88-633-9864
Facsimile: +81-88-633-9864

Abstract:

The purpose of this study is to propose a new method of image evaluation using statistical analysis. We used the Sign test and the Wilcoxon test to analyze the statistical significance of image differences. Using this method, we evaluated whether the small electrode of the DAP meter appears in the X-ray image. Two observed values, which were obtained by averaging all values under all exposure conditions, were compared. All the observation tests showed the same sign. Thus, the results proved that

the small electrode of the DAP meter is not present on the image. Using this method, it became possible to prove that the electrode was not depicted, which was impossible to determine using conventional methods. The method combining both the Sign test and the Wilcoxon test can be useful in image evaluation.

Keywords:
dose area product (DAP) meter, observer test, Sign test, Wilcoxon test, significant difference, small electrode

1 Introduction

The Sign test and Wilcoxon test are used to identify any statistically significant differences in binomial distribution [1]. In the Sign test, + and/or - signs are given by the magnitude of the value that is being examined. The P value is determined using the smaller code numbers as follows (Equation 1):

$$P = \left({}_{n}C_{0} + {}_{n}C_{1} + \cdots + {}_{n}C_{r}\right)\left(\frac{1}{2}\right)^{n} \qquad (1)$$

In Equation 1, the “n” represents the total code numbers to be compared, and “r” indicates the smaller number. If the binomial probability (P) is >0.05, then the null hypothesis cannot be denied. Thus, the presence or absence of a significant difference cannot be determined.

The Wilcoxon test is also used to determine the statistical significance of any differences, which are ranked according to the difference in the examined value. The + and/or - codes are given by the differences. The ranksum (T) is obtained from the value with the smaller code. The statistical significance of T is examined for the total code numbers “n” and compared using the Wilcoxon test table (when “n” is smaller than 25). If the ranksum T is >0.05, the null hypothesis cannot be ruled out and the difference is not considered to be statistically significant. Both the Sign test and the Wilcoxon test can prove that a there is a significant difference, but they cannot prove that there is no significant difference. The null hypothesis cannot be denied because neither test can distinguish whether there is no significant difference or whether the difference is not significant because the number of samples is insufficient.

In Japan, the entrance skin dose is used to assess the radiation exposure of a patient in the general imaging area. The diagnostic reference level (DRL) was published by the Japan Network for Research and Information on Medical Exposure (J-RIME) in 2015; however, the DRL also reflects the entrance surface dose [2]. In Europe and the United States, a dose area product (DAP) meter is used [3-8]. The DAP meters that are currently commercially available can be mounted on the movable diaphragm of the X-ray apparatus and a small electrode placed at the center can simultaneously estimate the air kerma. The DAP meter used in this study is shown in Figure 1. The small electrodes can clearly be seen. Although there is a risk of influencing the diagnosis if they are depicted on the photographed X-ray image, there have been no studies to show that the electrodes are not drawn. Only the X-ray absorption of the DAP meter has been discussed, without considering the existence of the small electrode itself [9]. The DAP meter used for evaluation is compliant to IEC 60580. The requirements of 4.8.5.4 of IEC 60580 2nd edition specifically describe concepts such as the X-ray transmittance of a DAP chamber. This requirement indicates that the quality of equivalent filtration of the ionization chamber shall not exceed 0.5 mm aluminum with a purity > 99 %. (The X-radiation generates an X-ray tube voltage of 70 kV with a percentage ripple < 10 %

and a total filtration of 2 mm aluminum.)

When visually evaluating images, we use statistical analyses to investigate the significance of differences. By allowing participants to observe images in which a DAP meter was installed and images in which a DAP meter was not installed, the absence of the small electrode on the image can be proven if it can be demonstrated that the difference is not statistically significant. However, while it is possible to prove a significant difference, no statistical methodology exists to prove that a difference is not significant. It is advantageous to obtain results that show that it is unnecessary to consider the influence of the small electrode on the X-ray image, as it is rejected. This proof implies that the DAP meter can be used freely. We analyzed the results of observer tests using two values (defined as the correct answer fraction, CAF) and proved that the small electrodes of the DAP meter do not appear on X-ray images by proving that there was a significant difference in the CAF. The two values were obtained by averaging all findings under all exposure conditions.

2 Materials and Methods

2.1. Creating and displaying an observed data set

We set up the DAP meter (PD-9100; Toreck Co., LTD. Yokohama, Japan) on the movable diaphragm of a general X-ray system (MRAD-A50S/70; Toshiba Medical

Systems, Nasu, Japan) and an image of the observed data set was photographed using our X-ray system. Table 1 shows the exposure conditions and Figure 2 indicates the geometric scheme for the observation image. The conditions in Table 1 (nine types of exposure conditions) include the maximum and minimum conditions in the clinical setting. Thus, if a significant difference were to be observed in this experiment, then it is recognized that there was no influence of the presence of the electrode on the X-ray image on usual examination. When the DAP meter was present, two images were obtained; when no DAP meter was present, one image was obtained. Each image was obtained under the same exposure condition with and without the DAP meter. A total of 27 observation image data sets were obtained.

The observation image was read with a CR system, REGIUS model 170 (Konica Minolta, Tokyo, Japan). By setting the reading to manual, the entrance X-ray dose is expressed linearly as a pixel value. The ImageJ software program (NIH, available at http://rsb.info.nih.gov/ij/) was used to ensure that all images had the same pixel value on the observation monitor. First, the original pixel value was converted exponentially. Second, the average value was adjusted to the average pixel value of condition C by the Divide function. Finally, it was returned to the logarithmic display. When we displayed our images, we set the window level to 100 and the window width to 500.

2.2 Observer test

Fifteen participants observed the images randomly. A RadiForce R22 (EIZO Co., Ltd. Hakusan, Japan) was used as an observation monitor. A black piece of paper with a square cut out was affixed in the same position as the small electrode and could be observed on the monitor. Before the experiment, we explained to the observer that the targeted small electrode was 4 cm × 4 cm in size on the image. We did not consider the stimulus-response matrix. Likewise, irrespective of whether stimulation was present or not, the right side of one 5 cm line segment was taken as the maximum value.

One participant observed all 27 images {(two signal (+) images + one signal (-) image) x 9 exposure conditions = 27 images}. If the observer felt that the small electrode (signal) was present, he/she placed a mark on the right side of a 5-cm line segment. If the electrode was not present, then the observer placed a mark on the left side. The position marked by the observer was displayed in length from the left end. Of the 27 observed values, the position corresponding to the far-right side of the line segment was regarded as the maximum value (defined as 1) of the participant. All other results were normalized with a maximum value of 1. Fifteen participants performed the same task, and the average value for each image was calculated. These average values are shown as observed values.

2.3. Statistical analysis

In our study, the Sign test and Wilcoxon test were used to determine the significance of differences. Briefly, the differences between the Sign test and the Wilcoxon test are as follows: The Sign test simply analyzed which result was significant, and this test is based on a binomial distribution. Therefore, in the Sign test, only the direction of the difference is taken into consideration. On the other hand, the Wilcoxon test not only shows the number, but it also ranks and displays the magnitude of both differences (difference). In the Wilcoxon test, the magnitude of the difference is also taken into consideration in the order of ranking; thus, its detection power is high.

In the conventional method, the presence or absence of the DAP meter is compared with the standardized value. In other words, each result was compared (show as the observed value) regardless of the presence or absence of the DAP meter. In the proposed method, two types of CAF are used. The CAF of the DAP meter (+) was the same as that of the conventional method; the other CAF was calculated as follows: 1 – {observed value of the DAP meter (-)}. If the other CAF was significantly higher than the CAF of DAP meter (+), it proved that the small electrode was not depicted.

3 Results

3.1 Sign test

The exposure conditions are listed on the left side of Table 2 (A to I). The results obtained by the conventional method are shown in the middle and the results obtained by the proposed method on the right. In the middle of Table 2, the observed value was used (for example, representative values for condition I were 0.205, 0.295). We compared the observed values obtained when the DAP meter was included with those obtained when the DAP meter was not included. The result was recognized as "+" when the observed value of the images including the DAP meter was higher than that of the images that did not include the DAP meter; while the value was recognized as "-" when the value of the images including the DAP meter was smaller than the images that did not include a DAP meter.

In the conventional method, the probability that "+" was three (and "-" becomes 6 at the same time) was 0.254, and the probability was greater than the level of significance (P=0.05). Thus, the null hypothesis could not be ruled out.

On the other hand, in the proposed method, when the images obtained with a DAP meter were "+" the result was equivalent to the “observed value” and when the images without a DAP meter were "-" the result was equivalent to the “1-observed value”. That is, “1-0.295 = 0.705 (condition = I)”. As nine "+" signs were shown, the P value was 0.002 based on equation 1 (the "-" sign is zero; r = 0). This probability was

<0.05. Thus, a significant difference was confirmed.

3.2 Wilcoxon test

Figure 3 shows two graphs of the output values for each condition. The magnitude of the difference for each ranking is shown in Table 3. There were nine exposure conditions and the code number was nine. Using the Wilcoxon test table, when the number of codes (n) to be compared is 9, the point at which T shows significance (P = 0.05) is 5.

In the conventional method, when the exposure conditions were three (C, D and I), the values of the images without the DAP meter became higher (Table 3). The ranksum T at this time was 17. The result (T = 17) was ≥5 and did not reach 0.05. Thus, it was not considered to be a significant difference.

In contrast, under the proposed method, the values of the images without the DAP meter were high for all conditions (right side of Fig. 3). The ranksum at this time was 0. This result (T = 5) is considered to reflect statistical significance at a significance level of 0.05. Based on these statistically significant results, it can be stated that the small electrode was not included in the image.

4 Discussion

The entrance surface dose is used to measure radiation exposure in general imaging areas in Japan. The published DRL also refers to the entrance surface dose, which is measured with an ionization chamber dosimeter. However, we believe that the dose can be more accurately measured by a method that considers the size of the X-ray radiation field, such as the method that is used in Europe and the United States. Commercially available DAP meters not only measure the area dose but also simultaneously estimate air kerma, which is useful for simply estimating the dose. It is also possible to keep the dosimeter attached to the X-ray apparatus and to measure exposure during actual imaging. Before DAP meters can be used in a clinical setting, there are many problems that must be solved, including how to handle the value of the area dose. In this study, as a first step, we investigated whether the small electrode of the DAP meter was depicted in X-rays.

We used the Sign test and Wilcoxon test. The Sign test only evaluates the number of signs. On the other hand, the Wilcoxon test includes both the sign and the magnitude of the sensitivity difference. Using the conventional method, neither method showed a significant difference. If we can prove that there is no significant difference, then it could be stated that the DAP meter is not shown; however, it is not possible to prove that there no significant difference using conventional statistical methods.

In the proposed method, two types of CAF were devised for the statistical analysis to prove that there was a significant difference between them. As a result, both observation tests showed the same sign. If the CAF of the image without the DAP meter was significantly higher, then the observer did not recognize the DAP meter in the image. That is, the analysis would prove that the small electrodes of the DAP meter were not shown on the image. Thus, the method described in the present study made it possible to prove that the leads were not depicted, which is impossible with conventional methods. The proposed method proved that neither the Sign test nor the Wilcoxon test showed the presence of the small electrode of the DAP meter in the image.

5 Conclusion

In our study, the Sign test and Wilcoxon test were used to analyze the statistical significance of differences. In the proposed method, two types of CAF are used, and significant differences were recognized in both tests. The proposed method demonstrated that the small electrodes of the DAP meter were not observed in the image.

Compliance with ethical standards

Conflict of interest

All authors declare that they have no conflicts of interest.

Human and Animal Rights

All study procedures involving human participants were performed in accordance with the 1964 Declaration of Helsinki. Furthermore, this study did not contain any animals.

Informed Consent

Our institutional review board approved the use of the image database and students of the Tokushima University in this observation study (authorization number: 2797). Informed consent for the study was obtained from all participants.

Acknowledgment

We are grateful to Mr. Yasushi Matsuda of TORECK CO., LTD. for improving the manuscript and for their valuable discussions.

References

1) Ichikawa K. Statistics for Bioscience. Tokyo: Nankodo; 1991. p.42-45, p.50-53, p.359. (In Japanese)

2) Diagnostic Reference Levels Based on Latest Survers in Japan – Japan DRLs 2015 –. http://www.radher.jp/J-RIME/report/DRLhoukokusyoEng.pdf. Accessed 26 Jun 2018.

3) Karambatsakidou A, Tornvall P, Salh N, Chouliaras T, Löfberg PO, Fransson A. Skin dose alarm levels in cardiac angiography procedures: is a single DAP value sufficient? Br J Radiol. 2005; 78(933): 803-809.

4) Chida K, Saito H, Otani H, Kohzuki M, Takahashi S, Yamada S, Shirato K, Zuguchi M. Relationship between fluoroscopic time, dose-Area Product, body weight, and maximum radiation skin dose in cardiac interventional procedures. AJR Am J Roentgenol. 2006; 186: 774-778.

5) Chida K, Kagaya Y, Saito H, Takai Y, Takahashi S, Yamada S, Kohzuki M, Zuguchi M. Total entrance skin dose: An effective indicator of maximum radiation dose to the skin during percutaneous coronary intervention. AJR Am J Roentgenol. 2007; 189: 224-227.

6) Chida K, Kaga Y, Haga Y, Kataoka N, Kumasaka E, Meguro T, Zuguchi M. Occupational dose in interventional radiology procedures. AJR Am J Roentgenol. 2013; 200: 138-141.

7) Falco MD, Masala S, Stefanini M, Bagalà P, Morosetti D, Calabria E, Tonnetti A, Verona-Rinati G. Effective-dose estimation in interventional radiological procedures. Rad Phys & Tech. 2018; 11: 149-155.

8) Balter S. Methods for measuring fluoroscopic skin dose. Pediar Radiol. 2006; 36 (Suppl 2): 136-140.

9) IEC 60580 2nd Edition. Medical electrical equipment - Dose area product meters. International Electrotechnical Commission; 2001.

Figure captions

Table 1: Exposure conditions

Table 2: Result of the Sign test. The exposure conditions are listed on the left (A to I). The results obtained by the conventional method are shown in the left and the results obtained by the proposed method are shown on the right.

Table 3: Magnitude of differences of each exposure condition. These signs of these differences are the same as those shown in Table 2.

Figure 1: Dose area product (DAP) meter. The small electrode placed at the center can simultaneously estimate the air kerma.

Figure 2: Geometry scheme for the observation image. Only the small electrode of the DAP meter is included in the image; the subject is not included.

Figure 3: Wilcoxon test results. The results obtained by the conventional method and the proposed method are shown on the left and right, respectively.

In the conventional method, the probability (observed value) of the DAP meter (-) was higher than that of the DAP meter (+) under the conditions C, D, and I.

Contrarily, in the proposed method, the CAF of DAP meter (-) was higher than that of DAP meter (+) under all conditions.

Figure1

Figure 1: The dose area product (DAP) meter. The small electrode placed at the center can simultaneously estimate the air kerma.

Figure2

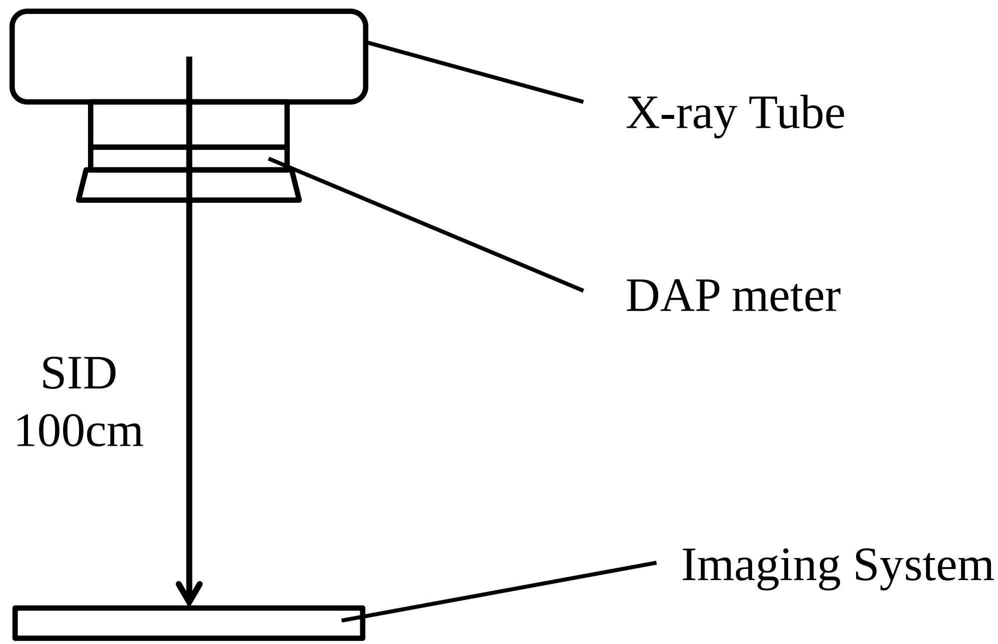

Figure 2: Geometry scheme for the observation image. Only the small electrode of the DAP meter is included in the image; the subject is not included.

Figure3

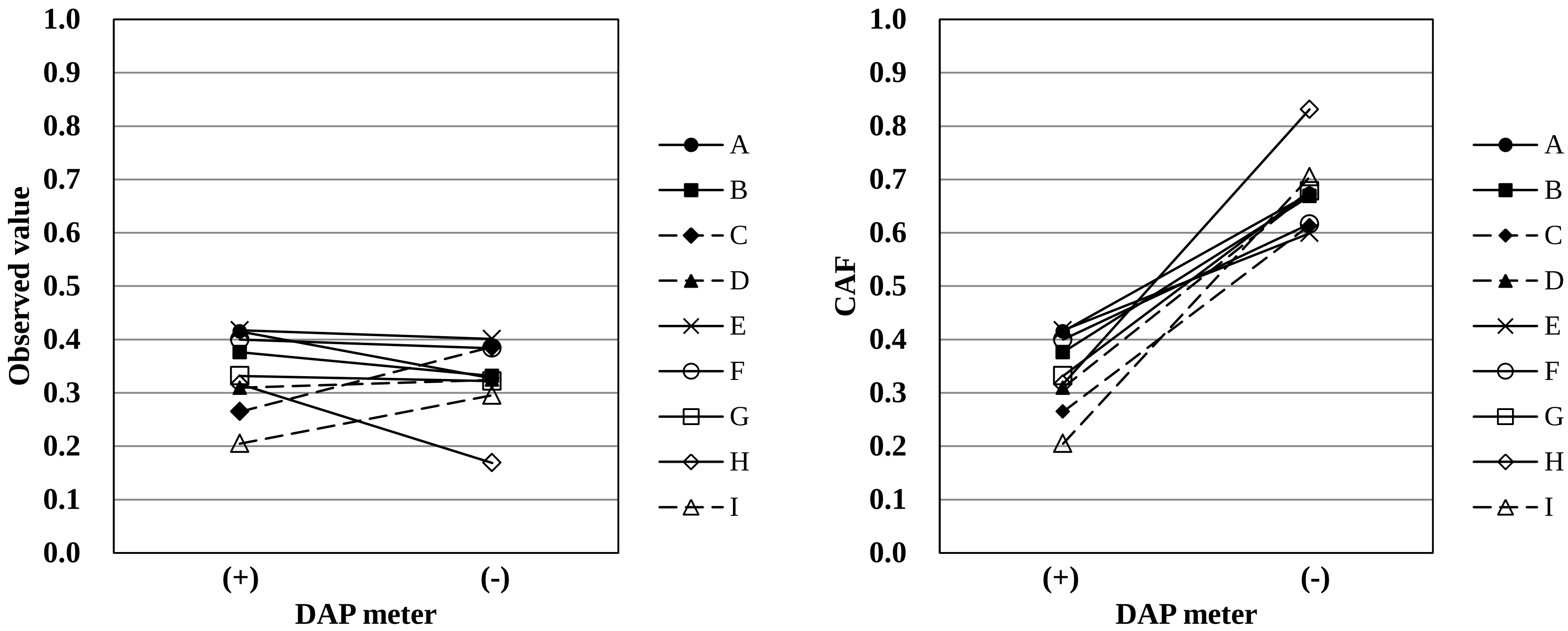

Figure 3: Observation results of Wilcoxon test; Left side was shown the conventional method and right side was figured the proposal method.

In the conventional method, the probability (observed value) of the DAP meter (-) were higher than those of the DAP meter (+) under the conditions C, D and I.
On the other hand, in the proposed method, the CAF of DAP meter (-) were higher than those of DAP meter (+) under all conditions.

Table 1: Exposure conditions

condition	Tube voltage (kV)	mAs
A	40	0.5
B	40	1.6
C	40	5
D	40	16
E	60	0.5
F	80	0.5
G	100	0.5
H	120	0.5
I	140	0.5

Table 3: The magnitude of differences of each exposure condition. These signs of these differences are the same as those shown in Table 2.

condition	DAP meter (+)	DAP meter (-)	difference	ranking
A	0.415	0.327	0.088	6
B	0.376	0.331	0.045	5
C	0.265	0.386	-0.122	8
D	0.309	0.324	-0.015	2
E	0.417	0.401	0.017	4
F	0.400	0.384	0.016	3
G	0.332	0.322	0.010	1
H	0.316	0.169	0.147	9
I	0.205	0.295	-0.091	7

Table 2: The result of Sign test. The exposure conditions are listed on the left (A to I). The results obtained by the conventional method are shown in the left and the results obtained by the proposed method are shown on the right.

condition	DAP meter		sign
	(+)	(-)	
A	0.415	0.327	-
B	0.376	0.331	-
C	0.265	0.386	+
D	0.309	0.324	+
E	0.417	0.401	-
F	0.400	0.384	-
G	0.332	0.322	-
H	0.316	0.169	-
I	0.205	0.295	+

condition	DAP meter		sign
	(+)	(-)	
A	0.415	0.673	+
B	0.376	0.669	+
C	0.265	0.614	+
D	0.309	0.676	+
E	0.417	0.599	+
F	0.400	0.616	+
G	0.332	0.678	+
H	0.316	0.831	+
I	0.205	0.705	+