

## 간스캔의 ROC 분석에 의한 진단적 평가

서울대학교 의과대학 내과학교실

이 명 철·문 대 혁·고 창 순

일본 국립방사선의학종합연구소

松本 徹·官野之男

= Abstract =

### ROC Analysis of Diagnostic Performance in Liver Scan

Myung Chul Lee, M.D., Dae Hyuk Moon, M.D. and Chang-Soon Koh, M.D.

Department of Internal Medicine, College of Medicine, Seoul National University, Seoul, Korea

Toru Matumoto, Ph.D. and Yukio Tateno, M.D.

National Institute of Radiological Sciences, Japan

To evaluate diagnostic accuracy of liver scintigraphy we analysed liver scans of 143 normal and 258 patients with various liver diseases. Three ROC curves for SOL, liver cirrhosis and diffuse liver disease were fitted using rating methods and areas under the ROC curves and their standard errors were calculated by the trapezoidal rule and the variance of the Wilcoxon statistic suggested by McNeil. We compared these results with that of National Institute of Radiological Science in Japan.

1) The sensitivity of liver scintigraphy was 74.2% in SOL, 71.8% in liver cirrhosis and 34.8% in diffuse liver disease. The specificity was 96.0% in SOL, 94.2% in liver cirrhosis and 87.6% in diffuse liver disease.

2) ROC curves of SOL and liver cirrhosis approached the upper left-hand corner closer than that of diffuse liver disease. Area ( $\pm$  standard error) under the ROC curve was  $0.868 \pm 0.024$  in SOL and  $0.867 \pm 0.028$  in liver cirrhosis. These were significantly higher than  $0.658 \pm 0.043$  in diffuse liver disease.

3) There was no interobserver difference in terms of ROC curves. But low sensitivity and high specificity of authors' SOL diagnosis suggested we used more strict decision threshold.

### 서 론

예민도(sensitivity)와 특이도(specificity)는 예, 아니오의 이원성인 검사결과를 가지는 임상검사의 좋고 나쁨의 평가에 이용된다<sup>1-4)</sup>. 그러나 연속적인 값을 가질 때에는 cutoff치의 선택에 따라 예민도와 특이도가 달라지기 때문에 여러 cutoff치에서의 양성율(true positive

rate, sensitivity)과 가양성율(false positive rate, 1-specificity)과의 관계를 도표로 나타낸 ROC 곡선(Receiver Operating Characteristic Curve)을 이용하게 된다<sup>5-9)</sup>. ROC 곡선은 질병의 유병율과 무관하며, cutoff치의 선택에 관여하는 편견을 배제할 수 있고, 판독자에 대한 정보를 알 수 있는 핵의학, 방사선 영상의 진단적 평가에 유용한 방법이다<sup>10-20)</sup>.

저자들은 간질환에서의 간스캔의 진단적 가치를 ROC 곡선의 분석으로 알아보고자 간내공간점유병변(space occupying lesion; SOL), 간경화증(liver cirrhosis;

본 논문은 1988년도 서울대학교병원 특진연구비 보조로 이루어 졌음.

LC), 기타 미만성 간질환(diffuse liver disease)에서 저자들의 ROC 곡선을 구하여 각 질환별 간스캔의 가치를 비교 분석하였으며 판독자간의 차이를 알고자 같은 환자를 대상으로한 松本등의 결과와 비교하였다<sup>21,22)</sup>.

### 대상 및 방법

대상은 일본 국립방사선의학종합연구소에서 부검, 수

술 등으로 간질환이 확인된 SOL 124명, 미만성 간질환 69명, 간경화증 61명 기타 4명과 정상인 143명이었다 (Table 1, 2).

간스캔은 이들에서 <sup>99m</sup>Tc sulfur colloid 2~5 mCi를 정맥주사하고 10~15분 후에 gamma camera로 전면부, 우측부, 후면부의 영상을 얻었다.

판독시에는 임상정보로서 환자의 성별, 연령, 체중, 간기능 검사, 초진소견을 참고로 하였으며 SOL, 간경화증, 미만성 간질환에 대하여 각각 1 Positive, 2 Probably positive, 3 Probably negative, 4 Negative의 4단계 확신도로 이들 질환의 유무를 판독하였다.

이들 각 질환에 대하여 cutoff치를 변화시켜 1 Positive

Table 1. Distribution of Cases with Confirmed Diagnosis

Normal	143
SOL	124
(Hepatoma)	18)
(Hepatoma + cirrhosis)	17)
(Metastasis)	76)
(Others)	13)
Diffuse liver disease	69
(Chronic hepatitis)	42)
(Acute hepatitis)	13)
(Others)	14)
Liver cirrhosis	61
Others	4
Total	401

Table 2. Methods of Confirmation

Autopsy	92
Surgery	161
Endoscopy	45
Biopsy	77
Angiography	53
CT	31
Clinical diagnosis	102

Table 3. Distribution of Level of Confidence, and True Positive (TP) and False Positive (FP) Ratios at Various Cutoff Levels

Diagnosis	Level of Confidence			
	1	2	3	4
SOL				
Positive	84	8	3	29
Negative	5	6	3	263
TPR	84/124 = 0.677	92/124 = 0.742	95/124 = 0.766	
FPR	5/277 = 0.018	11/277 = 0.040	14/277 = 0.051	
LC				
Positive	44	12	5	17
Negative	7	10	2	275
TPR	44/ 78 = 0.564	56/ 78 = 0.718	61/ 78 = 0.782	
FPR	7/294 = 0.024	17/294 = 0.058	19/294 = 0.065	
DLD				
Positive	9	15	4	41
Negative	3	15	6	121
TPR	9/ 69 = 0.130	24/ 69 = 0.348	28/ 69 = 0.406	
FPR	3/145 = 0.021	18/145 = 0.124	24/145 = 0.166	

(LC : Liver cirrhosis, DLD : Diffuse liver disease)

만을 양성으로 했을 때 1 Positive, 2 Probably positive 를 양성으로 했을 때, 1 Positive, 2 Probably positive, 3 Probably negative를 양성으로 했을 때의 양성률 (TPR), 가양성률(FPR)을 각각 구하고 이를 도시하여 3개의 ROC 곡선을 얻었으며, trapezoidal rule로 ROC 곡선 밑의 면적을 variance of Wilcoxon statistic으로 표준오차를 구하였다<sup>23)</sup>. 면적의 차이에 대한 검정은 면적의 차를 면적차의 표준오차로 나눈 값 즉  $z = (\text{Area 1} - \text{Area 2}) / \text{S.E.}(\text{Area 1} - \text{Area 2})$  이 1.96 이상일 때를 유의한 차가 있다고 판정하였다( $P < 0.05$ ).

성 적

SOL, 간경화증, 미만성 간질환의 진단에 대한 저자의 판독 확신도의 분포는 Table 3과 같았다. cutoff치의 변화에 따른 TPR는 SOL에서 0.677, 0.742, 0.766, 간경화증에서 0.564, 0.718, 0.782, 미만성 간질환에서 0.130, 0.348, 0.406이었으며 FPR는 SOL에서 0.018, 0.040, 0.051, 간경화증에서 0.024, 0.058, 0.065, 미만성

간질환에서 0.021, 0.124, 0.166이었다(Table 3). 저자와 일본 국립방사성의학종합연구소 11명의 TPR, FPR 를 도시한 것이 Fig. 1~3으로 SOL, 간경화증의 ROC 곡선의 미만성 간질환에 비하여 좌, 상에 위치하였으며 cutoff point의 ROC 곡선상의 분포는 SOL이 간경화증에 비하여 우측에 분포하는 경향을 보였다(Fig. 1, 2).

저자의 ROC 곡선의 밑의 면적과 표준오차가 SOL  $0.868 \pm 0.024$ , 간경화증  $0.867 \pm 0.028$ , 미만성 간질환  $0.628 \pm 0.043$ 으로 SOL, 간경화증 간에는 차이가 없었으나 미만성 간질환에서 유의하게 작았으며( $p < 0.01$ ), 11명의 의사의 평균 TPR, FPR로 구한 결과와는 차이가

Table 4. Area ( $\pm$  Standard Error) under the ROC Curves

Diagnosis	11 physicians	Author
SOL	0.868	$0.868 \pm 0.024$
LC	0.824	$0.867 \pm 0.028$
DLD	0.717	$0.628 \pm 0.043^{**}$

(\*\* :  $P < 0.01$ )

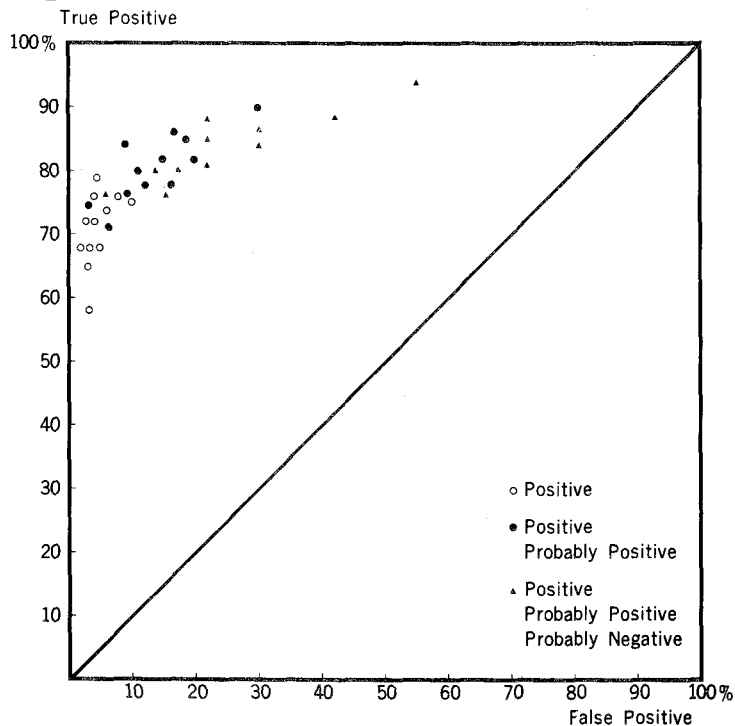


Fig. 1. ROC curve in patients with SOL (space occupying lesion).

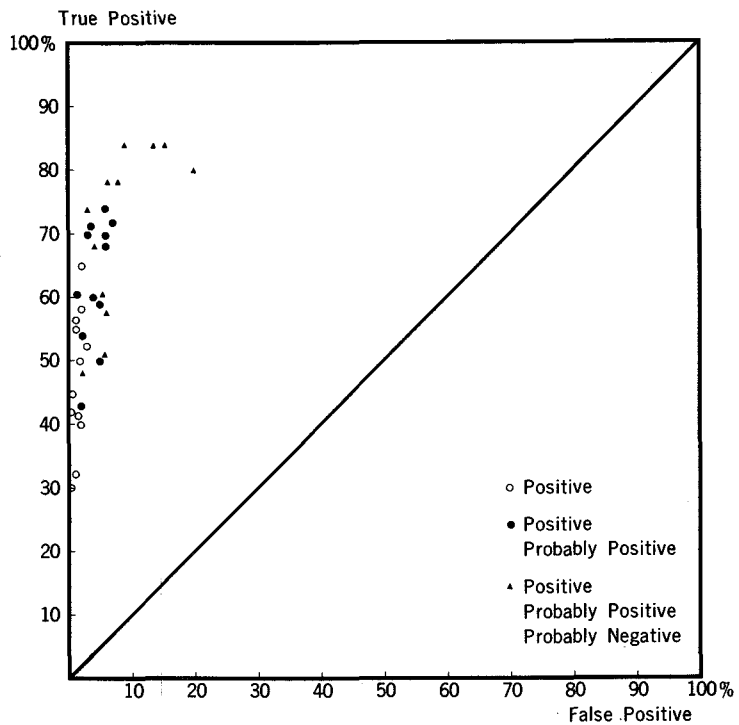


Fig. 2. ROC curve in patients with liver cirrhosis.

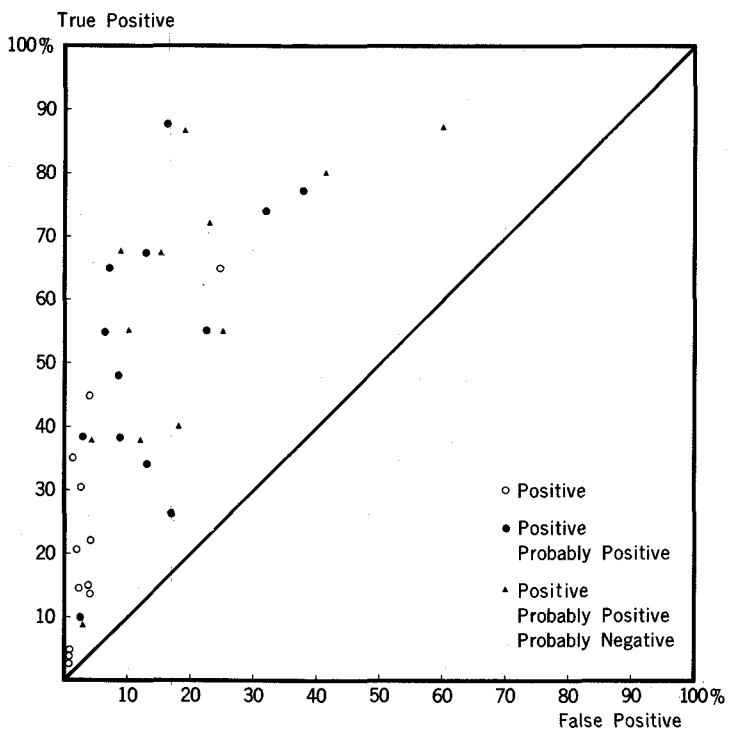


Fig. 3. ROC curve in patients with diffuse liver disease.

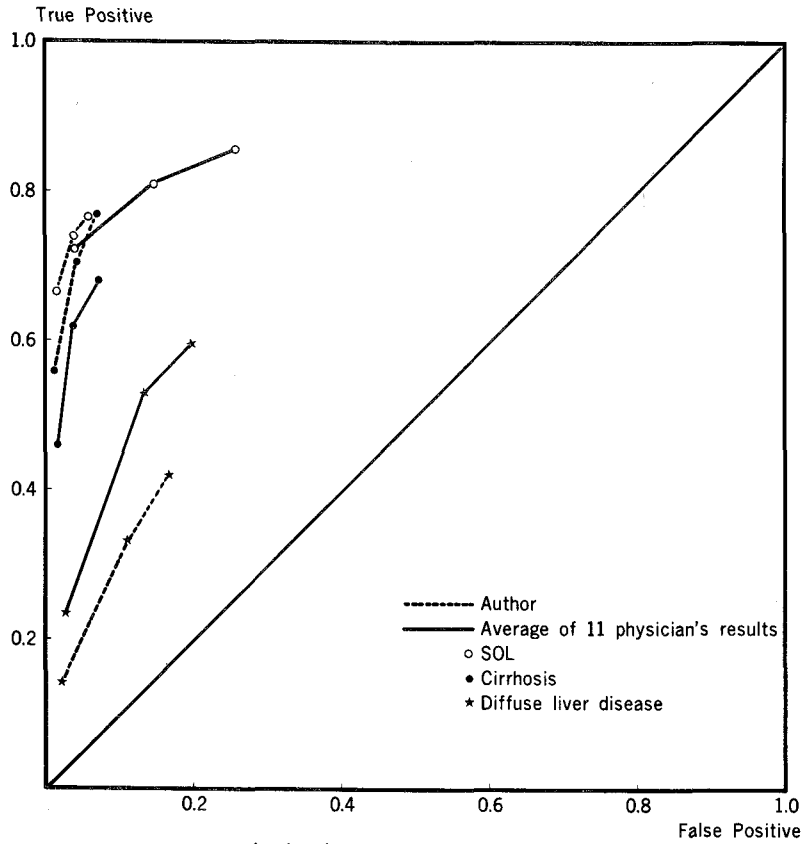


Fig. 4. Average curve for studies.

Table 5. Comparison of Sensitivity and Specificity of Liver Scan Between 11 Physicians and Author

Diagnosis	SN (11 physicians/ author)	SP (11 physicians/ author)
SOL	81.3/74.2	85.1/96.0
LC	62.7/71.8	96.3/94.2
DLD	55.6/34.8	86.1/87.6

(SN ; Sensitivity %, SP : Specificity %)

없었다(Fig. 4, Table 4).

확산도 2이하를 양성, 3이상을 음성으로 했을 때의 SOL, 간경화증, 미만성 간질환에 대한 저자의 예민도는 각각 74.2%, 71.8%, 34.8% 특이도는 96.3%, 94.2%, 87.6%로 SOL 진단에 있어 11명의 의사의 결과에 비하여 예민도는 낮았으나 특이도는 높은 경향을 보였다 (Table 5).

## 고 안

Fig. 1~4의 ROC 곡선의 SOL, 간경화증에서의 곡선이 미만성 간질환에 비하여 좌, 상에 위치하여 간스캔이 SOL, 간경화증의 진단에 유용함을 알 수 있었다. 이것은 간경화증을 포함한 미만성 간질환의 간스캔 소견이 비특이적인 것이지만 간의 좌엽대 우엽의 방사능비의 증가, 불규밀한 colloid 섭취, 비장비대, 골수섭취의 증가 등 간경화증에서의 간스캔 소견의 유용함을 말해 주는 것이라 하겠다<sup>24~26</sup>).

Table 4의 ROC 곡선 밑의 면적을 이용한 비교에서도 같은 결과를 얻을 수 있었다. 면적은 ROC 곡선에서 구할 수 있는 지표중 가장 널리 쓰이며 정상, 비정상군에서 임의로 하나씩 뽑은 두 표본치의 순위를 옮겨 나타낼 확률과 같다고 알려져 있다<sup>23</sup>. 면적이 0.5인 경우는 무용함 검사, 1은 이상적인 검사이며 표준오차를 구하여

통계적인 비교를 할 수 있다. 본 연구에서의 면적은 trapezoidal rule을 이용하여 구한 것으로 더우기 4단계의 확신도로 얻어진 값이어서 실제보다 작은 값을 갖는 것이 단점이다. 면적차의 표준오차의 계산에는 같은 환자를 대상으로 하였기 때문에 상관계수를 고려한 항을 빼 주어야 한다. 즉 실제 표준오차는  $\sqrt{SE^2(A1)+SE^2(A2)}$  보다 작은 값을 갖게 되나 SOL, 간경화증과 미만성 간질환의 면적차가 매우 커서 이 값을 그대로 사용하여서도 양자의 차이를 입증할 수 있었다<sup>27)</sup>.

핵의학 및 방사선영상의 판독에 있어 판독자는 그 자신의 성향, 기준, 사전 확률(prior probability)의 예측, 정확한 혹은 부정확한 판독의 가능한 모든 결과를 고려하게 되므로 판독자에 따라 cutoff치의 선택에 차이가 있게 된다<sup>5)</sup>. 실제 보고된 국소성 간질환에 대한 간스캔의 예민도는 60~98%, 특이도 64~99%, 기타 간질환의 경우는 예민도 40~96%, 특이도 19~88%로서 보고자에 따라 상이한 결과를 보이는 데<sup>28-36)</sup> 여기에는 환자대상군 및 검사방법의 차이, 표본 크기 등의 이유가 있겠지만 cutoff치의 차이도 고려할 수 있다<sup>37)</sup>. Fig. 1, 2에서 각 점은 TPR, FPR가 (0, 0), (1, 1)을 잇는 실제의 곡선 상에서 SOL은 우측에 간경화증은 좌측에 분포하였다. 이것은 이들 질환의 진단에 판독자가 서로 다른 기준 즉 간경화의 진단에는 FPR를 줄이기 위하여 엄격한 기준을 적용하며 SOL의 진단에는 FPR가 증가하더라도 TPR를 높이기 위하여 관대한 기준을 적용함을 말해 주는 것으로 SOL, 간경화증 진단의 false positive, false negative의 결과를 생각해 볼 때 당연하다 하겠으며 본 연구가 같은 대상, 방법을 사용한 결과라 할 때 이같은 차이가 Table 5의 SOL, 간경화증의 예민도, 특이도의 차이를 설명할 수 있을 것으로 생각되었다.

SOL의 진단에 있어 저자와 일본 국립방사선의학종합연구소의 ROC 곡선은 Fig. 4, Table 4에서와 같이 차이가 없었으나 저자의 예민도의 낮고 특이도가 높아 저자가 SOL의 진단에 보다 엄격한 기준을 사용함을 알 수 있었다.

### 결 론

저자들은 간질환에 있어서의 간스캔의 진단능을 알아 보고자 일본 국립방사선종합연구소에서 수술과 부검등으로 간질환이 확인된 258명과 정상인 143명에 대한 간스

캔을 SOL, 간경화증, 미만성 간질환에 대하여 4단계 확신도로 판독한 후 ROC 곡선을 구하였으며 이를 분석하여 다음의 결과를 얻었다.

1) SOL, 간경화증, 미만성 간질환에 대한 예민도는 각각 74.2%, 71.8%, 34.8% 특이도는 96.0%, 94.2%, 87.6%였다.

2) SOL, 간경화증의 ROC 곡선은 미만성 간질환에 비하여 좌, 상에 위치하였으며 ROC 곡선 밑의 면적(±표준오차)도 0.868±0.024, 0.867±0.028로 미만성 간질환의 0.628±0.043과 유의한 차가 있었다.

3) 저자와 일본 국립방사선종합연구소 ROC 곡선에는 차이가 없었으며 SOL의 진단에 저자가 엄격한 기준을 사용하는 경향을 보였다.

### REFERENCES

- 1) McNeil BJ, Adelstein SJ: *Determining the value of diagnostic and screening tests. J Nucl Med* 17:439, 1976
- 2) Patton DD: *Introduction to clinical decision making. Semin Nucl Med* 8:273, 1978
- 3) McNeil BJ, Keeler E, Adelstein SJ: *Primer on certain elements of medical decision making. N Engl J Med* 293:211, 1975
- 4) Katz MA: *A probability graph describing the predictive value of a highly sensitive diagnostic test. N Engl J Med* 291:1115, 1974
- 5) Metz CE: *Basic principles of ROC analysis. Semin Nucl Med* 8:283, 1978
- 6) Lusted LB: *General problems in medical decision making with comments on ROC analysis. Semin Nucl Med* 8:299, 1978
- 7) Swets JA: *ROC analysis applied to the evaluation of medical imaging techniques. Invest Radiol* 14:109, 1979
- 8) McNeil BJ, Hanley JA: *Statistical approaches to the analysis of Receiver Operating Characteristic (ROC) curves. Med Decis Making* 4:137, 1984
- 9) Gray R, Begg CB, Greenes RA: *Construction of Receiver Operating Characteristic curves when disease verification is subject to selection bias. Med Decis Making* 4:151, 1984
- 10) Turner DA, Fordham EW, Pagano JV, Ali A, Ramos MV, Ramchandran PC: *Brain scanning with the Anger multiplane tomographic scanner as a second examination. Radiology* 121:115, 1976
- 11) Turner DA, Ramchandran PC, Ali AA, Fordham

- EW, Ferry TA: *Brain scanning with the Anger multiplane tomographic scanner as a primary examination. Radiology 121:125, 1976*
- 12) Swets JA, Pickett RM, Whitehead SF, Getty DJ, Schnur JA, Swets JB, Freeman BA: *Assessment of diagnostic technologies. Science 205:753, 1979*
- 13) McNeil BJ, Sanders R, Alderson PO, Hessel SJ, Finberg H, Siegelman SS, Adams DF, Abrams HL: *A prospective study of computed tomography, ultrasound, and gallium imaging in patients with fever. Radiology 139:647, 1981*
- 14) Abrams HL, Siegelman SS, Adams DF, Sanders R, Finberg HJ, Hessel SJ, McNeil BJ: *Computed tomography versus ultrasound of the adrenal gland; A prospective study. Radiology 143:121, 1982*
- 15) Hessel SJ, Siegelman SS, McNeil BJ, Sanders R, Adams DF, Alderson PO, Finberg HJ, Abrams HL: *A prospective evaluation of computed tomography and ultrasound of the pancreas. Radiology 143:129, 1982*
- 16) Hermann GA, Herrera N, Sugiura HT: *Comparison of interlaboratory survey data in terms of Receiver Operating Characteristic (ROC) indices. J Nucl Med 23:525, 1982*
- 17) McNeil BJ, Hanley JA, Funkenstein HH, Wallman J: *Paired Receiver Operating Characteristic curves and the effect of history on radiographic interpretation. Radiology 149:75, 1983*
- 18) Sullivan DC, Coleman RE, Mills SR, Ravin CE, Hedlund LW: *Lung scan interpretation: Effect of different observers and different criteria. Radiology 149:803, 1983*
- 19) Miller TR, Goldman KJ, Epstein DM, Biello DR, Sampathkumaran KS, Kumar B, Siegel BA: *Improved interpretation of gated cardiac images by use of digital filters. Radiology 152:795, 1984*
- 20) Poon PY, Bronskill MJ, Henkelman RM, Rideout DF, Shulman HS, Weisbrod GL, Steinhardt MI, Dunlap HJ, Ginsberg RJ, Feld R Sutcliffe SB, Williams JI: *Mediastinal lymph node metastasis from bronchogenic carcinoma: Detection with MR imaging and CT. Radiology 162:651, 1987*
- 21) 松本 徹, 飯沼 武, 館野 之男, 町田喜久雄: 肝ツンテグラムの臨床的有効度の定良的評價—(1)方法論 0205 SOLマ 診断のROCミル 解析を中心に—. (核醫學第19巻 1號別刷)
- 22) 松本 徹, 飯沼 武, 館野 之男, 町田喜久雄: 肝ツンテグラムの臨床的有効度の定良的評價—(2)SOL 診断の醫師間變動の解析—. (核醫學第19巻 3號別刷)
- 23) Hanley JA, McNeil BJ: *The meaning and use of the area under a Receiver Operating characteristic curve. Radiology 143:29, 1982*
- 24) Christie JH, et al: *The correlation of clearance and distribution of colloidal gold in the liver as an index of hepatic cirrhosis. Radiology 88:134, 1967*
- 25) Simon TR, Neumann RL, Gorelick F, Riely C, Hoffer P, Gottschalk A: *Scintigraphic diagnosis of cirrhosis: A Receiver Operating Characteristic analysis of the common interpretive criteria. Radiology 138:723, 1981*
- 26) Shreiner DP, Barlai-Kovach M: *Diagnosis of alcoholic cirrhosis with right-to-left hepatic lobe ratio. J Nucl Med 22:116, 1981*
- 27) Hanley JA, McNeil BJ: *A method of comparing the areas under Receiver Operating Characteristic curves derived from the same cases. Radiology 148:839, 1983*
- 28) Drum DE: *Optimizing the clinical value of hepatic scintigraphy. Semin Nucl Med 8:346, 1978*
- 29) Nagler W, Bender MA, Blau M: *Radioisotope photoscanning of the liver. Gastroenterology 44:36, 1963*
- 30) Gollin FF, Sims TL, Cameron JR: *Liver scanning and liver function tests; A comparative study. JAMA 187:111, 1964*
- 31) Ferrier FL, Hatcher CR, Achord JL: *The value of liver scanning for detection of metastatic cancer. Am Surg 35:112, 1969*
- 32) Covington E: *The accuracy of liver photoscans. Am J Roentgenol 109:742, 1970*
- 33) Castagna J, Benfield JR, Yamada H: *The reliability of liver scans and function tests in detecting metastases. Surg Gynecol Obstet 134:463, 1974*
- 34) Dupriest RW, Hames JE, Roesch J: *A comparison of scintiscans and arteriograms for identifying metastatic intrahepatic tumors, Surg Gynecol. Obstet 136: 705, 1973*
- 35) Levin J, Geddes EW, Kew MC: *Radionuclide scanning of the liver in primary hepatic cancer; analysis of 202 cases. J Nucl Med 15:296, 1974*
- 36) Lunia S, Parthasarathy KL, Bakshi S: *An evaluation of <sup>99m</sup>Tc sulfur colloid liver scintiscans and their usefulness in metastatic workup; A review of 1424 studies. J Nucl Med 16:62, 1975*
- 37) Fritz SL, Preston DF, Gallagher JH: *ROC analysis of diagnostic performance in liver scintigraphy. J Nucl Med 22:121, 1981*