資料

胸部ファントムを用いた一対比較法の Thurstone 法と Scheffé 法(Nakaya)による病変検出能の評価研究

望月安雄

日本医療科学大学保健医療学部診療放射線学科

論文受付
2013 年 2 月 27 日論文受理
2013 年 11 月 25 日

Code No. 520

緒言

医用画像の画質評価は、物理評価と視覚評価で行われている。視覚評価法は、画像に含まれる因子を観察して総合的に画質の良否を判断できるため、臨床画像の診断能を評価する手段として重用されている。

視覚評価の一つである一対比較の評価方法は、Thurstone 法と Scheffé 法が用いられている。Thurstone 法は提示した 2 枚の観察試料(画像)を比較し、試料間の優劣だけを判断して個々の試料の観察結果を集計する。その集計結果で心理学的尺度値を求め、試料の刺激を数量化(尺度化)¹⁾する。しかし、Thurstone 法は試料間の優劣の差は評価の対象としないため、試料間の差がはっきりしていると、試料間の優劣を強制的に判断

するために評価が偏る欠点が指摘¹⁾されている。

Scheffé 法²⁾は 2 個のデータを比較・判断するときの基準を、5 段階や 7 段階で評点化して評価する。そのため、Scheffé 法は試料の組み合わせや提示する順序に関する効果も統計的に調べることができるため、Thurstone 法よりも細かな分析が可能になる²⁾。しかし、Scheffé 法では試料間の差が僅かな時、2 個のデータは同程度と同位評価で判断され、評価の結果が曖昧になる欠点が報告²⁾されている。

従来、胸部撮影での視覚評価は、胸部全体あるいは肺野や心陰影部の領域での描写について報告され、肺野部や心陰影部などの領域に存在する病変検出の評価の報告³⁾は少ない。本報告は、官能検査法で主観的判

Evaluation Study of Abnormal Detectability with Thurstone and Scheffé (Nakaya) of Paired Comparison Method Using Chest Phantom

Yasuo Mochizuki

Faculty of Health Science, Department of Radiological Technology, Nihon Institute of Medical Science

Received February 27, 2013; Revision accepted November 25, 2013

Code No. 520

Summary

Monitors are increasingly being used as diagnostic imaging devices. In this study, using an all-purpose liquid-crystal display (LCD), the rate of detection of abnormalities was investigated using Thurstone's and Scheffé's (Nakaya) paired comparison methods. A chest phantom was prepared as a test sample with acrylic and aluminum plates and intensities suggesting small adenocarcinomas. For the acquisition conditions for computed radiography, after setting the baseline at a dose at which the film density of the standard screen-film system at the same as those for the lung, costal bone, and mediastinum, 5 steps of 2-fold serial doses were then set: 1/4, 1/2, 1, 2, and 4. The test sample was observed by 10 students. On the Thurstone scale, detectability decreased with a decrease in the dose in the lung, costal bone, and mediastinum. When the significance of differences between the values at adjacent doses was investigated using the yardstick method, using Scheffé's method revealed a significant difference between the 4- and 2-fold doses and between the 1/2 and 1/4 doses in the pulmonary region. A significant difference was also noted between the baseline and 1/2 doses in the mediastinum. Changes in the order of the scale values may not occur in the intervals in which significant differences were noted using Scheffé's methods.

Key words: paired comparisons method, Thurstone method, Scheffé (Nakaya) method, scale interval value, significance, Mosteller, Kendall

Fig. 1 The images of the lung used to observe the paired comparison.
A small adenocarcinoma, used as a test sample is shown in the circle.

断の測定が易しく、心理的判断を数量化⁴⁾して比較評価ができる一対比較法の利点に着目し、胸部ファントムの肺野部・縦隔部・肋骨部に模擬病変を配置し、胸部撮影における病変描出に適した撮影線量推定の可能性について調べた。

観察する表示系は、デジタル画像装置の進展に伴って、ハードコピーからソフトコピーへとフィルムレス化が進み、画像診断はモニタに移行している現状に沿って、汎用液晶モニタ(liquid crystal display: LCD)に観察試料を表示して検出能(尺度値)を求め、統計的有意差検定を行った。また、本報告ではLCDに描写される模擬病変の微小なコントラストに配慮して、一対比較法の評価はThurstone法とScheffé法を用い、receiver operating characteristic(ROC)解析と比べて放射線画像評価での適用の報告^{1, 2, 4)}が少ない一対比較法の有用性について、得られた知見を報告する。

1. 実験方法

1-1 観察試料画像の作成

肺野部、縦隔部、肋骨部を模擬した胸部ブロックファントム(以下、胸部ファントム)と5種類の異なる肺腺癌ファントム⁵⁾(京都科学製: No. 8, 9, 11, 12, 13)を信号として用い、それぞれの肺腺癌ファントムで肺野部、縦隔部、肋骨部の模擬X線写真を作成した。また、胸部ファントムは、肺野部では4 cm、縦隔部では10 cm、肋骨部では4 cmの亚克力ファントムを用い、肋骨部に更に0.2 cmのアルミニウム板を使用することで、レーザイメージャ DRYPRO MODEL 771(コニカ・ミノルタ社製)で得られた肺野部の濃度は1.70、縦隔の濃度は0.45、肋骨部の濃度は1.15と増感紙-フィルム系での胸部単純X線画像での適正条件と同等な濃度³⁾を

得ることができた。このときの撮影条件は焦点-輝尽性蛍光体プレート間距離 200 cm、管電圧 120 kV、管電流 200 mA、撮影時間 32 msで行った。この撮影線量を基準に、4倍線量 120 ms(A)、2倍線量 63 ms(B)、基準線量 32 ms(C)、1/2倍線量 16 ms(D)、1/4倍線量 8 ms(E)と変化させ、5種類の肺腺癌ファントムと一緒に撮影して前述の各部の濃度になるように調整した。Fig. 1に計15枚の観察試料を作成した肺野部での観察試料の1例を示し、模擬肺腺癌ファントムは破線の円内のように配置した。

Computed radiography(CR)装置はRegius 170(コニカ・ミノルタ社製)を用いた。撮影条件の画像処理の強調方法は、肺野部・肋骨部でハイブリッド、縦隔部ではスタンダードで、強調内容はイコライゼーション(E)と周波数(F)を共にOFFとして、撮影線量以外はすべて同一条件とした。散乱X線の除去は、移動型のアルミニウムグリッドで格子比 10:1、グリッド密度 40 本/cm(三田屋製作所)を使用し、X線発生装置はMODEL R-20(株式会社 島津製作所)で撮影した。

1-2 一対比較法による観察実験

LCD(EIZO, Flex Scan L 565)の観察条件は、LCD表面への外部光の影響がない角度に配置し、デジタル照度計 T-1(ミノルタ製)で、LCD周辺の照度が約 180 lxに室内光を調節し、digital imaging and communication in medicine(DICOM)viewerにPOP-Net Essential(イメージワン社)を用いてLCDに表示して観察実験を行った。

観察は1人あたり照射時間の異なる5種類の画像(試料)を、順次2枚ずつ提示し、肺野部、縦隔部、肋骨部のそれぞれの部位で10通りの評価を行った。そして、

観察者 10 人の結果を胸部ファントムの各部位ごとに Thurstone 法と Scheffé 法で集計し、尺度値²⁾で表した。なお、本研究の観察者は、診療放射線学科を専攻している 10 名の学生で行い、実験で得られたデータについては、結果の公表や教育教材の資料として使用できるルールを設け、同意を得ている。

Thurstone 法と Scheffé 法での評価基準は、肺野部、縦隔部、肋骨部の試料ごとに提示した 2 画像で 5 種の“模擬病変の描写”の優劣を総合的に判断した。また、本実験では Scheffé 法は 5 段階評価で行い、良い：+2、やや良い：+1、同程度：0、やや悪い：-1、悪い：-2 の 5 点系でのカテゴリ尺度を用いた。一対比較法の観察時の注意点は、試料画像 A と試料画像 B を往復で評価をしないことが重要である(順序効果)。しかし、2 枚同時に提示し医用画像を評価する場合、対象画像を左右に往復して評価(診断)するのが一般的で、試料画像の提示における順序効果は除外できると考えられる。また、観察者も 10 名と少ないために、Scheffé 法は順序効果を考慮しない中屋氏の変法⁶⁾を用いた。よって、試料画像の提示位置による効果(位置効果)のみを考慮した。

実際の観察は、観察距離と観察時間は各観察者で任意とし、観察グループを 2 班に分けて、左側の試料画像と右側の試料画像を差し替えた 2 パターンをランダムに提示した結果を集計して評価を判断した。

1-3 Thurstone 法の尺度化と適合度検定(Mosteller)

Thurstone 法は、肺野部・縦隔部・肋骨部の各部位で撮影線量ごとに実験データを集計した度数(f_{ij})でクロス表を作成した。この値を観察者 10 名で除して得られた観測比率(P_{ij})を、汎用表計算ソフト(Microsoft Excel 2003: Excel)の NORMSINV 関数によって標準正規分布得点(Z-スコア： Z_{ij})として求め、得点を観察画像(試料)の A, B, C, D, E の列で平均して間隔尺度の推定値を取得した。この推定値を心理尺度のヤードスティック上の一本の数直線上に配置することによって、観察画像の尺度値順位と画像間距離の関係性を見ることができ、また、列で平均して間隔尺度の推定値の総和が 0 のときケース V⁷⁾が成立する。

Thurstone 法は観察試料での尺度化(値)が妥当であるかの適合度は Mosteller の方法⁷⁾で検定できる。方法は尺度値を推定した手続きを逆戻りして、尺度値から比率の推定値を求め、それを逆正弦変換して漸近的に χ^2 の統計量を算出する。具体的には Excel を用い、 Z_{ij} から NORMSDIST 関数で期待比率の $\hat{p}_{ij}$ を求め、観測比率の P_{ij} と $\hat{p}_{ij}$ を逆正弦値 θ_{ij} と $\hat{\theta}_{ij}$ (度)に変換すると、 $\theta_{ij} - \hat{\theta}_{ij}$ は漸近的に平均値 0、分散が $821/n_{ij}$ の正規分布

に従う次式が得られる。

$$\left(\frac{180}{\pi}\right)^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{n_{ij}}\right) = \frac{821}{n_{ij}} \quad \dots\dots\dots (1)$$

ただし n_{ij} は観察者数。

したがって、 χ^2 検定統計量と自由度(df)は次式によって与えられる。

$$\chi^2 = \sum_{i < j} \frac{n_{ij}}{821} (\theta_{ij} - \hat{\theta}_{ij})^2 \quad \dots\dots\dots (2)$$

$$df = \frac{(n-1)(n-2)}{2} \quad \dots\dots\dots (3)$$

なお、逆正弦変換は $\theta_{ij} = \sin^{-1} \sqrt{P_{ij}}$ と $\hat{\theta}_{ij} = \sin^{-1} \sqrt{\hat{p}_{ij}}$ とで算出し、式(2)で Thurstone 法で得られた尺度値の適合度を検定することができる。

一対比較法における観察者間の判断の一致性は、Kendall⁸⁾によって検定方法が与えられ、次式で偏差平方和(S)を求める。

$$S = \frac{n(n-1)t(t-1)}{4} + \sum_{i=j} f_{ij}^2 - n \sum_{i=j} f_{ij} \quad \dots\dots\dots (4)$$

ただし、 t ：観察画像の試料数、 n ：観察者数。

次に、検定統計量 χ^2 を計算する。

$$\chi^2 = \frac{4}{n-2} \left(S - \frac{n(n-1)(n-3)t(t-1)}{8(n-2)} \right) \quad \dots\dots\dots (5)$$

その時の自由度(df)は次式で得られる。

$$df = \frac{n(n-1)t(t-1)}{2(n-2)^2} \quad \dots\dots\dots (6)$$

なお、式(4)の 2 項および 3 項は観察者の度数(f_{ij})のクロス表から計算できる。

1-4 Scheffé 法の刺激間距離の検定

Scheffé 法は、福田ら⁹⁾が Excel で作成した浦の変法の表計算ソフトを参考に、中屋の変法で表計算を行った。肺野部・縦隔部・肋骨部の各部位で分散分析を計算し、ヤードスティック上の座標の刺激間距離は、主効果の平均値(平均嗜好度)を次式^{6, 10)}で算出する。

$$\hat{\alpha}_i = \frac{1}{m} (X_{\cdot j} - X_{i \cdot}) \quad \dots\dots\dots (7)$$

ただし、 t ：刺激の個数(観察画像の試料数)、 n ：観察者数。分散分析で主効果に統計的有意差を認めた時は、 $\alpha_i - \alpha_j$ の信頼区間を求める。そのために有意水準 ϕ

のステューデント化された範囲 q_ϕ の表^{6,9)}から値を求め、次式でヤードスティック(Y_ϕ)を計算する。ただし V_e は誤差の不偏分散。

$$Y_\phi = q \sqrt{\frac{V_e}{tn}} \dots\dots\dots (8)$$

信頼区間 $\hat{\alpha}_i - \hat{\alpha}_j - Y_\phi \leq \hat{\alpha}_i - \hat{\alpha}_j \leq \hat{\alpha}_i - \hat{\alpha}_j + Y_\phi$ を求め、各画像間の順位の妥当性を確認するために、有意水準 1%や 5%の信頼区間とステューデント化された値(Y_ϕ)と撮影線量(刺激)間の距離を比較できる。

2. 実験結果

Thurstone 法の結果、本実験では欠測値は認められなかった。また、尺度値となる各画像間の Z-スコアの列で総和が 0 となり、ケース V が成立している。Thurstone 法での間隔尺度値をヤードスティック直線上に示す(Fig. 2)。検出能は肺野部・縦隔部・肋骨部のすべてで、4 倍線量(A)、2 倍線量(B)、基準線量(C)、1/2 倍線量(D)、1/4 倍線量(E)と、撮影線量の高い順になった。

Mosteller の適合度検定を式(2)で計算した結果、観測者 10 名、観察画像 5 枚、自由度が 6 で χ^2 値は 5.38 を得た。有意水準 1%と 5%での $\chi^2(0.01, 6)$ は 16.81, $\chi^2(0.05, 6)$ は 12.59 と有意水準での χ^2 の数値が、計算で得られた χ^2 値の 5.38 よりも大きく有意差は認められなかった。

Kendall 検定の一致性の係数 W は、0~1.0 の範囲で 1.0 に近いほど一致度がよい。本実験では肺野で 0.439、縦隔で 0.564、肋骨で 0.514 となり、 χ^2 値は肺野で 17.56、縦隔で 22.56、肋骨で 20.56 が得られた。自由度が約 14、有意水準が 1%と 5%の $\chi^2(0.01, 14)$ で 29.14、 $\chi^2(0.05, 14)$ で 23.68 と各部位で χ^2 の計算値が小さいことから、観測者相互間で有意差がないと判断した。

肺野・縦隔・肋骨の各部位の Scheffé 法での分散分析表を Table 1 に示す。F 検定の結果、肺野部と縦隔部の主効果において有意水準 0.01 で有意差が認められた。主効果で有意差を認めたため、自由度 54 の有意水準 1%と 5%のステューデント化された q_ϕ 値を用いて求めたヤードスティック値(Y_ϕ)と信頼区間を Table 2 示す。この時 Y_ϕ の自由度 54 の q_ϕ 値は数表にないため、近傍の自由度 50 と 60 の q_ϕ 値から内挿して得られた $Y_{(0.01)}$ で 4.843 と $Y_{(0.05)}$ で 3.991 によって計算を行った。

Table 2 から肺野部では信頼区間の A と B および D と E 間、縦隔部では B と C 間において有意水準 5%で有意差を認めた。しかし、肋骨部ではすべての信頼区

Fig. 2 Interval scale values using Thurstone's methods are presented on a yardstick line.

Fig. 3 Linear regression and coefficient of contribution using Scheffé's methods.

SIG (*): the level of significance was set at 5%.

間で 0 を含み、有意差は認められなかった。Fig. 3 に撮影線量を刺激に用いた座標をヤードスティック上に示した。アスタリスク(*)は 5%での有意差を示す。

3. 考 察

3-1 Thurstone 法の尺度値と検定

ヤードスティック上に表示した Thurstone 法での間隔尺度値の検出能は、肺野部・縦隔部・肋骨部のすべてで、線量が最も高い 4 倍線量で優れ、2 倍線量、基準線量、1/2 倍線量と続き、1/4 倍線量で劣り、撮影線量

Table 1 Analysis-of-variance tables in the pulmonary field, lung, mediastinum, and costal bone using Scheffé's method

	Factorial variation	Sum of squares (S)	Degrees of freedom (fd)	Variance (V)	Ratio of variance (F)	F-test decision	F (0.05) *	F (0.01) **
Pulmonary	Total (T)	100	67					
	Main effect	4	34	8.51	20.39	**	2.54	3.69
	Main effect×personal	36	6	0.18	0.42		1.63	2.01
	Combined effect	6	4	0.68	1.62		2.27	3.16
	Error (e)	54	23	0.42				
Mediastinum	Total (T)	100	101					
	Main effect	4	30	7.42	10.84	**	2.54	3.69
	Main effect×personal	36	26	0.71	1.04		1.63	2.01
	Combined effect	6	9	1.47	2.15		2.27	3.16
	Error (e)	54	37	0.68				
Lungs Rib	Total (T)	100	123					
	Main effect	4	8	2.11	1.83		2.54	3.69
	Main effect×personal	36	44	1.21	1.05		1.63	2.01
	Combined effect	6	9	1.48	1.28		2.27	3.16
	Error (e)	54	62	1.15				

A significant difference (**) was noted in the main effect in the mediastinum using the F test ($p \leq 0.01$).

Table 2 Results of 95% and 99% confidence intervals at each dose using Scheffé's methods

Y_ϕ	Interval scale $\alpha_i - \alpha_j$	Confidence interval			
		upper (95%)	lower (95%)	upper (99%)	lower (99%)
Pulmonary	$\hat{\alpha}_A - \hat{\alpha}_B = 0.38$	0.74	0.02	0.82	-0.06
	($Y_{0.05} = 0.364$)				
	$\hat{\alpha}_B - \hat{\alpha}_C = 0.12$	0.48	-0.24	0.56	-0.32
	($Y_{0.01} = 0.442$)				
Mediastinum	$\hat{\alpha}_C - \hat{\alpha}_D = 0.20$	0.56	-0.16	0.64	-0.24
	$\hat{\alpha}_D - \hat{\alpha}_E = 0.42$	0.78	0.06	0.86	-0.02
Lugs Rib	$\hat{\alpha}_A - \hat{\alpha}_B = 0.18$	0.65	-0.29	0.75	-0.39
	($Y_{0.05} = 0.467$)				
	$\hat{\alpha}_B - \hat{\alpha}_C = 0.54$	1.01	0.07	1.11	-0.03
	($Y_{0.01} = 0.567$)				
Lugs Rib	$\hat{\alpha}_C - \hat{\alpha}_D = 0$	0.47	-0.47	0.57	-0.57
	$\hat{\alpha}_D - \hat{\alpha}_E = 0.16$	0.63	-0.31	0.73	-0.41
Lugs Rib	$\hat{\alpha}_A - \hat{\alpha}_B = 0.30$	0.91	-0.31	1.03	-0.43
	($Y_{0.05} = 0.605$)				
	$\hat{\alpha}_B - \hat{\alpha}_C = 0.12$	0.73	-0.49	0.85	-0.61
	($Y_{0.01} = 0.735$)				
Lugs Rib	$\hat{\alpha}_C - \hat{\alpha}_D = 0.06$	0.67	-0.55	0.79	-0.67
	$\hat{\alpha}_D - \hat{\alpha}_E = 0.02$	0.63	-0.59	0.75	-0.71

$df=54$: α (0.05) at q value 3.991, α (0.01) at q value 4.843

A significant difference in the confidence interval was noted between the 4- and 2-fold doses and between the 1/2 and 1/4 doses in the pulmonary region. Using Scheffé's method, a significant difference in the confidence interval was noted between the baseline and 1/2 doses in the mediastinum.

の高い順となった(Fig. 2). この結果は高線量では線量の増加によって量子モトルが低減し、微小な模擬病変の識別コントラストが向上したと考えられる¹¹⁾. また、透過 X 線量が少ない 1/4 倍線量(D)で評価が低い、この要因はノイズの影響によって観察者が識別するためのコントラストが低下したために、画像信号の肺腺癌ファントムを識別するのが困難だったと推察できる。また、Thurstone 法で撮影線量に対応した尺度値が得られたの

は、必ず 2 個の試料に優劣を強制的につけなければならない判断基準が要因と考えられる。

Fig. 2 の尺度値が妥当であるかを Mosteller の方法で適合度検定した結果、観測比率と期待比率の適合度は有意水準 1% で有意差が認められず、観察試料の変動成分に問題はなく、間隔尺度値は妥当であると判定できた。また、Kendall の一致性の検定で、観察者間で統計的有意差が認められなかった。したがって、試料作成

や計算が平易な Thurstone 法は、X 線画像の優劣(良否)を簡易に判断する視覚評価の方法として適している。なお、本実験では田畑ら¹²⁾が提案した Thurstone 法の尺度値での信頼区間の検定は行っていない。

3-2 Scheffé 法の尺度値と信頼区間の検定

Scheffé 法(Nakaya)の分散分析の結果、肺野部と縦隔部の主効果において、条件が強い有意水準 0.01 でも有意差が認められたが有意水準 0.01 の信頼区間の計算では有意差が認められなかった(Table 1)。尺度距離順位の信憑性を確認するために、スチューデント化された q_0 値を用いて、ヤードスティック上の隣り合う尺度(刺激)間の距離の信頼区間を有意水準 0.05(*)で求めた。その結果、肺野部では A と B および D と E 間、縦隔部では B と C 間において有意差を認めた。しかし、分散分析では $p \leq 0.01(**)$ でも有意差が認められたが、信頼区間の計算では有意水準 0.01 で有意差がみられなかった。これは尺度値(検出能の評価)と撮影線量との間は全体としては関係があるが、個々の尺度間では有意な差が出なかったと考えられる。したがって観察者数を増やすか、線量の増減幅を大きくすれば有意差が生じる可能性が推測できる。

3-3 間隔尺度と撮影線量について

Fig. 3 に示した Scheffé 法の尺度(刺激)間の距離で、信頼区間が肺野部の 4 倍線量(A)と 2 倍線量(B)、1/2 倍線量(D)と 1/4 倍線量(E)間および縦隔部の B と C 間の線量間では有意差を認めたが、縦隔部の C と D の尺度値は同じで、肋骨部では D と E の尺度値がほぼ同等であった。したがって、有意差が認められた区間では、尺度値の順位の入れ替わりは起こらないと推定できる。しかし、他の区間では有意差が認められず、これらの間隔尺度間(画像間)では、観察者にとって試料の検出能が同程度であったと類推できる。この有意差が認められなかった範囲では、尺度順位の入れ替わりが起こる可能性を否定できない。

心理的判断を数量化した一対比較法の尺度値を、統計的有意差検定で評価した結果、Scheffé 法では尺度間距離を肺野部・縦隔部・肋骨部のすべての領域で総合的に判断すると、C と D 間では検出能に差がないと考えられる。したがって、これらの線量間では病変検出能に差がない画像が得られると類推できる。すなわち、被ばく低減の可能性を示唆している。したがって、画質指数(image quality figure: IQF)¹³⁾やノイズ特性(Wiener spectrum: WS)¹⁴⁾と撮影線量の関係から、被ばく線量を低減できると提案されている報告と同様な結果が得ら

れた。しかし、撮影線量と一対比較の尺度値の関係からは、撮影線量と IQF 値や WS 値で推定されている“しきい値”のような一定の規則性は認められなかった。

Scheffé 法は比較判断に差が僅かな 2 個の試料間では“同程度”の判断基準を設けているために、試料間の刺激差が小さいと評価の検出力が低くなってしまう可能性⁴⁾があり、本実験でも縦隔部の C と D 間で同じ尺度値となった。以上から、Scheffé 法は順位付け尺度距離による検出能の過大評価や過小評価に注意することで、相対的な画像評価に適していると考えられる。また、ROC 解析や IQF をはじめ他の視覚評価法は、微小なコントラストの検出能を評価するのに適しているが、一対比較法は“描写された画像の優劣”で評価するため観察画像に制約が少なく、高コントラストの画像や高線量で取得した画像で検出能を判断できる利点がある。これらから、主観的判断の測定が容易な Thurstone 法と Scheffé 法は、画像の優劣を概観して把握する手段としての視覚評価法として有用である。

3-4 Thurstone 法と Scheffé 法の表計算について

本実験は Thurstone 法と Scheffé 法の解析を汎用表計算ソフトの Excel 2003 で行った。Thurstone 法は尺度値を表計算で簡単に求めることができる。しかし、Mosteller の適合度と Kendall の一致性を検定する表計算が別途で必要となる。

Scheffé 法は表計算で反復計算する部分が多く若干煩雑であるが、スチューデント化された q_0 値から撮影線量(刺激)間の距離の信頼区間を求めることで統計的有意差検定が評価できるため、一度表計算ソフトを作成すると総合的には平易であった。

Thurstone 法と Scheffé 法の表計算は、 q_0 値の自由度を数表で確認する以外は Excel にある関数を用いることで数表を必要としない。したがって一連の作業をパソコン上で処理できるため、官能評価を瞬時に施行できた。また、フリーソフトの R 統計¹⁵⁾に用意されている Thurstone 法と Scheffé の原法および中屋変法の 7 段階の評点法では、尺度値の表作成まで実行できる。なお、Thurstone 法では R 統計で、本実験の Excel での計算と同じ結果を得ている。

以上、ガンマ特性が 2.2 の汎用 LCD を用い、一対比較法の Thurstone 法と Scheffé 法によって尺度値で評価した検出能は、基本的に高線量の画像で評価が高く、低線量の画像で評価が低い妥当な結果となった。今後、観察者が識別できる最少の輝度差[just noticeable difference(JND, 最少弁別閾)]による Grayscale Standard Display Function(GSDF)に準拠した高精細

LCDを用い、一対比較法の尺度値で観察画像と撮影線量について検証を行い、輝度や画素の影響についての関係も顕かにすることが課題として残った。

4. 結 語

ガンマ特性が2.2の汎用液晶モニタを用い、一対比較法のThurstone法とScheffé法の集計値をヤードスティック上で尺度化し、撮影線量の増減による画像(試料)間の検出能の関係性を調べ、一対比較法の有用性について調べた。

Thurstone法ではMostellerの適合度(試料間の変動)とKendallの一致性(観察者間の変動)、Scheffé法ではヤードスティック Y_0 で信頼区間を計算して統計的有意差検定を行った。その結果、Thurstone法は間隔尺度から評価画像の概観を知るのに優れ、Scheffé法は5段階の評点で判断して結果を得ることができ、分散分析を用いることによって主効果だけでなく試料の組み合わせ

に対する効果なども統計的有意差検定で調べることができた。しかし、一対比較法は尺度値で順位は判断できるが、IQFやWS値のように線量と検出能に一定の規則性が認められず、一対比較法の尺度値での判断は、過大評価や過小評価にならないように注意することが必要である。

以上、官能評価法のThurstone法やScheffé法は、優劣を比較する画像(試料)作成や尺度距離によって検出能が平易に評価できるため、画像の良否を概観で簡易に判断するのに適した視覚評価法である。

謝 辞

日本医療科学大学保健医療学部診療放射線学科に在籍し、視覚評価実験の試料作成や観察および解析に協力いただきました学生の皆さんに感謝いたします。

参考文献

- 1) 大賀泰文, 辻本武士, 田畑洋二, 他. X線写真の主観的評価法—Thurstoneの一対比較法による尺度化の試み—. 日放技学誌 1989; 45(7): 831-839.
- 2) 中前光弘. 統計的官能検査法の理論と放射線技術科学への応用—Scheffé(シェッフェ)の一対比較法を中心に—. 日放技学誌 2010; 66(11): 1502-1507.
- 3) 望月安雄, 阿部慎司, 門間正彦, 他. 肺腺癌ファントムを用いたソフトコピー(LCD)とハードコピー(Film)の病変検出能の比較. 日放技学誌 2011; 67(4): 328-335.
- 4) 中前光弘, 田畑洋二, 大賀泰文, 他. Schefféの一対比較法による主観的評価法. 日放技学誌 1996; 52(11): 1561-1565.
- 5) 小野孝二, 赤羽恵一, 羽田道彦, 他. 視覚評価のための肺腺癌模擬病変ファントムの開発. 日放技学誌 2004; 60(9): 1301-1307.
- 6) 佐藤 信. 統計的官能検査法. 東京: 日科技連出版社, 1985.
- 7) Mosteller F. Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. Psychometrika 1951; 16 (2): 207-218.
- 8) Kendall MG, Smith B. On the method of paired comparisons. Biometrika 1939; 31: 324-345.
- 9) 福田忠彦 監. 人間工学ガイド 感性を科学する方法. 東京: サイエンティスト社, 2009.
- 10) 辺見秀一. 胸部CT画像の肺野および胸郭・縦隔の写真濃度に関する視覚的印象—視覚的主観評価による錯視現象の検証—. 日放技学誌 2004; 60(7): 983-992.
- 11) 望月安雄, 小川 互, 小野寺浩幸, 他. 増感紙-フィルム系のNEQ(u)と一対比較法の相関—物理評価と視覚評価の検討—. 日放技学誌 1998; 54(3): 363-370.
- 12) 田畑洋二, 大賀泰文, 角田充弘, 他. X線写真の主観的評価法—主観的尺度値の信頼性—. 日放技学誌 1993; 49(1): 13-17.
- 13) 望月安雄, 阿部慎司, 山口弘次郎, 他. Image quality figureの閾値によるComputed Radiographyでの適正線量の推定. 日放技学誌 2009; 65(4): 430-437.
- 14) 望月安雄, 新藤博明. Computed radiographyにおける撮影線量最適化のための簡易決定法の基礎研究. 日放技誌 2012; 59(2): 185-192.
- 15) 青木繁伸. Rによる統計解析. 東京: オーム社, 2009.

問合せ先

〒350-0435 入間郡毛呂山町下川原 1276

日本医療科学大学保健医療学部診療放射線学科 望月安雄