

생체의 돼지 폐를 이용한 인공 폐결절의 부피측정: 반자동 부피측정 프로그램과 영상의학과 의사의 수동측정간의 비교 연구¹

전주현 · 김진환 · 김성수 · 전호상² · 이현주³ · 박노혁⁴ · 조규성⁵

목적: 다채널전산화단층촬영을 이용한 폐암 검진이 증가함에 따라 직경 1 cm 미만의 소폐결절의 발견이 증가하고 있지만, 소폐결절에서 악성과 양성을 감별하기 어렵기 때문에 추적검사로 그 크기 변화를 관찰하는 것이 중요하다. 이 연구에서는 인공 결절을 이용하여, 컴퓨터 프로그램을 이용한 체적 진단(computer-aided volumetry, CAV) 결과와 영상의학과 의사의 수동 측정 결과의 정확도와 재현성을 비교 연구하였다.

대상과 방법: 공기로 팽창하고 고정된 돼지의 폐에 직경이 3 mm, 5 mm, 8 mm, 10 mm, 12 mm인 15개의 인공 결절을 삽입하여 다채널전산화단층촬영으로 촬영한 뒤 4명의 영상의학과 의사와 CAV가 측정한 직경과 부피를 비교하였다. 측정 재현성을 비교하고 상기 바이아스(bias)를 줄이고자 5일 간격으로 각 3번씩 측정하였고, 통계학적 분석을 위해 Paired T-test와 Bland-Altman plot을 사용하였다.

결과: 영상의학과 의사가 측정한 결절의 직경과, 체적 결과는 참값과 통계학적으로 유의하게 차이가 있는 것으로 나타났지만($p < 0.01$), CAV가 측정한 값은 참값과 유의한 차이가 없는 것으로 나타났다.

결론: CAV가 부피 측정에 있어 정확도와 재현성이 높으므로, 폐소결절의 양성 및 악성 결절의 감별 진단과 폐종양의 항암 치료 반응을 평가하는 데 있어 매우 유용하다 할 수 있다.

다채널전산화단층촬영(multi-slice computed tomography, 이하 MSCT)과 저선량전산화단층촬영(low-dose computed tomography, 이하 LDCT)을 이용한 폐암 검진이 증가함에 따라 폐결절의 발견이 증가하고 있으며 특히 직경 1 cm 미만의 소결절이 발견되는 빈도가 높아졌다. 이런 1 cm 미만의 소결절에서 양성과 악성 결절을 감별 진단하는 것이 중요하지만(1), 불규칙하거나 침상의 모양을 보이지 않는 경우가 많아 모양만으로 진단이 어려우므로 주기적인 추적 검사를 하여 폐결절의 성장 속도, 즉 부피가 두 배로 되는 기간(doubling time)을 아는 것이 가장 믿을 만한 소견이다.

그러나 World Health Organization (WHO) criteria (2, 3)와 Response Evaluation Criteria in Solid Tumors (RECIST) guideline (4) 모두 간접적인 결절의 단면적(WHO)과 단직경(uni-dimension, RECIST)의 변화를 기준

으로 하고 있다. 정확히 구(sphere)나 타원형이 아닌 종양에서 직경이나 단면적의 변화가 종양의 부피 변화 정도와 반드시 비례한다고 볼 수 없고, 폐결절이 불규칙하거나 침상의 모양일 경우 측정자에 따라 종양의 충실부위(solid portion)만 측정하거나, 또는 침상 부위(spiculation)를 포함할 수 있어 오차가 발생할 수 있다(interobserver variation). 더구나 폐결절의 추적 검사 시 한 사람의 관찰자가 직경을 측정하였다 하더라도 같은 결과를 보이지 않는다(intraobserver variation)(5). 따라서 결절의 부피를 정확히 측정할 수 있는 방법이 필요하다.

여러 연구에서 컴퓨터 프로그램을 이용한 체적 진단(computer-aided volumetry, 이하 CAV)이 폐결절의 부피를 측정하는데 정확성이 있으며, 특히 직경 1 cm 미만의 소결절에서 유용하다고 하였다(6, 7). 또한 재현성(reproducibility)이 높아 폐결절의 성장 정도와 폐암의 치료 효과를 정확히 알 수 있는 비침습적인 방법으로 보고되고 있다(1, 8, 9).

본 연구에서는 MSCT dataset를 이용하여 폐결절의 체적을 계산하는 알고리즘을 개발하였으며, 본 연구의 목적은 모형(phantom)을 이용하여 폐결절의 체적에 대한 최적 기준(gold standard)을 얻음으로써, 본 연구에서 개발한 컴퓨터를

¹충남대학교병원 영상의학과

²양산부산대학교병원 방사선종양학과

³서울대학교병원 영상의학과

⁴관동대 명지병원 영상의학과

⁵한국과학기술원 원자력 및 양자공학과

이 논문은 2009년 6월 23일 접수하여 2009년 10월 27일에 채택되었음.

이용한 체적진단 알고리즘의 정확도를 알아보고, CAD와 영상 의학과 의사의 정확도와 재현성을 비교 연구해 보고자 하였다. 또, 직경을 측정함으로써 발생하는 오차가 부피로 변환 할 경우 얼마나 증폭되는지 알아보았다.

대상과 방법

모형 및 인공 폐결절

공기로 팽창하고 고정된 돼지의 폐에 직경과 부피를 아는 15개의 인공 결절(synthetic simulated tumor, Kyoto Kagaku, Japan)을 삽입하였다. 각 결절의 직경은 3 mm, 5 mm, 8 mm, 10 mm, 12 mm이었으며 부피는 각각 14.13 mm³, 65.42 mm³, 267.95 mm³, 523.33 mm³, 904.32 mm³ 이었다. 직경에 따라 각각 3개씩 고정된 돼지의 폐에 삽입하였다. 인공 결절은 모두 평활하고 경계가 잘 그려지는 구의 형태를 지녔다(Fig. 1).

영상획득을 위하여 16-channel MSCT (Somatom Sensation 16, Siemens, Erlangen, Germany)를 이용하여 120 Kvp과 120 mA로, Field of view는 350 mm, 1 mm collimation, matrix 512 × 512 pixel, tube의 회전시간은 0.5초로 촬영하고, 촬영 후 1 mm 절편 두께로 재구성하여 DICOM 형식의 디지털 데이터로 저장하였다.

결절의 크기 측정

RECIST guideline은 측면영상(axial image)에서 최장 직경(longest diameter)을 측정하도록 하고 있으며 이에 따라 네 명의 영상의학과 의사가 Picture Archiving and Communication System (PACS) workstation 상에서 전자자를 이용하여 측면영상에서 15개의 결절의 가장 긴 직경을 측정하였다. 영상의학과 의사들은 결절의 크기와 각 크기에 따른 결절의 수 그리고 모양에 대한 정보가 전혀 없었으며, 각자 독립적으로 결절의 크기를 측정하였고, 상기 바이어스(recall

bias)를 최소화하기 위해 5일간의 간격을 두고 3번씩 측정하였다. 측정된 결절의 최장 직경을 기록하고 그에 기초하여 구의 부피를 계산하였다($v = 4/3\pi r^3$).

결절의 직경을 측정한 window level은 -700, window width는 1,500으로 모든 영상의학과 의사가 동일한 조건에서 측정하도록 하였다.

CAV Algorithm

본 연구에서 사용한 결절 인식 및 부피 계산 알고리즘은 결절이 포함되어 있는 단층영상들을 이용하여 결절의 부피를 산출하도록 개발되었다. 사용자가 결절이 관찰되는 하나의 단층 영상에서 결절의 위치를 지정하면 적응성 문턱(adaptive threshold) 값이 자동으로 계산된다. 이 적응성 문턱 값은 정해진 값(absolute threshold)을 임의로 적용하는 대신 해당 CT 영상을 구성하는 화소 값들의 분포도(HU value histogram)를 근거로 결정된다. 이 분포도에는 CT 영상에 포함된 조직들의 종류에 따라 서로 다른 지역 최댓값(regional peak value)을 가지는 여러 개의 포물선이 겹쳐 나타나게 되며, 이 중 가장 큰 최댓값을 가지는 영역이 결절에 해당된다. 본 연구에 사용된 문턱 값은 결절 분포의 최댓값과 그에 가장 인접한 분포의 최댓값의 평균값을 사용하였는데, 이는 한 영상 내에서 다른 물질 간의 구별을 명확히 할 수 있을 뿐만 아니라 여러 가지 촬영 조건들의 차이에서 기인하는 서로 다른 환자들의 영상 간 화소 값 차이에서 오는 상대오차를 최소화할 수 있다(Fig. 2). 이 문턱 값을 기준으로 해당 영상에서 결절의 단면 영역을 인식하게 되고, 그 후에 결절을 포함하면서 인접한 단층영상들을 차례로 처리하여 모든 결절 단면들을 확보한다. 마지막으로 이 영상들을 가상공간 내에 쌓으면 결절의 정확한 3차원 모형을 얻을 수 있으며 부피 계산도 가능해진다. 이때 첫 단층영상에서의 결절 위치 지정을 제외한 모든 과정은 컴퓨터에 의해 자동으로 진행된다(Fig. 3).

총 15개의 인공결절 중 13개를 대상으로 최장 직경과 부피

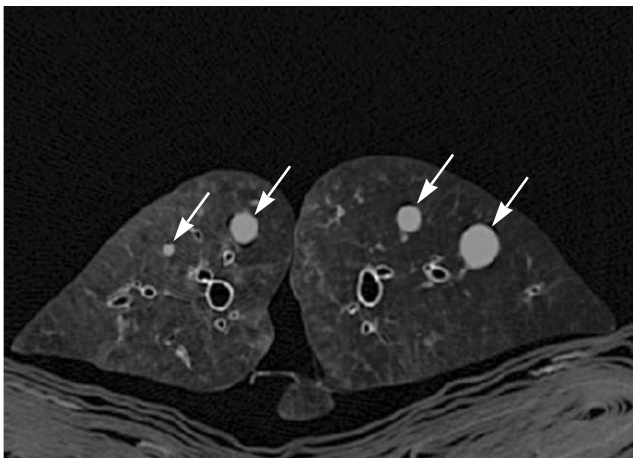


Fig. 1. Tumor phantoms. Axial image of artificial pulmonary nodules (arrows) in air inflated ex vivo porcine lung.

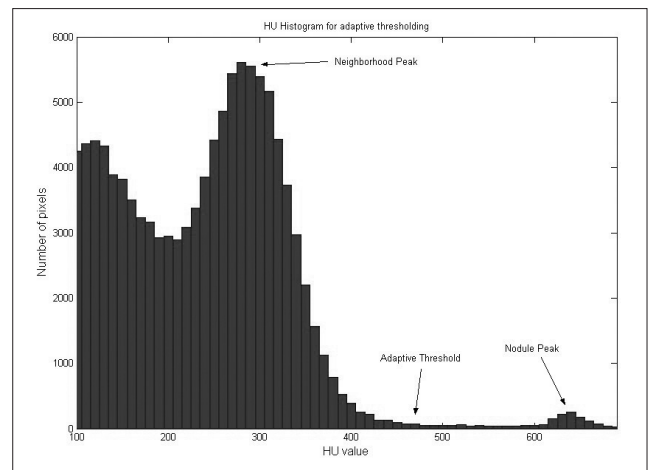


Fig. 2. Histogram for adaptive threshold. Adaptive threshold for each artificial nodule was estimated using mean value of HU value of nodule peak and that of neighborhood peak.

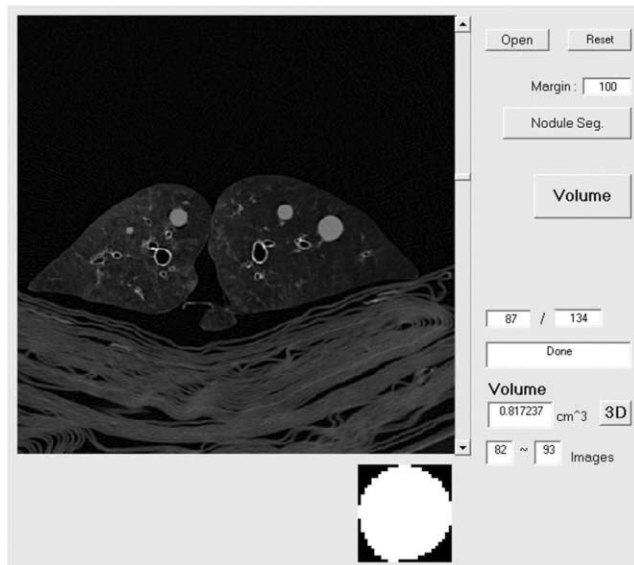
를 계산하였다. 8 mm와 12 mm 결절 중 각 하나씩은 인접한 혈관과의 하우스필드 수치 일치도가 높아 적응성 문턱 값으로도 구별할 수 없어 연구 대상에서 제외하였다.

통계 처리

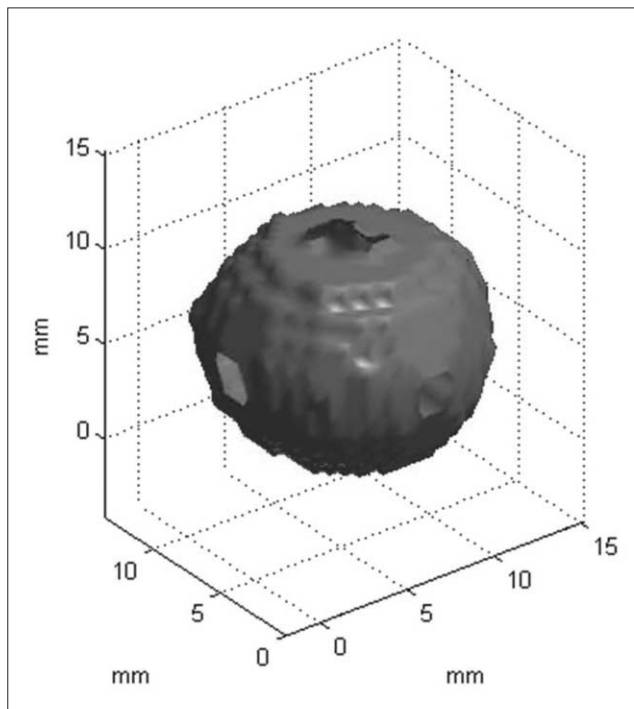
결절의 크기에 따라 4명의 영상의학과 의사가 측정된 결절의 직경과 각 결절별로 측정값의 표준편차를 계산하였고, 직경을 토대로 계산한 부피의 평균과 표준 편차를 개인별로 구하고 참값과의 편차를 백분율(percent(%))로 표시하였다. CAV로 측정된 결절의 부피 값도 평균과 표준 편차를 구하고 참값과의 편차를 %로 표시하였다. 또, 각 결절에 대하여 영상의학과 의

사의 측정값 평균을 구한 후 참값과의 편차를 계산하고 그 평균과 표준 편차를 구하였다. 참값과의 편차는 측정값에서 실제 값을 뺀 값을 실제 값으로 나누고 이 값에 100을 곱한 값으로 하였다. 그리고 편차의 평균의 유의성을 알아보기 위하여 paired-t test를 사용하였으며 유의 수준은 $p < 0.01$ 일 때 유의하다고 하였다.

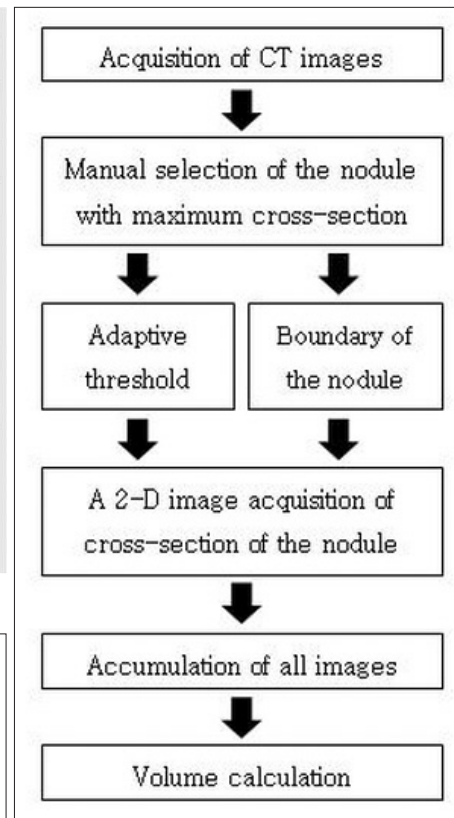
또한, 인공 결절의 실제 부피 값을 기준으로 하여 실제 부피에 대한 CAV 측정값의 차이와, 실제 부피에 대한 영상의학과 의사의 측정값의 차이를 Bland-Altman plot을 사용하여 비교하였다. 그리고 인공 결절의 부피가 클수록 오차의 절대값도 커지기 때문에, 실제 부피에 대한 각 측정값 차이를 백분율로



A



B



C

Fig. 3. Graphical user interface of semi-automated volumetry.

A. Selected images are shown on the right side of monitor. To start volumetry, click on the selected nodule, then press the volume button. The calculated volume is visible on the right-lower portion of monitor.

B. Three-dimensional image from 1 mm thickness MSCT data of an artificial nodule

C. The overall workflow for a nodule volume measurement.

환산하여 분석하였다. 통계분석은 MedCalc 프로그램 (MedCalc® ver 11.0, MedCalc Software, bvba)을 사용하였으며, 이 통계 방법에서 bias란 두 개의 측정방법으로 얻어진 정값의 차이의 평균을 의미하며, precision은 bias의 표준 편차를 나타낸다.

결 과

4명의 영상의학과 의사가 측정한 값은 한 명의 영상의학과 의사가 같은 결절을 측정하더라도 매번 다른 값이 측정되는 것으로 나타났고(Table 1), 측정한 직경의 평균값과 표준 편차

Table 1. Radiologist's Manual Measurement of Longest Diameter of Nodule. The Diameter Estimates and Standard Deviation of Each Artificial Nodule on MSCT Data were Calculated

Nodule No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
True Diameter(mm)	3	3	12	5	3	10	12	5	8	10	8	10	5	12	8
Radiologist 1															
1st	2.8	3.04	12.34	6.39	3.26	10.32	12.19	5.1	7.95	10.05	7.93	9.94	5.03	11.98	7.89
2nd	3.18	3.01	11.25	5.98	3.01	10.11	12.12	5.03	8.37	10.28	7.67	9.55	5.03	12.15	7.77
3rd	3.37	2.96	11.42	5.82	3.37	9.79	12.12	5.05	7.93	9.94	8.04	9.78	5.1	12.17	7.31
Standard Deviation	0.29	0.04	0.59	0.29	0.18	0.27	0.04	0.04	0.25	0.17	0.19	0.20	0.04	0.10	0.31
Radiologist 2															
1st	3.97	3.01	12.76	5.29	3.6	10.28	12.19	5.17	8.24	10.47	8.62	9.94	5.62	12.69	8.58
2nd	3.97	3.26	12.54	6.33	3.01	10.55	12.17	5.35	8.12	10.1	7.77	9.97	5.1	12.43	8.61
3rd	3.85	2.93	12.12	5.98	3.26	10.55	12.12	5.24	8.61	10.64	7.89	10.18	5.62	12.57	8.58
Standard Deviation	0.07	0.17	0.33	0.53	0.30	0.16	0.04	0.09	0.26	0.28	0.46	0.13	0.30	0.13	0.02
Radiologist 3															
1st	4.26	2.64	12.57	5.81	3.01	10.87	13.82	5.68	9.28	11.32	9.28	11.07	6.54	13.02	7.63
2nd	3.45	3.04	12.19	5.17	3.01	11.35	12.98	4.67	8.4	10.9	7.54	10.45	5.43	12.57	8.82
3rd	3.37	2.54	13.06	5.5	3.04	9.62	12.55	5.08	7.94	10.45	8.62	11	5.17	12.15	8.4
Standard Deviation	0.49	0.26	0.44	0.32	0.02	0.89	0.65	0.51	0.68	0.44	0.88	0.34	0.73	0.44	0.60
Radiologist 4															
1st	3.45	3.57	11.88	4.87	3.04	10.45	12.3	5.29	7.94	10.04	8.33	9.72	5.35	11.56	7.94
2nd	3.26	3.6	12.12	5.98	3.18	10.11	12.17	5.1	7.77	10.11	8.61	10.28	5.24	11.42	7.99
3rd	3.37	2.8	12.12	5.05	3.37	10.32	11.78	4.52	8.12	9.94	7.54	9.97	5.94	11.84	8.4
Standard Deviation	0.10	0.45	0.14	0.60	0.17	0.17	0.27	0.40	0.18	0.09	0.55	0.28	0.38	0.21	0.25

Table 2. Radiologist's Manual Measurement of Longest Diameter of Nodule. The Diameter Estimates, Standard Deviation and Percentages of Error of Artificial Nodules on MSCT Data were Calculated

True Diameter (mm)	3	5	8	10	12
Radiologist 1					
Diameter Estimates	3.11	5.39	7.87	9.97	12.12
Standard Deviation	0.20	0.53	0.34	0.25	0.07
% of Error	3.7%	7.8%	-1.6%	-0.3%	1.0%
Radiologist 2					
Diameter Estimates	3.43	5.52	8.46	10.30	12.36
Standard Deviation	0.43	0.41	0.22	0.27	0.24
% of Error	14.3%	10.4%	5.7%	3.0%	3.0%
Radiologist 3					
Diameter Estimates	3.15	5.45	8.41	10.78	12.85
Standard Deviation	0.51	0.53	0.59	0.54	0.57
% of Error	5.0%	9.0%	5.1%	7.8%	7.1%
Radiologist 4					
Diameter Estimates	3.29	5.26	8.03	10.10	11.85
Standard Deviation	0.26	0.47	0.21	0.22	0.34
% of Error	9.8%	5.2%	0.3%	1.0%	-1.3%
Average (Radiologist)					
Diameter Estimates	3.25	5.41	8.19	10.29	12.29
% of Error	8.2%	8.1%	2.4%	2.9%	2.5%

그리고 참값과의 편차를 결절의 크기에 따라 백분율로 표시하였을 때, 개인별 측정값과 측정값의 평균 모두에서 8 mm, 10 mm, 12 mm 결절보다 직경 3 mm 와 5 mm의 소결절에서 참값과의 차이가 더 크게 나타났다(Table 2).

CAV에서 측정한 결절의 부피는 같은 위치의 같은 결절에서 측정 시 매번 같은 값이 계산되어 결절 별 측정 오차가 일어나지 않았다. CAV로 측정한 부피와 4명의 영상의학과 의사가 측정한 직경에 근거하여 계산한 부피 값을 결절의 크기별로 분류하여 평균값과 참값과의 편차(percentage, %)를 비교한 결과(Table 3), CAV는 결절의 크기에 관계없이 편차가 비교적 고르게 보였으나, 영상의학과 의사가 측정한 결절의 부피 값은 직경 3 mm와 5 mm의 소결절에서 참값과의 차이가 더 크게 나타났다. 또, 측정된 직경 값의 오차보다 환산된 부피 값의 오차가 훨씬 크게 나타났다(Fig. 4).

폐결절의 부피 값은 CAV의 편차 평균이 -17.97이므로 영상의학과 의사가 측정한 값의 편차 평균인 30.61보다 작으며, 결절의 크기별 측정값 평균의 표준 편차는 CAV가 28.9, 영상의학과 의사가 27.3으로 비슷하였다. paired-t test를 통해 평균 차이의 유의도를 알아보았고, 참값과 CAV 측정값은 $p = 0.045$, 참값과 영상의학과 의사의 측정값은 $p = 0.002$ 로, 유의 수준을 $p < 0.01$ 로 보았을 때 CAV 측정값은 $p > 0.01$ 이므로 참값과의 평균 차이가 유의하다고 할 수 없다. 반면 영상의학과 의사의 측정값은 $p < 0.01$ 이므로 참값과의 평균 차이가 유의하다고 볼 수 있다(Table 4). 또, CAV는 참값보다 작게, 영상의학과 의사는 참값보다 크게 측정하는 경향이 관찰되었다.

Bland-Altman 분석에서도 실제 부피와 CAV 측정값의 $\text{bias} \pm \text{precision}$ 은 $-17.97 \pm 28.89 \text{ mm}^3$ 이고, 실제 부피와 영상의학과 의사의 측정값의 $\text{bias} \pm \text{precision}$ 은 $30.61 \pm$

27.31 mm^3 로 CAV로 측정한 값이 영상의학과 의사가 측정한 값보다 오차가 작은 것으로 나타났다. 실제 부피와 CAV 측정값 차이를 백분율로 환산한 $\text{bias} \pm \text{precision}$ 은 $-6.62 \pm 20.51\%$ 이고, 실제 부피와 영상의학과 의사의 측정값 차이를 백분율로 환산한 $\text{bias} \pm \text{precision}$ 은 $15.3 \pm 14.19\%$ 이며, 각각의 95% 신뢰구간(95% Confidence Interval)은 CAV 측정 시 -26.13%에서 11.13%, 영상의학과 의사가 측정 시 10.31%에서 33.29%이었다. 산점도의 기울기는 영상의학과

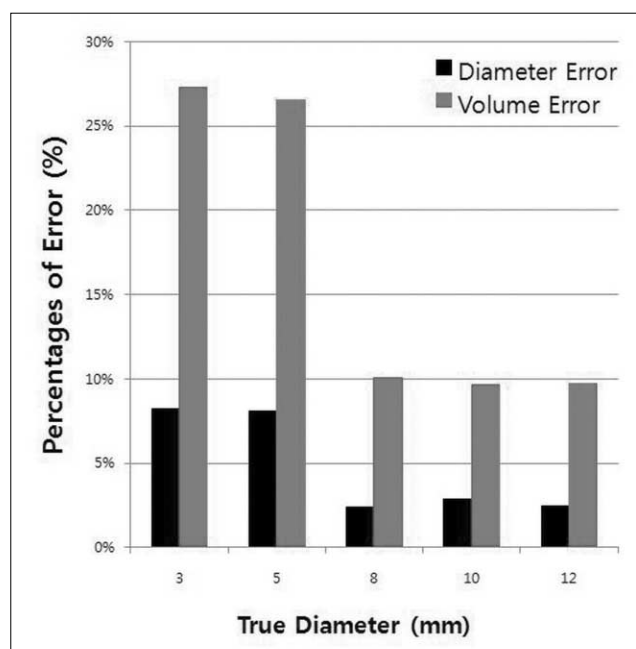


Fig. 4. Relationship between changes in uni-dimensional measurement and tumor volume.

Table 3. Volume Estimates and Percentages of Error of the Volume Measurements on the Basis of Diameter and Computer Aided Volumetry

True Volume (mm ³)	14.13	65.42	267.95	523.33	904.32
CAV					
Volume Estimates	14.09	62.72	236.02	512.74	839.42
% of Error	-0.3%	-4.1%	-11.9%	-2.0%	-7.2%
Radiologist 1					
Volume Estimates	15.76	82.05	255.10	519.16	932.11
% of Error	11.5%	25.4%	-4.8%	-0.8%	3.1%
Radiologist 2					
Volume Estimates	21.10	88.13	316.50	571.49	988.57
% of Error	49.3%	34.7%	18.1%	9.2%	9.3%
Radiologist 3					
Volume Estimates	16.37	84.72	311.48	655.80	1109.99
% of Error	15.9%	29.5%	16.2%	25.3%	22.7%
Radiologist 4					
Volume Estimates	18.69	76.16	270.64	539.90	869.73
% of Error	32.3%	16.4%	1.0%	3.2%	-3.8%
Average (Radiologist)					
Volume Estimates	17.98	82.76	288.43	571.59	975.10
% of Error	27.3%	26.5%	10.0%	9.6%	9.7%

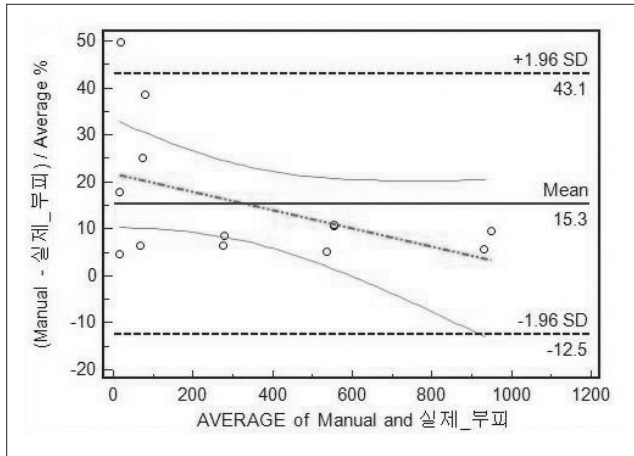


Fig. 5. The percentage of error of the volume estimates by radiologists in 15 synthetic nodules. The percentage of error become greater with decreasing nodule size.

Table 4. Overall Value of Standard Deviation and Error from True Volume

	CAV	Manual (All Radiologists)
Volume Error (mm ³)	-17.97	30.61
Standard Deviation	28.89	27.32
*p value	p=0.045	p=0.002

*Paired t-test, $p < 0.01$

의사가 측정 시에는 완전한 하강을 보여(Fig. 5) 인공 결절의 크기가 감소할수록 백분율 환산 오차 값이 증가하는 것으로 나타난 반면, CAV 측정 시 0에 가까운 값을 보여(Fig. 6) 결절의 크기에 따른 백분율 환산 오차 값의 변화가 거의 없는 것으로 나타났다.

고 찰

폐결절의 성장 여부를 결정하는데 있어 3차원 CAV가 최장경을 측정하는 것보다 더 정확하다고 할 수 있을 것이다. 예를 들어 직경 5 mm 결절의 경우, 전자자를 사용한다 하더라도 직경의 1 mm 차이는 부피로 환산하면 73%의 차이를 가져온다. 이러한 오차는 5 mm 미만의 소결절에서 더 크게 나타나며 이 정도 크기의 결절은 MSCT나 폐암 조기 검진에서 자주 발견되는 크기이다.

자동 부피 계산은 2차원(2D CAV)과 3차원(3D CAV)의 두 가지 방법이 있다. 2D CAV는 축상 영상에서 결절이 구(sphere)라는 가정하에 최장경을 측정하거나 타원형이라는 가정하에 두 방향의 직경을 측정하여 단면적을 알아낸 후 외삽법(extrapolation)으로 부피를 계산할 수 있다. 3D CAV는 화소의 감쇄(pixel attenuation)를 기준으로 하는 방법으로 불규칙한 모양의 폐결절의 경우에 Z-axis의 폐결절 데이터를 포함하므로 2차원 CAV에 비해 폐결절의 성장을 더 정확하게 측정할 수 있다. Yankelevitz 등(10)은 모형실험으로, 3차원 영상 추출(3D image extraction)을 이용하여 CAD로 $\pm 3\%$

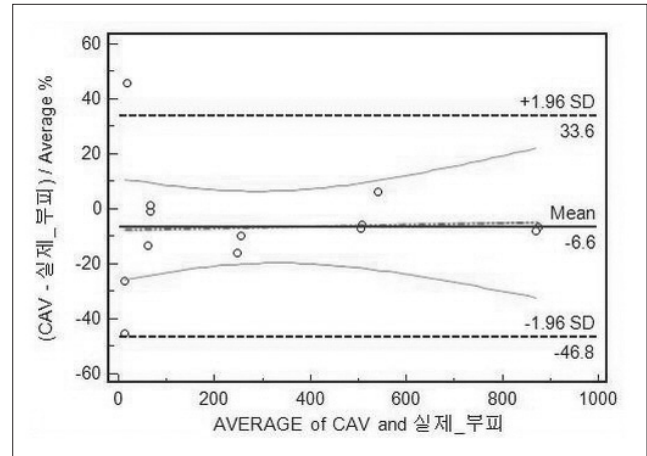


Fig. 6. The percentage of error of the volume estimates by CAV algorithm in 15 synthetic nodules. The percentage of error is nearly fixed at -6.6 % with all nodule.

의 오차를 보일 정도로 매우 높은 정확도(accuracy)를 보여주었으며 3D method가 2D method보다 더 정확하다고 보고하였다. Ko 등(1)은 합성 폐 실질을 갖는 흉부 모형에 인공 결절을 넣어 직경 약 2.5 mm 결절에서 부피오차(volume error)가 15%(1.2 mm³), 직경 약 4.9 mm 결절(부피 61 mm³)에서 부피 오차가 4%(2.1 mm³)를 보였다고 하였다. 본 연구에서 개발한 알고리즘은 3D CAV로, 직경 3 mm(부피 14.13 mm³) 인공결절은 참값과의 편차가 0.3%, 직경 5 mm(부피 65.42 mm³) 인공결절은 참값과의 편차가 4.1%를 보여 다른 연구와 비교하여 유사하거나 우수한 결과를 보였다.

폐결절의 성장 여부를의 기준은 정의상 부피를 측정하여야 하지만 실제 임상에서는 폐결절의 최장경을 측정하여 간접적으로 판정하고 있다. Jennings 등(11)은 직경 3 cm 이하의 폐결절의 성장 여부를 최장경, 단면적, 부피를 측정하여 각각의 정확도를 비교하였으며, 최장경을 사용한 경우 37%에서 부피값과 결과가 일치하지 않았으며, 54%에서 위음성을 보였다. 게다가 단면적을 기준으로 한 경우 최장경에 비해 정확도에 있어 장점이 없었다고 하였다. Tran 등(12)도 CT 상 전이성 폐결절의 직경, 단면적, 부피를 측정하여 각 측정값 사이에 일치율이 낮다(fair to poor)고 보고하였다. 본 연구에서도 각 참값과의 편차값이 직경 3 mm 결절의 경우 직경 8.2%와 부피 27.3%, 직경 5 mm 결절에서는 직경 8.1%와 부피 26.5%로 부피의 차가 직경의 차보다 크게 나타났으며, 직경 8 mm 이상의 결절보다 직경 5 mm 이하의 소결절들에서 참값과의 편차가 더 심한 것으로 나타났다.

본 연구에서 전자자를 이용하여 최장경을 측정하는 경우 다음과 같은 문제점이 발견되었다. 첫째, 캘리퍼(caliper)의 이동이 연속적이지 않고 일정 간격을 두고 움직이므로 미세한 움직임이 측정의 차이 즉 오차를 만들 수 있다. 둘째, 캘리퍼를 click하는 순간 길이가 변하기도 하였다. 셋째, 캘리퍼가 최장경에 대하여 직각으로 움직이지 않고 비스듬하게 측정될 수 있다. 게다가 현재 임상적으로 이용되고 있는 WHO criteria나 RECIST guideline은 불규칙하거나 침상인 결절은 크기나 부

피를 어떻게 측정해야 하는지 정확한 언급이 없다. 즉 불규칙한 모양의 경우, 측정자에 따라 그리고 측정할 때마다 최장경의 위치가 변할 수 있다. 또한, 침상의 결절인 경우, 침윤의 전 길이를 포함할 지 또는 종양의 충실성 부분만 측정해야 할지 정해져 있지 않다. 따라서 종양의 정확한 경계를 정하거나, 종양의 최장 직경과 그에 수직인 직경을 결정하는 것이 주관적으로 될 수밖에 없으며 관찰자 간 변이가 필연적으로 나타나게 된다(13). 본 연구에서 연구자 간 참값과의 편차는 직경 3 mm 결절의 경우, 직경은 3.7~14.3%, 부피는 11.5~49.3%, 직경 5 mm 결절의 경우, 직경은 5.2~10.4%, 부피는 16.4~34.7%로 나타났으며 5 mm 이하의 소결절에서 편차가 더 크게 보였다. 세 명의 영상의학과 의사는 모두 직경이 증가함에 따라 오차의 편차가 감소하였으나 한 명의 영상의학과 의사는 소결절에서 오차가 큰 점은 같았으나 8 mm 결절보다 10 mm와 12 mm 결절에서 오차의 편차가 더 심하여, 측정에 있어서 관찰자 간 변이를 잘 보여 주었다.

이에 반해 CAV는 오차 값의 변화가 적어 비교적 일정하게 나타났으며 참값과의 차이도 영상의학과 의사의 측정값보다 참값에 가까웠다. Das 등(14)도 CAV에 의한 폐결절 부피측정이 매우 정확하며 소결절의 성장 속도와 배가 시간(doubling time)을 결정하는데 매우 유용하다고 하였다. Revel 등(15)도 소프트웨어로 폐결절의 배가 시간을 500일을 기준으로 하여 악성 폐결절의 진단에 음성예측도가 98%라고 보고하였다.

항암제의 치료 효과 판정은 치료의 지속 여부, 약제의 교체를 결정하는 데 필요하며 크기 측정의 재현성이 중요하다. RECIST guideline에 따르면 부분 반응(Partial Response, PR)은 표적 병변의 최장경의 합이 베이스라인 최장경의 합보다 최소 30% 이상 감소한 경우이고, 병의 진행(Progressive Disease, PD)는 표적 병변들의 최장경의 합이 치료 시작 이후 기록된 가장 작은 최장경의 합에 비해 최소한 20% 이상 증가하거나, 한 개 이상의 새로운 병변이 나타나는 것이다. 본 연구에서 영상의학과 의사들의 측정값은 부피로 환산한 경우 참값과의 편차가 9.6 ~ 27.3%를 보였으며, 5 mm 이하의 소결절에서 가장 큰 편차를 보인 경우는 49.3%, 8 mm ~ 12 mm 크기의 결절에서 가장 큰 편차를 보인 경우는 25.3%의 편차를 보여 결절의 양성, 악성 감별뿐 아니라 치료 효과 판정에도 오진이 발생할 수 있음을 보여 주고 있다.

본 연구에는 몇 가지 제한점이 있다. 첫째, 소프트웨어로 인공 폐결절의 부피 계산을 하는 경우, 화소 감소 기준(threshold), 영상 두께(section thickness), 영상범위(Field Of View, FOV), 재구성 지표(reconstruction parameter), 조영 증강 여부(16-21)등의 영향을 받는다. 본 연구는 그에 따른 비교 연구가 되지 않았다. 둘째, 본 연구에 이용된 인공 결절은 모두 평탄한 경계를 갖는 구 형태의 결절로, 구형이 아닌 불규칙하거나 침상의 모양의 결절에 대한 연구나 결절의 음영에 따른 비교 연구가 되지 않았다. 셋째, 여러 연구에서 자동 부피 계산에 있어 폐결절의 분리가 영향을 미치며 특히 혈관과 붙어 있는 결절의 경우 혈관과 분리가 안되거나

혈관을 포함하여 부피를 계산하게 된다고 보고하였다(1, 22). 본 연구에서 개발된 프로그램은 혈관에 붙어 있는 폐결절 중 두 개의 결절에서 분리(segmentation)에 실패하였으며 이에 대한 알고리즘의 수정 보완이 필요하다. 넷째, 본 연구에서 사용된 CAV의 측정값은 paired t test에서 p 값을 일반적으로 적용하는 0.05를 기준으로 한다면 영상의학과 의사에 의한 측정값에 비해 정확도가 우수한 것은 사실이나, 두 방법 모두 통계학적으로 참값과 유의한 차이가 있는 것으로 나타나, 정확도를 높이기 위한 알고리즘의 개발이 더욱 필요할 것으로 생각한다. 마지막으로 window level -700, window width 1,500의 조건에서 결절의 크기 측정 시, 실제 크기보다 크게 측정되는 경향이 있어 편차가 증가할 수 있지만, 임상적으로 판독 시 보편적으로 사용하는 환경이기 때문에 그대로 적용하였다.

영상의학과 의사들이 손으로 측정한 결절의 직경과, 측정값을 이용하여 간접적으로 계산한 부피 값은 CAV보다 오차가 크며 결과가 일정하지 않아 상당한 관찰자 간 변이를 보이고 있다. 이에 반해 MSCT 데이터를 이용한 CAV는 결절의 직경이 아닌 부피를 직접 측정함으로써 정확도가 높으며, 같은 결절에 대한 측정의 재현성에 있어서도 유리하므로, 결절의 성장 속도를 이용한 양성 및 악성 결절의 감별 진단과 폐종양의 치료 반응을 평가하는데 있어 매우 유용한 검사이고 특히 직경 5 mm 이하의 소결절에서 더욱 유용할 것으로 보인다.

참 고 문 헌

1. Ko JP, Marcus R, Bomsztyk E, Babb JS, Stefanescu C, Kaur M, et al. Effect of blood vessels on measurement of nodule volume in a chest phantom. *Radiology* 2006;239:79-85
2. WHO. *WHO Handbook for Reporting results of cancer treatment*. WHO offset publication 48. Geneva, Switzerland: World Health Organization, 1979
3. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207-214
4. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein, et al. New guidelines to evaluate the response to treatment in solid tumors: european organization for research and treatment of cancer, national cancer institute of the united states, national cancer institute of canada. *J Natl Cancer Inst* 2000;92: 205-216
5. Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS, et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol* 2003;21:2574-2582
6. Marten K, Engelke C. Computer-aided detection and automated CT volumetry of pulmonary nodules. *Eur Radiol* 2007;17:888-901
7. Bolte H, Riede C, Müller-Hülsberg S, Freitag-Wolf S, Kohl G, Drews T, et al. Precision of computer-aided volumetry of artificial small solid pulmonary nodules in ex vivo porcine lungs. *Br J Radiol* 2007;80:414-421
8. Wormanns D, Kohl G, Klotz E, Marheine A, Beyer F, Heindel W, et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. *Eur Radiol* 2004;14:86-92
9. Bolte H, Riedel C, Jahnke T, Inan N, Freitag S, Kohl G, et al. Reproducibility of computer-aided volumetry of artificial small

- pulmonary nodules in ex vivo porcine lungs. *Invest Radiol* 2006;41:28-35
10. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology* 2000;217:251-256
 11. Jennings SG, Winer-Muram HT, Tarver RD, Farber MO. Lung tumor growth: assessment with CT-comparison of diameter and cross-sectional area with volume measurements. *Radiology* 2004;231:866-871
 12. Tran LN, Brown MS, Goldin JG, Yan X, Pais RC, McNitt-Gray MF, et al. Comparisons of treatment response classification between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. *Acad Radiol* 2004;11:1355-1360
 13. Schwartz LH, Ginsberg MS, DeCorato D, Rothenberg LN, Einstein S, Kijewski P, et al. Evaluation of tumor measurements in oncology: use of film-based and electronic techniques. *J Clin Oncol* 2000;18:2179-2184
 14. Das M, Mühlenbruch G, Katoh M, Bakai A, SalganiCoff M, Stanzel S, et al. Automated volumetry of solid pulmonary nodules in a phantom: accuracy across different CT scanner technologies. *Invest Radiol* 2007;42:297-302
 15. Revel MP, Merlin A, Peyrard S, Triki R, Couchon S, Chatellier G, et al. Software volumetric evaluation of doubling times for differentiating benign versus malignant pulmonary nodules. *AJR Am J Roentgenol* 2006;187:135-142
 16. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF, Bae KT. Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. *Radiology* 2005;235:850-856
 17. Harris KM, Adams H, Lloyd DC, Harvey DJ. The effect on apparent size of simulated pulmonary nodules of using three standard CT window settings. *Clin Radiol* 1993;47:241-244
 18. Van Hoe L, Haven F, Bellon E, Baert AL, Bosmans H, Feron M, et al. Factors influencing the accuracy of volume measurements in spiral CT: a phantom study. *J Comput Assist Tomogr* 1997;21:332-338
 19. Honda O, Sumikawa H, Johkoh T, Tomiyama N, Mihara N, Inoue A, et al. Computer-assisted lung nodule volumetry from multi-detector row CT: influence of image reconstruction parameters. *Eur J Radiol* 2007;62:106-113
 20. Petrou M, Quint LE, Nan B, Baker LH. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *AJR Am J Roentgenol* 2007;188:306-312
 21. Honda O, Johkoh T, Sumikawa H, Inoue A, Tomiyama N, Mihara N, et al. Pulmonary nodules: 3D volumetric measurement with multidetector CT-effect of intravenous contrast medium. *Radiology* 2007;45:881-887
 22. Gietema HA, Schaefer-Prokop CM, Mali WP, Groenewegen G, Prokop M. Pulmonary nodules: interscan variability of semiautomated volume measurements with multisection CT-influence of inspiration level, nodule size, and segmentation performance. *Radiology* 2007;245:888-894

Volumetry of Artificial Pulmonary Nodules in *Ex Vivo* Porcine Lungs: Comparison of Semi-automated Volumetry and Radiologists' Performance¹

Ju-Hyeon Jeon, M.D., Jin-Hwan Kim, M.D., Song Soo Kim, M.D., Hosang Jeon, Ph.D.²,
Hyun Ju Lee, M.D.³, Noh Hyuck Park, M.D.⁴, Gyuseong Cho, Ph.D.⁵

¹Department of Radiology, Chungnam National University Hospital

²Department of Therapeutic Radiology and Oncology, Pusan National University Yangsan Hospital

³Department of Radiology, Seoul National University Hospital

⁴Department of Radiology, Kwandong University College of Medicine, Myungji Hospital

⁵Department of Nuclear & Quantum Engineering, KAIST

Purpose: With the advent of MSCT, the detection rate of small pulmonary nodules is markedly greater. However, there is no definite diagnostic clue to differentiate between malignant and benign nodules, except for the interval growth in small nodule less than 1 cm in diameter. We evaluated the accuracy of computer-aided volumetry (CAV) and compared it with 4 radiologists' measurement.

Materials and Methods: Fifteen artificial nodules that were embedded in the ex vivo porcine lung were scanned by MSCT. The diameters and volumes of nodules were independently measured three times, at 5-day intervals, and by four radiologists as well as by CAV. We evaluated the accuracy of the measurements on the basis of the true diameter and volume of the nodules. Using a paired t-test and a Bland-Altman plot, we evaluated whether there was a statistically significant difference between the radiologists' measurements and the CAV.

Results: The accuracy of the manual measurements by radiologists revealed a statistically significant difference from the true diameter and volume of the artificial nodules ($p < 0.01$). Conversely, the accuracy of CAV did not show a statistically significant difference with the true nodule diameter and volume ($p > 0.01$).

Conclusion: The results of this study suggest that CAV is an accurate and useful tool to evaluate the volume of pulmonary nodules and can eventually be used to differentiate malignant and benign nodules as well as evaluate the therapeutic response of lung cancer.

Index words : Lung

Solitary Pulmonary Nodule

Lung Volume Measurements

Computer Aided Diagnosis

Address reprint requests to : Jin-Hwan Kim, M.D., Department of Radiology, Chungnam National University Hospital,
640 Daesa-dong, Jung-gu, Daejeon 301-721, Korea.
Tel. 82-42-280-7333 Fax. 82-42-253-0061 E-mail: michelan@cnu.ac.kr