

위암의 임상적 병기 설정을 위한 전산화단층촬영 및 초음파 내시경의 진단력 평가

¹서울대학교 의과대학 외과학교실, ²분당서울대학교병원 외과

신루미^{1,2} · 이주희² · 이문수² · 박도중^{1,2} · 김형호^{1,2} · 양한광¹ · 이건욱¹

목적: 위암의 수술 전 병기 설정은 환자의 예후를 결정하고 치료 방침을 결정하기 위해 매우 중요하다. 이에 저자들은 위암의 TN병기 설정을 위해 주요한 진단 도구인 전산화단층촬영과 초음파 내시경의 진단력에 대해 분석하였다.

대상 및 방법: 2003년 5월부터 2007년 12월까지 분당서울대병원에서 위암으로 위절제술을 시행 받은 1,206명의 환자 중, 본원에서 시행한 검사가 아니거나, 내시경적 점막 절제술 후 위절제술을 시행한 환자의 경우를 제외한 1,174명을 대상으로 후향적인 연구를 시행하였다. 두 검사의 수술 전 병기와 술 후 조직병리학적 병기가 일치하는 정도를 보기 위해서 코헨의 카파값(kappa value)과 ROC곡선하 면적(Area under the ROC curve, AUC)을 이용하였다.

결과: 1,174명의 환자의 평균 나이는 59.31±11.98세였고, 조기위암이 638예, 진행 위암이 536예였다. 1,117명의 환자에서 전산화단층촬영이 시행되었으며, 초음파 내시경은 839예에서 시행되었다. T병기를 진단하는데 있어 전산화단층촬영이나 초음파 내시경의 카파값은 각각 0.4039 (P=0.021), 0.4201 (P=0.026)로 중등도의 일치도(moderate agreement)를 보였으나, 전체 T병기를 진단하는데 있어서 통계적으로 유의한 차이는 없었다. 이외에도 조기위암에서의 점막 병변과 점막하 병변을 구분하는 초음파 내시경의 진단력은 정확도 58.92%, 카파값 0.206으로 fair agreement를 보여 예상보다 진단력이 낮음을 알 수 있었다. 이외에도 T2 이상의 병변과 T2 이하의 병변을 감별하기 위해 두 검사를 동시에 시행한 환자군에서 두 검사 중 높은 병기를 진단한 검사를 채택하였을 때 AUC가 0.84로 한 개의 검사를 시행했을 때보다 진단력이 증가하였다. N병기를 진단하는 데 있어서는 전산화단층촬영의 진단력이 초음파 내시경보다 더 높음을 알 수 있었다.

결론: 위암의 T병기를 진단하기 위해 전산화단층촬영과 초음파 내시경 중 어느 한 검사가 우월하다고 할 수 없었으나, T2 이상의 병변과 T2 이하의 병변을 구별하기 위해서는 두 검사를 모두 시행하여 높은 병기를 채택한 경우 진단력이 높았다. 또한, 림프절 전이 여부를 알기 위해서는 전산화단층촬영이 우월하였다. 점막 병변과 점막하 병변을 구별하기 위한 초음파 내시경의 진단력은 낮았다.

중심 단어: 전산화단층촬영, 초음파 내시경, 진단력, 병기, 위암

책임저자: 김형호, 경기도 성남시 분당구 구미동 300번지
분당서울대학교병원 외과, 463-707
Tel: 031-787-7099, Fax: 031-787-4055
E-mail: hhkim@snuh.org

투고일(2009년 9월 10일), 게재확정일(2009년 10월 27일)

서 론

현대 의학에서는 다양한 치료 방법들이 발달함에 따라 개인별 맞춤의학과 함께 덜 침습적인 치료가 선호되고 있다. 위암 치료에 있어서도 점막하 절제술(endoscopic submucosal dissection), 점막절제술(endoscopic mucosal resection), 복강경 하 위절제술(laparoscopic gastrectomy), 광범위 위절제술(radical gastrectomy), 제한적 위절제술(limited gastrectomy) 등 다양한 방법으로 개인별 맞춤 의학을 시행할 수 있게 되었다.(1) 하지만, 암의 진행 정도에 따라 적절한 치료 방법을 결정할 수 있기 때문에 수술 전 정확한 진단이 필수적이다. 특히 내시경의 발달과 더불어 증가하고 있는 점막 절제술과 점막하 절제술은 많은 조기위암의 치료에 이용되고 있으며,(2) 복강경 위절제술의 경우도 T2병기까지 적응증을 확대하려는 움직임이 있어서 수술 전 T2병기와 T3병기의 심달도(depth of invasion) 구별이 중요하다. 현재, 위암의 수술 전 병기 설정을 위한 주된 검사로서는 전산화단층촬영과 초음파 내시경이 사용되고 있으며, 각각의 진단적인 가치를 평가하기 위한 다양한 연구들이 행해지고 있으나, 보고자에 따라 차이가 있다(Table 1).(3) 또한 Kwee 등(4)이 위암의 점막 병변과 이 외 병변을 구별하는 데 있어서 초음파 내시경의 정확도에 대해 고찰한 논문의 의하면, 초음파 내시경의 민감도는 18.2~100%, 특이도는 34.7~100%까지의 다양한 범위를 보이고, 통계적으로 유의하게

Table 1. Summary of the studies evaluating the accuracy of preoperative imaging

Reference	Accuracy	CT*, %	EUS [†] , %
Botet et al.(5)	T staging	42	92
	N staging	48	78
	Overall	45	73
Habermann et al.(6)	T staging	76	86
	N staging	70	90

*CT = computed tomography; [†]EUS = endoscopic ultrasonography.

불균일성(heterogeneity)을 보인다. 그러나, Botet 등(5)과 Habermann 등(6)의 연구에 의하면 현재까지는 T병기와 N병기를 진단함에 있어 초음파 내시경이 조금 더 정확한 것처럼 보인다. 이에 T병기를 진단하는 데 있어서 초음파 내시경의 진단적 가치가 높이 평가되고 있고, 임상적으로도 점막 병변과 더 깊은 병변을 감별하기 위한 일차적 영상검사로 사용되고 있다. 그러나, 지금까지의 수술 전 병기 설정을 위한 초음파 내시경의 진단력을 평가하기 위한 연구는 주로 조기위암에 국한되었다. 이에 본 연구에서는 진행성 위암을 포함한 전체 위암에서의 초음파 내시경과 전산화단층촬영의 진단력(diagnostic performance)을 알아보기 위해 각 진단 도구인 전산화단층촬영과 초음파 내시경의 진단결과를 표준 검사(gold standard)인 술 후 병리결과와 비교하여 전반적으로 얼마나 정확하게 예측하였는 지와 세부적으로 T병기, N병기를 얼마나 정확하게 예측하였는 지를 기준으로 진단력을 평가하였다.

대상 및 방법

1) 대상

2003년 5월부터 2007년 12월까지 분당서울대병원에서 위암으로 확진되어 위절제술을 시행 받은 1,206명의 환자 중 본원에서 시행한 검사가 아니거나, 병변이 두 개 이상이거나, 내시경적 점막절제술 후 위절제술을 시행한 경우, 또한 고식적인 위절제술을 받은 환자와 수술 전 보조 항암요법을 받은 환자의 경우를 제외한 1,174예를 대상으로 후향적인 연구를 시행하였다. 총 1,174예 중 남자가 773예, 여자가 401예였으며 평균 나이는 59.37세(25~89)였다. 전체 1,174예 중에서 1,117예가 수술 전에 전산화단층촬영을 시행하였고, 839예가 초음파 내시경을 시행하였다. 수술 후 병리학적인 병기 설정에 따라 조기위암이 638예(54.3%)로 진행위암인 536예(45.7%)보다 많았고, Stage I이 732예(62.4%)였다. 위절제술을 시행한 예 중 총 717예(61.1%)가 개복술로 복강경을 이용한 위절제술보다 다소 많은 양상이었다. 술 후 조직병리학적인 T1병기는 619예(52.7%), T2병기는 342예(29.1%), T3병기는 177예(15.1%), T4병기는 36예(3.1%)였다(Table 2). T병기를 진단하기 위해 전산화단층촬영과 초음파 내시경을 모두 시행한 환자군(paired group)은 818명이었고, 그 중 N병기를 언급한 환자군(paired group)은 194명이었다. 술 전 초음파 내시경으로 진단된 T1병기는 524예(62.5%)로 이 중에서 점막 병변과 점막하 병변을 구분할 수 있는 환자 수는 340예였다

2) 연구 방법

위의 심달도에 따른 T병기는 Union international contrele cancer (UICC)의 TNM (1997)분류에 따른 정의로 구분하였다. 전산화단층촬영을 위해서는 16-channel MDCT (mX-

Table 2. Patients characteristics (n=1,174)

		Number	Percentage (%)
Sex	Male	773	65.8
	Female	401	34.1
Stage	I	732	62.4
	II	157	13.4
	III	133	11.3
	IV	151	12.9
EGC/AGC	EGC*	638	54.3
	AGC [†]	536	45.7
Pathologic T staging	T1	619	52.7
	T2	342	29.1
	T3	177	15.1
	T4	36	3.1
Operation	OG [‡]	717	61.1
	LAG [§]	457	38.9
Total		1,174	100

*EGC = early gastric cancer; [†]AGC = advanced gastric cancer; [‡]OG = open gastrectomy; [§]LAG = laparoscopy-assisted gastrectomy.

8000; Philips Medical Systems, Highland Heights, Ohio)를 사용하였고, 내시경 기종으로는 GIF-Q 260 (olympus, Tokyo, Japan)이었으며, 초음파내시경 GF-UM2000 (olympus, Tokyo, Japan)과 관측장비(EU-M2000, Olympus, Tokyo, Japan)를 이용하였다. 초음파 내시경은 분당서울대병원의 두 명의 소화기내과 교수가 시행했고, 전산화단층촬영의 판독 역시 영상의학과와 두 명의 교수가 시행하였다. 초음파 내시경과 전산화단층촬영 모두, 보통 위내시경으로 위암이 확진된 환자에 한해서 술 전 병기설정을 위해 시행하는 검사이므로, 두 검사를 시행 및 판독하기 전에 임상적인 정보는 제공되었다. T1병변은 점막과 점막하층에 국한된 경우이며, T2병변은 고유근층과 장막하층, T3병변은 장막층, 그리고 T4병변은 인근장으로 침범한 경우로 정의하였다. 또한 N병기의 경우, 내시경 초음파로는 장경이 8 mm 이以上이고 저에코(hypoechoic LN)를 보이며, 구별이 잘 되는 경우(well demarcated LN), 전산화단층촬영으로는 단축의 장경이 8 mm 이以上이거나 괴사(necrosis)가 있는 경우를 N+로 분류하였고, N-는 의미 있게 커진 림프절이 없는 경우로 분류하였다.

위의 방법으로 각각 전산화단층촬영과 초음파 내시경으로 진단된 예를 수술 후 조직 병리학적인 병기와 비교하여 T와 N병기를 정확히 예측한 정도를 분석하였고, 각 병기를 정확하게 예측한 정도를 진단력(diagnostic performance)이라 정의하였다. 그리고 T1병기에 있어서 점막 병변과 점막하 병변을 구분하는 초음파 내시경의 진단력에 대해서도 알아보았다. 두 검사를 동시에 시행한 환자군에서 두 검사 중

한 검사를 채택하여 진단한 경우 진단력의 차이가 있는 지에 대해서도 분석하였다.

통계 처리는 SPSS 12.0 (SPSS Inc. Chicago, IL, USA)과 SAS 9.1을 이용하였다. T병기를 진단하기 위한 전산화단층촬영과 초음파 내시경의 진단력(diagnostic performance)를 비교하기 위해 카파값(kappa value)과 카이제곱 검정을 사용하였다. 또한, T1병기의 점막 병변과 점막하 병변(T1a와 T1b) 진단을 위한 두 검사의 진단력을 비교하기 위해서 역시 카파값을 이용하여 정확히 예측한 정도를 보았다. 카파값은 최소 0에서부터 1까지의 일치도를 보는 값으로 다음과 같이 해석이 가능하다. <0.2인 경우 Poor agreement, 0.21~0.40 Fair agreement, 0.41~0.60 Moderate agreement, 0.61~0.80 Good agreement, 0.81~1.0 Very good agreement를 나타낸다. 두 검사를 동시에 시행한 환자군에서 T2 이상 또는 T2 이하의 병변을 구별하기 위한 정확도를 비교하기 위해서 ROC 곡선하 면적(area under the Receiver Operating Characteristic curve, AUC)을 사용하였다. AUC는 검사의 정확도를 나타내 주는 검사로서 AUC값에 따라서 0.9~1은 Excellent, 0.8~0.9 Good, 0.7~0.8 Fair, 0.6~0.7 Poor, 0.5~0.6 Fail을 의미하며 AUC가 큰 검사일수록 진단적인 정확도가 더 높다. N병기의 진단력을 보기 위해서는 두 검사를 동시에 시행한 환자군에서 각각의 True positive와 True negative의 합을 총 환자수로 나눈 값을 비교하였다.

결 과

1) T병기에 있어서 전산화단층촬영과 초음파 내시경의 진단적 일치도 비교

전산화단층촬영을 시행한 총 1,117예 중 735예(64.9%),

초음파 내시경을 시행한 총 839예 중 585예(69.7%)가 정확한 T병기를 예측했다. 이에 전체 T병기의 진단력을 보기 위한 카파값은 전산화단층촬영인 경우 0.4039, 초음파 내시경인 경우 0.4201로 두 검사법 모두 중등도의 일치도(moderate agreement)를 보였고, 초음파 내시경의 카파값이 다소 크긴 하였으나, 통계적인 유의성이 없었다(P=0.6274) (Table 3). 즉, T병기를 예측하는 데 있어서 초음파 내시경의 진단력이 더 높지 않다. 전산화단층촬영으로 과대평가(overstaging)된 경우는 172예(15.3%), 과소평가(understaging)된 경우는 220예(19.7%)였고, 초음파 내시경으로 과대평가된 경우는 154예(18.4%), 과소평가된 경우는 100예(11.9%)로 전산화단층촬영의 경우 과소평가된 예가 다소 많았고, 초음파 내시경인 경우에는 과대평가된 예가 많았다 이 외에도 각 병기별로 두 검사의 양성예측도(PPV)와 음성예측도(NPV) 및 정확도(overall accuracy)의 특징을 살펴 보았다. 두 검사 모두 낮은 T병기일수록 양성예측도가 높았고, 높은 T병기일수록 음성예측도가 더 높은 양상이었다. 특히하게도 두 검사 모두 정확도는 T2 병변의 경우 다른 병기에 비해 떨어지는 양상을 보였다.

이에 각각의 T병기별로 두 검사의 진단력을 비교하기 위해, 두 검사를 모두 실시한 집단에 대해 민감도와 특이도를 분석하였다(Table 4). T1병변에 대한 전산화단층촬영의 민감도는 92.4%로 초음파 내시경의 민감도인 83.1%보다 높았고(P<0.001), 특이도(CT 59%, EUS 76.3%, P<0.001)는 초음파 내시경이 더 높았다. 이 외에 T2병변에 대한 전산화단층촬영의 민감도는 22.2%로 초음파 내시경보다 낮은 값을 보였다.(P<0.001) T3병변의 경우 초음파 내시경의 민감도(CT 50%, EUS 54.4%, P=0.678)가 더 높았으나 이는 통계적으로

Table 3. Diagnostic performance of CT and EUS for T staging (n=1,117)

	CT T staging				Total	EUS T staging				Total
	cT1	cT2	cT3	cT4		eT1	eT2	eT3	eT4	
Pathologic T staging										
pT1	547	42	13	0	602	456	80	14	0	550
pT2	127	88	92	12	319	57	90	59	0	206
pT3	24	50	75	13	162	11	19	39	1	70
pT4	3	4	12	15	34	0	4	9	0	13
Total	701	184	192	40	1,117	524	193	121	1	839
Sensitivity	90.9	27.6	46.3	44.1		82.9	43.7	55.7	0	
Specificity	70.1	88	87.8	97.8		77.9	83.7	89.3	99.9	
PPV*	78	47.8	39.1	37.5		87.7	46.6	32.2	0	
NPV†	86.8	75.2	90.6	98.3		70.5	82	95.7	98.3	
Accuracy (%)	81.3	70.7	83.5	99.6	64.91	81.2	73.9	86.5	98.3	69.73
Overstaging (%)		172 (15.3%)					154 (18.4%)			
Understaging (%)		220 (19.7%)					100 (11.9%)			
Kappa		0.4039 (SE†: 0.021)					0.4201 (SE: 0.026)			

*PPV = positive predictive value; †NPV = negative predictive value; †SE= standard error.

유의하지 않았으며, T4병변에 대해서는 초음파 내시경으로 진단한 사례가 없어서 비교가 불가능했다.

두 검사를 모두 시행한 경우, 두 검사 중 한 검사의 결과를 채택했을 경우의 진단력의 차이에 대해서도 알아보았다. T병기를 진단함에 있어서 두 검사의 결과 중 낮은 병기를 채택하여 진단한 경우의 술 후 병리학적인 결과와 비교하였을 때 진단력은 병리학적인 결과를 정확히 예측한 비율(Percentage)로 계산을 하여 비교하였다. 낮은 병기를 채택한 경우의 진단력은 73.71%로 각 검사(CT 72.86%, EUS 70.17%)보다 높은 것을 관찰할 수 있었다(Table 5). 이에 반해, 두 검사 결과 중 높은 병기를 채택하여 진단한 경우의 진단력은 67.6%로 각 검사의 진단력보다 낮은 것을 관찰할 수 있었다.

Table 4. Sensitivity and specificity in CT and EUS (Paired group, n=818)

		CT*	EUS [†]	p-value
Sensitivity(%)	T1	92.4	83.1	<0.001
	T2	22.2	44.4	<0.001
	T3	50	54.4	0.678
	T4	41.7	0	
Specificity(%)	T1	59	76.3	<0.001
	T2	91.9	83.9	<0.001
	T3	91.2	89.9	0.308
	T4	99.3	99.9	0.125

*CT = computed tomography; [†]EUS = endoscopic ultrasonography.

마지막으로, 복강경 위절제술의 적응증인 T2 이하 병변을 구별해 내고, 점막 절제술 또는 점막하 절제술이 불가능한 T2 이상의 병변을 구별할 수 있는 진단력을 알아보기 위해 각 검사별 ROC 곡선하 면적(area under the ROC curve, AUC)을 비교하였다(Fig. 1). 우선, T2 이하의 병변을 구별해 내기 위해서는 초음파 내시경의 ROC 곡선하 면적이 0.824로 전산화단층촬영의 ROC 곡선하 면적인 0.819보다 큰 값을 보여 전산화단층촬영보다 초음파 내시경의 진단력이 더 높은 것을 알 수 있었다. 하지만, 두 검사를 동시에 시행하여 높은 병기로 진단한 경우의 ROC 곡선하 면적이 0.84로 각각의 검사를 시행한 것보다 더 정확한 것을 알 수 있었다. T2 이상의 병변을 구별하기 위한 ROC 곡선하 면적을 살펴

Table 5. Comparison of diagnostic performance in both CT and EUS (Paired group, n=818)

	N	Agreement (%)
CT_Tagr*	596	72.86%
EUS_Tagr [†]	574	70.17%
High_Tagr [‡]	553	67.60%
Low_Tagr [§]	603	73.71%

*CT_Tagr = agreement between CT and pathologic T; [†]EUS_Tagr = agreement between EUS and pathologic T; [‡]High_Tagr = agreement between Higher stage among the CT and EUS and pathologic T; [§]Low_Tagr = agreement between Lower stage among the CT and EUS and pathologic; ^{||}N = number of agreement.

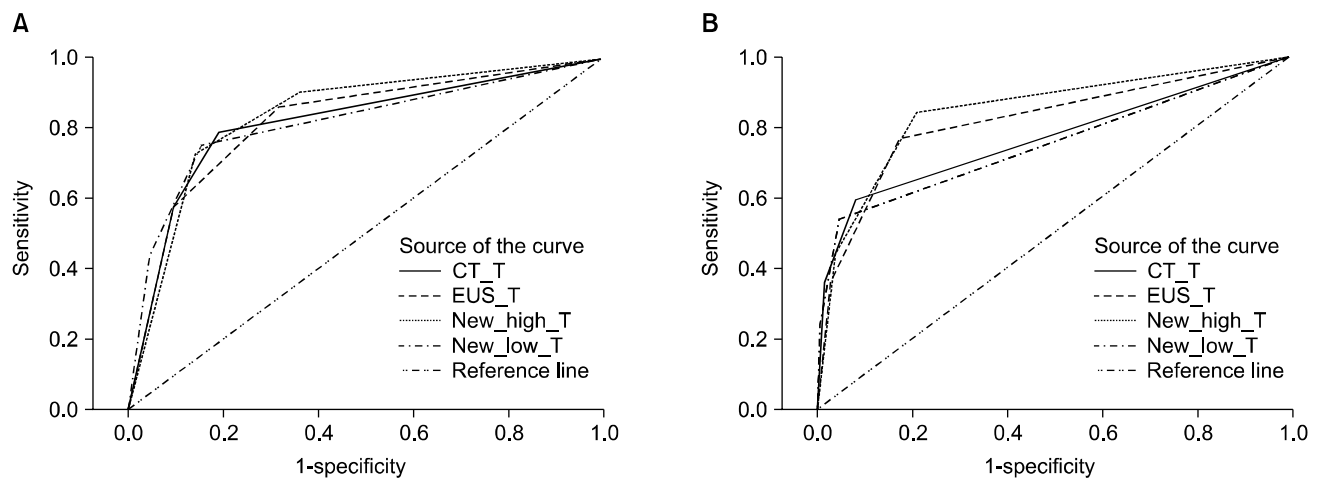


Fig. 1. Area under the ROC curve for discrimination of each T staging. (A) ROC curve for discrimination between T1, 2 and T3, 4. AUC of CT (CT_T) = 0.819; AUC of EUS (EUS_T) = 0.824; AUC of new_high_T* = 0.84; AUC of new_low_T[†] = 0.815 (*AUC of new_high_T = AUC when select higher stage among the CT and EUS in paired group. [†]AUC of new_low_T = AUC when select lower stage among the CT and EUS in paired group), (B) ROC curve for discrimination between T1 and T2, 3, 4. AUC of CT (CT_T) = 0.766; AUC of EUS (EUS_T) = 0.818; AUC of new_high_T* = 0.84; AUC of new_low_T[†] = 0.753 (*AUC of new_high_T = AUC when select higher stage among the CT and EUS in paired group; [†]AUC of new_low_T = AUC when select lower stage among the CT and EUS in paired group).

Table 6. Diagnostic performance of CT and EUS for N staging

		CT		EUS	
		N-	N+	N-	N+
Path	N-	563	104	61	38
	N+	227	263	39	59
Total		1,117		197	
Sensitivity (%)		53.67		60.2	
Specificity (%)		84.41		61.62	
Overall accuracy		74.95%		60.91%	

본 결과, 동일한 결과를 얻을 수 있었다. 즉, 두 검사 모두를 시행하여 높은 병기를 채택하였을 때의 ROC 곡선하 면적이 0.84로 전산화단층촬영(AUC of CT= 0.766)이나, 초음파 내시경(AUC of CT=0.818)보다 더 T2 이상의 병변과 T2 이하의 병변을 구별하는 진단력이 높은 것을 알 수 있다.

2) T병기에서 초음파 내시경의 점막 병변과 점막하 병변 구분에 있어서의 진단력

초음파 내시경에서 T1a와 T1b병기가 구별되는 340명의 환자들을 대상으로 카과값을 이용하여 점막 병변과 점막하 병변의 변별력 알아보았다. 정확도는 58.92%, 양성예측도는 71%, 카과값은 0.206 (P<0.001)으로 fair agreement를 보여 초음파 내시경의 점막 병변과 점막하 병변의 변별력은 비교적 낮았다(Table 6).

3) N병기에서 전산화단층촬영과 초음파 내시경의 진단력 비교

전산화단층촬영을 시행한 1,174에 중 N병기에 대해 기술한 1,157예와 초음파 내시경을 시행한 839예 중 N병기에 대해 기술한 197예를 분석하여 수술 전 임상적인 N병기를 진단함에 있어서 두 검사의 민감도와 특이도의 차이를 알아보았다(Table 6). CT의 민감도는 0.5367, EUS의 민감도는 0.6020이었으며 두 진단법의 민감도를 비교하기 위해 McNemar 검정을 수행한 결과, 유의수준 5%에서 두 진단법의 민감도는 유의한 차이가 있다고 할 수 없었다(P=0.7361). 전산화단층촬영의 특이도는 0.8441, 초음파 내시경의 특이도는 0.6162로 McNemar검정을 시행한 결과 두 검사의 특이도는 유의한 차이가 있었다(P=0.0061). 또한, 두 검사의 진단력의 차이를 알아보기 위해서 두 검사를 동시에 시행한 194예의 환자군에서 각각의 True positive와 True negative의 합을 총 환자수로 나눈 값을 비교하였다. 이에 전산화단층촬영은 진단력은 68.55%로 초음파 내시경(60.82%)보다 높았다. 결론적으로 N병기를 진단하기 위한 진단력은 전산화단층촬영이 더 우월하다.

고 찰

병기와 전신 상태에 따른 치료는 대부분의 암환자에서 보다 나은 결과(good outcome)을 위해서 매우 중요하다. 위암에 있어서 70세 이상의 고령 환자들의 비율이 증가함에 따라 수술 후 사망률(mortality)과 이환율(morbidity)이 증가하고 있다.(7) 보다 덜 침습적인 치료가 선호됨과 동시에 내시경적인 치료는 현재 조기위암의 치료에 널리 이용되고 있고, 최소한의 침습적인 치료로 생각되고 있다. Tada 등(8)은 1984년 Strip biopsy로 알려진 EMR technique을 개발하였고, Ono 등은 최근에 점막 절제술로 절제가 불가능한 30 mm 이상의 병변에 대한 점막하 절제술을 발전시킨 후, 많은 조기위암이 이 두 방법으로 치료되고 있다. 하지만, 조기 위암의 림프절 전이가 점막 병변의 경우 1~3%, 점막하 병변이 11~20%까지 보고 되고(9) 있어 수술 전 정확한 병기의 진단이 환자의 치료 방침을 결정하고, 추후 삶의 질에도 영향을 미칠 수 있어 점점 중요시 되고 있다.

위암의 T병기와 N병기를 진단하기 위해 많은 검사 방법들이 있다. 위내시경은 위암이 의심되는 환자들에게 있어 가장 기본적인 검사로 병변의 크기 및 위치 확인, 조직학적인 확진을 위해 필수적이다. 그러나 위내시경은 병변의 심달도 및 림프절 전이에 대한 정보를 제공할 수 없다는 한계를 가지고 있다. 이 내시경의 단점을 보완할 수 있는 검사가 바로 초음파 내시경이다. 초음파 내시경은 1979년 처음 도입되어,(10) 최근 급속한 발전을 보이고 있다. 초음파 내시경은 위벽의 각 층을 구별할 수 있기 때문에 조기위암의 심달도를 예측하기 위해 매우 유용한 검사로 간주되고 있다. 그러나 초음파 내시경 역시 시술자에 따른 편차가 있는 검사이고, M병기를 알 수 없으며, 근위 림프절의 전이만을 확인할 수 있다는 한계가 있다. 전산화단층촬영 역시 위암 환자의 수술 전 병기 설정에 있어, 원격 전이 여부를 판별하기 위해서 주로 사용되고 있고 최근에는 T병기의 진단을 위해서도 유용하다는 보고가 있다.(11)

Habermann 등(6)은 TN병기를 진단하는 정확도가 전산화단층촬영(T 76%, N 70%)과 초음파 내시경(T 86%, N 90%) 검사 모두 비슷하다고 보고하였다. 이와는 반대로 Botet 등(5)은 전산화단층촬영에 비해 초음파 내시경이 더 정확하다고 주장하였고, Puli 등(12)은 22개의 연구에 대해 meta-analysis를 시행한 결과 초음파 내시경은 조기위암의 점막 병변의 감별보다는 진행된 위암에서 더 정확하다고 결론지었다. Kwee와 Kwee(4)는 초음파 내시경의 진단력에 대한 기존의 18개의 연구를 분석해 본 결과, 각 자료들의 방법론적인 질(methodological quality)의 저하와 함께, 민감도와 특이도에 있어 불균질성(heterogeneity)을 관찰할 수 있었고 결론적으로 초음파 내시경의 정확도는 불분명하다고 주장하였다. 즉, 위암의 수술 전 병기 설정을 위한 초음파 내시경

과 전산화단층촬영의 정확도(overall accuracy or diagnostic efficiency)에 대한 많은 연구들이 진행되고 있으나, 현재까지 뚜렷한 결론은 없는 상태이다. 또한, 초음파 내시경에 대한 기존의 연구는 대부분 점막 절제술이나 점막하 절제술의 적응증을 감별하기 위해 조기위암에 국한되어 시행한 경우가 대부분이고, 진행성 위암에 대해 연구를 한 일부 보고가 있다 하더라도 대상 환자 수가 적다는 한계가 있다. 이에 본 연구는 다수의 환자를 대상으로 전산화단층촬영과 초음파 내시경의 진단력에 대해 알아 보았다는 데 의의가 있다.

지금까지 초음파 내시경의 진단력에 대한 연구들을 살펴보면 종양의 침습 범위가 점막에 국한되어 있는지를 감별하는 민감도와 특이도가 18.2~100% (median 87.8%), 34.7~100% (median 80.2%)로 매우 다양한 것을 알 수 있다.(4) 초음파 내시경의 T병기에 대한 정확도(overall accuracy)는 65~92% (T1 80~83%, T2 60~74%, T3 86%~100%, T4 25~83%), N병기에 대한 정확도는 65~90%로 보고되고 있다.(13-17) 최근에 Bentrem 등(18)이 R0 resection을 시행한 225명의 위암 환자에 대해 초음파 내시경의 정확도를 조사한 결과 T병기의 정확도는 57%이고, N병기의 정확도는 50%로 기존의 연구들보다 낮은 정확도를 보였다. 본 연구에서 T병기를 진단하기 위한 초음파 내시경의 정확도는 69.7% (T1 81.2%, T2 73.9%, T3 86.5%, T4 98.3%), N병기의 정확도는 60.91%로 Bentrem 등의 정확도보다 높은 것을 알 수 있었다.

전산화단층촬영에 있어서 T병기를 예측하는 정확도(overall accuracy)는 69~82%이고 MDCT (multi-detector row CT)의 도입으로 77~89%까지 향상되었다. N병기를 진단하는 데 있어서의 정확도는 59~78%로 보고되고 있다. (6,19-24) 이에 반해 본 연구에서 전산화단층촬영의 정확도는 T병기에서는 64.9%, N병기에서는 74.9%로 T병기를 진단하는 데 있어 정확도가 기존의 연구들보다 낮은 값을 보였다.

그러나, 전산화단층촬영과 초음파 내시경의 진단력을 비교하기 위한 카파값을 분석하였을 때, 두 검사 모두 중등도의 일치도를 보이며 T병기를 진단하는데 있어 통계적으로 유의한 차이가 없음을 확인하였다. 하지만, 복강경 위절제술의 적응증이 되는 T2 이하의 병변을 진단하거나, 점막 절제술이나 점막하 절제술의 비적응증을 감별하기 위한 T2 이상의 병변을 감별하기 위해서는 두 검사를 동시에 시행하여 높은 병기를 선택하는 경우가 진단력이 더 높았다. 두 검사 중 낮은 병기를 진단한 검사를 채택한 경우의 진단력이 더 높았음에도 불구하고, T2 이상의 병변 또는 T2 이하의 병변을 감별해 내는 진단력은 높은 병기를 진단한 검사를 채택한 경우가 더 높다는 것은 매우 흥미로운 사실이다. 이는 전체 분석 대상 중에서 T1병변이 차지하는 비율이 더 크기 때문에 두 검사 중 낮은 병기를 진단하는 경우를 채택하는 경우의 진단력이 더 높았을 것으로 생각한다. 또한,

T1병변에 대한 전산화단층촬영의 민감도가 높은 이유는 위 내시경 및 조직검사로 위암을 확정된 후 시행하기 때문인 것으로 생각한다. T2병변에 대한 전산화 단층 촬영의 민감도가 22.2%에 불과하고 초음파 내시경의 민감도가 44.4%라는 것을 감안할 경우 통계학적으로 검증하지는 않았지만, 개별적인 T병기를 예측하는데 있어서도 역시 초음파 내시경의 진단력이 더 우월할 것으로 생각된다. 이외에도 조기 위암에서의 점막 병변과 점막하 병변을 구분하는 초음파 내시경의 변별력은 정확도 58.92%, 양성예측도 71%, 카파 값 0.206 ($P < 0.001$)으로 Fair agreement를 보여 예상보다 진단력이 떨어짐을 알 수 있었다.

이 밖에도 두 검사 모두 T병기가 낮을수록 양성예측도가 높았고, 높을수록 음성예측도가 높았다. 또한 두 검사 모두 T2병변인 경우 정확도가 떨어지는 양상을 관찰할 수 있었다. T2병변의 진단에 있어서 초음파 내시경의 정확도가 떨어지는 이유는 점막하층 자체 감별의 어려움, 궤양에 의한 궤양 저부의 섬유화, 미세침윤 등으로 알려져 있다.(25) 박 등(26)은 T2병기의 경우 점막하 암의 섬유화, 궤양 등에 의한 과대평가와 미세침윤에 대한 과소평가, 장막하 암의 과대 평가 등으로 초음파 내시경의 정확도가 떨어진다고 주장한다.

본 연구는 진행 위암 환자들을 포함하여 다수의 환자를 대상으로 분석을 하였다는 장점을 가진 반면, 후향적으로 차트 리뷰를 통해서 자료를 수집한 한계를 가지고 있다. 이에 표현이 애매하거나 TN병기에 대해 정확한 언급이 없는 경우를 제외하여 정확한 자료를 수집하였다. 또한, TN병기에 대한 두 검사의 진단력을 비교하는 데 있어서, 무작위 배정이 아닌 후향적인 연구의 한계를 극복하기 위해 두 검사를 모두 시행한 환자들의 집단을 구별해서 분석하였다. 이 외에도 시술자 의존적으로 알려진 내시경 초음파를 두 명의 내시경 시술자가 시행하였다는 것에 반론을 제기할 수 있겠다. 그러나, 기존의 연구 결과에 따른 정확도와 유사한 정확도를 보이고, 모든 검사들은 시술자와 상관없이 정확한 진단을 내릴 수 있는 재현성(reproducibility)이 있어야 한다는 것을 감안했을 때 본 연구는 적절하다. 본 연구에서 T2병변에 대해 전산화단층촬영과 초음파 내시경 모두 다른 T병기와는 다르게 정확도가 떨어지는 것을 관찰할 수 있으나, 그 요인에 대해서는 따로 분석하지 않았다. 이에 이 연구 결과를 기반으로 두 검사의 T병기의 오차 이유 등에 대해 밝히는 연구들이 후행되는 것도 의미가 있겠다.

결 론

위암의 T병기를 진단하기 위해 전산화단층촬영과 초음파 내시경 중 어느 한 검사가 우월하다고 할 수 없었고 두 검사 모두 중등도의 일치도(moderate agreement)를 보여 진단적인 검사로서의 진단력이 낮았다. 또한 두 검사 모두 특

별히 T2병변을 진단하는 데 있어서 정확도가 낮았다. T병기 전반에 대한 진단력이 중등도이고 T2병변의 진단적인 정확도가 특별히 낮았던 것을 감안하였을 때, 복강경하 위 절제술을 위해 초음파 내시경을 이용한 T2와 T3병변의 구별을 적용하는 데 있어서 주의를 요해야 할 것이다. 또한, T1병변에 대해서 역시 점막 병변과 점막하 병변을 감별하기 위한 초음파 내시경의 진단력이 예상보다 낮으므로, 점막 절제술이나 점막하 절제술을 시행할 시 주의를 요해야 할 것이며, 무리한 시술은 지양해야 할 것이다. 그러나, 이러한 검사의 한계를 극복하기 위한 분석에서 한 환자에서 전산화단층촬영과 초음파 내시경을 동시에 시행하여 두 검사 결과 중 높은 병기를 채택하였을 때 T2 이상 또는 이하의 병변을 진단하는 데 있어 진단력의 향상을 꾀할 수 있었다. 이에 위암의 종류와는 무관하게 복강경하위절제술 또는 내시경적 시술을 시행하기 전에 두 검사를 모두 시행하는 것을 주장하는 바이다. 또한, 기술의 발달에 힘입어 초음파의 해상도와 분해능의 향상과 함께 개개의 시술자들의 재현성 및 정확도를 증가시키는 등의 노력으로 초음파 내시경의 진단적 정확도를 높여야 한다. 마지막으로, 림프절 전이 여부를 알기 위해서는 초음파 내시경보다 전산화단층촬영이 우월하였다. 이에 조기위암의 림프절 전이 여부를 확인을 위해 전산화단층촬영을 생략해서는 안 될 것이다.

REFERENCES

1. Nakajima T. Gastric cancer treatment guidelines in Japan. *Gastric Cancer* 2002;5:1-5.
2. Tada M, Murakami A, Karita M, Yanai H, Okita K. Endoscopic resection for early gastric cancer. *Endoscopy* 1993; 25:445-450.
3. Jensen EH, Tuttle TM. Preoperative staging and postoperative surveillance for gastric cancer. *Surg Oncol Clin N Am* 2007; 16:329-342.
4. Kwee RM, Kwee TC. The accuracy of endoscopic ultrasonography in differentiating mucosal from deeper gastric cancer. *Am J Gastroenterol* 2008;103:1801-1809.
5. Botet JF, Lightdale CJ, Zauber AG, Gerdes H, Winawer SJ, Urmacher C, Brennan MF. Preoperative staging of gastric cancer: comparison of endoscopic US and dynamic CT. *Radiology* 1991;181:426-432.
6. Habermann CR, Weiss F, Riecken R, Honarpisheh H, Bohnacker S, Staedtler C, Dieckmann C, Schoder V, Adam G. Preoperative staging of gastric adenocarcinoma: comparison of helical CT and endoscopic US. *Radiology* 2004;230:465-471.
7. Ahn HS, Lee HJ, Yoo MW, Kim SG, Im JP, Kim SH, Kim WH, Lee KU, Yang HK. Diagnostic accuracy of T and N stages with endoscopy, stomach protocol CT, and endoscopic ultrasonography in early gastric cancer. *J Surg Oncol* 2009;99: 20-27.

8. Tada M, Murata M, Murakami F, Shimada M, Mizumathi M, Arima K. Development of strip-off biopsy. *Gastroenterol Endosc* 1984;26:833-839.
9. Adachi Y, Shiraishi N, Kitano S. Modern treatment of early gastric cancer: review of the Japanese experience. *Dig Surg* 2002;19:333-339.
10. DiMagno EP, Buxton JL, Regan PT, Hattery RR, Wilson DA, Suarez JR, Green PS. Ultrasonic endoscope. *Lancet* 1980;1: 629-631.
11. Kim AY, Kim HJ, Ha HK. Gastric cancer by multidetector row CT: preoperative staging. *Abdom Imaging* 2005;30:465-472.
12. Puli SR, Batapati Krishna Reddy J, Bechtold ML, Antillon MR, Ibdah JA. How good is endoscopic ultrasound for TNM staging of gastric cancers? A meta-analysis and systematic review. *World J Gastroenterol* 2008;14:4011-4019.
13. Ganpathi IS, So JB, Ho KY. Endoscopic ultrasonography for gastric cancer: does it influence treatment? *Surg Endosc* 2006; 20:559-562.
14. Shimoyama S, Yasuda H, Hashimoto M, Tatsutomi Y, Aoki F, Mafune K, Kaminishi M. Accuracy of linear-array EUS for preoperative staging of gastric cardia cancer. *Gastrointest Endosc* 2004;60:50-55.
15. Tsendsuren T, Jun SM, Mian XH. Usefulness of endoscopic ultrasonography in preoperative TNM staging of gastric cancer. *World J Gastroenterol* 2006;12:43-47.
16. Willis S, Truong S, Gribnitz S, Fass J, Schumpelick V. Endoscopic ultrasonography in the preoperative staging of gastric cancer: accuracy and impact on surgical therapy. *Surg Endosc* 2000;14:951-954.
17. Xi WD, Zhao C, Ren GS. Endoscopic ultrasonography in preoperative staging of gastric cancer: determination of tumor invasion depth, nodal involvement and surgical resectability. *World J Gastroenterol* 2003;9:254-257.
18. Bentrem D, Gerdes H, Tang L, Brennan M, Coit D. Clinical correlation of endoscopic ultrasonography with pathologic stage and outcome in patients undergoing curative resection for gastric cancer. *Ann Surg Oncol* 2007;14:1853-1859.
19. D'Elia F, Zingarelli A, Palli D, Grani M. Hydro-dynamic CT preoperative staging of gastric cancer: correlation with pathological findings. A prospective study of 107 cases. *Eur Radiol* 2000;10:1877-1885.
20. Moschetta M, Stabile Ianora AA, Anglani A, Marzullo A, Scardapane A, Angelelli G. Preoperative T staging of gastric carcinoma obtained by MDCT vessel probe reconstructions and correlations with histological findings. *Eur Radiol* 2009.
21. Chen CY, Hsu JS, Wu DC, Kang WY, Hsieh JS, Jaw TS, Wu MT, Liu GC. Gastric cancer: preoperative local staging with 3D multi-detector row CT--correlation with surgical and histopathologic results. *Radiology* 2007;242:472-482.
22. Takao M, Fukuda T, Iwanaga S, Hayashi K, Kusano H, Okudaira S. Gastric cancer: evaluation of triphasic spiral CT and

- radiologic-pathologic correlation. *J Comput Assist Tomogr* 1998;22:288-294.
23. Hundt W, Braunschweig R, Reiser M. Assessment of gastric cancer: value of breathhold technique and two-phase spiral CT. *Eur Radiol* 1999;9:68-72.
24. Shimizu K, Ito K, Matsunaga N, Shimizu A, Kawakami Y. Diagnosis of gastric cancer with MDCT using the water-filling method and multiplanar reconstruction: CT-histologic correlation. *AJR Am J Roentgenol* 2005;185:1152-1158.
25. Park HB, Yeo HS, Kang MW, Oh KC, Cho JH, Cho HC, Chu JH, Kim WS, Kim MW. Esophagus, stomach & intestine. Endosc of endoscopic ultrasonography in gastric carcinoma. *Korean J Gastroenterol Endosc* 1997;17:8-14.
26. Park YS, Lee D, Lee DH, Kim NY, Jeong SH, Kim JW, Hwang JH, Lee SH, Kim JS, Jung HC, et al. Assessment of factors affecting the accuracy of endoscopic ultrasonography in T2 stage gastric cancer. *Korean J Gastroenterol* 2008;52: 86-90.

= Abstract =

Analysis of Diagnostic Performance of CT and EUS for Clinical TN Staging of Gastric Cancer

Ru Mi Shin, M.D.^{1,2}, Ju Hee Lee, M.D.², Moon Soo Lee, M.D.², Do Joong Park, M.D., Ph.D.^{1,2}, Hyung-Ho Kim, M.D., Ph.D.^{1,2}, Han-Kwang Yang, M.D., Ph.D.¹ and Kuhn Uk Lee, M.D., Ph.D., FACS¹

Department of Surgery, ¹Seoul National University College of Medicine, Seoul, ²Bundang Hospital, Seongnam, Korea

Purpose: Preoperative clinical staging of gastric cancer is very important for determining the treatment plans and predicting the prognosis. The previous reports regarding the accuracy of computed tomography or endoscopic ultrasound for the preoperative staging of gastric cancer have shown various outcomes. We analyzed the diagnostic performance of CT and EUS, which are important staging tools for the staging of TN gastric cancer.

Materials and Methods: We retrospectively analyzed 1,174 patients who underwent gastrectomy for gastric cancer at Seoul National University Bundang Hospital from May, 2003 to December, 2007. We derived the Kappa value to examine the agreement of the preoperative staging obtained from CT and EUS with the pathological staging.

Results: The mean age of the 1,174 patients was 59.31±11.98 years. Six hundred thirty seven patients had early gastric cancer and 536 had advanced gastric cancer. The diagnostic performance between CT and EUS for the T staging showed no significant difference between CT and EUS for the kappa values. The kappa values showed moderate agreement at 0.4039 (P=0.021) and 0.4201 (P=0.026), respectively. This suggests that there is no difference between the two examinations for the overall T staging. Analysis of the discrimination of mucosal and submucosal lesions with EUS showed an accuracy of 58.92% and a Kappa value of 0.206 (P<0.001), suggesting fair agreement and a lower diagnostic performance than expected. To differentiate lesions with stages higher than or equal to T2 or T3 from the lesion with stages lower than T2 or T3, respectively, adoption of the higher stage from the CT staging or the EUS staging showed a larger AUC of 0.84 than that from either stage alone. The CT-derived node stage had the higher diagnostic performance (68.55%) than that of the EUS-derived node stage (60.82%) for the node staging.

Conclusion: The CT-derived stage and EUS-derived stage showed comparable results for determining the T stage of gastric cancer. Yet the higher stage of the two stages from CT and EUS most accurately discriminated between those lesions with stages higher than T2 and those lesions with stages lower than T2. (**J Korean Gastric Cancer Assoc 2009;9:177-185**)

Key Words: Computed tomography, Endoscopic ultrasonography, Agreement, Gastric cancer