

조기 알츠하이머 치매의 뇌포도당 대사 감소 평가에서 오류발견률 조절법의 적용

서울대학교 의과대학 핵의학교실,¹ 신경정신과학교실,² 성균관대학교 인지과학전공³

이동수,¹ 강혜진,¹ 장명진,¹ 조상수,^{1,3} 강원준,¹ 이재성,¹ 강은주,¹ 이강욱,² 우종인,² 이명철¹

Application of False Discovery Rate Control in the Assessment of Decrease of FDG Uptake in Early Alzheimer Dementia

Dong Soo Lee, M.D.,¹ Hyejin Kang, M.S.,¹ Myung Jin Jang, M.S.,¹ Sang Soo Cho, M.S.,^{1,3}
Won Jun Kang M.D., Jae Sung Lee, Ph.D.,¹ Eunjoo Kang, Ph.D.,¹ Kang Uk Lee, M.D.,²
Jong In Woo, M.D.,² Myung Chul Lee, M.D.¹

Departments of Nuclear Medicine,¹ and Neuropsychiatry,² Seoul National University, and Interdisciplinary Program in Cognitive Science,³ Sungkyunkwan University, Seoul, Korea

Abstract

Purpose: Determining an appropriate thresholding is crucial for PDG PET analysis since strong control of Type I error could fail to find pathological differences between early Alzheimer' disease (AD) patients and healthy normal controls. We compared the SPM results on FDG PET imaging of early AD using uncorrected p-value, random-field based corrected p-value and false discovery rate (FDR) control. **Materials and Methods:** Twenty-eight patients (66±7 years old) with early AD and 18 age-matched normal controls (68±6 years old) underwent FDG brain PET. To identify brain regions with hypo-metabolism in group or individual patient compared to normal controls, group images or each patient's image was compared with normal controls using the same fixed p-value of 0.001 on uncorrected thresholding, random-field based corrected thresholding and FDR control. **Results:** The number of hypo-metabolic voxels was smallest in corrected p-value method, largest in uncorrected p-value method and intermediate in FDG thresholding in group analysis. Three types of result pattern were found. The first was that corrected p-value did not yield any voxel positive but FDR gave a few significantly hypometabolic voxels (8/28, 29%). The second was that both corrected p-value and FDR did not yield any positive region but numerous positive voxels were found with the threshold of uncorrected p-values (6/28, 21%). The last was that FDR was detected as many positive voxels as uncorrected p-value method (14/28, 50%). **Conclusions** FDR control could identify hypo-metabolic areas in group or individual patients with early AD. We recommend FDR control instead of uncorrected or random-field corrected thresholding method to find the areas showing hypometabolism especially in small group or individual analysis of FDG PET.

Key words: false discovery rate, Alzheimer dementia, FDG uptake, statistical parametric mapping

Received October. 11, 2003; accepted December. 16, 2003

Corresponding Author: Dong Soo Lee, M.D.

Department of Nuclear Medicine

Seoul National University College of Medicine

28 Yungun-Dong, Jongno-Gu, Seoul 110-744, Korea

Tel: 02-760-2501, Fax : 02-745-7690, E-mail: dsl@plaza.snu.ac.kr

서 론

뇌영상을 통계적 파라미터 매핑(Statistical Parametric Mapping; SPM)으로 분석할 때 통계적 추론방법으로 t검정을 한다.^{1,2)} 비교대상인 각 집단의 각 화소별 방사능값이나 MR 신호강도의 분포를 평가하여 t값으로 전환하고 역치를 정하여 역치 이상인 곳을 활성화된 화소 또는 이상부위라고 한다. 대개 화소의 수가 만개를 넘고, 이렇게 많은 수의 화소를 동시에 비교하면 반드시 위양성에 해당하는 화소가 드러나게 된다. 이상부위라고 찾은 화소 중에 어느 부분이 진짜 이상이 있는 곳인지 알 수 없다. 다중 비교에 따른 위양성의 증가를 해결하기 위해 계통적인 오류율(family-wise error rate)을 조절하는 방법이 도입되었다. Bonferroni 보정 방법은 그 중 가장 엄격하게 위양성을 조절하여 다중 비교 중 어느 것이든 위양성률이 지정한 값(α)을 넘지 않도록 한다. 또는 화소로 이루어진 데이터 입체가 공간에서 무작위 가우스장(random Gaussian field)을 이룬다고 가정하고 값의 크기와 범위가 벗어나는 부위를 찾기도 한다.^{1,3)}

그러나 이런 방법을 동원하면 통계검정의 검정력(power)이 떨어지므로 이상부위를 이상부위로 찾지 못하는 단점이 있다. 무작위 가우스장을 기준으로 이상 화소를 찾는 방법이 Bonferroni 보정방법보다는 덜 엄격하여 SPM에서는 보정 p 값의 디폴트로 쓰여왔다. 그러나 Bonferroni 방법이나 무작위 가우스장 보정 방법이나 모두 특히 검사 횟수가 많지 않거나 개인차가 큰 데이터로 통계비교할 때 검정력이 낮은 것이 문제이다.

최근 오류발견률 조절법이라는 사후(post-hoc) 분석방법이 대안으로 제안되었다.^{4,5)} 다중비교 후 얻은 p값을 크기순으로 나열하고 오류발견률이 지정한 기준보다 작은 화소를 고르는 방법으로서 원리상 제1형 오차율이 같을 때 계통적인 오류율을 조절하는 어떤 방법보다 통계 검정력이 높다. 다만 Genovese 등의 보고⁵⁾에 인용한 기능적 자기공명영상(functional magnetic resonance imaging; fMRI) 결과는 Bonferroni 보정 방법을 적용하였을 때보다 더

많은 활성부위를 찾지 못하였다.

이 연구에서는 오류발견률(false discovery rate; FDR) 조절법을 적용하여 F-18 fluorodeoxyglucose (FDG) 양전자단층촬영(positron emission tomography; PET) 영상을 SPM으로 분석하였을 때 다중비교 비보정(non-corrected) 방법과 무작위 가우스장 바탕의 보정 p값 적용 때와 비교하여 어떤 결과가 나타나는지 조사하였다. 대상 환자로 첫째, MR로 진단하기 어렵고 FDG PET에도 특징적인 소견을 찾지 어렵다고 알려진^{6,7)} 조기(early) 치매환자의 PET을 비교하는 과제를 선정하되, 둘째, 환자집단과 정상인집단 뿐 아니라 환자개인과 정상인집단을 비교하는 과제를 선택하여 자유도를 낮추고 낮은 자유도 때문에 검정력이 위협받도록 하여, 셋째, 오류발견률 조절방법이 SPM의 다중비교 보정(corrected)방법보다 얼마나 더 예민하게 이상 화소를 찾아낼 수 있는지 조사하였다.

대상 및 방법

연구대상

28명의 평균 66세(± 7)인 조기 알츠하이머 치매 환자(남자 15 여자 13)와 나이를 맞춘 18명(남자 10 여자 8)의 정상인(68 ± 6 세)을 대상으로 하였다. 정신과적 면담과 NINCDS/ADRDA의 알츠하이머 치매의 진단기준에 따라 조기 알츠하이머 치매여부를 진단하고 한국형 간이정신상태검사(Mini-Mental State Examination, 한국어판: MMSE-K)를 실시하였다. 환자의 MMSE-K 점수는 16.1 ± 4.1 , 정상대조군의 MMSE-K 점수는 27.8 ± 1.8 이었다.

FDG PET 영상촬영

ECAT EXACT 47 (Simence-CT1, Knoxville, USA) PET scanner (BGO 결정, 공간분해능= 6.1 mm, 축방향분해능= 4.3 mm 예민도= 214 kcps/ μ Ci/mi)를 사용하여 Ge-68 선원을 이용한 5분간의 투과 촬영을 실시하고 370MBq (10mCi)의 F-18 FDG를 정맥 주사하고 30분 후부터 25분간 방출 영상을 촬영하

였다. 방출 영상을 Shepp-Logan 필터를 이용하여 역투사 방식(차단 주파수 0.35)으로 재구성하여 투과 스캔에서 얻은 감쇠지도를 이용하여 감쇠 보정하였다. 얻은 영상의 각 화소의 크기는 2.1×2.1×3.4 mm, 영상 매트릭스의 크기는 128×128이었다.

PET 영상분석

모든 FDG PET 영상은 Analyze 형식으로 변환하여 MATLAB version 6 (Mathworks, Inc., MA, USA) 환경에서 사용하는 SPM99 소프트웨어 (Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK)로 영상 전처리와 통계분석을 하였다. 공간정규화는 MNI (Montreal Neurological Institute) 표준관에, 계수정규화는 전체 뇌 대신에 소뇌를 기준으로 삼았다. 개인과 정상인 집단 사이에 차이여부를 2 표본 비교 검정을 하였다.

통계추론시 1) 보정 안 한 p값 0.001보다 작은 화소, 2) 무작위 가우시안 보정한 p값 0.001보다 작은 화소, 3) 오류발견률 p값이 0.001보다 작은 화소를 각각 찾았다.

오류 발견을 조절법

오류발견률은, 발견된 이상 화소의 개수를 V라고 하고 그 중 실제로 이상이 있는 화소의 개수를 V_a , 실제로는 이상이 없는 화소의 개수를 V_i 라고

할 때 $FDR = \frac{V_i}{V_a + V_i} = \frac{V_i}{V}$ 이다. 만약

FDR 값이 0이라면 실제로 이상이 없는데 이상이 있다고 나타난 화소가 없다. 오류발견을 조절법에서는 가설검증을 하기 전에 오류발견율을 0.05로 한다면 정하고 각 화소의 값의 분포를 보아서 두 집단 간에 차이가 날 확률 p를 정한 후 이 확률 p를 크기 순으로 나열하여 결국 그 오류발견율일 때의 FDR(q)를 정한다. 이렇게 정한 q를 사용하면 수천 수 만개에 달하는 전체 화소 중 두 집단간 차이가 없다는 귀무가설이 기각된 화소 중에서 일정 비율이 위양성이다. 즉 찾은 V의 q곱($V \times q$) 만큼은 ‘평균

적으로’ 실제로 이상이 없음에도 불구하고 이상이 있다고 잘못 판단된 셈이 된다. 두 집단간 차이가 없다는 귀무가설을 기각한(즉 차이 있다고 본) 화소 중에 q비율만큼의 잘못 기각한(차이 없는데 차이 있다고 한) 화소가 있으므로 이를 오류발견률로 본다. 오류발견을 조절법은 잘못 발견한 화소 즉 위양성 화소를 미리 정한 만큼 허용하여 Bonferroni 보정이 모든 화소에 대해 위양성율을 정한 값보다 작도록 보정하는 것보다 덜 엄격한 기준을 쓰게 된다.^{4,5)}

오류발견을 조절법에서 귀무가설 기각을 결정하는 기준인 역치는 미리 지정하지 않고 화소의 값 분포로부터 구한 p값에 근거해서 정한다. 예를 들어 FDR 값 0.05로 일단 역치를 정하면 이 역치보다 작은 p값을 가진 화소들은 모두 귀무가설이 기각되고 따라서 차이(이상)가 있다고 본다. 요약하면

1. 우선 연구자가 0과 1사이의 FDR값을 정한다(예 =0.05).
2. 정상인과 환자 영상에서 N개의 화소에 대한 p값을 구하고 큰 값에서 작은 값으로 배열한다.

$$P(N) \geq P(N-1) \geq P(N-3) \geq \dots \geq P(1)$$
3. 제일 큰 p값인 N번째 화소부터 제일 작은 p값인 마지막 하나의 화소까지 유의수준 D_i 값을 차례로 계산한다.

$$i=N, N-1, N-2, \dots, 1$$

$$D_i = (i/N) \times q$$
4. 두 쌍의 값($P(i), D_i$)을 비교해서 $P(i) \leq D_i$ 의 관계가 성립하면 그 $P(i)$ 를 FDR 역치로 정한다.
5. $P(i)$ 보다 작거나 같은 p값을 가지는 화소들은 모두 귀무가설을 기각하므로 이상이 있는 화소이다.

즉 화소가 100개이고 각 화소에서 얻은 영상의 값의 분포로 보아 두 집단간 차이 날 확률이 임의로 Table 1과 같이 0.436, 0.429, 0.341 ... 0.013, 0.0006, 0.0003 같은 분포를 보이면 0.436은 $P(N)$ 에 해당하므로 D_N 값이 $(100/100) \times 0.05$ 이고 $P(N) \geq D_N$, 0.429는 $P(N-1)$ 에 해당하므로 D_{N-1} 값이 $(99/100) \times 0.05$ 이고 $P(N-1) \geq D_{N-1}$, 이와 같이 진행하면 0.0041은 $P(9)$ 에 해당하므로 D_9 값이 $(9/100) \times 0.05 = 0.0045$ 라서 이제 $P(9) < D_9$ 이라 9번

Table 1. Distribution of Probability of Each Voxel to Reject Null Hypothesis

Ho*	P(i)	Di	
N	0.436	0.0500	
N-1	0.429	0.0495	
N-2	0.341	0.0490	
N-3	0.232	0.0485	
N-4	0.206	0.0480	
:	:		
:	:		
10	0.0052	0.0050	
9	0.0041	0.0045	rejection
8	0.0038	0.0040	↓
7	0.0031	0.0035	
6	0.0024	0.0030	
5	0.0021	0.0025	
4	0.0016	0.0020	
3	0.0013	0.0015	
2	0.0006	0.0010	
1	0.0003	0.0005	

* Null hypothesis that the voxel is from the same population. P(i), Di, N. See texts.

부터 1번까지 9개의 화소가 이상이 있는 화소임을 발견한다. 이 9개의 화소 중에 평균적으로 5%가 위양성 화소이다. 이 때 만일 Bonferroni 보정을 하면 보정한 p값이 0.0005이므로 1번 화소만 이상이 있다고 발견된다.

만일 P(N)이 0.05보다 작으면 이 100개의 대상 화소는 모두 유의하게 이상이 있다. 만일 P(1)이 0.0005보다 크면 100개의 대상 화소 중 어느 것도 이상이 없다. 실제 P(i)의 분포는 비선형적이고 P(N)은 대개는 0.05보다 훨씬 큰 값이므로 단조롭게 증가하는 선과 0.05와 원점을 있는 선 사이의 어느 점에 해당하는 화소 i를 찾을 수 있고 이 i는 이미 크기 순으로 줄 새운 것이므로 1부터 i 까지가 이상이 있는 화소임을 나타내 준다.

결 과

1. 집단 분석 결과

조기 알츠하이머 치매 환자는 정상인과 집단비교 결과 세 기준 모두에서 양측 후 대상회와 시상 및 측두엽과 두정엽 영역의 FDG 섭취가 감소하였다. FDG 섭취감소 범위는 비보정 p값이 0.001인 역치일 때 가장 넓고, 오류발견율 0.001로 조절하였을 때 중간, 무작위 가우스장 보정 p값이 0.001일 때 가장 좁았다(Fig 1).

비보정 p값이 역치일때와 오류발견율 조절 기준일 때 후 대상회, 양측 시상 뿐 아니라 광범위한 전두엽 영역과 해마방회, 좌측 미상핵, 섬회와 양측 측두엽에 FDG 섭취가 공통적으로 감소하였다. 보정 p값을 기준 삼았을 때는 전두엽과 미상핵이나 섬회 (insular gyrus)의 FDG 섭취는 감소하지 않았고 측

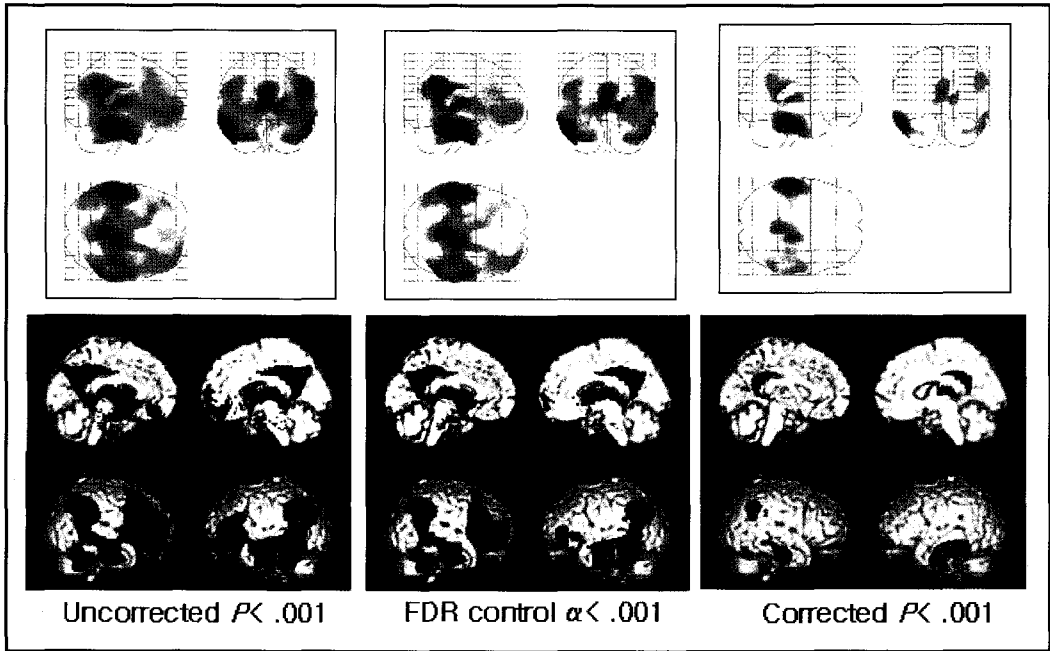


Fig 1. Results of group analysis; Hypometabolic brain areas in early Alzheimer's disease compare with normal control at three different thresholds.

두엽의 감소 범위도 작았다. 비보정 p값이 0.001인 역치일 때는 오류발견을 조절 기준일 때보다 FDG 감소부위가 같은 부위이나 범위가 조금 넓을 뿐 아니라 전대상회가 더 나타났다.

2. 개인 분석 결과

집단분석에서와 마찬가지로 세 가지 기준을 적용하였을 때 무작위 가우스장 보정 p값 때 가장 적은 수의 화소가, 오류발견을 조절시 중간정도, 비보정 p값 역치일 때 가장 많은 수의 화소가 발견되었다.

개인 분석 결과에 나타난 FDG 섭취감소 화소 수를 비교하여 3가지 경향이 있었다. 첫 집단은 오류발견을 조절 때 보정 p값을 적용하였을 때보다 많은 수의 화소가 발견되었으나, 비보정때보다는 적은 수의 화소가 발견된 집단은 8 예로 29%이었다(Fig 2A). 이 집단은 비보정 p값이 기준일 때 위의 두 방법에 비하여 훨씬 넓은 영역에서 FDG 섭취가 감소하였다. 두 번째 집단은 보정 p값을 적용하였을 때와 오류발견을 조절시 아무 곳도 나타나지 않은 집

단으로 비보정 p값 기준일 때만 비교적 넓은 부위의 FDG 섭취가 감소한 집단으로 6예로 전체 21%이었다(Fig 2B). 세 번째 집단은 오류발견을 조절시 비보정 p값을 적용하였을 때만큼 FDG 섭취감소부위가 발견된 집단으로 14예 즉 전체의 50%이었다(Fig. 2C). 그림 2에 위의 세 집단에 속하는 환자 중 대표적인 예를 각각 도시하였다.

고 찰

뇌기능 매핑시 통계분석방법으로 다중비교를 고려하여 무작위 가우스장 보정 방법을 흔히 쓴다. 그러나 PET이나 SPECT 영상으로 정상군과 환자군의 대사나 관류 차이를 찾을 때는 표본 숫자가 적어 무작위 가우스장보정 역치를 적용하면 이상부위를 찾지 못하는 경우가 흔하다.²⁾ 이럴 때 실제 차이가 없는 것인지, 무작위 가우스장 보정 방법이 검정력이 떨어지기 때문인지 알 수 없다. 개인과 정상인 집단을 비교하는 분석에서는 자유도가 더 적어지기 때

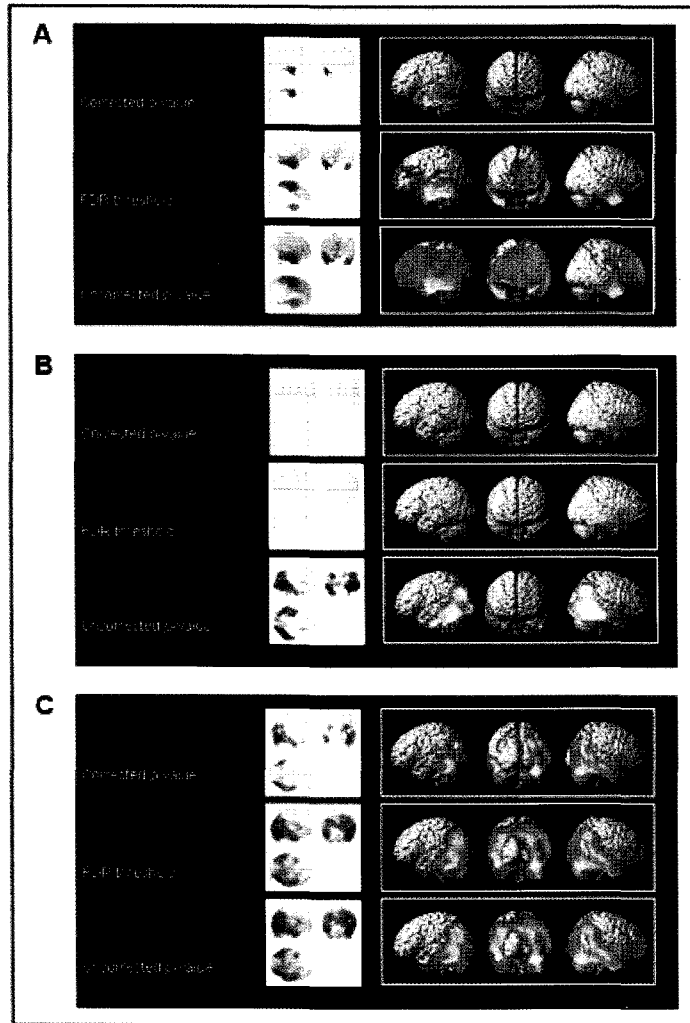


Fig 2. There are three types of results. (A) Corrected p-value threshold showed a few positive voxels but FDR (false discovery rate) control gave properly many and significant voxels (8/28, 29%). (B) Both corrected p-value threshold and FDR control did not yield any positive region but some positive voxels were found in uncorrected p-value threshold (6/28, 21%). (C) FDR control detected as many positive voxels as uncorrected p-value threshold (14/28, 50%).

문에 이런 어려움이 더욱 크다.

가우스 무작위장 보정 방법은 엄밀한 수학적 기초위에 도입한 기준으로 다중비교 통계추론방법 중에 가장 엄격한 방법인 Bonferroni보정 방법에 비해 검정력이 높다고 하였다.³⁾ 계통적 오류조절법 중에서 가장 검정력이 높을 것이라 제안된 오류발견을

조절법은 시뮬레이션에서는 Bonferroni 검정보다 검정력이 높았지만 실제 fMRI 데이터분석에서는 차이가 발견되지 않았다.⁵⁾ 우리는 이 연구에 든 예에서 분명한 차이를 확인하였다. 특히 무작위 가우스장 보정방법보다 오류발견을 조절법에 여러 의미있는 차이가 드러났다. 개인분석결과 전체의 반수에서

오류발견을 조절법에 가우스장 보정방법보다 많은 곳에서 이상이 나타났다. 원리상 적절한 수의 이상 화소를 찾는다고 알려진 오류발견을 조절법을 옳게 적용하였을 경우 PET 에 대한 집단이나 개인 분석에서, 같은 관찰로부터 더 많은 차이를 찾을 수 있으니 오류발견을 조절법을 쓰는 것이 좋을 것으로 생각된다.

오류발견을 조절법으로 발견된 이상 화소는 어떤 뜻을 지니는가? 오류발견을 조절법을 처음 제안한 Benjamini와 Hochberg⁴⁾는 새 치료방법의 효과를 보는 지표가 십 수개일 때 다중비교만큼 Bonferroni 보정하여 귀무가설을 모두 기각하지 못하여 치료방법의 효과가 있는데 없다고 잘못 나올 경우를 줄이기 위해 오류발견을 조절법을 쓰자고 하였다. 이때 오류발견을 조절의 결과로 하나나 두개의 지표에 대한 귀무가설이 기각되면 전체적으로 새 치료법의 효과가 인정되고 그 효과는 그 지표로 측정 가능했다고 해석하였다. 단 치료효과를 나타낸 지표가 위양성일 확률이 오류발견을 만큼이라 해석된다. Genovese 등⁵⁾은 가설이 화소 수만큼인 뇌매핑분석에 오류발견을 조절법을 도입하였고, 오류발견을 조절법은 역시 유전자 발현분석 즉 마이크로어레이 해석에도 쓰인다. 오류발견율은 발견한 이상 화소중에 '평균적으로' 오류발견율(예 0.05이면 5%) 만큼의 위양성 화소, 즉 이상이 없는데 이상이 있는 화소가 포함되었음을 뜻한다.

오류발견율은 발견된 화소 중 '평균적으로' 얼마만큼의 위양성 오류를 포함하고 있는가를 뜻한다.⁵⁾ 따라서 오류발견율보다 조금 더 많은 위양성을 포함할 수도 있고 적을 수도 있다. 그러나 전체화소중에 참 이상부위 화소가 차지하는 비율이 큰 경우는 위양성 비율이 오류발견율로 정한 값보다 훨씬 작을 수도 있다. 이런 점에서 오류발견율은 엄격한 기준에 속한다. 몇 안되는 부위의 활성화를 찾는 뇌 활성화 검사와 달리, 치매같은 질환에서 넓은 부위에 대사나 관류 이상을 보이는 FDG PET이나 SPECT 관류 영상을 대상으로 오류발견을 조절법으로 찾은 화소는 대부분 참이라고 보아도 좋다.

오류발견을 조절법을 적용하였을 때 위음성율이 이상 화소가 많을수록 급격히 줄어든다. 즉 이상이

있는데 못 찾을 가능성이 줄어든다. 무작위장보정법에 비하여 위음성율이 낮은 오류발견을 조절법 분석결과 아무 이상 화소가 발견되지 않았을 경우 정말 이상소견이 없다고 보아도 좋다. 따라서 우리가 제시한 개인 분석결과 오류발견을 조절법에 2/3가량은 이상소견이 발견되었는데 아무런 소견을 나타내지 않은 6예 환자는 정말 정상군과 차이가 없다고 보아도 좋다.

오류발견을 조절법은 찾은 이상 화소의 일정부분이 위양성이라는 사실을 함축하고, 위음성율이 Bonferroni 보정보다 적고 특히 이 연구의 결과에서 보듯 무작위장 보정보다도 적다는 점에서 장점이 있다. 따라서 PET 또는 SPECT 결과해석에 특히 개인분석에 쓰는 것이 좋을 것으로 생각된다. 오류발견을 조절법의 검정력은 더 높일 수 있다. 오류발견율의 분포상 평균값이 지정값(예 0.05)보다 작다고 정한 것은 실제로는 0.05에 정상화소/전체화소 비만큼의 곱보다 작은 것의 바깥 역치를 넓혀 잡았기 때문이다.^{4,8)} 따라서 정상화소/전체화소의 비를 추정하여 그만큼 보정하면 얻은 오류발견율이 0.05보다 실제는 작은 값이었음을 알릴 수 있거나 또는 0.05로 정하려 하면 실제 오류발견율은 0.05를 정상화소/전체화소 비로 나눈 값만큼으로 키울 수 있다.

오류발견을 조절법은 일종의 사후(post-hoc) 분석법이다.⁴⁾ SPM에서 우선 대상집단 사이에 각 화소별 t값을 구하여 t지도를 만들고 t값에 상응하는 p값을 줄 세우는 작업을 한다. 이 부분이 데이터기반(data-driven) 또는 적응적(adaptive)인 부분이다. 비교값인 D_i 를 각 화소가 독립적이라고 가정하였을 때는 $\sum_{i=1}^N \left\langle \frac{1}{i} \right\rangle$ 로 나누어 보정하여야 한다.⁵⁾ 이 보정값은 1보다 크므로, D_i 값이 이 연구에서 적용한 값보다 작아지고 적은 수의 화소를 찾게 된다. 이렇게 D_i 값이 이 연구에 적용한 값보다 작아지지만 실제 PET이나 뇌영상데이터에 오류발견을 조절법을 적용할 때는 화소들이 서로 독립적이지 않고 양의 상관(non-negative correlation)을 가지며, 대상영상을 편평화하여 이 상관관계가 증가하게 되므로 조금 더 낮은 기준인 1을 사용하여도 좋다고 한다.⁵⁾

이 연구에서는 Genovese 등⁵⁾이 뇌영상해석에 오

류발견을 조절법을 적용하는 방법을 상세히 설명한 것을 PET 영상 해석에 어떻게 적용하는지 제시하고 다시 풀어 설명하였다. 오류발견을 조절법은 SPM99에는 제공되지 않았으나 SPM2에서는 서브루틴으로 선택하여 적용할 수 있고, 발견한 이상 화소의 평균 몇 %가 위양성이라고 제시하면 되므로 그 결과 제시가 쉬우며, 부작용장 보정이나 Bonferroni 방법보다 상대적으로 검정력이 높으므로 PET이나 SPECT 영상을 분석하는데 사용하는 것을 권장한다.

요 약

목적: 오류발견을 조절법을 PET 영상분석에 이용하면 다중비교에 따르는 위양성을 줄이면서 동시에 검정력을 높일 수 있다. 조기 알츠하이머 치매 환자에서 오류발견을 조절법을 적용하였을 때와 비보정역치, 무작위 가우스장 보정역치를 적용하였을 때 FDG PET에 나타난 포도당 대사 감소영역이 어떻게 달라지는지 조사하였다. **방법:** 28명의 평균 66세 (± 7)인 조기 알츠하이머성 치매 환자와 연령을 matched 18명의 정상인(68 ± 6 세)의 FDG PET 영상을 SPM99 소프트웨어로 분석하였다. 환자군과 정상군의 차이와 각 환자와 정상군의 차이를 각각 비보정역치 $p < 0.001$, 무작위장 보정역치 $p < 0.001$, 오류발견을 조절법에 의한 오류발견을 0.001일 때를 정하여 이 세 통계적 역치에서 각각 뇌 포도당 대사감소 영역을 결정하였다. **결과:** 집단 분석결과 비보정역치를 사용하였을 때 가장 넓은 영역에서, 보정역치를 사용하였을 때 가장 좁은 영역에서, 오류발견을 조절법을 적용하였을 때 중간크기의 영역에 대사가 감소하였다. 개인분석결과 비보정역치 경우 발견된 대사감소 화소보다 오류발견을 조절시 많은 화소가 나타난 경우(8/28, 29%)와 보정역치 경우와 오류발견을 조절시에는 대사감소 부위가 나타나지 않고 비보정역치 경우에만 넓은 부위에 대사감소부위가 나타난 경우(6/28, 21%), 그리고 보정역치보다 오류발견을 조절시에 훨씬 많은 화소가 이상부위로 나타나서 비보정역치 경우에 근접하는 넓이를 찾을 수 있는 경우(14/28 50%)이었다. **결론:** 조기 알츠하

이머 치매 환자의 FDG PET을 오류발견을 조절법으로 분석한 결과 집단분석이나 개인분석 모두 대사감소부위를 잘 찾을 수 있었다. 집단의 크기가 작은 환자의 집단분석이나 특히 개인분석의 경우 오류발견을 조절법을 이용하여 FDG PET을 분석하는 것이 좋을 것이라 제안한다.

참고 문헌

- 1) Lee DS, Lee JS, Kim KM, Chung J-K, Lee MC. Functional brain mapping using $H_2^{15}O$ positron emission tomography (1); statistical parametric mapping method. *Korean J Nucl Med.* 1998;32:225-237.
- 2) Kim YK, Lee DS, Lee SK, Chung CK, Chung JK, Lee MC. ^{18}F -FDG PET in localization of frontal lobe epilepsy: comparison of visual and SPM analysis. *J Nucl Med.* 2002;43:1167-74.
- 3) Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Map.* 1996;4:58-73.
- 4) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B.* 1995;57: 289-300.
- 5) Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage.* 2002;15: 870-8.
- 6) Herholz K. FDG PET and differential diagnosis of dementia. *Alzheimer Dis Assoc Disord.* 1995;9: 6-16.
- 7) Sakamoto S, Ishii K, Sasaki M, Hosaka K, Mori T, Matsui M et al. Differences in cerebral metabolic impairment between early and late onset types of Alzheimer's disease. *J Neurol Sci.* 2002;200:27-32.
- 8) Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics.* 2003; 19:368-75.