

집단발병조사의 이론과 실제

- 암질환 -

박병주, 배종면, 안윤옥
서울 대학교 의과대학

이 덕 형
보건 사회부

전염병과 같은 급성 질환의 집단발생에 비해, 발생 및 사망 수준이 희귀한 암은 집단으로 발생하는 경우를 알아차리기는 상대적으로 힘들다. 그러나, 생활수준의 향상으로 인한 영양상태의 개선과 전염성 질환의 감소로 평균수명이 연장되고 노령인구가 증가하면서 암환자가 증가하고 있다. 이에 따라 암의 환경적 원인에 대하여 국민의 경각심이 고조되고 있는 가운데, 연도별 암 발생률 및 사망률에서 갑작스런 추이의 변동이 있거나, 예상치 않은 암의 발생 혹은 사망은 정보전달의 확산등에 힘입어 종종 사회문제로 제기되기도 한다. 따라서 긴 잠복기와 다양한 원인에 의해 희귀하게 발생하는 암이 특정지역에서, 특정 시기에 집단발병양상을 보인다면 이를 확인하고 그 원인을 규명하기 위한 체계적인 역학조사의 수행이 필요하게 된다. 그림 1 은 암 역학조사 수행도의 한 예이다 (안윤옥 등, 1993).

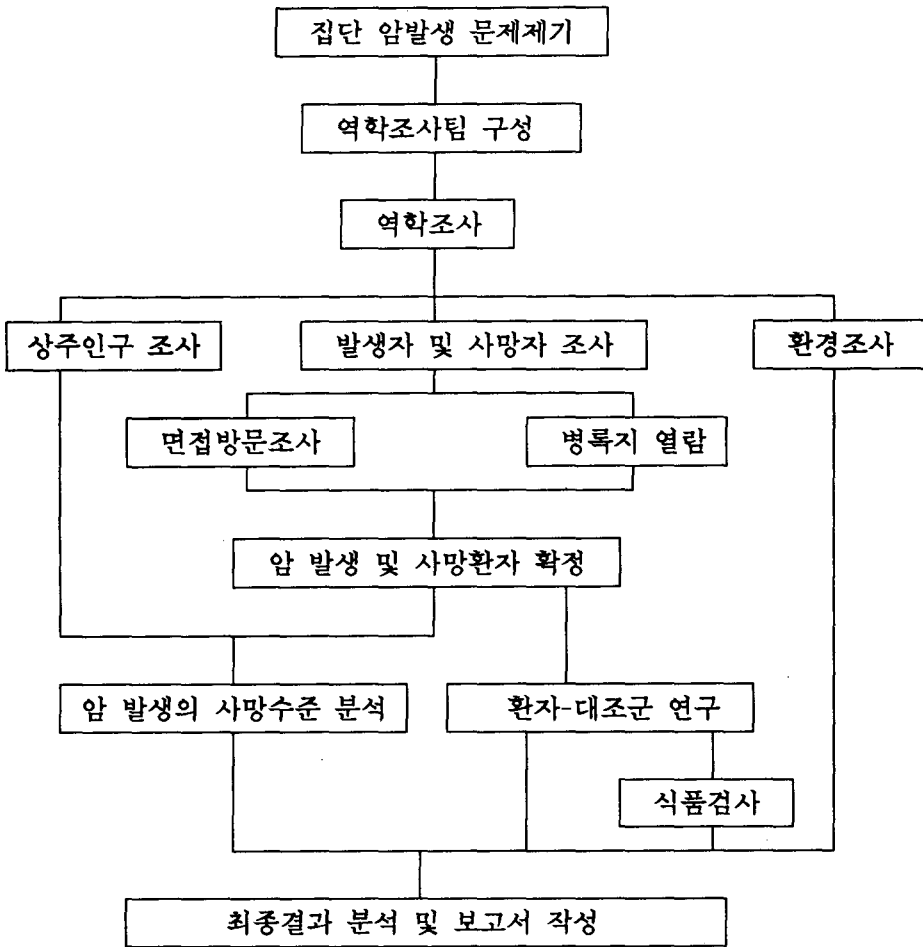


그림 1. 암발생의 역학조사 흐름도

1. 집단 발생유무 확인 : 발병률 및 사망률 조사

한 지역에서 특정기간에 암의 집단발병이 있다고 문제제기가 될 때, 우선 발생된 질병이 실제 암인지를 확인하는 것이 가장 중요하다. 왜냐하면 암은 주로 고령층에서 발생하는데, 노인에서는 다른 만성질환에 의한 증상과 구별하기 힘들어 진단의 오류등이 개입할 가능성이 높기 때문이다. 예를 들면, 폐암과 만성호흡기계 질환, 위암과 만성소화기계 질환, 간암과 간경변등의 유사 증상을 가진 질환들이 혼재되면서 병원에서 확진을 받지 않은 경우는 오분류될 수 있기 때문이다. 또한 암은 발생률/사망률이 낮기 때

문에, 발생지역의 주민의 성별, 연령별 분포에 따라 환자 발생 혹은 사망의 내용이 크게 차이 날 수도 있다.

따라서 암의 집단발생이 있다고 추정된다면 무엇보다도 먼저 발생된 환자의 진단명을 확인하여서, 그 집단에서의 암 발생률 혹은 사망률을 구하는 것이 중요하다. 이때 인근지역에서 암발생률과 사망률을 산출하여 비교하므로써 조사결과의 타당성을 높일 수도 있다.

가. 암 발생자 및 사망자 확인

암 발생자 혹은 사망자로 알려진 환자에 대해서는 사회인구학적 내용 - 연령, 성별, 직업등 - 뿐만 아니라 확진여부 및 확진내용을 확인할 필요가 있다. 암 발생 및 사망자를 확인하는 방법으로는 행정체계 이용, 면접 방문과 설문서조사, 그리고 병록지 조사 확인, 암등록자료와의 비교등이 있으며 이는 서로 보완적으로 사용된다.

우선 해당 지역의 사망신고자료등의 행정자료를 이용하여 가장 신속하게 암 발생 및 사망자를 확인할 수 있다. 그러나 암질환인 경우에는 확인 명단에서 사망신고의 미비등으로 누락 가능성이 있으므로, 여기서 나온 자료는 현지에서 조사연구진에 의해 확인과정을 받아야 한다. 명단에서 누락된 대상자들은 지역 상주인구를 조사하는 가운데 최근 각종 질병 발생자 및 사망자를 같이 조사하므로써(부록 1), 누락대상자를 확인할 수 있으며, 물론 확인과정을 거쳐야 한다.

현지에서 시행한 대상자 확인과정에서 대상자가 살아있다면 본인을 직접 방문면접하면서 그동안의 임상경과, 진단받은 시기와 병원, 치료과정 등을 확인하는 것이 중요하다. 그러나 여기에도 기억의 오류, 자신의 질병에 대한 인지부족 등이 있을 수 있으므로 반드시 병록지 확인과정이 동반되어야 한다. 만약 본인이 사망하였다면 배우자, 친척, 혹은 접촉 가능한 이웃인을 통해 간접적으로 사망직전의 환자상태를 비롯한 임상적 사항등을 알아보는 것도 도움이 되겠다.

병록지 확인은 면접을 통해 알아낸 진단받았던 병원을 방문하여 진단명, 진단시기, 병리조직학적 소견 등을 확인한다(부록 2). 그러나, 진단받은 병원이 여러군데라면, 최종 진단을 받은 병원을 찾아 병록지를 확인하여 조직검사 소견을 알아내는 것이 중요하다. 만약 병원파악이 안되는 경우에는 그 지역에서 해당 암에 대해 주민들이 어떻게 의료기관을 이용하고 있는가를 살펴보는 것이 중요하다. 예를 들면 암을 진단할 수 있는 2차 또는 3차병원을 중점으로 암 발생자 및 사망자의 진단내용을 우선 확인해 볼

수 있다. 혹은 지역내 특정질병을 다루는 병원이 있다면, 병록지 확인에 도움이 될 수 있다.

병록지 확인 과정에서 만약 “추정”이라는 진단명으로 나올 경우는 임상소견 등을 종합하여 임상적으로 판단할 필요도 있다. 또한 해당질병이 아니더라도 임상적으로 비슷한 질병현상을 가진 다른 질병에 대해서도 확진 내용을 확인하여 해당질병이 아님을 규명하는 것이 중요하겠다. 그러나 환자 개개인의 진찰권 번호가 체계적으로 부여되지 않았다거나, 병록지 보관이 제대로 안되는 사례 등 병원 병록지 관리 체계에 문제점이 있을 수 있는데, 이를 경우에는 병원 방문일이나 이름의 가나다순으로 접근하여야 될 경우도 있다.

혹 환자가 다른 지역에서 진단을 받았다면, 국립의료원에서 주관하고 있는 암등록사업의 등록 명단과 대조 확인하여 볼 수도 있다. 그러나 등록명단에 누락된 경우가 많아, 만약 전산화자료를 이용한 환자 대조작업시 없다하여도 암환자가 아니라고 할 수 없다.

이상의 암 발생/사망자를 확인하는 과정에서, 항상 고려해야 할 것은 실제 대상자가 조사에서 누락될 가능성이 있는지 여부이다. 발생/사망률이 낮은 암인 경우 환자 본인의 호소가 없을 경우 일회적 사건으로 묻혀버릴 수가 있으며, 질병명을 알지 못한 채 사망(노환)하거나, 실종 등에 의한 누락 가능성 때문이다. 이러한 경우에 대해서는 아래의 상주인구조사를 하면서, 정해진 기간내에 일어난 모든 질환자 및 사망자를 조사하거나 설문서에 기입하도록 하여 해당질병이 의심되는 경우는 모두 확인하는 것이 중요하겠다.

나. 상주인구 조사

이상의 암 발생자 및 사망자 확인과 함께, 이의 발생/사망률을 구함으로써 집단발생 여부를 확인하여야 한다. 이를 위해서는, 해당지역의 정해진 기간내 연령별, 성별 구조에 대한 자료를 확보하여야 하며, 그 지역의 해당질병의 연령별, 성별 발생/사망률을 직접 산출할 수 있을지를 확인해야 한다. 또한 그 지역의 암발생률 및 사망률의 수준을 평가하는 방법으로 그 지역에서 산출된 발생률 및 사망률의 표준화를 위해서는 인구센서스를 통한 전국 성별, 연령별 인구구조자료, 사망통계자료 및 지역 암등록자료 등을 확보해야 한다.

해당지역의 암 발생률 및 사망률을 적절히 산출하기 위하여는 그 지역에서 실제로 거주하고 있는 상주인구의 연령별, 성별 구조를 알아내는 조사가 중요한데, 우리 나라

에서는 기존의 행정자료가 부정확하여 - 특히 시골인 경우- 시간과 인력이 많이 든다. 만약 기존의 행정상 구해져 있는 인구 분포 자료를 이용한다면, 실제 상주인구와 차이가 날 수 있다는 사실을 염두에 두어야 한다. 해당지역에 주소지로 등재되어 있으나 실제 다른 지역에서 사는 경우, 이동인구에 대한 전출입 신고의 행정기록 미정리, 출생 및 사망자 신고가 제때 이루어 지지 않은 경우 등으로 차이가 생길 수 있기 때문이다. 대부분의 경우는 실제 상주인구에 대한 조사가 불가피한데, 조사지역의 넓이와 연구진의 인력을 고려하여 이루어져야 한다. 연구진이 많거나 좁은 지역이라면 해당지역 주민을 가가호호 방문할 수 있으나, 그렇지 않은 경우라면 가구별로 미리 작성된 설문지를 배부하여 상주인구를 기입하게 하는 설문지 방법을 이용할 수 있다(부록 1). 그러나 가구별 설문지 조사방법시 고령이나 文盲 등의 이유로 직접 설문서를 기록할 수 없는 가구원만 사는 경우가 있는데, 이때는 이웃이나 해당 지역 반장에 의해 기록하게 한다. 설문지 회수도 일단 반장이 취합하여 일차로 누락여부를 확인·보완토록 하고, 이를 다시 이장이 반별로 미회수 여부를 확인케 하면서, 현지 사정을 잘 아는 주민과 함께 가장 최근에 정리된 전출입 명부를 기초로 실제 상주하는 대상을 일일이 확인하는 과정을 병행한다면 그 지역의 실제 상주인구조사 결과에 신뢰성을 높일 수 있다.

따라서 효율적인 상주인구 조사가 이루어지기 위해서는 해당지역의 행정력 동원과 주민의 협조가 필수적이다. 주민의 협조를 구하기 위해 연구시작전 연구의 필요성을 홍보하고, 구체적인 협조사항을 알리기 위해서 주민과의 대화 모임을 갖는 것이 좋다. 이때 주민들이 가지고 있는 질병과 그 원인에 대한 선입견등을 인식할 수 있으며, 이를 연구진행시 참조가 되도록 하는 것이 중요하다. 또한 조사가 끝나고 현지를 떠나기 전에 중간결과를 요약하여 주민들에게 알리는 것도 필요하다. 한편 행정력의 동원은 중앙정부에서 지방행정기관으로 적극적인 사전협조가 이루어지도록 선행조치가 필요하지만, 실제 조사를 성공적으로 수행하기 위하여는 현지의 최소 기본 행정단위인 반장과 이장의 적극적인 협조가 필수적이다.

다. 암 발생 및 사망수준 비교

조사대상지역의 상주인구조사와 암환자 확인과정이 끝나면, 암의 발생/사망 수준이 어느 정도인지를 비교·확인한다(Hennekens et al,1987).

우선 전체 사망자와 50세 이상의 사망자 분률인 비례사망지수(PMI)를 산출하여 다른 집단과 비교할 수 있다. 또한 암은 발생이 매우 낮은 질환이므로 환자수와 상주인구수만으로 포아송분포 도표(Table 1)를 참조하여 본격적인 조사시행 이전에 현장에서

Table 1. 95% and 99% significance factors for the ratio of an observed value of a Poisson variables to the expectation (Nelder 1964)

x*	95%		99%		x	95%		99%	
	SMR** _U	SMR _L	SMR _U	SMR _L		SMR _U	SMR _L	SMR _U	SMR _L
1	.180	39.53	.135	199.6	60	.777	1.31	.723	1.43
2	.277	8.26	.216	19.42	70	.791	1.28	.740	1.39
3	.342	4.85	.273	8.88	80	.803	1.26	.754	1.36
4	.391	3.67	.318	5.95	90	.814	1.24	.766	1.33
5	.428	3.09	.353	4.63	100	.822	1.23	.777	1.31
6	.450	2.73	.383	3.00	120	.836	1.21	.794	1.28
7	.485	2.49	.409	3.43	140	.847	1.19	.807	1.26
8	.508	2.32	.431	3.11	160	.857	1.17	.818	1.24
9	.527	2.18	.450	2.88	180	.864	1.16	.827	1.22
10	.544	2.08	.467	2.69	200	.871	1.15	.835	1.21
11	.559	2.00	.483	2.55	250	.883	1.13	.851	1.18
12	.573	1.94	.497	2.43	300	.893	1.12	.863	1.16
13	.585	1.88	.510	2.33	350	.901	1.11	.873	1.15
14	.596	1.83	.522	2.25	400	.907	1.10	.880	1.14
15	.606	1.79	.533	2.18	450	.912	1.10	.887	1.13
16	.616	1.75	.543	2.11	500	.916	1.09	.892	1.12
17	.625	1.72	.552	2.06	600	.923	1.08	.901	1.11
18	.633	1.69	.561	2.01	700	.929	1.08	.908	1.11
19	.640	1.66	.569	1.97	800	.933	1.07	.913	1.09
20	.647	1.64	.577	1.93	900	.937	1.07	.918	1.09
21	.654	1.62	.584	1.90	1,000	.940	1.06	.922	1.08
22	.660	1.60	.591	1.87	1,200	.945	1.06	.929	1.08
23	.666	1.58	.598	1.84	1,400	.949	1.05	.934	1.07
24	.672	1.56	.604	1.81	1,600	.952	1.05	.938	1.06
25	.678	1.55	.610	1.79	1,800	.955	1.05	.941	1.06
26	.682	1.53	.615	1.76	2,000	.957	1.04	.944	1.06
27	.687	1.52	.621	1.74	2,500	.962	1.04	.950	1.05
28	.692	1.50	.626	1.72	3,000	.965	1.04	.954	1.05
29	.696	1.49	.631	1.71	3,500	.967	1.03	.958	1.04
30	.700	1.48	.635	1.69	4,000	.969	1.03	.960	1.04
35	.719	1.44	.656	1.62	5,000	.973	1.03	.964	1.03
40	.734	1.40	.674	1.56	6,000	.975	1.02	.967	1.03
45	.747	1.37	.689	1.52	7,000	.977	1.02	.970	1.03
50	.758	1.35	.702	1.40	8,000	.978	1.02	.972	1.03

SMR** = x*/m (x*: Observed numbers, m: Expected numbers)

암의 집단발병 유무를 간편하게 판단할 수 있다. 예를 들면, 특정 암 발생률이 연간 10만명당 40명이라고 알려진 가운데, 인구 1,000명의 지역에서 5년 간 암환자가 6명 (=x)이 발생하였다면, 관찰가능한 환자는 2명 (1000명 * 5 * 40/100,000)이다. 앞 도표에서 95% 신뢰구간 내에 x=6인 경우의 예상 환자 (=m)는 2.19명 (=6/2.73)에서 13.3명 (=6/0.450)이다. 이렇게 구해진 신뢰구간에는 기존의 발생률에서 나온 관찰 가능한 환자 2명이 포함되지 않으므로, 이 지역에서 최근 5년간에 암환자 발생은 평소시의 발생수준과는 다르며, 집단 발생의 가능성이 있다고 추측해 볼 수 있다.

한편, 암환자와 지역 주민의 연령별 성별 분포와, 해당지역의 발생률/사망률이 알려져 있다면 직접표준화 방법을 이용하여 발생 및 사망 수준을 비교할 수 있다. 그러나, 대부분의 역학조사는 해당지역의 발생률 및 사망률이 알려져 있지 않은 경우가 많아서, 전국의 발생/사망률을 통한 간접 표준화 방법을 이용한다. 특히 암질환같이 발생률이 낮은 경우에는 5 - 10년간의 발생사항을 묶어서 비교하는 것이 유용할 수 있는데, 이때는 누적발생률을 이용한 아래의 수식(5년을 예로 들때)에 따라 표준화 사망비(SMR)를 구할 수 있다(WHO 1987).

$$SMR = \frac{N}{E} = \frac{\sum D_i}{\sum L_i C_i + \sum D_i C_i}$$

- N : 과거 5년간 총 사망자수
- E : 전국 표준률에 의하여 산출된 과거 5년간 총 사망예상자수
- D_i : i age group 에서 과거 5년간 사망자수
- L_i : i age group 에서 현재 거주자수
- C_i : i age group 에서 과거 5년간 누적 발생률

이의 95% 신뢰구간은 Vandenbrouche (1984) 의 공식으로 구한다.

$$\frac{[\sqrt{\text{observed event} \pm (1.96 \times 0.5)}]^2}{\text{expected event}}$$

또한 암질환인 경우에는 국내의 실정으로는 아직까지 노인층에서 병원에서의 진단율이 낮아 확진이 잘 안되는 경우가 많고, 사망원인의 경우에도 ‘노환’등의 원인미상이 많아서, 75세 이상의 대상자를 제외하여 발생 및 사망수준을 알아보기도 한다.

2. 해당 질병의 관련요인 가설 설정 및 규명

역학조사의 궁극적인 목적이 발생의 원인을 규명하여 향후 질병 발생을 예방하는 것이므로, 발생 및 사망 수준을 비교하여 기존의 수준에 비해 유의한 차이가 있다면 해당 암의 원인을 규명하도록 해야 한다(Kelsey et al, 1986). 물론, 발생 및 사망수준을 확인하는 과정에서 예비적으로 원인규명을 동시에 행할 수도 있다.

원인규명을 위한 가설설정은 우선 해당 암에 대한 지식뿐만 아니라 그 지역의 특이한 환경조건, 생업등을 고려하여 정해진다. 혹은 주민들이 특정 암에 대한 선입감등을 갖고 있다면 이를 규명하기 위한 가설설정도 고려해 볼 수 있다. 이때 질환자의 성별, 연령별 분포 뿐만 아니라 지역별 분포를 지도에 표시하여 본다면 원인 규명의 실마리를 찾을 수도 있다. 아래 도표 2 처럼 성별로 발생 차이가 뚜렷하여, 여성에서 암환자의 발생이 없었다는 점은 환경적 요인에 의한 발암가능성이 낮다고 유추할 수 있다. 또한 그림 2 처럼 암환자의 분포도에서 지역별로 밀집된 양상을 보이지 않고 있다.

표 2. 5년간 암의 성별 연령별 발생 및 사망자 수

연령	발생		사망	
	남	여	남	여
45 - 49	1	0	0	0
50 - 54	0	0	0	0
55 - 59	3	0	0	0
60 - 64	2	0	2	0
65 - 69	1	0	0	0
70 +	1	0	0	0
합계	8	0	2	0

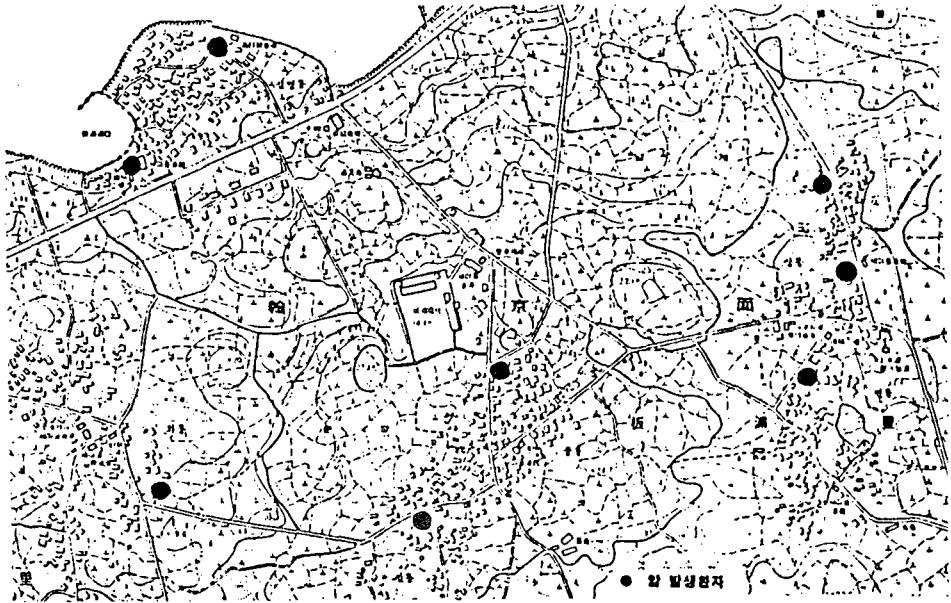


그림 2. 암환자의 지역 분포도

원인 규명은 환자-대조군 연구를 통해 이루어질 수 있으며, 특히 암인 경우는 발생이 희귀하므로 가장 효율적인 방법이 될 것이다. 대조군은 확인된 암환자 수에 따라 주로 이웃 건강인에서 선정한다. 대조군 선정 방법은 같은 지역에 거주하는 주민 가운데 건강한 자를 모두 대조군으로 삼는 방법, 성별/연령별로 빈도를 비슷하게 하여 뽑는 빈도 짝짓기 (frequency matching) 선정 방법, 또는 환자 개개인에 대해 성별/연령별로 1:1 짝짓는 개별 짝짓기 (individual matching)가 있다. 암환자 수가 많지 않은 경우에는 짝짓기 비를 1:3 내지 1:4 까지 높힘으로써 분석시 통계적 검정력을 높히는 방법을 고려할 수도 있다.

그러나, 지역주민에서 건강인을 대조군에 선정할 경우에 대조군에 대한 암이환 여부를 확인하기 위하여 필요한 임상적 검사를 시행하여야 한다. 현실적으로 직접적인 임상 검사의 시행이 불가능한 경우에는 역학조사에 참여한 의사가 사전에 작성한 임상증세에 관한 설문서(부록 3)로 본인을 직접 면접하여 해당 암질환을 가지고 있는지 여부를 판단하여야 한다.

환자-대조군 연구를 수행하기 위한 설문서에는 개인별 생활습관, 음용수, 식품등의 환경요인 등이 포함된다. 설문조사 방법에는 미리 작성한 설문서(부록 3)를 통해 방문 면접으로 직접 조사하는 방법과 자기 기입식으로 주민이 직접 작성하도록 하는 방법이 있겠다. 자기기입식인 경우는 조사팀의 인력이 부족한 경우에 이용할 수 있으나, 설문서 응답시 주민들이 질문을 다양하게 이해하여 응답할 수 있으므로 설문서의 신뢰도가 떨어지는 자료를 모을 수 있고, 행정력의 적극적인 협조 없이는 제대로 수행하기가 어려운 단점이 있다. 반면 방문면접방법은 인력과 시간을 많이 투입하게 되나, 조사자료의 신뢰도를 높힐 수 있고, 선정된 건강대조군에서 실제 건강한지를 역학조사자가 직접 확인해 볼 수 있다는 장점이 있다. 또한 주민과 직접 대화하므로써 주민의 긍정적인 협조를 얻어낼 수 있다.

이상의 수집된 자료들은 짝짓기를 하기 않고 대조군을 선정한 경우에는 조건화 되지 않은 로지스틱 분석(unconditional logistic regression)을 시행하고, 짝짓기가 된 경우라면 조건화된 로짓모델에 따른 다변량 분석(conditional logistic regression)을 통해 자료를 분석하여, 암발생과 유의하게 관련되는 위험요인을 파악하고, 가설을 검증하게 된다 (Breslow & Day, 1980).

환경적 요인과의 관련성을 규명하기 위해서는 수질 및 식수검사가 행하여 진다. 수질 검사는 우선 지역 주민의 주요 식수원을 알아내어 상수를 채취하는 것이 중요하며, 하수는 지역환경에 따라 채취장소를 선정하여야 한다. 이렇게 채취한 시료의 검사는 국립보건원 수질과에 의뢰할 수 있다. 한편 식품검사는 환자-대조군 연구를 시행하는 과정에서 환자가 발생한 가정과 대조군으로 선정된 주민의 가정에서 특정식품을 채취하여 발암성 물질을 함유하는지 여부를 국립보건원과 해당 연구기관에 검사를 의뢰하여 분석해 볼 수 있겠다.

참고문헌

1. 안윤옥, 박병주, 배종면, 이덕형, 김동진, 김준환, 강종원, 정용진, 이동모. 북제주 군 판포리 지역 암발생에 관한 역학 기초조사. 한국역학회지 1993;15(2):185-195.
2. Breslow NE, Day NE. Statistical methods in cancer research. volume 1 - The analysis of case-control studies. IARC, Lyon. p248-79, 1980
3. Hennekens CH, Buring JE. Epidemiology in medicine. Little, Brown & Co., Boston/Toronto. p66-98, 1987.
4. Kelsey JL, Thompson WD, Evans AS. Methods in observational epidemiology. Oxford University Press, New York. p212-53, 1986.
5. Nelder JA. Queries and notes: Significance factors for the ratio of a Poisson variable to its expectation. Biometrics 1964;20:639-43.
6. Vandenbrouche JP. A shortcut method for calculating the 95 per cent confidence interval of the standardized mortality ratio. Am J Epi 1984;115:303-304.
7. WHO. Cancer incidence in five continents. Volume V. IARC, Lyon. p787-9, 1987.