

《原 著》

シンチグラム処理における定量的評価の試み

福久健二郎* 飯沼 武* 永井 輝夫**

要旨 電子計算機による RI 像処理の研究は、1964年ごろから急速に進展した。しかし、これらの処理法のうち、どれが最適かを客観的に判定するのは容易ではない。この客観的方法として総得点法、ROC 解析、McNeil らの方法が知られているが、本報ではこれらの特性を検討する目的で IAEA の電算機シミュレーテッド・ファントムを用い、それぞれの方法で相互比較した。処理法はガウス分布のフィルタとオプティマム・フィルタを行い、同一条件でプロッタに表示し、4名のシンチグラム専門医に読影願った。また、判定法は主に IAEA が示す方法を用いた。

結果として、いずれの比較方法によても、個人差や処理法による差は明確になるが、得点法では FN の評価法、ROC 解析では FP がない場合の比較に問題があるほか、誤診の重大さをどう個々の臨床評価で考慮するかが大切であることを示した。

I. 緒 言

近年、電子計算機を用いて RI 像を処理し、画質改良を行なうことによって癌などの疾患の正診率向上をはかる試みが各種行なわれてきた¹⁻³⁾。RI 像の画質改良の目的は、シンチレーションカメラなどのイメージング装置の解像力に由来する像のぼけと、放射線計数に由来する雑音を除去し、可能なかぎり診断すべき情報を描出し得る像に修正することであり、この両者を同時に処理する方法が良好な成績を得ることも知られている²⁻⁶⁾。

しかしながら、真に画像処理の効果を判定するためには、複数の既知の病変を含む像を用い、処理法または表示法を変化させて複数の読影者による読影結果を統一的方法で解析評価する必要がある。このうちの既知の病変を含む像の収集が最も困難な問題の一つであり、実際のシンチグラムでは肝などの限られた部位につき、シンチグラム撮

影と手術・剖検などの病変確認とがほぼ同時期に施行され、情報記録が完全であることが要求される⁷⁾。第二には処理した像の再表示法による読影差の問題があり、とくに読影者の経験的要素が与える影響は大きい。第三には読影結果の判定方法ならびに相互比較の方法の問題があり、可能な限り客観的かつ実用的な方法を採用することが必要である。

以上の観点より、われわれは RI 像処理法の客観的評価法の実験として、IAEA の国際協力プログラム⁸⁻¹⁰⁾で使用したきわめて単純な電算機シミュレーテッド・ファントムを 2 種類の処理法で処理して同一条件でカーブ・プロッタに表示し、4 名の核医学専門医に読影していただき、これを 3 種の比較方法で解析して相互の問題点などを考察したので報告する。

II. 方 法

1. シミュレーテッド・ファントム

ファントムは肝臓をシンチレーションカメラで撮像して得たシンチグラムを模したもので、詳細は文献 8), 9) に述べられているので概略をふれるに止める。

Z_{ij} を像の中心位置から X 方向 i 絵素分、Y 方

* 放射線医学総合研究所

** 群馬大学医学部放射線医学教室

受付：54 年 6 月 13 日

最終稿受付：54 年 10 月 30 日

別刷請求先：千葉市穴川 4-9-1 (番 260)

放射線医学総合研究所

福 久 健二郎

向 j 絵素分だけそれぞれ離れた位置 r_{ij} におけるカウントとすると、

$$Z_{ij} = \begin{cases} \alpha - \beta r_{ij}^2 + \gamma & (0 \leq r_{ij} \leq 50) \\ \gamma & (50 \leq r_{ij}) \end{cases} \quad \dots \dots \dots (1)$$

ただし、 $\alpha = 100$ (カウント)

$\beta = 0.04$ (カウント)

$\gamma = 5$ (カウント)

なる式で表わされる幾何学的図形を作成する。

ここに絵素間隔 2.5 mm, 全絵素数を 128×128 とすると Fig. 1-a, b のような深椀をふせた形状となる。1-a は上からみた形状と椀の範囲, 1-b は中心を通る断面を示す。中央部の最大カウントはバックグラウンド ($\gamma = 5$ カウント) を含めて 105 カウント ($1,680$ カウント/ cm^2) である。このような基礎図形に対し、位置、半径および深さを変化させた hot または cold の円筒状の欠損を導入する。欠損部分における欠損の総カウント N は、

$$N = \pm \epsilon \sqrt{\sum_i \sum_j Z_{ij}} \quad \dots \dots \dots (2)$$

ただし、 ϵ : 欠損部分における絵素のカウント数から求めたカイ自乗値、

$\sum_i \sum_j Z_{ij}$: 欠損部分での欠損導入前の総カウント数

ここに、 ϵ を 0.0σ から 30.0σ (σ は欠損導入前

のカウントの標準偏差) まで、半径を 2.5 mm (1 絵素) から 12.5 mm (5 絵素) まで 2.5 mm きざみで変化させる。

このような欠損を作成した像にシンチグラム特有の像のぼけを導入するため、 $\sigma = 6.25 \text{ mm}$ (2.5 絵素, FWHM = 14.8 mm) の 2 次元ガウス分布のウェートを PSF として重畠積分する。次にこれに雑音を導入するため、得られたカウントを平均値とするポアソン乱数を発生させてとのカウントとし、磁気テープに収録する。IAEA はこのような像を 24 枚 (第 1 回および第 2 回の作業) 作成し、プログラムに参加した世界各国約 20 の施設に送った。各施設でそれぞれの方法で処理、再表示し、読影した欠損位置を報告することにより相互比較が行なわれた。

2. 処理および表示法

処理法は、 $\sigma = 1.5$ 絵素 (3.25 mm) の 2 次元ガウス分布のウェートをもつフィルタによる単純平滑化(以下 GF 法と略す)と、 $\omega = 0.4$ ライン/絵素、 $r = 2.5$ 絵素による最適化フィルタ^{5,9)}によるぼけ補正を加味した平滑化(以下 OF 法と略す)を用いた。また、表示法はカーブ・プロッタによる等高線表示法により^{11,12)}、10 カウントからはじめて 5 カウントきざみのレベルで表示した。Fig. 2-a, b はシミュレーテッド・ファントム 1 枚目のもの

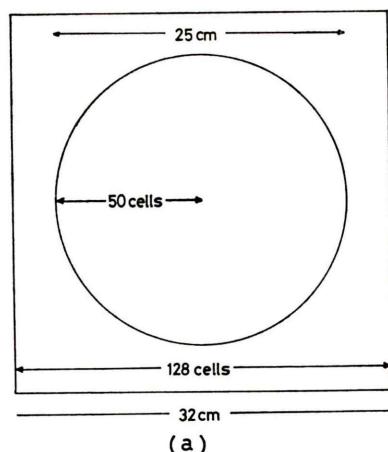
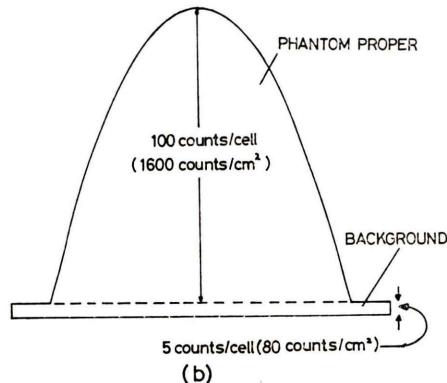


Fig. 1 Idealized count distribution of the computer-simulated phantom by IAEA.
(a) plan view and (b) profile.



を2-aはGF法で、2-bはOF法で処理して表示した。このように処理、再表示した像48枚を適当に順序を入れかえ、4名のシンチグラム専門医に読影願った。

3. 読影条件

読影にあたっては、次の条件を付けた。

- 1) 全画像のうち半数はhot、半数はcoldの欠損がある。hotとcoldが混在する像はない。hotかcoldかは各画像の右上に記入されているとおりである(Fig. 2 参照)。
- 2) 1つの像の4つの象限につき、それぞれ1個または0個の欠損がある。従って、像あたり最大4個、最小0個の欠損がある。
- 3) 欠損を含む象限は、全象限の約3分の2である。
- 4) 読影に際し、次の4段階の確信度を、その読影した欠損範囲とともに記入する。1:確実にあり、2:ありの疑い、3:ないかかもしれない、4:なし。

Fig. 3-aは読影医B、3-bは読影医Cによる同一ファントムを同一処理、表示したものと読影した結果で、3-aは象限(1)のみが欠損ありと読影されたのに対して、3-bでは4象限とも欠損ありと読影されている。

このような回答から中心位置とその確信度をカードにパンチして電算機に入力し、IAEAからの正解と比較して解析した。Fig. 3の場合、象限(1)と(2)は3-bの回答が正解に近く、象限(3)は欠損はあるけれども3-bの回答とは離れた位置にあり、象限(4)は欠損なしで3-aが正解である。

疾病診断には結果として正診と誤診があり得る。疾患有を正しく有とした場合(true positive、館野、飯沼の定義した用語では有病正診と呼ばれ¹³⁾、以下、TPと略す)はもちろん正診であるが、疾患のないものを無しとした無病正診(true negative(TN))もまた正診である。これに対して誤診には疾患が無いのに有とした無病誤診(false positive(FP))と、疾患が有るに無しとした有病誤診(false positive(FN))とがある。本報のごとくシンチグラム静態像の欠損検出に際しても同様のことがい

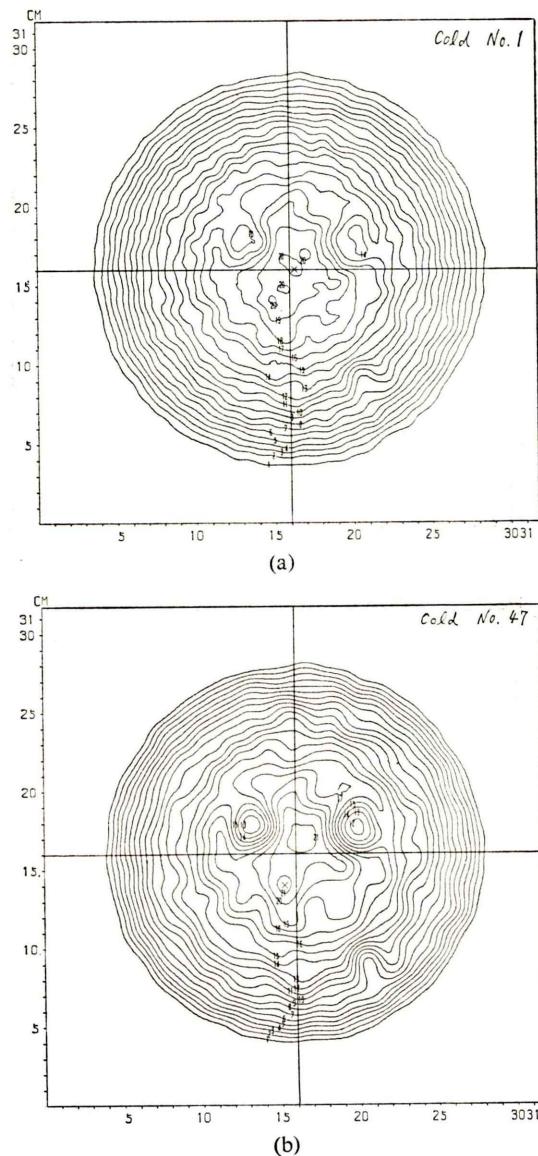
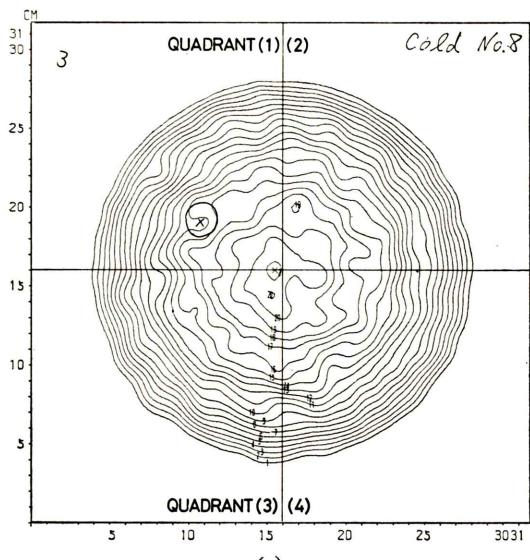


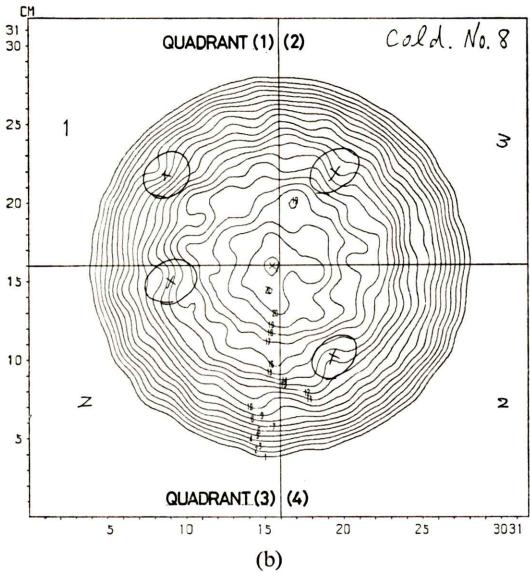
Fig. 2 Contour map displays of phantom No. 1. Processed by (a) Gaussian filter (GF-method), and (b) 'Optimum filter' (OF-method)

える。しかし、ここでTPの条件を定義しておく必要がある。

Fig. 4はIAEAが行なったTPの条件を示す⁸⁾。斜線入りの円を真の欠損、黒塗りの円を回答とするとき、象限(1)のように欠損がないのに有とした場合はFP、象限(2)のように欠損があるのに



(a)



(b)

Fig. 3 Results for same processed image by two different physicians, (a) observer B and (b) C.

無しとした場合は FN, 象限(3)のように欠損の中心と回答の中心とのなす角が 30° 以上のとき, または欠損の中心から 30mm 以内に回答が入っていないときは FP と FN との複合誤診とし, 結局, 象限(4)のごとく条件の範囲内に回答がある場合は TP, 欠損なしを無しとした場合を TN とし, 正解とする. このように 30mm の半径と 30° の角

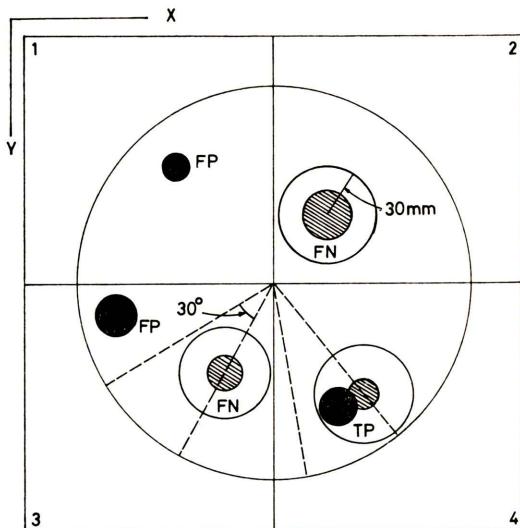


Fig. 4 A model plan of the rejection procedures employed by IAEA. The oblique-lined circles show the true lesions and the filled circles show interpretations.

度による判定条件を本報では(30° & 30mm)法と呼ぶ。これに対して(10° & 10mm)法、Fig. 4の象限(3)の複合誤診をFPだけとした(30° & 30mm, FP)法、さらに条件を緩めて欠損の存在する象限に有の回答をしただけでTPとする場合(象限法と呼ぶ)などを検討したが、主に(30° & 30mm)法による判定結果を解析した。

4. 結果の比較方法

IAEA の第 1 回相互比較では、次の得点法が用いられた。

S: 総得点

t_p : TP の数 t_n : TN の数

f_p : FP の数 f_n : FN の数

ここに、 a, b, c, d はそれぞれ TP, TN, FP, FN に対する評価のウェートであり、 $a=1, b=1, c=-1, d=-2$ とする方法がとられたが、第 2 回の作業では比較のため $d=-1$ による計算も行なわれた。

第2の方法はROC解析(Receiver Operating Characteristics)^{14,15)}で、第2回および第3回の

作業ではこの方法が主に用いられた。

ROC の場合は、読影者の確信度が要求される点が得点法と異なっている。

さらに、McNeil らの提案する評価方法¹⁶⁾についても検討した。この方法は(3)式に示す t_p , t_n , f_p , f_n をつかって、

$$E = (t_p + t_n) / (t_p + t_n + f_p + f_n) \dots \dots \dots (6)$$

の 5 つの式であらわそうとするもので, P_1 は predictive value of positive results, 館野, 飯沼¹³⁾ は有徵正診と呼んでいるもので, 有とした回答に対する真の欠損有の割合を示す. P_2 は predictive value of negative results, または無徵正診とよばれ, 無しとした回答に対する真の欠損無しの割合を示す. E は P_1 と P_2 の全回答における正解の割合で, これは別に,

$$L = (t_p / (t_p + f_p)) / (t_n / (t_n + f_n)) \dots \dots \dots (10)$$

なる尤度比によって示す場合もある.

S_1 は有病正診率, true positive ratio または sensitivity と呼ばれ, 真の欠損がどの程度検出されたかの割合を示し, S_2 は無病正診率, true negative ratio または specificity と呼ばれ, 真の欠損無しがどの程度検出されたかを示す.

本報においては以上3つの相互比較法を用いて検討した。

III. 結 果

1. 得点法

Table 1 に 4 名の専門医による読影結果を (30° & 30m/m) 法で判定し、処理法および hot と cold とにわけて TP, TN, FP, FN の数を確信度別に示した。正解の欠損数は hct が 30 個、 cold が 38 個、従って欠損なしの象限は 28 個である。Table 1 から、個人差はあるけれども各読影医とも CF 法の処理像のほうが TP を多く検出し、また、 hct より cold のほうが高い確信度で検出されているこ

とがわかる。とくに GF 法の hot 像は検出数が少なくて確信度も低く、この種の像が読影しにくいことを示している。個人別には読影医 A, B は、C, D より慎重に読影しているために FP が少なく、TP も少ない。

次に、前述の各種判定法を用い、確信度3(ないかもしれない)までの正解をTPとして得点法で比較した結果をTable 2に示す。Table中、(i)は(3)式のdを-1とし、(ii)はd=-2と評価した場合を示し、(30° & 30mm)法では(i)の得点は最高71(読影医A)、最低43(読影医D)ですべて正の得点であるが、(ii)では最高25、最低-5で、OF法がすべて正の得点、GF法は読影医Cを除いて負の得点となり、明らかにOF法の欠損検出数が多いことがわかる。積極的な読影をした読影医Cの場合、TPが多いかわりに負の要素となるFPが多く、そのためにOF法の得点はあまり良くないがGF法におけるFNの数が少なく、これが(ii)の評価法で最高点を得た原因と考えられる。Table 2には(ii)の評価法による(10° & 10mm)法、(30° & 30mm, FP)法、象限法による得点も示したが、とくに読影医Dでは、これらの判定条件が緩くなるにつれて得点も向上しており、後述のごとく像の再表示法に問題があることを示している。

2. ROC 解析

Fig. 5-a, b, c に (30° & 30 mm) 法による判定結果を ROC 曲線で示した。5-a は処理法には無関係に hot と cold とを読影医別に示し、5-b は hot, cold には無関係に GF 法と OF 法による結果を読影医ごとに示し、また、5-c は読影医 C および D の結果を処理法と hot, cold にわけて示した。読影医 A および B の GF 法による結果には FP が皆無のため、Fig. 5-b に示すことができなかった。これらの結果では、個人差もさることながら、得点法では明白には示せなかった hot と cold との差が示され、また、OF 法の読みすぎによる FP の増加も明示されている。なお、一般に RCC 曲線は横軸に欠損なしに対する FP の割合をパーセントで示すことが多いが、ここでの判定

Table 1 Classification of observed results into 4 categories, hot-OF, cold-OF, hot-GF and cold-GF, respectively

Physicians	Degree of confidence	Optimum Filter (OF-method)						Gaussian Filter (GF-method)					
		hot lesions			cold lesions			hot lesions			cold lesions		
		TP	TN	FP	FN	TP	TN	FP	FN	TP	TN	FP	FN
A	1	11			19			2				13	
	2	5			5			6				5	
	3	2			4			2			1	3	
B	4	18			12			9			18	19	
	Total	18	18		12			28	9	2	10	18	1
											20	21	10
C	1	14			21			4				15	
	2	2			2			2				2	
	3	3			3			1			5	2	
D	4	17			11			10			11	18	
	Total	19	17	1	11			26	10	1	12	11	18
											19	19	10
	1	17			17			8				20	
	2	4			6	1	2	3	1		4	2	
	3	4			7	3	2	11	5		4	3	
	4	9			1			2	1		16	11	
	Total	25	9	13	5	31	2	14	7		16	5	14
											25	9	2
	1	13			15			4				10	
	2	3			1	1		10	4		7	5	
	3	7			10	3	5	10	3		2	1	
	4	11			3			1	3		15	16	
	Total	23	11	11	7	30	1	14	8		13	15	4
											17	22	8

方法が、FP と FN を同時に含む場合があるために FP が 100% を超えることもあり得るので、FP の数によって示した。

3. McNeil の方法

Table 3 に、処理法別および hot, cold 別に確信度 3 (ないかもしない)までの正解を TP として、

Table 2 Total score using several rejection methods.

Physicians	(30° & 30 mm) rejection (ii)**				(30° & 30 mm) rejection (i)*	(10° & 10 mm) rejection (ii)**	(30° & 30 mm, FP) rejection (ii)**	quadrant rejection (ii)**				
	OF-method		GF-method									
	cold	hot	cold	hot								
A	15	12	-3	-11	13	71	11	15				
B	13	13	-9	-9	8	68	-12	8				
C	5	11	6	3	25	54	-3	49				
D	1	9	-10	-5	-5	43	-53	14				
								39				

Note (i)* total score = $1 \cdot t_p + 1 \cdot t_n - 1 \cdot f_p - 1 \cdot f_n$

(ii)** total score = $1 \cdot t_p + 1 \cdot t_n - 1 \cdot f_p - 2 \cdot f_n$

where t_p ; number of true positive

t_n ; number of true negative

f_p ; number of false positive

f_n ; number of false negative

Table 3 Calculated results by values defined by McNeil.¹⁶⁾

Kind of image	Categories of decision	Physicians			
		A	B	C	D
OF-method hot lesions	P_1	100.0%	95.0%	65.8%	67.6%
	P_2	60.0	60.7	90.0	78.6
	E	75.0	75.0	70.8	70.8
	S_2	60.0	63.3	96.2	88.5
	S_2	100.0	94.4	40.9	50.0
OF-method cold lesions	P_1	93.3%	96.3%	68.9%	68.2%
	P_2	50.0	47.6	66.7	25.0
	E	77.1	75.0	68.8	64.6
	S_1	75.7	70.3	96.9	90.9
	S_2	81.8	90.9	12.5	6.7
GF-method hot lesions	P_1	90.9%	100.0%	76.2%	78.5%
	P_2	48.6	48.6	59.3	48.4
	E	58.3	60.4	66.3	58.3
	S_1	34.5	36.7	59.3	44.8
	S_2	94.7	100.0	76.2	78.9
GF-method cold lesions	P_1	100.0%	100.0%	92.6%	88.0%
	P_2	37.0	34.5	42.9	34.8
	E	64.6	60.4	70.8	62.5
	S_1	55.3	50.0	67.6	59.5
	S_2	100.0	100.0	81.8	72.7

P_1 ; predictive value (positive) $t_p/(t_p+f_p)$

P_2 ; predictive value (negative) $t_n/(t_n+f_n)$

E ; efficiency $(t_p+t_n)/(t_p+t_n+f_p+f_n)$

S_1 ; sensitivity $t_p/(t_p+f_n)$

S_2 ; specificity $t_n/(t_n+f_p)$

McNeil ら¹⁶⁾の方法で計算したものを示す。ここでも個人差はあるが、読影医 A と B, C と D の共通点がとくに明白で、前者は P_1 と S_2 が高く、後者は P_2 と S_1 が高い。これらは Table 1 で示した読影の慎重、積極性を示すものであるが、ROC 曲線に比して直接的な評価を行なうパラメータとして比較することはむずかしい。

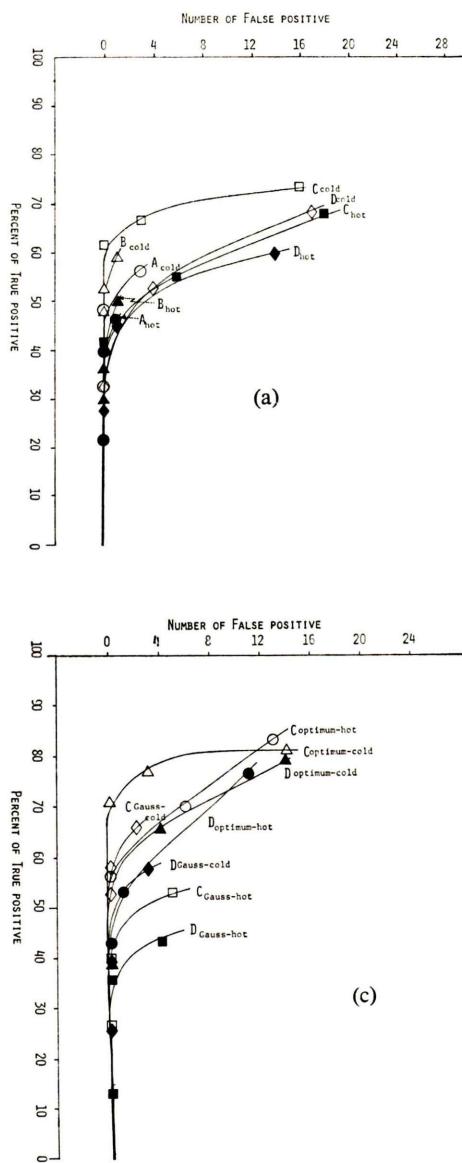
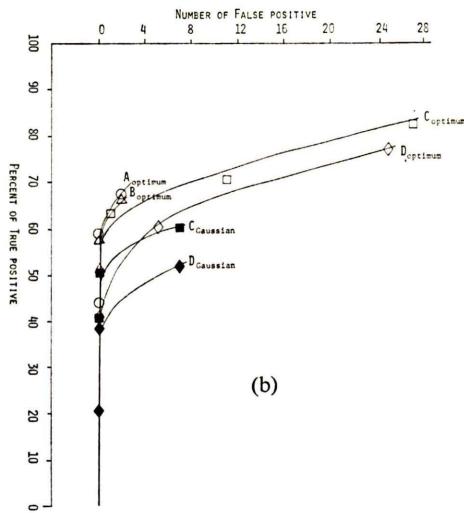


Fig. 5 ROC curves for various set of results

4. 読影に及ぼす欠損深さおよび半径の関係の検討

以上の相互比較とは別に、もとの欠損の深さおよび半径の大きさが読影に及ぼす影響を調べるために、Fig. 6 に示すように、縦に 2.5 mm きざみの半径、横に像の欠損を導入する前のカウントに対する欠損部分の全カウント分散値 σ の χ^2 値 ((2) 式における ϵ) を 6.0σ きざみにとり、確信度付きで正解数を示した。白丸は FN、黒丸は確信度 1 (確実にあり) の TP, 2/3 黒丸は確信度 2 の TP, 1/3 黒丸は確信度 3 の TP を表わし、読影医 C の成績を示した。その結果、欠損深さが大きくなるほど TP が増加し、確信度も高くなる。しかし、半径では cold では確信度に関係がありそうであるが hot では関連性があまりみられない。また、



- (a) Open marks show cold lesions and filled marks show hot lesions for four physicians.
- (b) Open marks show results by OF-method and filled show by GF-method for four physicians.
- (c) Hot-lesions processed by OF-method (circle), cold-OF (triangle), hot-GF (square) and cold-GF (rhombus) for physician C (open marks) and D (filled marks).

OPTIMUM FILTER	IMAGE Radius (mm)	CHI-SQUARE		0.0 ~ 6.0	6.0 ~ 12.0	12.0 ~ 18.0	18.0 ~ 24.0	24.0 ~ 30.0
		- 2.5	2.5 ~ 5.0					
HOT	- 2.5	○			○ ●	● ● ●		
	2.5 ~ 5.0	○			○ ●	● ● ●		
	5.0 ~ 7.5	○ ●	○ ○	○ ● ●	○ ● ●	● ● ●		
	7.5 ~ 10.0	○ ●		○ ●	○ ●	● ●		
	10.0 ~	○ ○ ○		○ ○ ○	○ ○ ○	● ●	● ●	● ●
COLD	- 2.5	○				● ● ●		
	2.5 ~ 5.0	○ ○		○ ○	○ ○	● ●	● ●	● ●
	5.0 ~ 7.5	○ ○		○ ○	○ ○	● ●	● ●	● ●
	7.5 ~ 10.0	○ ○		○ ○	○ ○	● ●	● ●	● ●
	10.0 ~		● ● ●	● ● ●	● ● ●	● ● ●	● ● ●	● ● ●
GAUSSIAN FILTER	- 2.5					●		
	2.5 ~ 5.0	○		○ ○	○ ○	○ ○ ●		
	5.0 ~ 7.5	○ ○	○ ○ ○	○ ○ ○	○ ○ ○	● ● ●		
	7.5 ~ 10.0	○ ○		○ ○	○ ○	● ●		
	10.0 ~	○ ○ ○		○ ○ ○	○ ○ ○	● ●	● ●	● ●
COLD	- 2.5	○				○ ○ ○		
	2.5 ~ 5.0	○ ○		○ ○	○ ○	● ●	● ●	● ●
	5.0 ~ 7.5	○ ○		○ ○	○ ○	● ●	● ●	● ●
	7.5 ~ 10.0	○ ○		○ ○	○ ○	● ●	● ●	● ●
	10.0 ~		○ ○ ○	○ ○ ○	● ●	● ●	● ●	● ●

Fig. 6 Distribution of chi-square significance (6.0 σ interval) and radius (2.5 mm interval) for the true lesions and results by physician C. Open circles show false negative, 2/3 open circles show the true positive with confident degree of 3 (possible negative), 1/3 open circles show TP with 2 (probably positive) and filled circles show TP with degree 1 (definitely positive).

OF 法では 12σ 以上の欠損はほとんど検出されているが、GF 法では 18σ を超えないとき読影できないことがわかる。なお、同様のことを像の中心から欠損中心までの距離について調べたが、半径以上に無関係であることがわかった。

IV. 考 案

前節までに述べた IAEA 作成の電算機シミュレーテッド・ファントムを用いた読影実験の結果では、3種類の比較法のどれを用いても次のような一般的な結果を得た。

(1) OF 法のほうが GF 法より正診率を向上する。(2) 本報のようなファントムでは cold の欠損のほうが hot より正診率が高い。(3) 読影の難易性は欠損の深さ (S/N 比) に最も関係が深い。

(2) の結果は、本報のファントムが高カウントの部分に欠損を導入した、肝臓、脾臓のシンチグラムに近い形状であるためと考えられ、もし、 ^{99m}Tc パーテクネート投与の脳シンチグラムの

ように低カウントの部分に欠損を導入した形状にシミュレートしたならば、別の結果が得られたかも知れない。また、このファントムでは、1象限1個の欠損という、実際的ではない形状を用いているが、TP を認定する基準を検討するためには必要なことであり、IAEA の第3回の相互比較に用いたファントムではこうした条件をやめて、より臨床例に近いものとしている。

本報に述べた、カーブプロッタによる等高線表示は、臨床医にとってなじみの薄いものであり、あらかじめいくつかの示唆が必要であったし、また、欠損位置の決定が難かしかったであろうことは、Table 2 で判定条件を緩めるほど得点が高くなつたことでも明らかであるが、われわれの電算機システム (TOSBAC 3400 モデル 31) の仕様上やむを得ない表示法である。しかし、医師の経験的先入観のない表示法と解すれば、それなりの意味があるかも知れない。いずれにせよ、同一の処理像 (あるいは無処理の像) を相異なる数種類の

表示法で表示し、相互比較によって最適化を追求することも重要であると考える。

相互比較の方法として、得点法、ROC解析およびMcNeilらの計算法を用いた。得点法ではFPよりFNの負のウェートを大きくすべきである。すなわち、無回答のときはTNとFNとの数で得点が決まるが、もし欠損なしの像が圧倒的に多い場合に、FPとFNに同じウェートを与えると、無回答でも正の得点を得ることになり、読影を目的として撮像したRI像を全く読影しないほうが成績が良いという矛盾を生じ得る。しかし、(3)式における $c=-1, d=-2$ の値が適當か否かは、そのシンチグラムの確定診断への依存度あるいは誤診の重大性に応じて検討される必要があろう。本報の結果ではTPが多くてFNが少ない場合、すなわちかなり積極的な読影が好成績を示すことが実証された。このことはシンチグラムの画質が悪く、読影が割合慎重に行なわれている現状ではあまり高い得点が期待できないことを意味する。

ROC曲線は読影者の心理的な読影差を低減するために確信度というウェートを付して読影し、結果をTP率とFP率とで2次元的に曲線に近似して示す点では直観的比較観察が可能であり、きわめて有用な方法である。この場合の確信度は、本報では4段階によったが、われわれの何回かの試み^{9,10,17)}でも5段階が限度で、それ以上は無意味のようである。そこにROC解析の限界もある。Fig.5-a～cは計算結果をもとに適当に曲線に近似しているが、たとえばFig.5-bのOF法による処理像の読影医AとCとの曲線のどちらが良い成績といえるかは判別し得ない。なぜなら、Aの曲線が途中で切れ、Cと直接比較できる要素はAの曲線との相対的位置までであり、このままAの曲線がCの上に伸びるという保証は全くないからである。とくに、このFig.5-bでは、GF法処理像の、AとBとの読影結果にFPがなかったという事実は、慎重な読影を行なうとROC解析が不可能になることを示している。

これに対してMcNeilらの方法は、積極的読影によってsensitivityが高くなり、慎重な読影に

よってspecificityが高くなる。ともに読影態度を示す重要なファクターであり、本法が他の2法よりも効果的比較方法であると思われるが、この場合においてもsensitivityとspecificityの両者ともに大きい場合を除いては、読影結果としての比較は単純ではなく、得点法におけると同様に誤診の重大性を考慮に入れた評価法を検討すべきあることを示している。

ここで得られた結果が、ただちに実際のシンチグラムに適用できるものではないが、ぼけ修正を加味したフィルタ(OF法)が欠損認識における正診率向上の可能性を示したので、確定診断のついたデジタル・シンチグラムについてこの種の研究を推進することは価値あることと思われる。とくに、電算機で処理した像をX線フィルムに撮影できるマイクロ・ドット・イメージヤまたはその同等の装置を用いて表示した像を対象として研究を行なうことにより、正診率を向上する技術の検討が推進し得るものと考える。われわれは、手術によって結果の判明している肝臓のデジタル・シンチグラムのデータ・ベースを開発中であり¹⁷⁾、これについてなんらかの報告を試みたい。

V. 結論

RI像処理法の効果を客観的に評価するための実験として、IAEAが作成した電算機シミュレーテッド・ファンтомを用い、2種類の処理法で処理し、4名の専門医が読影した結果を3種類の比較方法で検討した。

その結果、いずれの比較方法でも全員単純フィルタよりもぼけ修正を導入したフィルタによる処理のほうが成績が良く、後者の効果があることが確認された。比較方法ではROC曲線が直観的に正診率を比較できるので有利であるが、慎重な読影ではFPが出現しないために比較が不可能の場合がある。McNeilの方法は読影医の態度(慎重か積極的か)をはかるパラメータとして有用であるが、総得点法とともに相互比較は必ずしも単純ではない。従って本報の結果に限れば、ROC曲線で比較したうえで補助的にsensitivityとspeci-

ficity の値を計算することで相互比較と読影医の読影態度が定量的に示し得ると考える。しかし、処理法および表示法の最適化は、必ずしも一元化し得るものではなく、臓器ごと、あるいは診断の目的ごとに検討すべきかも知れない。その場合においても本報の考え方や比較方法は十分役立つものと考える。

本報の一部は第17回日本核医学会総会で発表した。本研究においてシンチグラム読影にご協力いただき、また、各種のご助言を賜わった千葉大医学部 内山 晓、慶應大医学部 久保敦司、東京大医学部 町田喜久雄および東京女子医大 山崎統四郎の各先生（以上五十音順）に深謝する。

文 献

- 1) Brown DW: Digital computer analysis and display of the radioisotope scan. *J Nucl Med* **5**: 802-806, 1964
- 2) Tauxe WN, Chaapel DW: Contrast-enhancement of scanning procedures by highspeed computer. *J Nucl Med* **6**: 326, 1965
- 3) Brown DW: Digital computer analysis and display of the radionuclide scan. *J Nucl Med* **7**: 740-753, 1966
- 4) Nagai T, Iinuma TA: Computer-focusing for area scans. *J Nucl Med* **9**: 507-516, 1968
- 5) Tanaka E, Iinuma TA: Approaches to optimal processing in radioisotope imaging. *Phys Med Biol* **15**: 683-694, 1970
- 6) Kirch DL, Brown DW: Recent advances in digital processing static and dynamic scintigraphic data. Proc. 2nd Symp. Sharing of Computer Programs and Technology in Nuclear Medicine, Oak Ridge, Tennessee, USAEC Rep., CONF-720430: 27, 1972
- 7) 福久健二郎, 館野之男, 飯沼 武, 他: 電子計算機による肝シンチグラムおよび肝手術所見の登録システムの開発, 核医学 **14** (第17回日本核医学会総会抄録集) 219, 1977
- 8) IAEA Co-ordinated research programme on the intercomparison of computer-assisted scintigraphic techniques. Progress Report. Medical Radioisotope Scintigraphy 1: IAEA, Vienna, 727, 1973
- 9) Iinuma TA, Fukuhisa K, Tanaka E: Improvement of scintigraphic image quality by means of the "Optimum filter" and other filtering techniques. International Congress Series No. 339, Radiology, Proc. of 13th International Congress of Radiology. Excerpta Medica Amsterdam **2**: 617, 1974
- 10) 福久健二郎, 飯沼 武, 松本 徹, 他: IAEA シンチグラフィー技術の国際相互比較プログラムについて, 日本ME学会医用画像処理研究会資料, MI 15-3, 1977
- 11) Loken MK, Ponto RA, Kush GS: Quantification of image studies obtained with a scintillation Anger Camera. *Radiology*, **95**: 505-516, 1970
- 12) 有水 昇, 飯沼 武: 核医学大系第3巻, 核医学イメージング診断法, pp. 110, 実業公報社, 東京, 1977
- 13) 館野之男, 飯沼 武: 誤診などに関する用語の提案, 医学のあゆみ **97**: 424-426, 1976
- 14) Green DM, Swets JA: Signal detection theory and Psycho-physics. Wiley, New York, 1966
- 15) Metz CE: Empirical evaluation of imaging procedure in terms of information content and receiver operating characteristic curve. *J Nucl Med* **13**: 453, 1972
- 16) McNeil BJ, Adelstein SJ: Determining the value of diagnostic and screening tests. *J Nucl Med* **17**: 439-448, 1976
- 17) 飯沼 武, 福久健二郎: 電子計算機による RI イメージ形成のシミュレーション. 日本医放学会雑誌 **31**: 1270-1285, 1972

Summary

An Analytical Study of Scintigram Interpretation

Kenjiro FUKUHISA*, Takeshi A. IINUMA* and Teruo NAGAI**

**National Institute of Radiological Sciences, Chiba*

***Department of Radiology, Gunma University, Maebashi*

Various image processing methods are available for transforming scintigraphic image quality. However, it is not so easy to determine the best processing method in view of diagnosis of disease by physicians. In order to know the effect of image quality on the physician's interpretation, 24 images of computer-simulated phantom which was established in the first coordinated programme on the "Intercomparison of Computer-assisted Scintigraphic Techniques" sponsored by IAEA, are processed by using the 'Optimum Filter' with medium frequency enhancing effect (OF-method) and the Gaussian filter with simple smoothing effect (GF-method), and are displayed as a contour map with same display factors by a curve plotter of computer system. The processed and displayed 48 images in random order are viewed by four physicians who are specialists in scintigraphic diagnosis, with the suggestions as follows:

- (1) The images are consisted of 48 pictures.,
- (2) Each images is divided into four quadrants, each of which could contain either one lesion or none. Therefore, each images could contain from 0 to 4 lesions,
- (3) All 1esions in the phantom are either cold or hot, but there are no mixed lesions of both in a picture,

- (4) The probability of presence of a lesion in a given quadrant is taken to be about 2/3, and
- (5) The lesions observed should be entered with following figure of rating from 1 to 4 indicating the degree of confidence; 1:definitely positive, 2:probably positive, 3:possibly negative and 4:definitely negative.

The results of their answers are analyzed with (1) total score method, (2) ROC curves and (3) values defined by McNeil.

Among the results, the OF-method gives significantly better results compared with the GF-method, though the former are increased number of false positive compared with the latter. The cold lesions are more detectable compared with hot lesions of the phantom. The cold lesions are similar to a scintigram of human liver administrated 99m-Tc-colloid.

Further investigation is being directed what kind of display can give the best results by using various displaying devices for same processed images, as the display method in this study seems to be not suitable for detecting the lesions in scintigraphic images.

Key words: Sintigraphic image processing, Computer-simulated phantom, Image display, ROC curve