Journal of Preventive Medicine and Public Health
November 2010, Vol. 43, No. 6, 543-548

doi: 10.3961/jpmph.2010.43.6.543

원전 코호트 연구의 적정 대상규모와 검정력 추정

조인성[1], 송민교[1], 최윤희[2], 이충민[3], 안윤옥[1,3]

[1]서울대학교 의과대학 예방의학교실, [2]서울대학교 의학연구협력센터, [3]서울대학교 의학연구원 원자력영향역학연구소

Power Estimation and Follow-Up Period Evaluation in Korea Radiation Effect and Epidemiology Cohort Study

In Seong Cho[1], Minkyo Song[1], Yunhee Choi[2], Zhong Min Li[3], Yoon-Ok Ahn[1,3]

[1]Department of Preventive Medicine, College of Medicine, Seoul National University;
[2]Medical Research Collaborating Center, Seoul National University Hospital/Seoul National University;
[3]Institute of Radiation Effect & Epidemiology, Seoul National University Medical Research Centre

Objectives: The objective of this study was to calculate sample size and power in an ongoing cohort, Korea radiation effect and epidemiology cohort (KREEC).
Method: Sample size calculation was performed using PASS 2002 based on Cox regression and Poisson regression models. Person-year was calculated by using data from '1993-1997 Total cancer incidence by sex and age, Seoul' and Korean statistical informative service.
Results: With the assumption of relative risk=1.3, exposure:non-exposure=1:2 and power=0.8, sample size calculation was 405 events based on a Cox regression model. When the relative risk was assumed to be 1.5 then number of events was 170. Based on a Poisson regression model, relative risk=1.3, exposure:non-exposure=1:2 and power=0.8 rendered 385 events. Relative risk of 1.5 resulted in a total of 157 events. We calculated person-years (PY) with event numbers and cancer incidence rate in the non-exposure group. Based on a Cox regression model, with relative risk=1.3, exposure:non-exposure=1:2 and power=0.8, 136 245PY was needed to secure the power. In a Poisson regression model, with relative risk=1.3, exposure:non-exposure=1:2 and power=0.8, person-year needed was 129517PY. A total of 1939 cases were identified in KREEC until December 2007.
Conclusions: A retrospective power calculation in an ongoing study might be biased by the data. Prospective power calculation should be carried out based on various assumptions prior to the study.

Key words: Cohort studies, Epidemiologic studies, Power, Sample size, Research design
J Prev Med Public Health 2010;43(6):543-548

서 론

역학 연구를 진행함에 있어 연구자가 연구를 타당하게 수행하기 위해서는 연구 계획단계에서 연구가설과 평가 변수를 먼저 명확히 선정하는 것이 필요하다. 또한 연구자가 수용 가능한 오류 안에서 연구 가설을 통계적으로 확증하기 위해서는 필요한 대상수가 얼마인지 계산하는 대상수 산출의 과정을 거쳐야 한다. 대상수 산출을 위해 필요한 정보들은 통계 가설(귀무가설과 대립가설), 단측검정 혹은 양측검정여부, 예상되는 효과 크기, 제 1종 오류와 제 2종 오류가 기본적으로 필요하게 되며, 그 외에도 가설 증명에 적절한 연구 설계, 연구 설계에 따른 적절한 통계분석 방법과 이에 따른 부수적인 값들이 필요하게 된다 [1,2].

연구자가 연구 가설을 검정하기 위해 사용되는 통계적 가설 검정은 귀무가설과 대립가설을세우는 것에서부터 시작한다. 그 후 모아진 자료와 가설과의 일치성을 비교하거나 각 가설에서 자료의 결과가 나왔을 확률을 계산하여 가설 검정을 하게 된다. 대립가설에 대해서만 일치도를 보게 되면 논증에서의 일반원칙인 엄밀성(general principle of parasimony)에 위배되게 된다 [2]. 따라서 대립가설뿐 아니라 귀무가설에 대해서도 자료와의 일치성 정도를 같이 검토하는 것이 필요하다. 일치성을 검정하게 될 때 제 1종 오류를 5%로 설정하는 것은 5%의 오류가 있을수 있음을 감수하고 귀무가설을 부정하여 상대적으로 연구자의 가설을 지지하겠다는 것이다. 제 2종 오류를 20%로 설정하는 것은 20%의 오류가 있을수 있음을 감수하고 귀무가설을 채택하여 상대적으로 연구자의 가설을 지지하지 않겠다는 것이다. 즉 연구대상수 산출에서 제 1종 오류와 제 2종 오류를

책임저자 : 안윤옥 (서울특별시 종로구 대학로 103번지, 전화 : 02-740-8322, 팩스 : 02-747-4830, E-mail: yoahn@snu.ac.kr)
접수 : 2010년 6월 9일, 채택 : 2010년 10월 12일

Table 1. Korea radiation effect and epidemiology cohort

	Near	Intermediate	Far
Participants (1992-2005)	11383 (male:4496, female:6887)	10357 (male:4459, female:5898)	14489 (male:6060, female:8429)
Person-year during follow-up (1993-2009)	113438PY (male:48256, female:65182)	99734PY (male:44838, female:54896)	129306PY (male:52421, female:76885)
Identified cancer incidence (up to Dec. 2007)	609 (male:344, female:265)	617 (male:361, female:256)	713 (male:436, female:277)
Radiation related cancer (up to Dec. 2007)	366 (male:219, female:147)	363 (male:215, female:148)	438 (male:284, female:154)
Standardized incidence rate of caner (100000PY)	male:387.0 female:233.8	male:377.1 female:252.7	male:376.4 female:195.9

PY: person-years.

5%, 20%로 설정한다는 것은 유의수준 5%에서 귀무가설을 검정하되 참인 대립가설을 채택하지 못할 확률을 20%미만으로 하기 위한 최소의 자료 수를 계산한다는 의미이다.

이러한 연구대상수 산출을 거쳐 (또는 연구대상수 산출을 하지 않고) 연구를 진행하는 도중에 연구자가 어느 시점까지 진행된 연구가 과연 어느 수준의 통계적 검정력을 갖는지 궁금할 수 있다. 하지만 분석 계획에 따라 분석을 하는 것이 아니고 중간에 여러 번 분석을 하게 될 경우 제 1종 오류의 증가를 유발하게 된다 [3]. 따라서 분석 계획에 따라 분석을 하는 것이 중요하다. 또는 연구가 끝나는 시점에서 통계적으로 유의하지 않은 분석결과(귀무가설을 부정할 수 없는 결과)가 나왔을 때, 그 결과가 실제적인 상황을 나타낸 것인지, 아니면 검정력이 부족해서 그런 것인지 알기 어려운 경우, 검정력 계산을 고려하게 된다.

본 연구에서는 실제로 진행 중인 원전코호트의 자료를 이용하여 대상수 산출과 검정력을 계산하여 그 타당성에 대해 고찰하였다. 연구에 사용된 자료는 원전코호트 자료로 원전코호트의 목적(대립가설)은 '원전 주변에 거주하게 되면 암발생 위험도가 상승하게 되는가' 를 확인하기 위함이다. 이를 위해 원전 주변지역 주민 및 대조지역 주민을 비교하고, 원전종사자 및 대조종사자를 비교하기로 하였으며 비교 암종은 모든 암종과 전리방사선 관련 암(위암, 간암, 폐암, 갑상샘 암, 백혈병, 골암, 유방암)으로 계획하였다. 원전코호트의 구성은 원전지역 주민 코호트(원전 주변 지역군, 근거리 대조군, 원거리 대조군)와 종사자 코호트(원전종사자군, 대조 종사자군)로 구성되어 1992년부터 2005년까지 기반코호트가 구축되었다. 기반조사자료는 거주지역을 비롯하여 대상자의 연령, 생활수준, 결혼 상태 등의 인구학적 요인, 흡연, 음주, 운동 등의 생활습관 요인, 직업력, 질병과거력, 가족력 등의 요인과 생화학적 및 이학적 검사를 수집하였고 추적관찰조사를 통하여서 암의 발생과 암으로 인한 사망 등을 수집하고 있다. 암 발생 추적관찰조사 결과 주변지역, 근거리 및 원거리 대조지역 주민 코호트에서 2007년 말까지의 암 발생자는 1939건으로 확인되었다 (Table 1). 그런데 원전코호트는 계획단계에서 적정대상규모를 산정하지 않았었고 검정력을 계산하지 않은 채 연구가 진행 되었다. 따라서 현재의 결과로 통계적분석을 하는데 충분한 대상규모인지 판단할 근거가 없으며, 적정대상규모가 없어 연구의 종료 시점을 정할 근거 또한 없는 셈이기에 현 시점에서 원전코호트 연구의 추적관찰규모(관찰대상자 수 및 관찰기간)을 재 확인하고 통계적으로 유의한 검정력을 확보하는 것은 연구결과의 도출, 연구 종료 시점의 지정 및 연구의 질 평가에 있어서 매우 중요한 내용이다. 이러한 문제점을 해결하고 근거를 제시하기 위해 본 연구가 진행되었다.

방 법

원전 코호트는 원전 주변지역 주민 및 대조지역 주민을 대상으로 하는 코호트와 원전종사자 및 대조 종사자를 대상으로 하는 코호트 2개가 합쳐져 있는 형태이다. 이번 연구는 원전 주변지역 주민 및 대조지역 주민을 대상으로 하는 코호트를 대상으로 진행하였다.

연구 계획 당시의 암 발생률을 예측하기 위해 '1993-1997 서울시 전체 암의 성별, 연령군별 발생률' 을 이용하였다. 이유는 연구 계획 당시인 1990년경의 암발생자료에 대한 통계자료가 없어서 가장 근접한 통계자료로 '1993-1997 서울시 전체 암의 성별, 연령군별 발생률' 을 이용하였다 (Table 2). 각 지역의 인구자료는 통계청의 1990년 행정구역별 연령별 인구 구조를 사용하였다. 이는 연구계획시점인 1990년의 인구구조를 기반으로 계산하기 위함이다. 암 발생률 예측시 통계청 자료의 연령구분(20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60세 이상)에 적합하도록 '1993-1997 서울시 전체 암의 성별, 연령군별 발생률' 을 수정하여 적용하였다. 비노출군과 노출군에서의 암발생

Table 2. 1993-1997 Incidence of total cancer by age group, gender, Seoul

Age	Incidence of cancer (per 100000 person-year)	
	Male	Female
0 - 4	18.2	14.4
5 - 9	12.2	8.3
10 - 14	11.2	10.6
15 - 19	13.3	13.4
20 - 24	14.2	22.4
25 - 29	20.6	37.8
30 - 34	38.0	80.8
35 - 39	69.5	124.8
40 - 44	137.8	182.2
45 - 49	247.5	260.3
50 - 54	445.2	334.5
55 - 59	730.1	454.9
60 - 64	1156.7	616.6
65 - 69	1717.4	798.0
70 - 74	2317.5	996.5
75 - 79	2674.7	1089.9
80 - 84	2610.4	1069.0
Over 85	2299.9	843.4
Total	186.7	163.9

률을 직접표준화법을 이용하여 산출하였다. 각 코호트는 원전 주변지역 주민 코호트, 근거리 대조지역 주민 코호트, 원거리 대조지역 주민 코호트를 1:1:1로 모을 계획이므로 노출군(원전 주변지역 주민 코호트)과 비노출군(대조지역 주민 코호트)을 1:2로 계산하였다. 코호트 내 지역별 인원은 원전 주변지역 주민 코호트는 고리, 울진, 영광, 월성에서 원전으로부터 반경 5 km이내 거주 주민, 근거리 대조지역 주민 코호트는 원전으로부터 반경 5 km이상부터 30 km 이내 거주 주민, 원거리 대조지역 주민 코호트는 양평, 함안, 충주지역 거주주민으로 코호트별 각 지역에서 같은 수로 모을 것으로 계획하였다. 위험도는 노출에 따른 지역간의 암 발생을 비교하는 것이기 때문에 식이분석에서와 같이 노출이 광범위하게 일어나고 있다고 가정하고 효과도 식이와 비슷하리라 가정하였다. 대부분의 영양역학의 위험도에 대한 점추정치는 0.5-2.0로 나오므로 본 연구에서 1.3과 1.5를 예상위험도로 설정하였다 [4]. 대상수산출에는 Cox regression과 Poisson regression에 따라 여러 조건하에서 산출 하였다. 통계 방법을 여러 가지 사용한 것은 연구의 목적이 확실하지 않을 경우 가능한 것을 모두 계산하여 그 중에 보수적으로 대상수가 많은 것을 사용하기 위해서이다. 대상수산출은 PASS 2002 (NCSS, LLC. Kaysville, Utha, USA)를 사용하였다. 산출 된 대상수에 직접 표준화법으로 계산한 예상 암발생률을 적용하여 남녀를 동수로 모을 것을 가정할 때 예상되는 인년을 계산하였다.

1. Cox regression: Time to event data

코호트 연구나 임상 시험에서는 연구에 참여하는 시기가 다르거나, 중도절단으로 관찰 기간이 각기 달라진다. 이들 연구에서는 time to event data가 모인다 즉, 사건 발생 여부가 아닌 , 사건이 발생하기까지의 시간이 결과변수가 된다. 카플란마이어 곡선을를 이용하여 범주형 설명변수의 각 범주에서 생존곡선 비교가 가능하며 로그-랭크 검정을 통해 생존곡선의 유의한 차이를 평가 한다. 하지만 여러 개의 설명변수가 존재할 경우 multiple regression model이 필요하다. time to event data의 분석에는 Cox regression model이 의학연구에서는 가장 널리 사용되고 있다. Cox medel은 생존 시간의 평균보다 위험률에 초점을 맞춘 분석법이다 [5-8].

2. Poisson regression: Frequencies and rates

우리가 관찰하고자 하는 사건이 각각의 사건 생길 때까지의 시간이 독립적이고 지수 분포를 따르는 질병이라면 Poisson regression을 사용할수 있다. Poisson regression 분석은 단위 시간당 사건의 평균 발생 건수를 종속변수로 하여 각 설명변수와의 관련성이 상대위험도로 추정된다 [5-9].

결 과

Cox regression model 에서 상대 위험도=1.3, 노출군:비노출군=1:2 검정력 0.9일 경우 560의 사건수가 필요하다고 나왔으며 같은 조건에서 검정력이 0.8일 경우 405의 사건수가 필요하다고 나왔다. 상대 위험도가 1.5일 경우 검정력 0.9이면 235사건수, 검정력 0.8이면 170사건수였다. Poisson regression model 은 상대위험도가 1.3이고 나머지는 같은 조건일 경우 각각 524, 385로 Cox model보다 적은 수의 대상수가 산출 되었다. 상대위험도가 1.5일 경우 검정력 0.9이면 212사건수, 검정력 0.8이면 157사건수였다 (Table3).

'1993-1997 서울시 전체 암의 성별, 연령군별 발생률' 에서 20대 이상의 모든 암 발생률은 서울 남자 182.7명, 여자 151.9명을 나타냈다. 지역별 예상 암발생률을 구하여 노출군에서의 모든 암 발생률은 10만인년당 남자 344.7명, 여자 253.7명, 비노출군에서의 모든 암발생률은 10만인년당 남자 346.9명, 여자 247.6명이었다 (Table 4).

대상수 산출과 비노출군에서의 암발생률을 근거로 남녀를 같은 수로 모집할 경우 각 대상수를 모으기까지 필요한

Table 3. Event of cancer incidence by model, relative risk and exposure:non-exposure ratio

Model	RR	Power	Events		
			Exp: Non-exp 1:1	Exp: Non-exp 1:2	Exp: Non-exp 1:3
Cox regression	1.3	0.9	498	560	664
		0.8	360	405	479
	1.5	0.9	209	235	278
		0.8	151	170	201
Poisson regression	1.3	0.9	473	524	612
		0.8	345	385	451
	1.5	0.9	193	212	247
		0.8	142	157	183

RR: relative risk, Exp: exposure group, Non-exp: Non-exposure group.

Table 4. Expectation of incidence of cancer by areas

Area	Incidence of cancer (per 100 000 person-year)	
	Male	Female
Exposure	344.7	253.7
Non-exposure	346.9	247.6

Table 5. Expectation of person-year

Model	RR	Power	Person-year		
			Exp: Non-exp 1:1	Exp: Non-exp 1:2	Exp: Non-exp 1:3
Cox regression	1.3	0.9	167 531	188 389	223 375
		0.8	121 107	136 245	161 140
	1.5	0.9	70 309	79 056	93 521
		0.8	50 798	57 189	67 618
Poisson regression	1.3	0.9	159 121	176 278	205 882
		0.8	116 061	129 517	151 720
	1.5	0.9	64 927	71 319	83 093
		0.8	47 770	52 816	61 563

RR: relative risk, Exp: exposure group,
Non-exp: Non-exposure group.

인년는 Cox regression model에서 상대위험도 1.3, 노출군:비노출군 1:2, 검정력 0.8 일경우 136,245인년, Poisson regression model 에서 상대위험도 1.3, 노출군:비노출군 1:2, 검정력 0.8 일경우 129517인년이었다 (Table 5) .

고 찰

연구를 진행하면서 검정력을 계산하는 경우는 크게 2가지 경우이다. 첫번째 연구의 종료 시점을 결정하고 싶을 때와 두번째 분석한 결과의 검정력을 알고 싶을 때이다. 모든 자료가 모아진 상황에서라면 자료를 분석하고 분석되어 나온 자료를 가지고 검정력을 계산할 수 있다. 하지만 이렇게 후향적 검정력 계산(retrospective power, post hoc power, observed power)을 하는 것에 대해서는 논란이 많이 있다. 주된 이유로는 이미 모아진 자료를 가지고 검정력을 구하게 되면 자연히 자료가 가진 특성과 유사하게 검정력이 계산되기 때문이다. 의미가 있다고 나온 결과에서의 검정력은 높게 나오게 되고 의미가 없다고 나온 결과에서의 검정력은 낮게 나온다는 것이다. 다시 말해서 '낮은 검정력 때문에 분석결과가 의미가 없다.' 고 판단하거나 혹은 '검정력이 높은 분석이다.' 와 같이 검정력 자체로 연구 결과를 해석하는 것은 잘못된 것이라는 이야기다. 많은 통계학자들이 다른 방법으로 여러가지 검정력 계산법을 고안하였다. 신뢰구간으로 검정력을 계산한다거나 연구 모델에 따른 특정 수치를 이용해 검정력을 계산하는 방법이다. 하지만 이들 방법도 위에 명시한 논란에서 자유롭지 못하다 [6, 9-18]. 따라서 본 연구에서는 모아진 자료를 근거로 검정력을 계산하지 않고 독립적인 근거를 이용한 대상수 산출을 하였으며 이를 바탕으로 인년을 계산하였다. 독립적인 근거를 사용하기 위해 모아진 자료에서 나온 비교위험도를 사용하지 않고 처음에 연구 계획시 설정한 예상위험도를 사용하였다. 또한 암발생률도 현 시점에서의 통계자료나 모아진 자료에서 계산하여 이용하지 않고 연구계획 당시의 1990년경의 암발생자료를 사용하기 위해 '1993-1997 서울시 전체 암의 성별, 연령군별 발생률' 과 통계청의 1990년 행정구역별 연령별 인구 구조를 사용하여 암발생률을 계산하여 사용하였다. 이를 원전코호트 자료와 비교함으로써 검정력을 추정하였다. 원전코호트의 원전 주변지역 주민 및 대조지역 주민을 대상으로 하는 주변지역, 근거리 및 원거리 대조지역 주민 코호트에서 확인된 암 발생자는 1939건으로 연구계획의 조건인 상대위험도 1.5, 노출군과 비노출군이 1:2로 모아질 때 Cox model에서 검정력 80%이상 확보하는 대상수 170을 넘어서기 때문에 본 연구는 검정력을 확보했다고 판단하면 안된다. 다음의 조건을 고려해야 한다. 상대위험도의 크기가 달라질 수 있다. 상대위험도가 1에 가까울수록 필요한 대상수는 많아지게 된다. Cox model에서 상대위험도가 1.3으로 1에 가까워 졌을 때, 검정력 0.8, 노출군과 비노출군이 1:2로 모을 때의 필요사건수인 405사건수, 같은 조건에서 검정력이 0.9일때의 필요사건수인 560건으로 늘어남을 확인할 수 있다. 현재 모아진 자료는 노출군과 비노출군의 모집이 1:2보다는 1:3에 더 가깝다. 이 경우에도 검정력 0.8일 때 필요사건수 479건, 검정력 0.9일 때 필요사건수 664건으로 늘어나게 된다. 즉 예상 상대위험도가 1.3, 유의수준 5%이하, 검정력 90%이상 확보하기 위한 최소 대상수가 664건이다. 또한 인년을 비교하여도 Cox model에서 상대위험도 1.3이고 노출군:비노출군이 1:2, 검정력 0.8일 때 136 245인년과 검정력 0.9 일 때

188,389인년으로 계산되었으며 노출군과 비노출군이 1:3일 경우에도 검정력 0.8에서 161 140인년과 검정력 0.9에서 223 375인년을 넘는 342 248인년을 관찰 하였다. 전체 암보다 발생수가 적은 전리방사선 관련 암의 발생수도 2007년 말까지 1167건으로 필요 사건수를 넘긴것으로 판단된다.

위에서 대상수 산출의 기본요소인 예상 상대위험도의 변화와 노출군과 비노출군의 비율을 변화할 수 있기 때문에 여러 가지 조건하에 대상수를 산출하고 현재와 유사한 상황을 비교하였다. 이외에도 대상수 산출에서 추적 관찰의 손실(follow up loss)을 고려해봐야 한다. 관찰기간동안 10%의 추적 관찰의 손실이 예상된다면 이를 더 모아야 한다. 또한 다른 보정요인을 제외하고 대상수를 계산하였다. 하지만 실제 다른 보정요인이 존재할것이고 방사선에 의한 암 발생 설명력을 1보다 낮을 것이다. 이럴 경우에도 필요한 대상수는 늘어나게 될 것이다. 인년을 계산할 때 다음의 요소를 더 생각해봐야 한다. 각 지역의 연령구조를 코호트가 반영 할 것이라고 예상하였고, 남녀의 모집인원은 동수, 지역별 모집인원도 동수 등 여러 가정하에 인년을 계산하였다. 하지만 연구진행에서 이를 지키려 노력하였음에도 연령구조가 다르게 모집되고, 지역별 모집인원이 달라지는 등 여러가지에서 예상과 다른 수치를 보이게 되었다. 따라서 인년을 직접 비교하는 것에는 대상수 비교보다 많은 바이어스가 개입된다고 생각할 수 있다.

이론적으로 코호트를 계획할 때 필요 사건수를 얻기 위한 필요 인년을 계산하고 이를 이루기 위해 필요한 인원과 관찰 기간을 계산하게 된다. 이는 인원을 고정하여도 관찰기간을 늘리면 필요 사건수를 얻을수 있다는 논리로 설명될수 있다. 하지만 암과 같이 40대 이상에서 호발하고 발생률이 낮은 질환에서는 현실적으로 관찰 기간이 제한적일 수 있다. 즉 암의 발생을 70을 얻기 위해 100명을 70년 관찰해야 한다면 40대 이상을 모은 코호트에서 이런 관찰기간을 바라는 것은 불가능하다. 즉 처음부터 인원을 늘려야 한다.

일반적으로 연구대상수를 산출시 제 2종오류를 0.2로 설정하여 구한 연구상수를 연구에서 모아진 자료와 단순비교를 한다. 하지만 이렇게 단순비교하여 대상수 이상을 모으면 검정력이 80%이상으로 높아진다고 판단하는 것은 잘못된 판단이다. 대상수에 영향을 끼치는 요인은 제 1종오류, 제 2종오류, 상대위험도의 크기등 다양한 요인이 영향을 끼치게 된다. 특히 대상수 산출시에 상대위험도의 크기또한 연구자가 예상하여 정한다. 즉 연구 계획시 정한 상대위험도의 크기를 증명할 때 필요한 대상수의 산출인 것이다. 하지만 연구를 진행하면 상대 위험도가 예상보다 클 수도 있고 작을 수도 있다. 본 연구에서는 상대위험도가 예상보다 작아 처음 예상 위험도인 1.5보다 낮은 1.3을 계산하여 비교하였다.

현재 구축되어 있는 코호트 자료를 가지고 여러 조건에서의 대상수를 산출하고 이를 비교하며 검정력을 추론해볼 수 있지만 이때 또 한가지 고려해야할 점이 있다. 그것은 상대위험도의 크기와 검정력의 상관관계이다. 어떠한 연구에서 실제적으로 의미가 있는 상대위험도의 크기가 최소 1.5라고 할 때 실제 자료에서 추정된 상대위험도는 1.5보다 작은 1.3이었음에도 통계적으로 유의한 결과를 얻었다면 연구 결론은 어떻게 내려져야 하는가이다. 이런 경우 과도한 대상수 모집으로 인해 검정력이 높아져 실제로 의미가 없는 결과가 통계적으로 유의하게 나온 것은 아닌지 고민해봐야 한다. 통계적으로 의미가 있다고 해서 실제로 효과의 차이에 의미가 있다는 얘기는 아니기 때문이다. 연구자가 주장하고 싶은 차이가 실제적인 의미가 있는지 자료가 이를 지지하는지 예상된 차이가 나왔는지 확인하여야 한다.

또한 모아진 자료가 선택 바이어스에 의해 과도한 사건수가 모인 것은 아닌지 확인해봐야 한다. 그 외에도 암과 같이 빈도가 낮고 한정된 사람에게서 발생할 수 있는 암은 한정되어 있다고 볼 수 있다. 이런 상황에서는 추적관찰만으로는 얻을 수 있는 사건수가 한정 될 수 있다. 암의 종류를 구분하지 않았지만 암의 종류에 따라 경중의 차이가 있을 수 있고 이것이 군별 다른 비율로 포함되게 될 경우 추적 관찰기간에 영향을 끼칠 수 있다. 암의 종류에 따라 잠복기와 임상전단계(질병전단계)가 다르게 된다. 코호트 분석시 잠복기와 임상전단계(질병전단계)를 고려하여 연구설계를 하여야 한다. 본 연구에서 대상수 산출시 중도절단을 고려하지 않고 대상수를 산출하였으나 연구를 진행하게 되면 중도절단이 생기게 되므로 이를 고려하여야 한다.

연구대상수 산출과 검정력은 귀무가설이 거짓일 때 이를 받아들이지 않을 확률을 적절히 확보하기 위해서 매우 중요하다. 그러나 연구가 진행중일 때 하는 후향적 검정력 연구(retrospective power, post hoc power, observed power)와 체계적 문헌고찰 같은 후향적 연구, 분석(retrospective analysis)를 하기 위한 검정력 연구는 다름을 인지하여야 한다.

연구대상수 산출과 검정력 계산은 연구계획단계에서 수행되어야 하며 연구 중반에 이를 수행해야 할 경우 그 연구와 독립적인 근거를 이용하여야 한다. 만약, 그 연구의 중반까지의 자료를 근거로 검정력 분석(power analysis)을 하는 경우, 최종 연구 결과에 대한 제 1종 오류는 5%보다 커질 것이다. 본 연구에서는 독립적인 근거를 사용하여 다양한

상황에서의 연구대상수를 산출하였으며 초기단계처럼 가정하여 Cox model을 이용하여 원전코호트 연구의 검정력을 추정하였다. 처음 예상치, 노출군과 비노출군이 1:2, 상대위험도 1.5, 검정력 0.8에서 필요한 사건수를 구하였다. 하지만 모아진 자료와 유사한 상황을 다시 가정하여 필요한 사건수를 구하여 이를 원전코호트 추적조사를 통하여 현재까지 확인된 암발생자수(1939건), 전리방사선 관련 암으로 확인된 사건수(1167건)와 비교하여 검정력을 추정하였다. 이는 비교위험도가 1.3고 나머지 상황이 모아진 자료와 유사한 상황임을 가정하고 분석을 하였을 때 나온 결과의 타당성을 지지한다.

참고문헌

1. Ahn YO, Yoo KY, Park BJ, Kim DH, Bae JM, Kim DH, et al. *Epidemiology: The Principles and Applications*. Seoul: Seoul National University Press; 2005 p. 205-222. (Korean)
2. Ahn YO, Yoo KY, Park BJ. *Manual for Medical Statistics*. Seoul: Seoul National University Press; 2005 p. 183-242. (Korean)
3. Armitage P, McPherson CK, Rowe BC. Repeated significance tests on accumulating data. *J Roy Stat Soc Ser A* 1969; 132(2): 235-244.
4. Willett W. *Nutritional Epidemiology*. 2nd ed. New York: Oxford University Press; 1998. p. 1-18.
5. Bender R. Introduction to the use of regression models in epidemiology. *Methods Mol Biol* 2009; 471: 179-195.
6. Breslow NE, Day NE. Statistical methods in cancer research. Volume II--The design and analysis of cohort studies. *IARC Sci Publ* 1987;(82): 1-406.
7. Singh H, Mahmud SM. Different study designs in the epidemiology of cancer: case-control vs. cohort studies. *Methods Mol Biol* 2009; 471: 217-225.
8. Xue X, Hoover DR. Statistical methods in cancer epidemiological studies. *Methods Mol Biol* 2009; 471: 239-272.
9. Brown CC, Green SB. Additional power computations for designing comparative Poisson trials. *Am J Epidemiol* 1982; 115(5): 752-758.
10. Blettner M, Ashby D. Power calculation for cohort studies with improved estimation of expected numbers of deaths. *Soz Praventivmed* 1992; 37(1): 13-21.
11. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L Erlbaum Associates; 1988. p. 1-18, p. 531-542.
12. Goodman SN. Confidence limits vs power, calculations. *Epidemiology* 1994; 5(2): 266-268.
13. Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med* 1994; 121(3): 200-206.
14. Gordon I. Sample size estimation in occupational mortality studies with use of confidence interval theory. *Am J Epidemiol* 1987; 125(1): 158-162.
15. Hoenig JM, Heisey DM. The abuse of power: the pervasive fallacy of power calculations for data analysis. *Am Stat* 2001; 55(1): 19-24.
16. Howe GR, Chiarelli AM. Methodological issues in cohort studies. II: Power calculations. *Int J Epidemiol* 1988; 17(2): 464-468.
17. Lubin JH, Gail MH. On power and sample size for studying features of the relative odds of disease. *Am J Epidemiol* 1990; 131(3): 552-566.
18. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. *Annu Rev Psychol* 2008; 59: 537-563.