

## 암과 식이요인의 관련성 파악에서의 가능한 바이아스

-서울코호트의 예를 중심으로-

<sup>1</sup>가천의과대학교 예방의학교실, <sup>2</sup>국제백신연구소, <sup>3</sup>한림대학교 의과대학 예방의학교실,  
<sup>4</sup>제주대학교 의과대학 예방의학교실, <sup>5</sup>성균관대학교 의과대학 예방의학교실,  
<sup>6</sup>울산대학교 의과대학 예방의학교실, <sup>7</sup>서울대학교 의과대학 예방의학교실

김대성<sup>1</sup> · 구혜원<sup>2</sup> · 김동현<sup>3</sup> · 배종면<sup>4</sup>  
신명희<sup>5</sup> · 이무송<sup>6</sup> · 이충민<sup>7</sup> · 안윤옥<sup>7</sup>

### Biases in Evaluation of Relation between Diet and Cancer

-An Example from Seoul Cohort Study-

Dae-Sung Kim<sup>1</sup>, Hye-Won Koo<sup>2</sup>, Dong Hyun Kim<sup>3</sup>, Jong-Myon Bae<sup>4</sup>  
Myong Hee Shin<sup>5</sup>, Moo-Song Lee<sup>6</sup>, Chung Min Lee<sup>7</sup> and Yoon-Ok Ahn<sup>7</sup>

<sup>1</sup>Department of Preventive Medicine, Gachon Medical School,

<sup>2</sup>International Vaccine Institute, <sup>3</sup>Department of Social Medicine,

Hallim University College of Medicine, <sup>4</sup>Department of Preventive Medicine

Cheju National University College of Medicine,

<sup>5</sup>Department of Preventive Medicine, Sung Kyun Kwan University College of Medicine

<sup>6</sup>Department of Preventive Medicine, University of Ulsan College of Medicine

<sup>7</sup>Department of Preventive Medicine, Seoul National University College of Medicine

## 서 론

Doll과 Peto (1981)은 35%의 암이 식이요인에 기인한다고 보고한 바가 있다. 그러나 이러한 수치는 미국에서도 불확실한 것으로 받아들여지고 있으며 학자에 따라서 10%에서 70%까지 식이요인이 관여하는 정도에 대하여 다양한 주장을 하고 있으며, 또한 식이의 어떠한 부분이 암의 발병과 관여되는지는 명확히 밝혀진 바가 없어 아직 많은 연구가 필요한 분야이다(Willet, 1994). 더구나 식이습관은 나라마다 상이함은 물론 총 암 발생률 및 부위별 암발생현황 등의 암의 발생양상 또한 나라마다 다양하다. 따라서 효과적인 암예방 전략을 수립하기 위해서는 해당 지역에서의 독자적인 연구가 필수적이라고 할 수 있다.

미국의 National Academy of Science (Committee

on Diet, Nutrition, and Cancer, 1982)는 암예방을 위하여 식이습관과 관련된 내용으로 다음과 같이 권고하였다.

- 전체칼로리의 30~40% 이하로 지방분섭취를 줄일 것
- 과일, 야채, 곡류섭취를 늘일 것
- 염장식품과 훈제식품의 섭취를 최소화할 것
- 식품의 제조, 처리, 저장과정에서 발암물질의 오염을 최소화할 것
- 음식에서의 mutagen을 발견하려는 노력과 그 발암성을 확인하는 연구를 지속하고 가능하면 mutagen은 제거하거나 최소화 할 것
- 알콜섭취는 적당한 수준을 유지할 것

또한 2000년도에 American Institute of Cancer Research에서 발표한 The AICR Diet & Health Guidelines for Cancer Prevention 에 따르면 일반인

들에 대한 암예방 권고사항은 다음과 같다.

- 식물성 성분이 풍부한 음식을 섭취할 것
- 야채나 과일을 많이 섭취할 것
- 적당한 체중과 육체적 활동을 유지할 것
- 술은 건강에 해가 되지 않을 수준에서 적당히 마실 것
- 지방분과 염분이 적은 음식을 섭취할 것
- 음식의 조리과 보관은 안전하게 할 것
- 금연할 것

약 20 년의 격차를 두고 발표된 암예방 권고사항이지만 그 동안 많은 역학적 연구가 있었음에도 불구하고 그 내용은 총론적인 수준에 머물고 있음을 알 수 있다. 권고사항은 매우 구체적이어야 함에도 불구하고 그렇지 못한 이유는 역학적 연구결과들을 일반인들이 이해가능하고 실행가능한 언어로 표현하는데 무리가 있다는 것 외에도 다음과 같은 이유가 있을 것이다. 첫째, 식이과 관련된 연구결과들은 정량적으로 표현하기 힘들다는 것이다. 이는 식이요인을 파악하는 방법이

계된 역학적 연구들은 반정량적 빈도설문서(semi-quantitative food frequency questionnaire)를 적용하고 있으므로 그 식이 섭취 정도는 중앙값 혹은 사분위수 등의 상대적 섭취량으로 표현되므로 해석에 있어서 정량적인 적용이 어렵다. 둘째, 구체적인 식이 항목에 대한 연구는 연구가 수행된 지역, 식이상태의 파악방법, 통계적 처리방법 등에 의하여 그 연구결과가 상당히 상이하게 나타난다.

특히 식이 항목을 선정하는 연구단계에서 실제의 식이상태를 파악하기 위하여 많은 종류의 식이항목을 조사항목으로 선정하게 되는데, 이로 인하여 예측하기 힘든 바이아스들이 발생 할 수 있다. 각각의 식이 항목은 서로 상관관계를 가지고 있으므로 통계학적으로 이를 처리하여야 하며, 각 식이의 섭취가 다른 식이 섭취의 흡수, 대사, 배설에 영향을 미칠 수 있으므로 하나의 식이 항목과 질병과의 관련성에 대한 연구결과를 해석하기 위해서는 각별한 주의를 기울여야 한다. 일반적으로 이러한 이유로 인하여 다변량 분석이 수행하

표 1. 한국에서 수행된 식이요인과 암발생과의 관련성에 대한 연구 결과

| Authors            | Foods                         | Cancer site | OR    | 95% CI    |
|--------------------|-------------------------------|-------------|-------|-----------|
| Shin, 1995         | Green vegetable               | Breast      | 0.5   | 0.2~1.1   |
| Shin, 1995         | Boiled vegetable              | Breast      | 0.4   | 0.2~0.7   |
| Shin, 1995         | Dairy products                | Breast      | 0.4   | 0.2~0.9   |
| Shin, 1995         | Soybean                       | Breast      | 0.4   | 0.2~0.9   |
| Shin, 1995         | Garlic                        | Breast      | 0.4   | 0.2~0.9   |
| Kim, 1997          | Meat                          | Large bowe  | 12.57 | 1.1~6.0   |
| Kim, 1997          | Red meat                      | Large bowel | 2.89  | 1.2~7.1   |
| Lee et al., 1995   | Mung bean pancake             | Stomach     | 0.1   | 0.05~0.4  |
| Lee et al., 1995   | Tofu                          | Stomach     | 0.2   | 0.1~0.8   |
| Lee et al., 1995   | Spinach                       | Stomach     | 0.1   | 0.02~0.3  |
| Lee et al., 1995   | Cabbage                       | Stomach     | 0.2   | 0.02~0.3  |
| Lee et al., 1995   | Broiled fish                  | Stomach     | 7.7   | 2.4~24.7  |
| Lee et al., 1995   | Soybean paste stew            | Stomach     | 5.5   | 2.5~12.1  |
| Lee et al., 1995   | Hot pepper-soybean paste stew | Stomach     | 4.2   | 1.5~12.0  |
| Park, 1998         | Total meat                    | Stomach     | 0.3   | 0.1~0.5   |
| Park, 1998         | Salted fish                   | Stomach     | 1.7   | 0.4~1.3   |
| Park, 1998         | Citrus fruits                 | Stomach     | 0.4   | 0.1~0.5   |
| Park, 1998         | Fruits juice                  | Stomach     | 0.5   | 0.3~0.9   |
| Park, 1998         | Garlic                        | Stomach     | 0.6   | 0.4~0.9   |
| Yun and Choi, 1998 | Jinseng                       | All         | 0.40  | 0.28~0.56 |
| Yun and Choi, 1990 | Jinseng                       | All         | 0.56  | 0.46~0.69 |

가지고 있는 근원적인 문제로서 대부분의 잘 설계 되는데 교란변수의 선정과 상호작용(interac-

tion)의 검정이 적절하지 않은 경우 연구결과의 타당성을 보장하기 힘들게 된다(Willett, 1990).

본 연구에서는 다수의 식이 항목으로 파악된 식이요인들과 암과의 관련성을 파악하기 위하여 수행된 코호트 연구의 예를 사용하여 분석단계에서 반드시 고려하여야 할 교란바이아스의 유형과 실제 적용을 검토하고자 한다.

채류, 유제품, 콩류와 마늘이 유방암의 보호요인으로 나타났다(Shin, 1995). 대장암의 경우 육류, 특히 붉은살 육류가 대장암의 위험요인으로 나타났다(Kim, 1997). 위암의 경우 녹두전, 두부, 시금치, 양배추, 과일류, 인삼 등이 보호요인으로 나타나고 있으며, 구운생선, 된장찌개 등이 위험요인으로 나타났다(Lee et al., 1995; Park, 1998).

**한국에서 수행된 역학적 연구들**

식이요인과 암발생과의 관련성에 대하여 우리나라에서 수행된 연구결과는 그리 많지 않다(표 1). 식이관련 연구는 주로 위암을 중심으로 이루어졌으며(Lee et al., 1995; Park 1998) 소수의 유방암(Shin, 1995), 대장암(Kim, 1997) 관련 논문이 있을 뿐이다.

연구결과들을 전반적으로 살펴보면 유방암과

**서울코호트의 분석결과**

1992년에 14,533명의 서울시 중년남성을 대상으로 코호트를 구축하였다. 당시 흡연, 음주, 식이습관, 과거병력, 가족질병력 등에 관한 정보를 수집하였으며 이후 전국암등록자료, 사망자료, 의무기록지 직접확인 등을 통하여 암환자 발병을 진행하고 있다. 이상의 자료를 통하여 1993년부터 1998년까지 6개년 동안의 암발병자를 확보하여

**표 2.** 1992년 서울코호트 구축시 자가기입식 설문조사용 식이항목

| 범 주            | 항 목   |
|----------------|---|
| 빵종류            | 흰빵, 식빵, 토스트; 도넛츠, 파배기; 기타빵(팥빵, 크림빵, 샌드위치, 곰보빵, 야채빵 등); 버터(빵에 발라먹는); 마이가린(빵에 발라먹는); 쥬, 꿀, 시럽(빵에 발라먹는)  |
| 곡류             | 쌀밥(혼식); 현미밥; 비빔밥; 볶음밥; 떡(인절미, 송편, 절편, 시루떡, 가래떡 등); 생선초밥, 김초밥; 만두; 국수(칼국수, 우동 등); 냉면, 모밀 등; 라면; 피자; 감자(찐 것, 조리); 튀김감자(볶은감자, 감자전 등);  |
| 육류, 생선류        | 쇠고기(불고기, 갈비, 수육 등); 곰탕, 설렁탕, 도가니탕, 해장국 등; 돼지고기(족발포함), 햄, 돈까스; 돼지삼겹살, 베이컨; 닭고기(철면조, 오리); 개고기; 간(소, 닭, 돼지, 개); 쏘세지; 생선회; 절인생선(고등어, 굴비, 갈치); 신선, 냉동, 말린생선(구이, 조리, 볶음 등); 가공생선(통조림, 맛살, 어묵 등); 오징어, 낙지; 새우, 조개, 굴; 젓갈(직접섭취) |
| 야채, 채소 반찬류     | 김치(깍뚜기, 오이 등); 고춧가루없는 백김치(동치미, 오이지 등); 녹색야채(풋고추, 미나리, 숙갓, 달래 등); 상추(생으로 먹는); 양상추(생으로 먹는); 생으로 먹는 녹황색야채(오이, 당근, 가지, 도라지 등); 데쳐서 먹는 야채, 나물(시금치, 콩나물, 버섯, 고사리); 마요네즈(야채첨가)   |
| 찌개류, 해조류 콩, 계란 | 김치찌개; 된장찌개; 생선찌개(매운탕); 미역, 다시마, 파래(국, 무침); 김(김초밥 제외); 계란(닭, 오리, 메추리); 콩(콩자반, 삶은콩); 두부, 군두부, 두부전(콩비지 포함); 녹두빈대떡  |
| 과일             | 토마토; 귤; 자몽, 오렌지; 사과; 복숭아; 자두; 바나나; 배; 참외, 멜론; 수박; 딸기; 포도; 키위; 오렌지쥬스; 토마토쥬스; 야채쥬스; 과일쥬스(포도, 딸기, 복숭아 등)   |
| 간식, 후식         | 아이스크림; 떠먹는 요거르트; 케익, 초코렛, 캔디, 비스킷; 땅콩, 호두, 잣; 우유; 두유; 치즈  |
| 차종류            | 커피; 홍차; 녹차, 숙차; 콜라, 사이다   |
| 기타             | 마늘(양념제외, 직접섭취); 양파(양념제외, 직접섭취); 인삼(술, 차제외), 인삼다린 것, 인삼젤리, 홍삼정; 잡채; 짬아지(단무지, 무지 등)   |

관련된 위험요인은 명확히 밝혀지지 않았으며 야 뒤 식이습관에 따른 암발병의 양상을 파악하였다.

연구에 사용될 코호트는 1992년 당시 서울지역에 근무하는 공무원 및 사립학교교직원 의료보험 피보험자 40세에서 59세의 남자 중에서 무작위 추출한 14,533명에 대하여 우편설문을 실시하여 식이습관을 비롯하여 육체적 활동, 음주습관, 흡연 습관, 식이습관, 체격, 사회경제적 상태, 직업, 과거의 질병력, 가족의 질병력 등을 파악하였다. 연구대상의 연령분포는 1992년을 기준으로 40대 초반이 34.2%로서 가장 많았으며 평균 연령은 48.7세(표준편차: 5세)이었다. 표 2는 설문조사시 파악한 식이에 대한 정보이며 각 개별 항목에 대하여 1년 동안 섭취한 평균횟수와 한번 섭취할 때의 평균적인 양을 파악하였다.

서울 코호트 구성원 14,533명에 대한 1993년 1월부터 1996년 12월까지의 의료보험관리공단 수진자료에서 모든 종류의 암(ICD-9: 140~209 혹은

들에 대하여 수진 기관과 수진 날짜를 파악하였다. 이 중 한의원을 제외한 2차 이상의 의료기관에 대하여 해당 병원에 방문하여 의무기록지를 열람하였으며 암 발병 확인을 위하여 고안된 구조화된 요약양식에 필요한 정보를 기입하였다. 의무기록조사에서 파악되지 않는 암발생환례를 확인하기 위해서 중앙암등록자료와 서울시지역암등록자료를 획득하여 연구대상에서의 암발생례를 확인하였으며, 또한 전국사망자료를 이용하여 의사진단서가 첨부된 암사망례를 파악하였다. 이상의 암 추적조사의 완전성을 Capture-recapture 방법으로 추정한 결과 전체 암발생례의 89.9%가 파악된 것으로 추정되었다(Kim et al., 1999).

교란바이아스가 존재하기 위한 필수적인 조건은 변수들이 서로 통계적 관련성이 있어야 한다. 서울코호트의 81개의 식이항목들에서 상관관계가

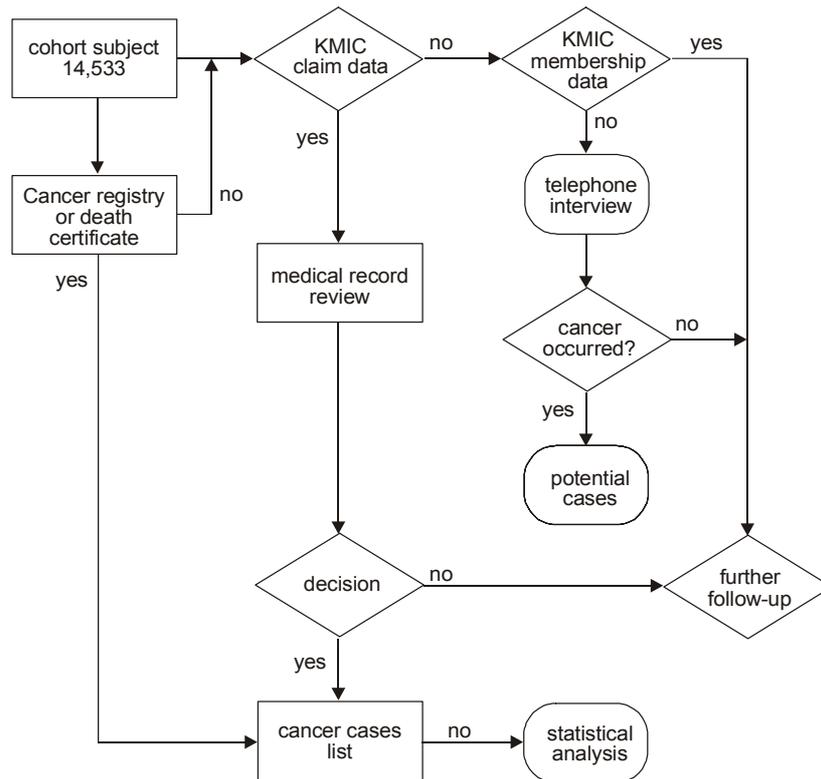


그림 1. 서울코호트에서의 암발생 추적조사전략. ICD-10: C00~D09)으로 1회 이상 진료 받은 사람

존재하는지를 파악하기 위하여 스피어만 상관분석을 수행한 결과는 표 2와 같다. 김치와 같은 흔히 먹는 음식을 제외한 대부분의 식이항목들이 p-value 0.0001 이하의 상관관계를 나타내고 있다. 흔히 먹는 음식 역시 p-value가 0.0001 이하의 상관관계를 나타내는 항목이 과반수를 차지함을 알 수 있다(표 3). 그러나 스피어만 상관계수가 0.4를 넘는 경우는 거의 없었다(표 4). 상추나 녹색야채 등에서 다른 식이항목과 높은 상관관계를 나타내

고 있으며, 과일과 관련된 식이 항목들도 다른 식이항목과 비교적 높은 상관관계를 나타내고 있다. 육류의 섭취는 상대적으로 다른 식이항목과의 상관관계가 낮음을 알 수 있다. 전체적으로 대부분의 식이항목들은 서로 통계적으로 유의한 상관관계가 존재하고 있으나 그 상관관계의 강도는 그리 높지 않음을 알 수 있다.

상관관계는 양의 방향과 음의 방향으로 나눌 수 있으며 본 자료에서 파악된 상관관계의 방향

표 3. 서울코호트 설문자료에서의 식이항목간의 스피어만상관분석결과 산출된 p-value의 크기에 따른 식이항목의 수(총 81개 항목)

| 식이항목             | 유의수준에 따른 식이항목의 수 |        |       |       |    |
|------------------|------------------|--------|-------|-------|----|
|                  | <0.0001          | <0.001 | <0.01 | <0.05 | ns |
| 쇠고기(불고기, 갈비, 수육) | 79               | 1      |       |       |    |
| 돼지고기, 햄, 돈까스     | 72               | 2      | 3     | 1     | 2  |
| 삼겹살, 베이컨         | 72               | 2      |       | 1     | 5  |
| 염장생선             | 79               | 1      |       |       |    |
| 젓갈               | 80               |        |       |       |    |
| 김치               | 46               | 11     | 2     | 7     | 14 |
| 녹색야채             | 80               |        |       |       |    |
| 상추               | 79               | 1      |       |       |    |
| 된장찌개             | 70               | 2      | 1     |       | 7  |
| 굴                | 78               |        | 1     |       | 1  |
| 키위               | 76               |        | 1     | 2     | 1  |

표 4. 서울코호트 자료에서의 식이항목 간의 스피어만 상관계수의 크기에 따른 식이항목의 수

| 식이항목             | 상관계수에 따른 식이항목의 수 |      |      |      |
|------------------|------------------|------|------|------|
|                  | >0.4             | >0.2 | >0.1 | ≤0.1 |
| 쇠고기(불고기, 갈비, 수육) | 1                | 13   | 57   | 9    |
| 돼지고기, 햄, 돈까스     | 1                | 5    | 31   | 43   |
| 삼겹살, 베이컨         | 1                | 4    | 29   | 46   |
| 염장생선             | 1                | 19   | 47   | 13   |
| 젓갈               | 1                | 12   | 49   | 18   |
| 김치               | 0                | 3    | 12   | 65   |
| 녹색야채             | 2                | 27   | 37   | 14   |
| 상추               | 2                | 36   | 33   | 9    |
| 된장찌개             | 1                | 12   | 38   | 29   |
| 굴                | 1                | 20   | 45   | 14   |
| 키위               | 0                | 21   | 45   | 14   |

표 5. 식이 항목간의 스피어만 상관분석 결과 음의 상관관계를 나타내는 항목

| 식이항목             | 음의 상관관계를 보이는 항목과 그 상관계수 |
|------------------|-------------------------|
| 쇠고기(불고기, 갈비, 수육) | -                       |
| 돼지고기, 햄, 돈까스     | 현미밥(-0.004), 녹차(-0.020) |
| 삼겹살, 베이컨         | 밥(-0.022), 녹차(-0.050)   |
| 염장생선             | -                       |
| 젓갈               | -                       |
| 김치               | 19개 항목 (-0.03~-0.001)   |
| 녹색야채             | -                       |
| 상추               | -                       |
| 된장찌개             | 흰빵(-0.003), 버터(-0.002)  |
| 굴                | -                       |
| 키위               | 밥(-0.008), 김치(-0.004)   |

표 6. 두 변수의 양의 상관관계로 인하여 교란바이아스가 나타나는 경우-I; 교란변수인 위험요인을 통제하지 않은 상태에서는 위험요인이 아닌 변수가 위험요인인 것처럼 나타날 수 있다

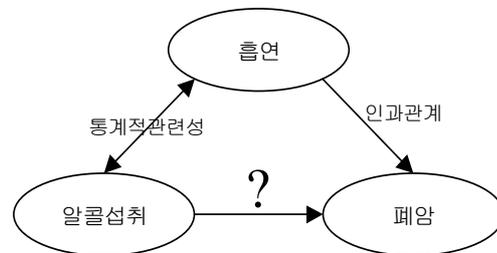
| 교란변수<br>흡연(true risk factor) | 알콜섭취(false risk factor) |             | 단변수 분석 |
|------------------------------|-------------------------|-------------|--------|
|                              | -                       | +           |        |
| -                            | 1.0 (a)                 | 1.0 (b)     | 1.0    |
| +                            | 10.0 (c)                | 10.0 (d)    | 10.0   |
| 단변수 분석                       | toward 1.0              | toward 10.0 |        |

\*: cell내의 숫자는 단위 인구당 발생률을 의미

을 정리하면 표 5와 같다. 일부의 낮은 상관관계를 보이는 경우를 제외하고는 대부분은 양의 상관관계를 나타 내고 있음을 알 수 있다.

대부분의 식이항목이 서로 상관관계를 가지고 있으므로 식이항목과 암과의 관련성 평가에서 어느 정도의 교란바이아스가 있을 것으로 예상된다. 기본적인 교란바이아스를 도식화하면 다음과 같다. 예를 들어 알콜섭취가 폐암에 미치는 영향을 파악하고자 할 때 교란변수인 흡연을 반드시 통제하여야 한다. 흡연은 폐암의 원인이면서 알콜섭취와 상관관계를 가지고 있기 때문이다. 이러한 경우 교란변수를 통제하지 않은 상태에서 단변수 분석을 수행하면 알콜섭취와 폐암은 인과적 관련성을 가지고 있는 것으로 나타나게 된다.

이상의 관계는 표 6와 같이 나타낼 수 있다. 표에서 나타나는 바와 같이 흡연의 대응위험도는 10.0이며 알콜의 대응위험도는 1.0이다. 폭로분율



이 0.5이라고 가정하면 알콜섭취군과 비섭취군은 위험도가 5.5로서 대응위험도는 1.0이 되지만 두 변수가 서로 강한 상관관계를 가지고 있을 경우 분할표의 a와 d쪽에 연구대상이 몰리게 된다. 따라서 알콜 섭취군에서의 위험도는 5.5에서 떨어져서 10에 가까워지고 비폭로군에서의 위험도는 5.5에서 떨어져서 1에 가까워진다. 따라서 흡연이라는 교란변수를 보정하지 않은 상태에서는 알콜섭

표 7. 두 변수의 양의 상관관계로 인하여 교란바이아스가 나타나는 경우-II; 교란변수인 위험요인을 통제하지 않은 상태에서는 실제 위험요인의 대응위험도가 실제보다 과장되어 나타날 수 있다

| 교란변수<br>true risk factor (A) | True risk factor (B) |            | 단변수 분석     |
|------------------------------|----------------------|------------|------------|
|                              | -                    | +          |            |
| -                            | 1.0 (a)              | 2.0 (b)    | toward 1.0 |
| +                            | 2.0 (c)              | 4.0 (d)    | toward 4.0 |
| 단변수 분석                       | toward 1.0           | toward 4.0 |            |

\*: cell내의 숫자는 단위 인구당 발생률을 의미

표 8. 두 변수의 양의 상관관계로 인하여 교란바이아스가 나타나는 경우-III; 교란변수인 위험요인을 통제하지 않은 상태에서는 실제 보호요인의 대응위험도가 실제보다 과소추정될 수 있으며 위험요인도 유의성이 사라질 수 있다

| 교란변수<br>true risk factor (A) | True protective factor (B) |            | 단변수 분석     |
|------------------------------|----------------------------|------------|------------|
|                              | -                          | +          |            |
| -                            | 1.0 (a)                    | 0.5 (b)    | toward 1.0 |
| +                            | 2.0 (c)                    | 1.0 (d)    | toward 1.0 |
| 단변수 분석                       | toward 1.0                 | toward 1.0 |            |

\*: cell내의 숫자는 단위 인구당 발생률을 의미

취의 대응위험도는 1보다 커지게 된다.

우리가 파악하고자 하는 변수가 실제로 질병과 관련이 있으며 교란변수와 관련성이 있다는 상황을 가정해보자. 표 7에서 변수 A와 B는 둘 다 질병과 관련이 있는 요인이다. 두 요인의 대응위험도의 참값은 2.0이며 서로 교호작용은 존재하지 않는다고 가정할 때 두 변수가 서로 상관관계가 존재하지 않으면 위험요인 B가 없는 군에서의 위험도는 1.5이며 위험요인 B가 있는 군에서의 위험도는 3.0으로 대응위험도가 2.0으로 산출된다. 그러나, 두 변수에서 양의 상관관계가 있을 때는 a와 d의 가중치가 커짐에 따라 위험요인 B가 없는 군에서의 위험도는 1.5에서 1사이에 위치하게 되고, 위험요인 B가 있는 군에서의 위험도는 3.0에서 4.0사이에 위치하게 되어 대응위험도가 2.0보다 커지게 된다. 이는 변수 A의 입장에서도 마찬가지여서 서로 보정을 하지 않은 상태에서는 대응위험도가 참값보다 과대추정된다.

표 8의 경우는 반대상황으로서 질병의 보호요인을 평가하고자 할 때 실제로 보호요인임에도

불구하고 이 요인과 상관관계를 가지는 위험요인을 통제하지 않은 경우 통계적 유의성을 상실할 수 있음을 나타내고 있다. 즉, 변수 B의 경우 대응위험도의 참값은 0.5이지만 위험요인인 변수 A와 양의 상관관계를 가지고 있으므로 해서 단변수 분석 결과 보호요인을 가지고 있는 군과 그렇지 않은 군에서 공통적으로 위험도가 1.0에 수렴하게 되어 대응위험도가 과소추정될 수 있음을 보여주며 이러한 현상은 위험요인인 변수 A에서도 공통적으로 일어나게 된다. 즉, 교란바이아스로 인하여 통계적 유의성을 상실하는 예를 개념적으로 나타내고 있다.

서울코호트의 6년간의 추적조사결과 위암과 식이항목의 대응위험도는 표 9와 같다. 개고기섭취(OR=1.23)와 젓갈섭취(OR=1.16)는 위암의 위험요인으로 나타났으며 이외의 육류섭취는 통계적 유의성이 없는 것으로 나타났다. 야채와 과일에 대한 분석에서는 과일쥬스(OR=0.82)를 제외한 대부분의 항목에서 유의한 변수가 관찰되지 않았다. 유제품 군에서 아이스크림(OR=0.56)과 요거트(OR=

표 9. 서울코호트에서의 식이항목과 위암의 대응위험도: polychotomous logistic regression 결과(adjusted for age and socioeconomic status)

| 식이항목   | 대응위험도 | 95% 신뢰구간    |
|--------|-------|-------------|
| 쇠고기    | 0.912 | 0.776~1.072 |
| 돼지고기   | 1.104 | 0.946~1.290 |
| 삼겹살    | 0.879 | 0.730~1.060 |
| 개고기    | 1.234 | 1.001~1.520 |
| 염장생선   | 1.131 | 0.979~1.305 |
| 젓갈     | 1.158 | 1.015~1.322 |
| 김치     | 0.944 | 0.840~1.061 |
| 녹색야채   | 0.939 | 0.837~1.054 |
| 황색야채   | 0.874 | 0.761~1.002 |
| 된장찌개   | 0.960 | 0.838~1.099 |
| 과일쥬스   | 0.821 | 0.685~0.985 |
| 아이스크림  | 0.555 | 0.396~0.778 |
| 요거트    | 0.821 | 0.704~0.957 |
| 땅콩, 호두 | 0.812 | 0.672~0.981 |
| 녹차     | 0.884 | 0.782~1.000 |
| 양파     | 0.840 | 0.746~0.946 |

표 10. 두 개의 유의하지 않은 독립변수를 조합으로 구성된 로지스틱모델. 의심되는 위험요인과 보호요인을 짝으로 구성함(adjusted for age and socioeconomic status)

| 독립변수 | 대응위험도 | 95% 신뢰구간  |
|------|-------|-----------|
| 쇠고기  | 0.73  | 0.45~1.19 |
| 녹색야채 | 0.72  | 0.43~1.21 |
| 쇠고기  | 0.72  | 0.44~1.16 |
| 황색야채 | 0.77  | 0.47~1.24 |
| 삼겹살  | 0.82  | 0.50~1.34 |
| 녹색야채 | 0.69  | 0.42~1.14 |
| 삼겹살  | 0.81  | 0.49~1.31 |
| 황색야채 | 0.73  | 0.46~1.17 |
| 염장생선 | 2.00  | 1.12~3.60 |
| 녹색야채 | 0.61  | 0.37~1.03 |
| 염장생선 | 2.00  | 1.11~3.60 |
| 황색야채 | 0.61  | 0.38~0.99 |
| 된장찌개 | 0.72  | 0.45~1.16 |
| 녹색야채 | 0.70  | 0.41~1.17 |
| 된장찌개 | 0.89  | 0.56~1.42 |
| 황색야채 | 0.73  | 0.45~1.18 |

0.82)가 위암의 보호요인으로 나타났으며, 기타 땅콩, 호두 섭취(OR=0.81), 양파 섭취(0.84)가 위암의 보호요인으로 나타났다.

표 10의 분석결과는 표 8의 개념을 적용하기 위하여 단변수분석결과 통계적으로 유의하지 않은 가능한 위험요인과 가능한 보호요인을 포함한 다변량분석을 수행한 결과이다. 염장생선과 황색야채를 포함한 분석모델에서 단변수분석에서 보이지 않았던 유의성(OR=2.00 and 0.61)이 나타남을 알 수 있으며 염장생선과 녹색야채를 포함한 분석모델에서도 이와 유사한 현상(OR=2.00 and 0.61)이 나타남을 알 수 있다.

### 고 찰

식이를 측정하는 방법으로는 식이항목별 측정과 영양소를 측정하는 방법이 있다. 식이항목별 측정 중 24시간 회상법(short-term recall), 식이일기, 및 식이빈도설문법이 있으며 영양소를 측정하는 방법은 생체시료를 이용한 실험실적 방법과 설문서로부터 재구성하는 방법이 있다. 각각의 방법마다 장단점이 있으나 식이빈도 설문법이 과거

의 식이폭로를 비교적 정확하며 이로부터 미세영양소 섭취량의 산출도 가능하므로 역학적 연구에서 비교적 흔하게 사용된다. 또한 새로운 가설을 구상할 수 있다는 장점이 있으며, 암예방 지침을 만들기 위해서도 공중보건적 관점에서는 영양소보다는 식이항목이 유리하다.

본 자료에서는 식이항목들이 서로 통계적으로 유의한 상관관계를 가지고 있음을 보여주었다. 식이항목들간의 상관관계들이 연구결과에 어떤 영향을 미치는지에 대한 연구는 많지 않다. Elmstahl와 Gullberg (1997)는 식이항목들간의 상관관계가 존재함으로 인하여 다중공선성의 문제가 발생하므로 식이와 관련된 다변량 분석시에는 변수의 수를 가능하면 줄이는 것을 권고하였다. 그러나 교란바이아스로 인한 연구결과 오류는 다중공선성의 문제와는 별개의 문제이다. 본 연구에서 나타난 바와 같이 교란바이아스로 인해서 단변수 분석에서 실제 위험요인이 과소평가될 수도 있으며 실제 보호요인이 식이항목간의 상관관계가 다양한 오류를 발생시킬 수 있음을 알 수 있다.

식이자료와 암발생과의 관련성에 대한 분석시 독립변수의 상호관련성을 상관분석을 통하여 검

토함으로서 가능한 교란작용이 어떻게 일어날것 인지를 예측하는 것이 중요하다. 비록 단변수 분석에서 유의하지 않게 나타났다고 하더라도 교란 바이아스로 인해서 유의하지 않게 나타날 가능성을 있음을 고려하여 적절한 교란변수를 선택하여 참값을 추정할 수 있다.

### 참고 문헌

- 1) American Institute of Cancer Research. <http://www.aicr.org/report2.htm>.
- 2) Committee on Diet, Nutrition, and Cancer: Assembly of Life Sciences, National Research Council. Diet, Nutrition, and Cancer. Washington, DC: National Academy Press, 1982.
- 3) Kim DS, Lee MS, Kim DH, Bae JM, Shin MH, Lee CM, Koo H, Kang W, Ahn YO. Evaluation of the completeness of cancer case ascertainment in the Seoul Male Cohort Study: Application of the capture-recapture method. *J Epidemiol* 1999; 9: 146-154.
- 4) Elmstahl S, Gullberg B. Bias in diet assessment methods--consequences of collinearity and measurement errors on power and observed relative risks. *Int J Epidemiol* 1997; 26: 1071-1079.
- 5) Lee JK, Park BJ, Yoo KY, Ahn YO. Dietary factors and stomach cancer: a case-control study in Korea. *Int J Epidemiol* 1995; 24: 33-41.
- 6) Yun TK, Choi SY. A case-control study of ginseng intake and cancer. *Int J Epidemiol* 1990 Dec; 19(4): 871-876.
- 7) Yun TK, Choi SY. Non-organ specific cancer prevention of ginseng: a prospective study in Korea. *Int J Epidemiol* 1998; 27(3): 359-364.
- 8) Willett W. Nutritional epidemiology, New York: Oxford University Press, 1990.
- 9) 김동현. 일과 및 여가시간에서의 육체적활동량과 대장암 발생위험과의 관련성 연구. 서울대학교 대학원 1997.
- 10) 박혜성. 식이가 위암발생에 미치는 영향에 관한 연구. 한국역학회지 1998 33 pp 82-101.
- 11) 신명희. 한국 여성에서의 식이습관과 유방암 발생 위험과의 관련성에 관한 환자-대조군 연구. 서울대학교 대학원 1995.