



Title	映像診断における診断精度及び統計的総合映像診断に関する研究-特に肝細胞癌に関して-
Author(s)	川平, 幸三郎
Citation	日本医学放射線学会雑誌. 1985, 45(5), p. 741-751
Version Type	VoR
URL	https://hdl.handle.net/11094/15508
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

映像診断における診断精度及び統計的総合映像診断に関する研究

—特に肝細胞癌に関して—

九州大学医学部放射線科学教室（主任：松浦啓一教授）

川平幸三郎

（昭和59年10月22日受付）

（昭和59年12月13日最終原稿受付）

Diagnostic Accuracy of Medical Imagings and An Integrated Statistical Approach to Diagnosis —With Special Attention to Hepatocellular Carcinoma—

Kosaburo Kawahira

Department of Radiology, Faculty of Medicine, Kyushu University
(Director: Prof. Keiichi Matsuura)

Research Code No. : 500, 207

Key Words : *Angiography, Computed tomography, Ultrasound,
Tc-99m phytate scintigraphy, Diagnostic accuracy,
Hepatocellular carcinoma, Integrated medical
imagings*

Twenty-two patients with hepatocellular carcinoma and 36 without were clinically studied using angiography, computed tomography (CT), ultrasound (US) and Tc-99m phytate scintigraphy (RI). Diagnoses were confirmed by biopsy, surgery or autopsy, in all 58 patients.

The judgement of multi-readers was obtained, independently. For a diagnosis of hepatocellular carcinoma, angiography was the most useful. Sensitivity, specificity and accuracy of this modality were 0.67, 0.85 and 0.78, respectively. Specificity and accuracy of US were 0.84 and 0.67, respectively, thereby US was the second best modality. CT was more useful than than RI for a diagnosis of hepatocellular carcinoma, in this series.

Diagnostic accuracy of the 4 modalities in the diagnosis of liver mass were also discussed.

To integrate these 4 images for a diagnosis of hepatocellular carcinoma, a multivariate analysis was made. The diagnostic accuracy of the system was considerably higher than that of any single modality.

1. 緒言

最近の映像診断法の著しい進歩により、各種疾患において、何種類もの映像診断法を駆使することが可能となった。一つの検査法で確実に診断が出来ない場合、つぎからつぎへと検査を行うことも多い。はたして各種検査で得られた情報を有効に利用しているかという点になるとはなはだ心もとないのが現状である。

そこで本研究では肝疾患の映像診断法である肝動脈造影、コンピュータ断層法、超音波断層法および肝シンチグラフィ（以下それぞれAG、CT、USおよびRIと略す。）について各検査法の有用性を検討した。各検査法は個々に、または他と比較して有用性が報告^{1)~13)}されているが、4種類に関して同じ対象について検討した報告はない。さらに、個々の検査法から確診を得られなくても、

各検査法全てを行った場合に総合的に診断精度を向上させる多変量解析法の導入についても検討したので報告する。

II. 研究対象および方法

(A) 対象

昭和54年1月から約24カ月間に九州大学医学部附属病院放射線科を受診し、ほぼ同時期にAG, CT, USおよびRIの各検査が全てなされ、同時に生検、手術または剖検で確診がなされた58例を対象とした。このうち肝細胞癌症例は22例であった (Table 1)。

(B) 映像診断法並びに検討映像

超音波検査は Ausonic 社製 U.I. Octoson (3.0 MHz, マトリックス256×256), PHO/SONIC 社製 Searle SM (3.5MHz, マトリックス256×256) および東芝製 SAL-20A (3.5MHz), 検者は経験5年以上の医師である。映像は35mmフィルムまたはマルチフォーマット・カメラ (100mmフィルム) に撮影されたものとした。

肝シンチグラムは^{99m}Tc-フチン酸, 約2mCi 静注後シンチカメラ (カウント数400K counts) で得、患者の前面、後面および左右の側面像を1枚の大陸フィルムに記録されたものとした。

コンピュータ断層法は東芝 TCT-60A (9秒スキャン, マトリックス320×320) または AS & E 製 R-450 (5.0秒スキャン, マトリックス256×256) を用いた。経口造影剤 (ガストログラフィン8ml + 水300ml) の投与後、経静脈性造影剤投与を行わず10mm幅, 10mm間隔連続スキャンで検査した。ウィンドウレベル50HU, ウィンドウ幅400HUでしめされる観察条件の像とした。映像はマルチ

フォーマット・カメラで6ツ切フィルムあたり4映像記録されたものとした。

血管造影は Seldinger 法によった。腹腔動脈造影では76% meglumine and sodium diazotritate 50-70ml を毎秒8-10ml の速度で注入し、肝動脈造影では40-50ml を毎秒5-7ml 注入した。撮影は前1枚, 造影後毎秒2枚2秒間, 毎秒1枚3秒間, 2秒に1枚を14秒間の計19秒間15枚, または毎秒1枚6秒間, 2秒に1枚16秒間の計22秒間15枚おこなった。肝動脈造影の動脈相, 実質相および静脈相の3枚を選び、対象とした。検査は放射線科医が行い、映像は大角フィルムに記録したものとした。

以上のように映像の大きさはすべて日常臨床に用いられたものであるが、実際に読影にもちいたフィルムは管理の都合上これらのコピー・フィルム (サクラ QD または KODAK X-OM AT duplicating film) によった。

(C) 検査法の有用性の検討

過去の症例を多数の医師に再読影させ、各検査法毎に診断精度を求めた。読影は3つの施設 (九大, 熊大, 長大) の放射線科医によって行われた (Table 2)。読影に携わった医師の数は検査手技によって幾分かちがうが19人から40人であった。

読影に参加した人数のうちRIの読影者が最も少なかったが、これはRIでは肝腫瘍の検出はできて質的診断は不可能との考えが一般的であり、肝細胞癌に関しては診断を辞退した人が多かったためである。

読影に際しては全く臨床情報を与えることなく、各症例につき肝腫瘍の有無について+, +, ±, -, -の決定を求め、更に診断名を付けさせた。

Table 1 Materials

	No. of Cases
Hepatocellular Carcinoma	22
Cholangiocarcinoma	6
Liver Metastasis	2
Hemangioma	2
Liver Cyst	2
Liver Cirrhosis	9
Normal	15
Total	58

Table 2 Radiologists Participated

	Liver Mass	Hepatocellular Carcinoma
AG	26	26
CT	40	40
US	34	31
RI	31	19

Some doctors participated only in detecting liver mass.

なお症例は検査法ごとにランダムに配列した上で各人別個に読影させた。念のためフィルムに記された患者名はあらかじめ隠しておいた。

肝腫瘤は肝細胞癌、胆管細胞癌、転移性肝癌、嚢胞、海綿状血管腫の合計34症例でみられた。これらの延べ読影回数(34×読影者数)のうち正しく肝腫瘤と診断された割合をもって肝腫瘤診断に関する有病正診率(sensitivity)とし、肝腫瘤以外の症例の延べ読影回数(24×読影者数)のうち正しく肝腫瘤以外と診断された割合をもって肝腫瘤診断に関する無病正診率(specificity)とした。また全症例の延べ読影回数(58×読影者数)のうち肝腫瘤または肝腫瘤以外と正しく診断された症例の割合をもって正確度(accuracy)とし、これらをAG, CT, US, RIについて求めた。

さらに肝細胞癌症例の延べ読影回数(22×読影者数)のうち正しく肝細胞癌と診断された割合、肝細胞癌以外の症例の延べ読影回数(36×読影者数)のうち正しく肝細胞癌以外と診断された割合をもって各々肝細胞癌診断に関する有病正診率、無病正診率とした。

また全症例の延べ読影回数(58×読影者数)のうち正しく肝細胞癌または肝細胞癌以外と読影された割合をもって肝細胞癌診断に関する正確度とし、これらをAG, CT, US および RI について求めた。

このほか全症例の延べ読影回数のうち確定診断通りに読影された割合を overall accuracy としこれを求めた。

読影には放射線科医がたずさわった。経験が読影結果に及ぼす効果について検討するために読影者の経験を日常臨床の場で作成したレポート数で

3群に分けた。大体スタッフに相当する301件以上作成者からなる第1群、診断のトレーニング中に相当する51件以上300件作成者からなる第2群、初学者に相当する50件以下作成者からなる第3群とした。各群毎に腫瘤診断に関する有病正診率、無病正診率、正確度および overall accuracy を求め検討した。

また同様に各群毎に肝細胞癌に関する有病正診率、無病正診率、正確度および overall accuracy を求め検討した。

(D) 総合的診断のための多変量解析法

症例につき肝細胞癌と診断した読影者の全体に占める割合を求め肝細胞癌所見陽性率(以下、所見陽性率と略す。)とした。

対象症例を同種類の4変量(AG, CT, US および RI による所見陽性率)母集団からの標本と考え、観測された症例が肝細胞癌または肝細胞癌以外に分類されるかを決定する方法として判別分析^{20)~22)}を用い、純粋に映像診断のみでどれくらいの診断精度を有するかを検討した。

すなわち肝細胞癌との判断は、標本数が多い場合正規分布するものとして判別分析をおこなった。判別分析は付録の方法に基づき BASIC でプログラムを組みコンピュータでおこなった。なお $\vec{\mu}_1$, $\vec{\mu}_2$, Σ が未知であるから推定値で代用した。

III. 研究結果

(A) 診断精度

(1) 肝腫瘤に関する診断 (Table 3)

肝腫瘤に関する有病正診率はAGで0.83とずば抜けて高く、RIが0.61と有意にCTより高かった ($p < 0.05$).

無病正診率はAGが0.83とやはり最も高く、つ

Table 3 Diagnostic Accuracy of Medical Imagings for Liver Mass

	No.	AG	No.	CT	No.	US	No.	RI
Sensitivity	884	0.83* ¹	1360	0.54	1156	0.57	1054	0.61* ³
Specificity	624	0.83* ¹	960	0.76* ^{2,3}	816	0.65	744	0.67
Accuracy	1508	0.83* ¹	2320	0.63* ²	1972	0.60	1798	0.64* ⁴

*¹: $p < .05$, AG v CT., $p < .05$, AG v US., $p < .05$, AG v RI.

*²: $p < .05$, CT v US.

*³: $p < .05$, CT v RI.

*⁴: $p < .05$, US v RI.

Table 4 Diagnostic Accuracy of Medical Imagings for Hepatocellular Carcinoma

	No.	AG	No.	CT	No.	US	No.	RI
Sensitivity	572	0.67* ¹	880	0.36	682	0.39	418	0.38
Specificity	936	0.85* ²	1440	0.80* ³	1116	0.84* ^{4,*5}	684	0.74
Accuracy	1508	0.78* ¹	2320	0.63	1798	0.67* ^{4,*5}	1102	0.59
Overall Accuracy	1508	0.53* ¹	2320	0.40* ^{3,*4}	1798	0.33	1102	0.33

*¹: p<.05, AG v CT., p<.05, AG v US., p<.05, AG v RI.*²: p<.05, AG v CT., p<.05, AG v RI.*³: p<.05, CT v RI.*⁴: p<.05, CT v US.*⁵: p<.05, US v RI.

Table 5 Diagnostic Accuracy of Medical Imagings for Liver Mass according to Diagnostic Experience

		Group 1		Group 2		Group 3	
		No.	>301	No.	300>>51	No.	50>
AG	Sensitivity	204	0.84	238	0.79	442	0.84
	Specificity	144	0.90* ³	168	0.93* ²	312	0.74
	Accuracy	348	0.86* ³	406	0.85* ²	754	0.80
CT	Sensitivity	442	0.60* ¹	442	0.50	476	0.54
	Specificity	312	0.77* ³	312	0.84* ^{1,*2}	336	0.66
	Accuracy	754	0.67	754	0.64	812	0.63
US	Sensitivity	68	0.75* ^{1,*3}	306	0.53	782	0.56
	Specificity	48	0.89* ³	216	0.74* ²	552	0.59
	Accuracy	116	0.81* ^{1,*3}	522	0.62	1334	0.58
RI	Sensitivity	374	0.63* ¹	340	0.55	340	0.66* ²
	Specificity	264	0.71* ³	240	0.73* ²	240	0.57
	Accuracy	638	0.66	580	0.62	580	0.62

*¹: p<.05, Group 1 v Group 2.*²: p<.05, Group 2 v Group 3.*³: p<.05, Group 3 v Group 1.

Table 6 Diagnostic Accuracy of Medical Imagings for Hepatocellular Carcinoma according to Diagnostic Experience

		Group 1		Group 2		Group 3	
		No.	>301	No.	300>>51	No.	50>
AG	Sensitivity	132	0.79* ³	154	0.69	286	0.61
	Specificity	216	0.91* ¹	252	0.87	468	0.81
	Accuracy	348	0.86* ^{1,*3}	406	0.80* ²	754	0.73
	Overall Accuracy	348	0.61* ³	406	0.57* ²	754	0.49
CT	Sensitivity	286	0.45* ^{1,*3}	286	0.33	308	0.31
	Specificity	468	0.76	468	0.73	504	0.90* ^{2,*3}
	Accuracy	754	0.64* ¹	754	0.58	812	0.68* ²
	Overall Accuracy	754	0.47* ³	754	0.42* ²	812	0.33
US	Sensitivity	44	0.70* ^{1,*3}	176	0.36	462	0.37
	Specificity	72	0.87	288	0.78	756	0.86* ²
	Accuracy	116	0.81* ^{1,*3}	464	0.62	1218	0.67
	Overall Accuracy	116	0.53* ^{1,*3}	464	0.32	1218	0.31
RI	Sensitivity	132	0.28	132	0.33	154	0.40* ³
	Specificity	216	0.81* ³	216	0.76* ²	252	0.66
	Accuracy	348	0.61	348	0.62	406	0.56
	Overall Accuracy	348	0.28	348	0.40* ^{1,*2}	406	0.31

*¹: p<.05, Group 1 v Group 2.*²: p<.05, Group 2 v Group 3.*³: p<.05, Group 3 v Group 1.

いでCTが0.76と有意に高かった。USとRIには有意差はみられなかった。正確度はAGが0.83と高く、USが0.60と有意に低かった。CTとRIには有意差はなかった ($p < 0.05$)。

(2) 肝細胞癌に関する診断 (Table 4)

有病正診率はAGで有意に高く、CT, US, RI間に有意差は見られなかった。無病正診率ではAGはUSと有意差はなかったが、CT, RIよりも有意に高かった。RIではCTより有意に低かった。正確度ではAGが最も高く、つぎにUSがRI, CTより高かった。overall accuracyはAGが最も高く、ついでCTが高かった ($p < 0.05$)。

(3) 読影経験の違いと有病正診率との関係 (Table 5, 6)

肝腫瘍診断ではAGで有病正診率に経験による差がないものの、無病正診率、正確度は第3群で有意に低かった ($p < 0.05$)。

CTでは有病正診率が第1群は第2群に比較し有意に高かったが、第3群とは差がなかった。一方無病正診率は第3群で低かったが、正確度としては差はみられなかった ($p < 0.05$)。

USでは有病正診率が第1群で高く、無病正診率は第3群で有意に低かった。正確度は第1群で有意に高かった ($p < 0.05$)。

RIでは有病正診率が第2群で低かったが、第3群とは差がみられなかった。無病正診率は第3群で低かったが、正確度としては差がみられなかった ($p < 0.05$)。

肝細胞癌診断に関してはAGとUSで経験と共に有病、無病正診率、正確度は高い傾向を示した。

CTでは有病正診率は経験と共に高値をしめたものの、無病正診率が逆に経験者で低かった。正確度は第1群で第2群より高かった。しかし第3群とは差がなかった ($p < 0.05$)。

RIでは経験と共に有病正診率が低くなり、無病正診率は第1, 2群で高かった。正確度については差がなかった ($p < 0.05$)。

RIを除き経験は豊富なほど肝腫瘍、肝細胞癌の診断精度は高くなる傾向にあるといえた。またoverall accuracyはRIを除き経験が豊富な程高値を示した ($p < 0.05$)。

(B) 多変量解析による総合診断

(1) 所見陽性率について

肝細胞癌と肝細胞癌以外に分け、AG, CT, USまたはRIにより肝細胞癌としての所見がみられた割合、すなわち所見陽性率をみると (Fig. 1), 肝細胞癌群で所見陽性率は肝細胞癌以外群に比べAG, CT, USに関して有意に高かった。RIに関しては有意差はみられなかった。

(2) 所見陽性率による肝細胞癌の判別

1) 説明変数の選択

肝細胞癌群と肝細胞癌以外群の所見陽性率の平均値に最も差が見られたのはAGである。その他の検査法での肝細胞癌とそれ以外の判別は一見できそうにない。しかし、AGの所見陽性率に対するCT, US, およびRIの所見陽性率の相関係数は各々0.43, 0.60および0.37であり、低いながら相関があった。判別に最も有効な説明変数に対して相関を示す変数を説明変数として加えることは判別式の効率を上昇すると言われており²¹⁾、4種の検査法の所見陽性率を説明変数として判別式を求めることには充分意味がある。そこで4種の検査法の所見陽性率を説明変数として判別式を求めた。

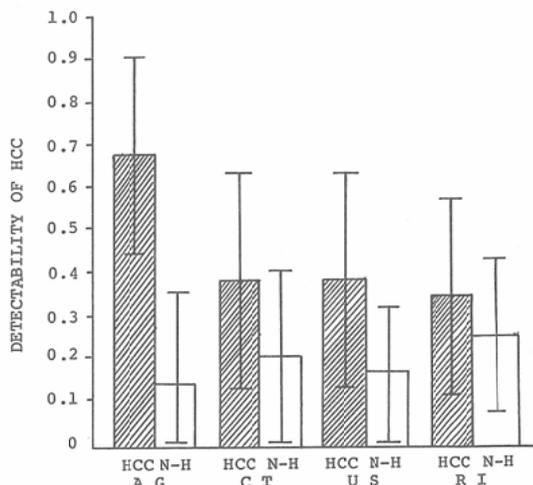


Fig. 1 Detectability of hepatocellular carcinoma by multi-reader Average and standard error are shown.

HCC: Hepatocellular carcinoma

N·H: Other than HCC

2) 肝細胞癌とそれ以外の判別

2群の各平均所見陽性率はAG, CT, USまたはRIについてつぎの通りであった。

$$\begin{bmatrix} \rightarrow \\ \mu_1 & \mu_2 \end{bmatrix} = \begin{bmatrix} 0.6736 & 0.1525 \\ 0.3586 & 0.2020 \\ 0.3895 & 0.1569 \\ 0.3386 & 0.2619 \end{bmatrix}$$

AG, CT, USおよびRIに関する級内分散・共分散行列 Σ は

$$\begin{bmatrix} 0.04625 & 0.01438 & 0.01236 & 0.01587 \\ 0.01438 & 0.05053 & 0.01644 & 0.00574 \\ 0.01236 & 0.01644 & 0.03487 & 0.00783 \\ 0.01587 & 0.00574 & 0.00783 & 0.04121 \end{bmatrix} \quad (56D.F.)$$

逆行列 Σ^{-1} は

$$\begin{bmatrix} 27.7238 & -5.1148 & -5.4089 & -8.9363 \\ -5.1148 & 24.3778 & -9.7773 & 0.4319 \\ -5.4089 & -9.7773 & 35.9658 & -3.3888 \\ -8.9363 & 0.4319 & -3.3888 & 28.2911 \end{bmatrix}$$

と推定された。

判別式は $G_i(\vec{x}) =$

$$a_i \times A + b_i \times C + c_i \times U + d_i \times R + \text{CONST.}_i \quad (i=1, 2), \text{但し } A, C, U, R \text{ は各々の AG, CT, US および RI の所見陽性率.}$$

肝細胞癌群 ($i=1$) について, $a_1, b_1, c_1, d_1, \text{CONST.}_1$ は各11.7080, 1.6345, 5.7117, 2.3948, 5.7541と計算された。同様に肝細胞癌以外 ($i=2$) については係数は各々0.0056, 2.7234, 1.9556, 5.6022, -1.1622と計算された。

たとえば, AG, CT, US, RI に関して所見陽性率が各0.88, 0.6, 0.1および0.21である症例について検討すると, 肝細胞癌群について式値 $G_1(\vec{x})$ は6.6037, 肝細胞癌以外について式値 $G_2(\vec{x})$ は1.8488となるので $G(\vec{x})$ は正となり, この症例は肝細胞癌に分類されることになる。

内部標本58例についてこの判別を実行すると肝

細胞癌22例のうち18例, 肝細胞癌以外36例のうち32例は正しく診断された。したがって4種の検査法を併用した場合の有病正診率, 無病正診率および正確度は各0.82, 0.89, および0.86ということになり, 肝細胞癌診断に関して, 最高の診断精度を示したAG(有病正診率=0.67, 無病正診率=0.85, 正確度=0.78)より診断精度は向上した。

また301件以上レポートを作成した, いわゆる経験豊富な医師のみによる読影結果から症例の所見陽性率を算出し, 判別式に代入すると, 有病正診率は0.86, 無病正診率が0.94となり正確度は0.91と向上した。

肝細胞癌と肝細胞癌以外の2群に関してマハラノビスの汎距離²⁰⁾²¹⁾ (Dm^2) を計算してみると6.5553となり,

$$F = \frac{22 \times 36}{(58-2) \times 58} \times \frac{58-4-1}{4} \times 6.5553$$

$= 21.18 > F_{0.01}(4, 53) = 3.83$ であるので2群は充分判別できるものと思われた。

また $\sqrt{Dm^2}/2 = 1.2802$ より正確度の推定値は0.90となる。内部標本ではかなり推定値に近い判別ができた。

IV. 考 察

(A) 診断精度

肝腫瘍診断および肝細胞癌診断に関して有病正診率, 無病正診率, 正確度をAG, CT, US, RIの各検査法について求めた。

なお診断に際しては確信度を付けさせたが, 読影者により+, +, ±, -, --のうち, 殆どの症例について+, --と割り切る人と逆に±を好む人があった。集計にあたっては+以上をありとすることにした。

肝腫瘍診断ではAGで有病正診率, 無病正診率, 正確度のいずれにおいても高値を示した。

CT, US, RIでは有病正診率は0.54から0.61の範囲であり高くなかった。また正確度はいずれも約0.60前後であり, これも高くなかった。

肝細胞癌診断に関してAG, CT, USおよびRIでは有病正診率はそれぞれ0.67, 0.36, 0.39, 0.34であった。RIの読影は質的診断が困難であると辞退した人を除く人々によってなされたが, やはり

有病正診率は低かった。一方 CT や US では質的診断をほぼ全員が行ったものの、有病正診率は RI と同程度であり、低い。CT や US では肝細胞癌に特徴的所見の報告^{4)~6)}があるものの、CT や US で質的診断を行うことは難しいのだろう。

肝腫瘍と肝細胞癌診断に関して AG が優れていることは間違いなさそうであるが、問題はここで得た診断精度が普遍的に正しいかどうかと言う点である。

そこで本研究で求めた診断精度に影響をおよぼすと思われるものについて検討を加えてみた。なお本研究では自施設のものを読影した場合に特に高い診断精度を示すことは無かったので次の項目に絞って考察する。

(1) 診断経験の違い

優れた医学的手法は学習によりその診断精度の向上が期待できる。本研究では経験の度合をレポートの作成数で分けたところ肝腫瘍診断に関して正確度に有意差のみられたのは AG, US においてであり、CT, RI では差はみられなかった。

有病正診率は AG, CT, US にて最も経験豊富な読影者群で高い値を示した。RI では次に高い値を示したが最高値とは差がなく、経験豊富なほど

高い値を示すものと言える。

肝細胞癌診断に関しては AG で正確度が経験と共に忠実に向上した。CT, US, RI では正確度は最も経験の少ない読影者でかなり高く、かならずしも経験とともに向上したわけではなかった。有病正診率は経験と共に向上したが、無病正診率もかなり低下したからである。赤木らの報告¹⁵⁾では RI では著明な学習効果がなかったが、CT では肝細胞癌の診断に関して学習効果がみられたというが、これとは一部矛盾した。

このように読影経験の違いで診断精度が微妙に異なるから、一般に読影者層の構成が変るとともに診断精度は変動するものと思われる。従って本研究で得られた診断精度の絶対値が普遍的であるとは言い難い。

(2) 再読影結果との比較

一般に2度以上読影すると読影結果に変動を生ずる。再読影により診断精度がどのように変動するかを検討するために過去に行った検討結果²³⁾と比較した。そこでは、ここに用いた肝細胞癌のうち12例、肝細胞癌以外の症例のうち22例を対象とした。AG, CT, US および RI の順に50, 54, 49 および52人の医師が読影に参加したが、読影者の約

Table 7-a The Result of Dual Reading: Hepatocellular Carcinoma Case

Case No.	AG			CT			US			RI		
	1st	2nd	D	1st	2nd	D	1st	2nd	D	1st	2nd	D
1	0.94	0.88	0.06	0.80	0.60	0.20	0.69	0.10	0.59	0.48	0.21	0.27
2	0.84	0.85	-0.01	0.93	0.73	0.20	0.63	0.52	0.11	0.83	0.37	0.46
3	0.50	0.85	-0.35	0.07	0.05	0.02	0.65	0.52	0.13	0.81	0.63	0.18
4	0.76	0.62	0.14	0.72	0.58	0.14	0.51	0.39	0.12	0.52	0.21	0.31
5	0.88	0.96	-0.08	0.41	0.13	0.28	0.78	0.65	0.13	0.71	0.84	-0.13
6	0.88	0.92	-0.04	0.67	0.60	0.07	0.51	0.48	0.03	0.79	0.47	0.32
7	0.88	0.62	0.26	0.72	0.53	0.19	0.63	0.45	0.18	0.73	0.42	0.31
8	0.86	0.88	-0.02	0.65	0.55	0.10	0.37	0.19	0.18	0.63	0.58	0.03
9	0.94	0.85	0.09	0.44	0.23	0.21	0.18	0.13	0.05	0.37	0.05	0.32
10	0.24	0.12	0.12	0.67	0.30	0.37	0.76	0.26	0.50	0.31	0.05	0.26
11	0.84	0.77	0.07	0.39	0.48	-0.09	0.88	0.68	0.20	0.67	0.37	0.30
12	0.54	0.50	0.04	0.37	0.15	0.22	0.10	0.06	0.04	0.08	0.05	0.03
Average	0.76	0.74		0.57	0.41		0.56	0.37		0.58	0.35	
S.D.	0.22	0.24		0.24	0.23		0.24	0.22		0.23	0.25	

D=Difference [Detectability (1st) - (2nd)]

AG: $F_s = \frac{z^2 \times n}{s^2} = 0.495 < F_{0.01}(1,11) = 9.33$

CT: $F_s = 23.11$

US: $F_s = 13.67$

RI: $F_s = 22.45$

Table 7-b The Result of Dual Reading: non-Hepatocellular Carcinoma Case

Case No.	AG			CT			US			RI		
	1st	2nd	D									
1	0.44	0.00	0.44	0.50	0.13	0.37	0.24	0.10	0.14	0.77	0.37	0.40
2	0.14	0.08	0.06	0.13	0.05	0.08	0.53	0.23	0.30	0.00	0.16	-0.16
3	0.16	0.00	0.16	0.30	0.13	0.17	0.13	0.03	0.10	0.25	0.16	0.09
4	0.20	0.19	0.01	0.13	0.15	-0.02	0.20	0.10	0.10	0.04	0.05	-0.01
5	0.02	0.00	0.02	0.09	0.10	-0.01	0.10	0.00	0.10	0.46	0.37	0.09
6	0.02	0.04	-0.02	0.19	0.03	0.16	0.41	0.26	0.15	0.15	0.32	-0.17
7	0.06	0.04	0.02	0.17	0.00	0.17	0.42	0.00	0.42	0.38	0.21	0.17
8	0.48	0.38	0.10	0.22	0.48	-0.26	0.24	0.00	0.24	0.48	0.42	0.06
9	0.02	0.04	-0.02	0.22	0.10	0.12	0.35	0.35	0.00	0.02	0.05	-0.03
10	0.14	0.08	0.06	0.26	0.10	0.16	0.24	0.13	0.11	0.90	0.74	0.16
11	0.06	0.04	0.02	0.00	0.05	-0.05	0.13	0.16	-0.03	0.06	0.11	-0.05
12	0.00	0.23	-0.23	0.59	0.43	0.16	0.35	0.45	-0.10	0.25	0.21	0.04
13	0.06	0.38	-0.32	0.30	0.18	0.12	0.20	0.16	0.04	0.21	0.42	-0.21
14	0.50	0.08	0.42	0.52	0.23	0.29	0.13	0.10	0.03	0.85	0.68	0.17
15	0.14	0.04	0.10	0.37	0.05	0.32	0.20	0.00	0.20	0.44	0.36	0.08
16	0.06	0.00	0.06	0.26	0.23	0.03	0.63	0.55	0.08	0.04	0.16	-0.12
17	0.56	0.58	-0.02	0.37	0.13	0.24	0.39	0.23	0.16	0.37	0.42	-0.05
18	0.02	0.00	0.02	0.02	0.08	-0.06	0.18	0.13	0.05	0.04	0.00	0.04
19	0.34	0.08	0.26	0.78	0.48	0.30	0.45	0.19	0.26	0.48	0.21	0.27
20	0.10	0.08	0.02	0.26	0.08	0.18	0.12	0.03	0.09	0.08	0.21	-0.13
21	0.04	0.00	0.04	0.13	0.13	0.00	0.29	0.10	0.19	0.08	0.11	-0.03
22	0.66	0.42	0.24	0.91	0.78	0.13	0.82	0.58	0.24	0.19	0.26	-0.07
Average	0.19	0.13		0.31	0.19		0.29	0.18		0.30	0.27	
S.D.	0.21	0.17		0.23	0.19		0.19	0.17		0.27	0.19	

D = Difference [Detectability (1st) - (2nd)]

AG: $F_s = 3.21 < F_{0.01}(1,21) = 7.95$

CT: $F_s = 14.09$

US: $F_s = 27.01$

RI: $F_s = 0.587$

半数は本研究における読影者と重複していた (Table 7a, b).

肝細胞癌と肝細胞癌以外に分け、症例毎に肝細胞癌と診断した読影者の割合 (所見陽性率) をもとめ、前回と今回の差を検定したところ、CT, US では有意差がみられたが、AG では有意差がなかった。RI では肝細胞癌以外の症例で前回と有意差が見られなかった。

読影を繰り返しても所見陽性率の変動が少ない検査法は安定性の点で、優れた検査法と言える。この点でも AG は優れた検査法とすることができる。

RI では肝細胞癌以外の症例に関して所見陽性率は安定していると考えられた。

再読影で安定した結果をうる事が出来る検査法では本研究の様にして求めた診断精度が普遍性を持つ可能性はある。

つぎに本研究で得た診断精度を他の報告と比較してみたい。

高島らは infusion hepatic angiography²⁾を併用した血管造影の有用性を報告¹⁾しているが、有病正診率は95%であったという。蓮尾は単純 CT で58例の肝細胞癌症例中51例に病変を検出したとい (有病正診率は88%)、血管造影の手法を併用した CT での肝細胞癌の検出率は更に高率であったと報告¹¹⁾している。しかしこれらは多数の読影者による検討方法とは研究方法が異なり、そのまま比較することはできない。

肝腫瘍診断に関しては多人数読影による方法の報告がある。CT, US, RI についての遠山らの報告¹²⁾では、肝腫瘍診断に関して有病正診率、無病正診率および正確度が CT で各0.83, 1.00, 0.96, US で各0.83, 1.00, 0.96, RI で各1.00, 0.91, 0.93であったという。US はともかく比較的客観

的な検査法といわれる CT や RI に関しても値が大きく異なっている。この差の原因については明確でないが次の2つのことが理由として推定される。一つは検査手技の違いである。遠山らは CT で肝腫瘍が疑われた場合造影剤を追加している。経静脈性造影剤を使った方が単純 CT の情報を上回るとの報告¹⁴⁾もあり、造影剤の使用が診断精度を高くしている可能性はある。第二は読影対象の選び方の問題である。本研究では正常といえども血管造影は全例施行され、また生検も多数例で行われていることから、肝腫瘍などと紛らわしい症例が多く含まれた可能性がある。しかしこの他にも読影者の力量の差、機種の違い、検者の力量の差も複合的に働いてこのような差となったものと推察される。

RI による肝腫瘍の診断については、松本らエフィカシー委員会の研究¹⁵⁾がある。11人の読影者による肝腫瘍診断の有病正診率は0.85無病正診率は0.80であったという。本研究に比して高いが、松本らが読影に際して、年齢、性、触診所見、肝機能成績を読影者に与えた点は本研究のやり方と大きく異なっているので、本研究とは比較できない。

以上の検討から臨床研究では条件が一律でなく、診断精度の絶対値を求めることはできない。しかし本研究のように同一の対象を用いて診断精度を比較することは意味があるものと思われる。

(B) 多変量解析による総合映像診断

(a) 所見陽性率について

対象症例が肝細胞癌と診断可能か否かという点にのみ注目してみると、所見を探しやすく、且つ肝細胞癌らしく認識しやすい病変ほど、それが何であれ、多くの読影者が肝細胞癌と診断することになる。逆に病変部がはっきりしないか、はっきりしていても典型的な肝細胞癌にほど遠いと感じられるものでは肝細胞癌と診断する読影者の数は減少するものと思われる。そこで肝細胞癌の所見陽性率は前者では高くまた後者では低くなる。

所見陽性率は一般に、なんらかの診断に際し医師の判断の一致度をみることに過ぎない。Spitzer と Fleiss は医師の判断の一致度に偶然性が及ぼ

す影響について述べているが、判断の一致に偶然性の影響がかなりはいることを示唆している¹⁸⁾。

Fleiss は多数の医師の診断の一致が偶然かどうかを調べるのに kappa 統計量をもって表現できるとしている¹⁹⁾。本研究の場合、全58症例について AG, CT, US および RI で各26, 40, 31および19人の医師の一致に関して kappa 統計量は0.4557, 0.2556, 0.2577および0.2232でまた kappa/SE はそれぞれ11.22, 9.70, 7.61, および3.62となりいずれも限界値1.96よりも大きく、診断医の一致度は偶然性による一致度より高度に有意に大きいとすることができる。

評価が主観尺度であればなかなか偶然による一致の程度を越えることはない。本研究の場合、評価は単なる主観尺度に基づくと言うよりも客観的尺度に基づくと言えるのかもしれない。

実際本研究で肝細胞癌症例では所見陽性率は全体に高値を示し、肝細胞癌以外では所見陽性率は低値を示す傾向がみられた。この傾向は AG で著明であり、CT, US および RI ではそれほど明瞭でない。

いずれにしろ肝細胞癌診断に関して優れた映像診断法ほど、また客観的な映像診断法ほど、多数の読影者が肝細胞癌とそれ以外を明瞭に判別出来ることになるが、この所見陽性率をもって映像上の所見を数量化できた。

(b) 判別式

医用映像上みられる所見についてはその数量化が難しいため、判別式による研究はないようである。

映像診断の領域ではいずれも Bayes の定理に基づく単一映像についての研究であり、脳腫瘍を CT で診断する研究¹⁶⁾や胸部所見に関する研究¹⁷⁾がある。

本研究では58例について AG, CT, US および RI の肝細胞癌の所見陽性率を説明変数として判別式を求めたが、かなり良好な判別を得た。単に4種類の映像診断の結果のみでもかなり良好な総合診断が可能であった。

なお、肝細胞癌12例、肝細胞癌以外22例の2回の読影結果 (Table 7-a, b) から、AG に関しては

所見陽性率に有意差がなかったもののその他では差がみられ、所見陽性率は普遍性に問題があった。しかし、所見陽性率を用いてこの34例を判別してみると、2回とも肝細胞癌の11/12 (0.92) で等しく正しい診断ができ、肝細胞癌以外については、それぞれ19/22 (0.86), 20/22 (0.91) とほぼ同じ様に正しい診断ができた。2回にわたる読影結果を用いた判別結果がかなり良く一致した理由の第一は判別に最も寄与するAGの所見陽性率が安定しているためと思われる。

V. 総括

1. 各検査法の有用性を検討するため客観的映像評価法として、4種類の検査を施行した診断確定例を過去にさかのぼり収集し、3つの施設の多数の医師が読影するという研究を行った。

肝腫瘍診断では3つの施設の放射線科医がAG, CT, US およびRIについて各26, 40, 34, 31人が参加しつぎの知見を得た。

- 有病正診率はAG, CT, US, RIでそれぞれ0.83, 0.54, 0.57, 0.61であった。
- 無病正診率はAG, CT, US, RIでそれぞれ0.83, 0.76, 0.65, 0.67であった。
- 正確度はAG, CT, US, RIでそれぞれ0.83, 0.63, 0.60, 0.64であった。

肝細胞癌の診断では3つの施設の放射線科医がAG, CT, US およびRIについて各26, 40, 31, 19人が参加しつぎの知見を得た。

- 有病正診率はAG, CT, US, RIでそれぞれ0.67, 0.36, 0.39, 0.34であった。
- 無病正診率はAG, CT, US, RIでそれぞれ0.85, 0.80, 0.84, 0.74であった。
- 正確度はAG, CT, US, RIでそれぞれ0.78, 0.63, 0.67, 0.59であった。

2. 所見陽性率を説明変数として純粹に映像診断のみに基づく多変量解析により総合診断を試みた結果、有病正診率0.82, 無病正診率0.89, および正確度0.86と良好であった。

VI. 結語

AG, CT, US およびRI検査が同時期になされ診断の確定した肝細胞癌22例、肝細胞癌以外の36例を多数の医師に読影させた結果、CT, US また

はRIによる肝細胞癌の質的診断は可能ではあるが難しいことが多く、より正確な診断はAGによる必要があると思われた。またAG, CT, US およびRIによる読影結果から多変量解析による総合診断を試みたが、内部標本で良い結果が得られたことを述べた。

稿を終るにあたって、本研究の御指導、御校閲をいただいた九州大学教授 松浦啓一先生に心より感謝の意を表します。また本研究の統計処理に関して多大の御示唆、御意見を頂いた九州大学講師 池田正人先生、さらに御助言をいただきました九州大学助教授 増田康治先生、九州大学助教授 西谷 弘先生に御礼を申しあげます。最後になりましたが、各種検査データをこころよく提供していただいた九州大学の諸兄、また読影に御協力いただいた熊本大学教授 高橋睦正先生、長崎大学助教授 林 邦昭先生はじめ両大学ならびに九州大学の諸兄に深く感謝の意を表します。本研究の一部は厚生省助成金(福田班)の援助を受けた。

付録

確率密度はデータがn次元多変量正規分布すると仮定する場合、

$$F(\vec{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp(-1/2 \cdot Q)$$

但し $Q = [(\vec{x} - \vec{\mu}) \Sigma^{-1} (\vec{x} - \vec{\mu})]$

Σ : 分散・共分散行列

Σ^{-1} : Σ の逆行列

$|\Sigma|$: Σ の行列式

$(\vec{x} - \vec{\mu})'$: $(\vec{x} - \vec{\mu})$ の転置行列

$\vec{\mu}'$, \vec{x}' : $\vec{\mu}$, \vec{x} の転置ベクトル

\vec{x} : 変数ベクトル

$\vec{\mu}$: 平均ベクトル

n: 次元数(この場合4)

関数値がある値より大きいか小さいかによって分類を行うとき、この関数を判別関数という。正規分布に従っている2つのクラスの分類を行うための線型判別関数G(x)はつぎのようにして求める。

得られた標本が確率密度を F_1, F_2 である2つの母集団のどちらかの標本であるとし、それぞれからの標本である確率を π_1, π_2 ($\pi_1 + \pi_2 = 1$)と仮定すると

$$\begin{aligned} G(\vec{x}) &= \log \frac{\pi_1 F_1(\vec{x})}{\pi_2 F_2(\vec{x})} \\ &= (\vec{\mu}_1 - \vec{\mu}_2)' \Sigma^{-1} \vec{x} \\ &\quad - (1/2) \times (\vec{\mu}_1' \Sigma^{-1} \vec{\mu}_1 - \vec{\mu}_2' \Sigma^{-1} \vec{\mu}_2) \\ &\quad + \log \pi_1 - \log \pi_2 \end{aligned}$$

さらに2つ母集団それぞれからの標本である確率が等し

い ($\pi_1 = \pi_2$) と仮定すれば $G(\vec{x}) = G_1(\vec{x}) - G_2(\vec{x})$ となる。

$$\text{ここで } G_1(\vec{x}) = \vec{\mu}_1 \Sigma^{-1} \vec{x} - 1/2 \times \vec{\mu}_1 \Sigma^{-1} \vec{\mu}_1$$

$$G_2(\vec{x}) = \vec{\mu}_2 \Sigma^{-1} \vec{x} - 1/2 \times \vec{\mu}_2 \Sigma^{-1} \vec{\mu}_2$$

$G_1(\vec{x}) > G_2(\vec{x})$ のとき $G(\vec{x}) > 0$ となり第 1 の母集団からのものに分類され、また $G_1(\vec{x}) < G_2(\vec{x})$ のとき $G(\vec{x}) < 0$ となり第 2 の母集団からのものに分類されることになる。すなわち、 $G_1(\vec{x})$ 、 $G_2(\vec{x})$ に標本値を代入して大きな関数値をとる方からの標本であるとすれば、誤分類が少なく良い判別が期待できる。

文 献

- 1) 松井 修, 高島 力: 各種画像診断の限界—肝臓。画像診断, 3: 681—688, 1983
- 2) Takashima, T. and Matsui, O.: Infusion hepatic angiography in the detection of small hepatocellular carcinoma. Radiology, 136: 321—325, 1980
- 3) Kido, C., Sasaki, T. and Kaneko, M.: Angiography of primary liver cancer. Am. J. Roentgenol., 113: 70—880, 1971
- 4) Kamin, P.D., Bernardino, M.E. and Green, B.: Ultrasound manifestations of hepatocellular carcinoma. Radiology, 131: 459—461, 1979
- 5) Itai, Y., Nishikawa, J. and Tasaka, A.: Computed tomography in the evaluation of hepatocellular carcinoma. Radiology, 131: 165—170, 1979
- 6) Kunstlinger, F., Federle, M.P., Moss, A. and Marks, W.: Computed tomography of hepatocellular carcinoma. Am. J. Roentgenol., 134: 431—437, 1980
- 7) Scherer, U., Rothe, R., Eisenburg, J., Schildberg, F.-W., Meister, P. and Kissner, J.: Diagnostic accuracy of CT in circumscribed liver disease. Am. J. Roentgenol., 130: 711—714, 1978
- 8) 油野民雄, 桑島 章, 中嶋憲一, 多田 明, 小泉 潔, 今堀恵美子, 立野育郎: 核医学検査の進歩と問題点—肝臓イメージング—. 臨放, 28: 561—566, 1983
- 9) Broderick, T.W., Gosink, B., Menuck, L., Harris, R. and Wilcox, J.: Echographic and radionuclide detection of hepatoma. Radiology, 135: 149—151, 1980
- 10) Rossi, P. and Gould, H.R.: Angiography and scanning in liver disease. Radiology, 96: 553—562, 1970
- 11) 蓮尾金博: 肝細胞癌における Computed Tomography の診断価値に関する研究。日本医放会誌, 42: 130—151, 1982
- 12) 遠山淳子, 石垣武男, 石川 勉, 丹羽幸吉, 伴野辰雄, 水谷雅子, 牧野直樹, 水谷弘和, 鎌田憲子, 今輩倍康行, 佐久間貞行: 肝腫瘍における肝シンチグラフィ, 超音波断層撮影, コンピュータ断層撮影の診断能。画像診断, 2: 882—889, 1982
- 13) 松本 徹, 飯沼 武, 館野之男, 町田喜久男: 肝シンチグラフィの臨床の有効度の定量的評価—方法論と SOL 診断の ROC 解析を中心に—. 核医学, 19: 51—64, 1982
- 14) Berland, L.L., Lawson, T.L., Foley, W.D., Melrose, B.L., Chintapalli, K.N. and Talor, A. J.: Comparison of pre-and postcontrast CT in hepatic masses. Am. J. Roentgenol., 138: 853—858, 1982
- 15) 赤木弘昭, 金崎美樹, 平井和三, 山崎紘一, 坂田恒彦, 河合武司, 関本 寛, 末澤慶昭, 岡崎 進, 石田博文, 白川恵俊, 前田裕子, 新宅敬夫, 坂谷知泰, 石丸徹郎, 福田徹夫: 電子計算機の利用とその評価—核医学を中心として—. 日本医放会誌, 42: 1005—1040, 1982
- 16) 小牧専一郎: 原発性脳腫瘍の CT パターンの解析およびそれによる自動診断の可能性の検討。日本医放会誌, 39: 1055—1069, 1979
- 17) 小川史顕: 胸部 X 線診断情報処理システムに関する研究。日本医放会誌, 39: 862—877, 1979
- 18) Spitzer, R.L. and Fleiss, J.L.: A reanalysis of reliability of psychiatric diagnosis. Brit. J. Psychiat., 25: 341—347, 1974
- 19) Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychological Bulletin, 76: 378—382, 1971
- 20) Kendall, M.G.: A course in multivariate analysis. Charles Griffin & Co. Ltd., 1968 (浦 昭二, 竹並輝之訳: 多変量解析の基礎。サイエンス社, 1972)
- 21) 高根芳雄, 柳井晴夫: 現代人の統計—多変量解析法。朝倉書店, 1977
- 22) Rao, C.R. and Slater, P.: Multivariate analysis applied to differences between neurotic groups. Brit. J. Psych. (Stat. Sect.), 2: 17—29, 1949
- 23) 松浦啓一, 平田経雄, 西谷 弘, 飯尾正宏, 町田喜久雄, 蜂屋順一, 高島 力, 久田欣一, 高橋睦正, 多田信平, 川上憲司, 本保善一郎, 林 邦明, 金子昌生, 佐久間貞行, 松川 明, 有水 昇, 倉恒匡徳, 田坂 皓, 有井穂積: 主要臓器診断に対する放射線科学的検索の体系化に関する研究 (00437031)。昭和56年度科学研究費補助金 (総合研究 A) 研究成果報告書, 1982