

역학		번호: J - B - 22			
제 목	국문	Protein Array를 통한 조직 특이적 암 생체지표 개발을 위한 군집 분석 적용			
	영문	Application of Cluster Analysis to Develop Organ Specific Cancer Biomarkers in Protein Array			
저 자 및 소 속	국문	최윤희, 이민종 ¹⁾ , 김우호 ¹⁾ , 유근영, 양미희 서울대학교 의과대학 예방의학교실, 1) 서울대학교 의과대학 병리학교실			
	영문	Yunhee Choi, Minjong Lee ¹⁾ , Wooho Kim ¹⁾ , Keun-Young Yoo, Mihi Yang Department of Preventive Medicine, Seoul National University College of medicine, 1) Department of Pathology, Seoul National University College of medicine			
분 야	역 학	발 표 자		발표형식	포스터
진행상황	연구중 → 완료예정시기: 2003년 08월				
<p>연구목적 Protein Array를 이용하여 생산되는 다변량 데이터를 활용하여, 조직 특이적 암 생체지표를 개발하여 예방 및 치료의 효율성을 높이는데 기여하기 위하여 적절한 통계 기법인 군집분석법을 개발, 적용한다.</p> <p>연구방법 Protein Array: 각 장기별 array slide에 약 60명의 암 환자 (총 537명: 간암, 60 명; 위암, 55 명 등)조직에서 채집한 약 60 tissue cores (2.0 mm in diameter)를 정착하여 4 개의mucin (Muc1, Muc2, Muc5AC, Muc6)과 7 개의cytokeratin (CK7, CK8, CK13, CK14, CK18, CK19, CK20)으로 모두 11 개의 생체지표에 대하여 monoclonal antibody를 이용, immunohistochemistry을 실시하였다. 염색정도에 따른 단백질발현정도는 현미경 관찰로 정량화하였다. 통계분석: 11개의 생체지표의 발현양상을 조합하였을 때, 각 조합이 하나의 장기를 특정 지을 수 있는지 알아 보기 위하여 군집분석을 하였다. 데이터수가 다양하므로 계층적 방법보다는 K-means를 사용하였고, 적절한 군집 수를CCC (Cubic clustering criterion) 과 Pseudo F 통계량을 동시에 고려하여 선택하였다. 또한, 군집분석의 효율성을 높이기 위해 11개의 생체지표에 대한 Variable Selection 이 Pseudo F 통계량을 기준으로 시행되었다.</p> <p>연구결과 및 고찰 Pseudo F 와 CCC 값이 군집수가 4, 6, 8 일 때 국소적 최대값을 가졌다. 이 세가지 가능한 군집 수 중 군집에 대한 가장 효과적인 정보를 주는 군집수는 6이었다. 생체지표 중 variable selection을 통해, Muc2, CK8, CK18이 최종 분석에서 제외되었다. 그 결과, 췌장 암의 경우, Muc1, CK7, CK13, CK19발현이 높고 Muc6, CK14발현이 낮았으며, 간암의 경우, 모든 생체지표발현이 미흡하여 위 지표가 간암의 진단에는 유용하지 않는 것으로 생각된다. 위암의 경우, Muc5, CK7, CK13, CK19의 발현이 높고, 혈액, 맹장암 등에서는 CK13, CK19, CK20의 발현이 높았다. 한편, 6개 보다 많은 군집 수를 사용하면 더 많은 장기가 구분되어지지 않고 오히려 아무 정보도 주지 못 하는 불필요한 군집이 늘어남을 발견하였다. 이것으로부터 군집분석에 사용된 8개 지표를 12개의 장기에 대한 지표로 한꺼번에 사용되는 것은 부적절함을 보여준다.</p>					