

보건정보통계학회지 제39권 제2호
ISSN 2287-3708(Print) ISSN 2287-3716(Online)
Journal of Health Informatics and Statistics (JHIS)
2014; 39(2): 32-43

한국 여성의 유방암 발생에 대한 코호트 효과

조호진¹⁾, 주우현¹⁾, 김윤남²⁾, 배종면³⁾, 남정모^{1,4)†}

¹⁾연세대학교 의학통계학과, ²⁾연세의료원 임상시험센터

³⁾제주대학교 의학전문대학원 예방의학교실, ⁴⁾연세대학교 의과대학 예방의학교실

Cohort Effects of Female Breast Cancer Incidence in Korea

Ho Jin Cho¹⁾, Woo Hyun Joo¹⁾, Youn Nam Kim²⁾, Jong Myon Bae³⁾,
Chung Mo Nam^{1,4)†}

¹⁾Department of Biostatistics, Yonsei University College of Medicine

²⁾Division of Clinical Data Management Research, Clinical Trials Center, Severance Hospital, Yonsei University Health System

³⁾ Department of Preventive Medicine, School of Medicine, Jeju National University

⁴⁾Department of Preventive Medicine, Yonsei University College of Medicine

Abstract

Objectives: The purpose of the study is to review various methods in age-period-cohort (APC) analysis and to provide a guideline to choose adequate method evaluating age, period, and cohort effects. We investigated age, period, and cohort effects of breast cancer incidence between 1999 and 2011 in Korea.

Methods: Data on female breast cancer incidence from 1999 to 2011 were drawn from the Korean national statistical office. The 5-year period of data units (1999-2003, 2004-2008, and 2009-2011) and 5-year age interval (30-34-80-84) were used to calculate 13 birth cohorts. The graphical approach, constrained generalized linear model (CGLM) approach, median polish approach and intrinsic estimator (IE) approach were used to estimate age, period, and cohort effects.

Results: The age and period effects existed significantly in CGLM, median polish, IE approaches. The breast cancer incidence increased along with age and period. However, there was a difference in cohort effect. For CGLM, positive cohort effects for recent cohort emerged significantly, but for the other methods, no significant effects shown.

Conclusions: While previous studies have used the CGLM method, CGLM depends on arbitrary parameter constraints. Therefore, we suggest median polish approach or IE approach for analyzing APC models to obtain more accurate results.

Keywords: APC model, Cohort effect, Breast cancer, CGLM, Median polish, Intrinsic estimator

[접수일: 2014년 10월 30일, 수정일: 2014년 12월 18일, 게재확정일: 2014년 12월19일]

† Corresponding Author: Chung Mo Nam, PhD

Department of Preventive Medicine / Department of Biostatistics, Yonsei University College of Medicine,
50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea. Tel: +82-2-2228-1871

E-mail: CMNAM@yuhs.ac

* This study was supported by funding from the Korean Foundation for Cancer Research in 2013 (no. 2013-2).

1. 서론

유방암은 한국 여성암 중 갑상선암 다음으로 가장 많이 발생하는 질병으로 최근 계속적으로 그 발생률이 증가하고 있다. 그러나 한국 여성의 유방암의 발생률은 나이가 증가하면서 계속 증가하는 것이 아니라 40-50세 사이에서 가장 높고 그 이후 연령에서 다시 감소하는 경향성을 보이고 있다 [1]. 하지만 대부분의 외국 사례를 보면 유방암 발생률은 나이가 증가할수록 증가하는 경향성을 보이며 60세 이상에서 발생률이 가장 높았다 [2-5]. 국내와 외국의 유방암 발생률 분포가 상이하기 때문에 어떠한 이유에서 이런 차이가 발생하는지 그 이유를 찾는 것은 매우 중요한 문제이다. 외국 자료를 이용한 유방암 발생률에 대한 연령-기간-코호트(age-period-cohort, APC) 분석 결과는 제시되어 있지만 국내 유방암에 대한 결과는 제시된 바 없다 [6,7]. 그러므로 본 논문을 통해 APC 분석을 실시해 발생률에 대한 출생 코호트의 효과가 존재하는지 알아보고자 한다.

APC 모형이란 연령, 기간, 출생코호트에 따라 발생률이나 사망률의 시간적 경향성을 보기 위해 시행된다. 주요 목적은 시간이 지남에 따라 각각의 효과들이 발생률 등에 영향을 미치는지를 파악하는 것이다. 각각의 효과를 살펴보면, 연령 효과는 특정 나이에 관계된 효과로, 나이가 들어감에 따라 어떠한 원인에 대한 노출이 축적되거나 생리학적인 변화로 인한 효과라고 할 수 있다. 대부분의 경우 동일한 코호트 내에서 시간이 지남에 따라 연령 효과에 의해 특정 질병에 걸리는 비율이 높아지는 경향을 보인다. 기간 효과는 어떤 특정한 시점에서 널리 퍼져있는 위험요인에 노출되는 등의 환경의 변화에 따라 발생률이 연령이나 코호트에 관계없이 나타나는 효과이다. 코호트 효과는 태어난 시점에 따른 효과이다. 같은 출생 코호트에 속한 사람들끼리 시간이 지남에 따라 같은 위험을 갖는다고 볼 수 있다. 어떤 새로운 환경에서

연령별로 다른 양상을 보일 경우 이를 코호트 효과라고 정의한다. 예를 들어, 특정 시점에서 전 연령에서 병에 걸리는 비율이 크게 증가하는 전염병이 발생했다면 이를 기간 효과로 보고 특정 연령에서 특히 병에 많이 걸렸다면 이를 코호트 효과로 본다는 것이다 [8].

위의 효과들을 고려해 APC 연구 모형을 세울 때 연령, 기간, 코호트 효과가 결과 변수와 선형적 관계에 있다고 할 수 있다. 하지만 이러한 경우 추정에 대한 식별(identification)의 문제가 발생한다. 세 변수가

코호트 = 기간 - 연령(또는, 기간 = 연령 + 코호트)라는 완전한 선형 관계를 가지기 때문에 공선성이 존재하여 연령, 기간, 코호트 효과를 동시에 추정할 수 없는 문제가 발생한다. 이러한 문제점 때문에 추정을 위해 추가적인 제한이 필요하다.

본 논문에서는 식별의 문제를 해결할 수 있는 방법들 중 APC 연구모형을 구축할 때 역학 분야에서 가장 많이 쓰이는 분석 방법인 제약된 일반선형회귀모형(constrained generalized linear model, CGLM) 방법의 한계점을 언급하고, 이를 보완할 수 있는 통계학적 방법론을 제시한다. 또한 소개한 방법들을 실제 우리나라 유방암 발생률 자료에 적용하여 각 방법들의 장단점을 알아보고 우리나라 여성 유방암 발생률에 코호트 효과가 영향을 미치는지, 외국 사례와의 차이점이 있는지 등을 알아보려고 한다.

2. 연구 방법

1) 분석 자료

본 연구는 통계청 국가통계포털의 24개 암종/성/연령(5세)별 암발생자 수, 발생률 자료를 이용하였다. 사용한 자료는 1999년부터 2011년까지 한국 30세 이상 여성의 3개의 5년 단위 기간(1999-2003, 2004-2008, 2009-2013)과 연령은 5세 간

격(30-34세~80-84세) 유방암 발생률이다 (Table 1) [9]. APC 모형을 분석하기 위해서는 연령과 기간을 동일한 구간(예, 5년 단위)으로 묶어야 하나 아직까지 2012-2013년의 발생률 자료가 없기 때문에 2009-2011년 자료의 평균치로 2009-2013년 자료를 대체하였다. 또한 30세 미만에서는 유방암이 거의 발생하지 않기 때문에 30세 이상의 자료만 사용하였고, 84세까지의 자료를 사용하였다. 이 자료를 이용해 13개의 코호트를 계산하여 분석에 사용하였다.

발생률은 통계청에서 제시한 연령표준화발생률을 사용하였다. 이 때 표준인구는 우리나라 2000년 주민등록연앙인구를 사용하였다. 국제질병분류법(ICD-10)에 따라 유방암은 C50(유방의 악성 신생물)으로 정의하였다 [9].

Table 1. Age-specific incidence for female breast cancer, 1999-2011, Korea (per 100,000 persons)

Age	Period		
	1999-2003	2004-2008	2009-2011
30-34	20.32	24.76	27.97
35-39	40.74	51.14	61.50
40-44	67.40	90.42	104.97
45-49	84.86	121.58	144.93
50-54	78.30	104.12	133.07
55-59	65.48	89.84	112.67
60-64	50.70	81.82	102.8
65-69	39.42	59.22	81.13
70-74	31.90	47.58	65.67
75-79	25.04	36.86	46.10
80-84	22.02	28.16	35.87

2) 분석 방법

(1) 그래프를 이용한 방법(graphical approach)

그래프를 이용한 방법은 연령, 기간, 코호트 효과를 시각적으로 찾아내는 방법이다. 기간에 따른 연령별 발생률과 출생 코호트에 따른 연령별 발생

률을 도표로 나타낸다. 같은 기간이나 코호트 내에서 연령에 따라 발생률이 달라지는 경우 이를 연령효과로 보고, 모든 연령 그룹에서 기간에 따라 발생률이 달라지는 경우 이를 기간효과로 본다. 또한 연령별 발생률이 기간에 따라 평행하지 않은 경우나 발생률이 출생 코호트의 모든 연령대에서 상승되는 경우 이를 코호트 효과로 보는 것이다. 이 방법을 통해 대략적인 연령, 기간, 코호트 효과의 경향성을 쉽게 볼 수 있지만 수치적으로 정확한 값이나 유의성 등을 파악하기는 어렵다 [10,11].

(2) 제약된 일반선형회귀모형(CGLM)

제약된 일반선형회귀모형은 일반적으로 가장 널리 쓰이고 있는 방법으로 연령, 기간, 코호트 효과 중 최소 한 가지 이상의 모수를 제약하는 방법이다. 먼저 발생률에 대한 모형을 다음과 같이 정의한다.

$$\text{Log}(Y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ij} \\ (i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = m - i + j) \cdots \text{식}(1)$$

여기서 Y_{ijk} 는 해당 칸의 발생률을 뜻하고 α_i 는 m 개의 연령 효과 중 i 번째 효과, n 개의 기간 효과 중 j 번째 효과, γ_k 는 $m+n-1$ 개의 코호트 효과 중 k 번째 효과를 뜻한다. 이 때, 발생률 Y_{ijk} 가 포아송(poisson) 분포를 따른다고 가정한다. 이 방법은 코호트 효과가 발생률에 미치는 직접적인 영향을 볼 수 있고 사용이 편리하다는 장점이 있다. 하지만 완전한 공선성이 발생하기 때문에 최소 1개 이상의 모수를 제한하여야 하고 어떤 모수를 제약하는지에 따라 결과가 매우 민감하게 변화하기 때문에 이러한 점이 문제가 되고 있다 [12,13].

본 연구에서는 다음과 같은 세 가지 서로 다른 제약을 두고 추정치를 구하고자 한다. Figure 1을 통해 첫 번째 연령과 두 번째 연령의 효과가 같다고($\alpha_1 = \alpha_2$) 두는 연령 제한 모델, 첫 번째 기간과

두 번째 기간의 효과가 같다고($\beta_1 = \beta_2$) 두는 기간 제한 모델, 첫 번째 코호트와 두 번째 코호트의 효과가 같다고($\gamma_1 = \gamma_2$) 두는 코호트 제한 모델을 설정하고 각각의 경우 추정된 회귀식의 계수를 비교해본다.

(3) Median Polish 방법

Median Polish 방법은 코호트 효과를 연령과 기간의 교호작용이라고 정의하고 다음과 같이 코호트 효과를 추정한다. 먼저 연령과 기간에 대한 다음과 같은 이요인모형을 구축한다.

$$\begin{aligned} \text{Log}(Y_{ijk}) &= \mu + \alpha_i + \beta_j + \epsilon_{ij} \\ (i &= 1, 2, \dots, m, j = 1, 2, \dots, n) \end{aligned}$$

여기서 ϵ_{ij} 는 i 번째 연령 효과와 j 번째 기간 효과로 설명되지 않는 오차항이다. 분할표를 이용하여 각각의 행과 열의 중앙값을 반복적으로 제거하여 잔차를 추정한다. 오차항은 다음과 같은 회귀식을 통해 나타낼 수 있다.

$$e_k = \mu + \gamma_k + e_{ijk}$$

이 때, e_k 를 Median Polish의 잔차라 하면 이는 μ 라는 절편, $m+n-1$ 개의 코호트 효과 중 k 번째 효과를 뜻하는 γ_k , 임의의 오차항 e_{ijk} 로 나누어 쓸 수 있다. Median Polish 방법은 γ_k 의 추정치와 유의성 검정을 통해 코호트 효과를 예측하고 검정한다 [8,10,13-15]. 이 방법의 경우 코호트 효과가 발생률에 직접적으로 미치는 영향을 나타내기 어렵지만 변수를 제한하는 방법이 아니므로 결과가 변수에 의존적이지 않고 동일하다.

(4) Intrinsic estimator (IE) 방법

Yang et al. [16,17]이 제안한 Intrinsic estimator (IE) 방법은 선형회귀모형에서 역행렬이 존재하지 않는 경우에 추정치의 식별 문제를 해결하기

위해 Moore-Penrose의 일반화 역행렬을 이용하는 방법이다. APC 모형 식(1)의 Moore-Penrose 추정치를 행렬과 벡터를 이용하면 다음과 같다.

$$\hat{b} = (X^T X)^- X^T Y$$

여기서 X 는 연령, 기간, 코호트 독립변수로 이루어진 계획행렬이다. 이 때 $(X^T X)$ 행렬의 고유치가 0에 대응되는 고유벡터 부분공간과, 이 공간과 직교하는 부분공간으로 분해할 수 있다. 이러한 모수 공간의 직교 분해를 통해 다음과 같은 식이 도출된다.

$$\begin{aligned} \hat{b} &= B + sB_0 \\ B &= (I - B_0 B_0^T) \hat{b} \end{aligned}$$

여기서 B_0 는 고유치가 0에 대응되는 고유벡터이다. B_0 벡터는 온전히 연령과 기간의 함수로, 결과 변수와는 관계가 없다. B 는 s 가 0일 때 기하학적 투사(geometric projection)를 만족하는 값, 즉 IE가 된다.

IE는 기간이 고정되었을 때 불편성을 가지고 제약된 일반선형회귀모형보다 작은 분산을 가지는 장점을 가진다 [12]. 또한 Median Polish 방법과 마찬가지로 변수에 관계없이 동일한 결과를 얻을 수 있다.

제약된 일반선형회귀모형 방법은 R 3.1.1을 사용하여 분석하였고 Median Polish 방법의 경우 R 패키지 'medpolish'를 사용하였다. IE 방법은 Stata version 13, apc_ie 패키지를 이용하여 분석하였다.

3. 연구 결과

1) 그래프를 이용한 방법(graphical approach)

기간과 출생 코호트에 따른 연령별 유방암 발

생률을 도표화 하였다 (Figure 1). 먼저 Figure 1의 기간에 따른 연령별 유방암 발생률 도표를 보면 기간이 증가할수록 유방암 발생률이 증가하는 것을 볼 수 있다. 또한 같은 기간 내에서는 연령이 증가함에 따라 발생률이 증가하다가 50세 이후로 감소하는 경향을 보이고 있다. 또한 모든 기간에서 발생률의 패턴이 동일하므로 이 그래프 상에서 코호트 효과는 없는 것으로 보인다.

출생 코호트에 따른 연령별 유방암 발생률 도표

에서는 최근 코호트로 올수록 발생률이 증가하다가 다시 감소하는 경향성이 있음을 볼 수 있다. 또한 40-50세 사이의 기울기가 가장 가파르므로 이 연령대의 발생률이 가장 높다고 볼 수 있다. 대부분의 경우 같은 코호트 내에서는 나이가 증가할수록 발생률이 증가하는 경향을 보이고 있음을 알 수 있다. 이를 통해 각 효과들의 통계학적 유의성은 알 수 없지만 연령, 기간, 코호트 효과가 유방암 발생률에 대략적으로 어떤 영향을 미치는지를 알 수 있다.

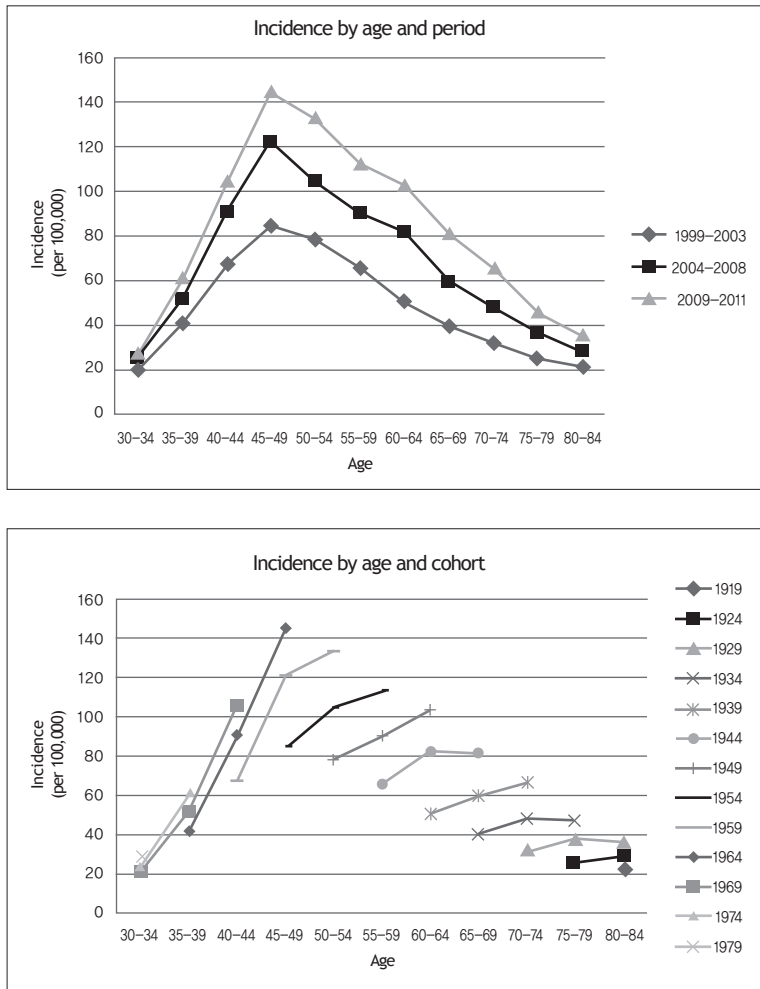


Figure 1. Age-specific incidence for female breast cancer by period and cohort, 1999-2011, Korea.

2) 제약된 일반선형회귀모형(CGLM)

연령, 기간, 코호트 효과 모두 제한을 어떻게 설정하느냐에 따라 추정치가 매우 달라짐을 알 수 있다 (Figure 2). Figure 2의 결과를 보면 어떤 변수를 제약하느냐에 따라 추정되는 회귀 계수의 값과 경향이 매우 큰 차이를 보인다 [12,13]. 따라서 본 연구에서는 어떤 모수에 제한을 두는 것이 좋은지 알아보기 위해 연령, 기간, 코호트에 대한 6가지 서로 다른 모형을 구축하고 우도비 검정을 시행하였다 (Table 2). Table 2에서 연령을 포함한 모형의 디비언스(deviance)가 다른 효과를 포함한 모형의 디비언스보다 매우 작기 때문에 연령

이 유방암 발생과 밀접한 연관이 있다는 것을 알 수 있다. 기간 효과를 포함한 모형의 디비언스 값이 가장 크기 때문에 기간 효과를 제한하는 것이 좋지만 본 연구에서는 기간 효과가 3개이므로 제약을 하는 데 어려움이 있다. 따라서 본 연구에서는 코호트 효과를 제한하기로 하였다. Figure 1을 참고하여 첫 번째 코호트 효과와 두 번째 코호트 효과가 같다고 두었다 ($\gamma_1 = \gamma_2$). 본 자료의 경우 이론적 분산보다 실제 추정된 분산이 커지는 과분산(overdispersion)이 발생하여 퀴시-포아송(Quasi-Poisson) 분포를 사용해서 과분산 문제를 해결하였다.

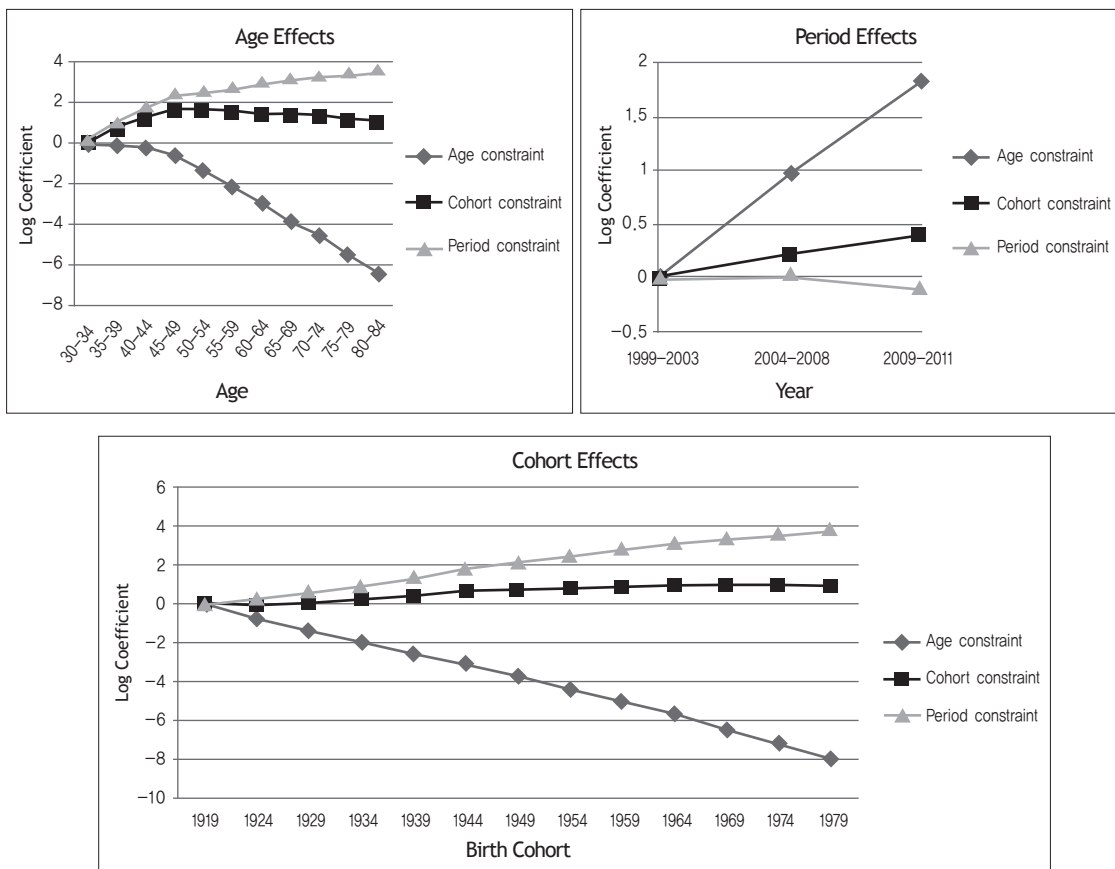


Figure 2. Age, period, and cohort effects estimated from the CGLM with three identifying constraints.

*CGLM: constrained generalized linear model

위의 모형을 적용하여 분석한 결과는 다음과 같다 (Table 3). 연령 효과를 보면 30-34세에 비해 그 이후 연령에서 모두 유의하게 증가하였다. 30대부터 계속 증가하여 45-49세에서 가장 높았고 그 이후 연령에서는 감소하는 경향을 보였다. 기간 효과 또한 1999-2003년에 비해 그 이후 기간에서 발생률이 유의하게 증가하는 것을 확인할 수 있다. 코호트 효과는 첫 번째 코호트와 두 번째 코호트는 같다고 제한하였으므로 이들 간에는 차이가 없고 그 이후 코호트에서는 첫 번째 코호트에

비해 선형적으로 유의하게 증가하였다. 즉, 나중에 태어난 코호트일수록 유방암 발생률이 높아지는 경향을 보인다고 할 수 있다. 이 자료에서 연령 효과의 계수 범위는 2.083에서 5.115, 기간 효과는 1.272에서 1.460, 코호트 효과는 1.12에서 2.618정도임을 알 수 있다. 이를 통해 유방암 발생률의 변화는 연령 효과가 가장 큰 영향을 미치고, 그 뒤를 이어 코호트 효과가 영향을 미치는 것으로 결론 내릴 수 있다.

Table 2. Likelihood ratio (LR) test of model fit using the CGLM estimator: Korean female breast cancer incidence 1999-2011

Model*	Deviance	DF	LR test	(d.f.)	p-value
A	111.71	22	111.42	(13)	<0.001
P	469.88	30	469.59	(21)	<0.001
C	180.18	20	179.90	(11)	<0.001
AP	4.51	20	4.22	(11)	0.963
AC	1.51	10	1.22	(1)	0.269
PC	79.06	18	78.77	(9)	<0.001
APC	0.29	9			

*Included independent variables (A: age, P: period, C: cohort)
CGLM: constrained generalized linear model

Table 3. Constrained generalized linear approach: age-period-cohort effects on breast cancer incidence in Korea, 1999-2011

Variables	Risk ratio	95% Confidence interval	
		Lower	Upper
Age (30-84 years)			
30-34	Reference		
35-39	2.08	1.86	2.33
40-44	3.65	2.99	4.46
45-49	5.11	3.82	6.84
50-54	4.93	3.36	7.72
55-59	4.58	2.84	7.38
60-64	4.47	2.53	7.91
65-69	3.97	2.04	7.71
70-74	3.83	1.79	8.82
75-79	3.30	1.40	7.76
80-84	2.88	1.14	7.32
Period (1999-2011)			
1999-2003	Reference		
2004-2008	1.27	1.16	1.40
2009-2011	1.46	1.21	1.76
Cohort			
1919	Reference		
1924	1.00		
1929	1.12	0.98	1.27
1934	1.28	1.03	1.58
1939	1.52	1.12	2.07
1944	1.86	1.24	2.78
1949	2.05	1.25	3.37
1954	2.18	1.21	3.94
1959	2.43	1.23	4.81
1964	2.55	1.17	5.53
1969	2.57	1.08	6.14
1974	2.62	1.00	6.87
1979	2.51	0.87	7.24

3) Median Polish 방법

Median Polish 방법으로 분석한 결과는 다음과 같다 (Table 4). 연령 효과와 기간 효과는 앞의 제약된 일반선형회귀모형의 결과와 동일한 결과를 보였다. 그러나 제약된 일반선형회귀모형의 결과와 다르게 코호트 효과는 존재하지 않았다. 추정된 Median Polish 잔차 그래프를 통해서도 코호트에 따른 특별한 경향성을 보이지 않는 것을 확인할 수 있다 (Figure 3).

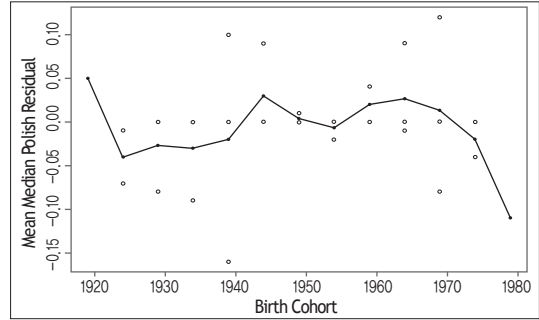


Figure 3. Median polish approach: nonadditive influences of age and period by birth cohort on female breast cancer incidence, 1999-2011.

Table 4. Median polish approach: age-period-cohort effects on breast cancer incidence in Korea, 1999-2011

Variables	Risk ratio	95% Confidence interval	
		Lower	Upper
Age (30-84 years)			
30-34	Reference		
35-39	2.10	1.59	2.77
40-44	3.60	2.78	4.66
45-49	4.81	3.74	6.19
50-54	4.32	3.35	5.57
55-59	3.67	2.83	4.75
60-64	3.22	2.48	4.19
65-69	2.46	1.88	3.23
70-74	1.99	1.50	2.63
75-79	1.48	1.10	1.99
80-84	1.18	0.86	1.61
Period (1999-2011)			
1999-2003	Reference		
2004-2008	1.40	1.25	1.56
2009-2011	1.74	1.57	1.94
Cohort			
1919	Reference		
1924	0.91	0.79	1.06
1929	0.93	0.80	1.07
1934	0.92	0.80	1.07
1939	0.93	0.81	1.08
1944	0.98	0.85	1.13
1949	0.95	0.83	1.10
1954	0.95	0.82	1.09
1959	0.97	0.84	1.12
1964	0.98	0.85	1.13
1969	0.96	0.84	1.11
1974	0.93	0.80	1.09
1979	0.85	0.72	1.02

4) Intrinsic estimator (IE) 방법

IE 방법을 적용하여 분석한 결과는 다음과 같다 (Table 5). 연령 효과부터 살펴보면 유방암의 발생률이 40세부터 유의하게 증가하다가 45-49세 사이에서 가장 높았으며 그 이후 감소하는 경향을 보였다. 기간 효과는 최근 기간으로 올수록 증가하는 경향을 보였다. 코호트 효과는 Median Polish 방법과 마찬가지로 통계적으로 유의하지 않았고, 추정된 로그변환한 회귀계수와 코호트 간의 그래프에서도 Median Polish 방법과 비슷한 경향을 보였다 (Figure 4).

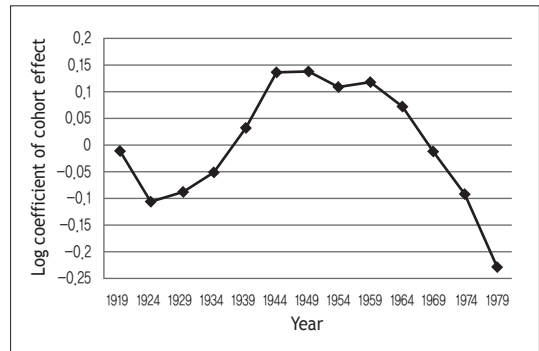


Figure 4. Intrinsic estimator coefficient estimates of the cohort effects using Korean female breast cancer incidence, 1999-2011.

Table 5. Intrinsic estimator approach: age-period-cohort effects on breast cancer incidence in Korea, 1999-2011

Variables	Risk ratio	95% Confidence interval	
		Lower	Upper
Age (30-84 years)			
30-34	0.48	0.36	0.63
35-39	0.91	0.77	1.09
40-44	1.46	1.26	1.69
45-49	1.85	1.60	2.15
50-54	1.63	1.39	1.91
55-59	1.37	1.16	1.63
60-64	1.22	1.02	1.46
65-69	0.99	0.82	1.18
70-74	0.86	0.72	1.04
75-79	0.68	0.56	0.82
80-84	0.54	0.42	0.70
Period (1999-2011)			
1999-2003	0.74	0.69	0.79
2004-2008	1.04	0.97	1.10
2009-2011	1.31	1.23	1.39
Cohort			
1919	0.99	0.66	1.48
1924	0.90	0.68	1.18
1929	0.92	0.75	1.12
1934	0.95	0.78	1.16
1939	1.03	0.84	1.26
1944	1.15	0.94	1.40
1949	1.15	0.95	1.39
1954	1.11	0.93	1.33
1959	1.13	0.96	1.32
1964	1.07	0.92	1.25
1969	0.99	0.84	1.16
1974	0.91	0.73	1.14
1979	0.80	0.51	1.23

4. 고찰 및 결론

앞의 분석 결과를 통해 CGLM, Median Polish, IE 방법 모두에서 연령 효과와 기간 효과가 유의하게 존재한다는 것을 확인할 수 있다. 하지만 코호트 효과의 경우 세 모형 간에 차이가 발생하였다. CGLM의 경우 코호트 효과가 유의하므로 늦게 태어난 코호트일수록 유방암 발생률이 높다고 할 수 있다. 반면 Median Polish 방법과 IE의 경우를 보면 코호트 효과가 유의하지 않으므로 코호트 효과가 존재하지 않는다고 할 수 있다. 즉, 유방암의 발생률을 증가시킨다고 알려져 있는 많은 요인들이 모든 연령 그룹에서 비슷한 정도의 영향을 미쳤고 태어난 시기는 유방암 발생률을 높이는 데에 큰 영향을 미치지 않는다고 할 수 있다.

이처럼 같은 자료를 사용하였지만 어떤 방법을 사용하느냐에 따라 코호트 효과의 존재 여부가 차이가 발생하는 것을 볼 수 있다. 하지만 모든 자료에서 CGLM과 Median Polish, IE 방법의 결과가 항상 다르게 나오는 것은 아니므로 CGLM 방법을 사용하는 것이 좋지 않다고 할 수는 없다. 그러나 CGLM 방법을 사용할 경우 CGLM 방법의 단점을 인지하고 어떠한 변수를 제한해야 하는지 신중하게 결정해야 한다. 본 논문에서는 위의 결과를 통해 CGLM 방법 보다는 나머지 두 개의 방법을 사용하는 것이 각각의 효과가 미치는 영향을 정확하게 볼 수 있다고 판단되기 때문에 Median Polish 방법이나 IE 방법을 통해 APC 분석을 하는 것을 추천한다.

또한 외국의 유방암 발생률 APC 분석 결과를 보면 대부분 코호트 효과가 유의하게 존재한다고 결론내리고 있다 [6,18-22]. 하지만 대부분이 CGLM 방법을 사용하여 분석을 한 결과이므로 다른 방법을 사용한 결과와도 비교를 해 보는 것이 좋을 것으로 생각된다. 외국 결과와의 차이가 분석 방법에 의한 차이인지 한국 여성의 발생률이

45-49세에서 가장 높아지고 다시 감소하는 경향성 때문인지, 이러한 경향성은 왜 발생하는지 등을 추후에 더 연구해 볼 필요가 있다. 더불어 우리나라 유방암 발생률 자료도 기간 효과의 개수가 3개로 매우 작기 때문에 분석 시 제한점이 있었다. 그러므로 추후 발생률에 대한 지속적인 연구가 필요할 것으로 보인다.

References

- [1] Jung K, Won Y, Oh C, Lee D, Lee J. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2011. *Cancer Research and Treatment* 2014; 46(2): 109-123.
- [2] Cancer Research UK. Available at <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/incidence/uk-breast-cancer-incidence-statistics> [accessed on 16 December 2014].
- [3] American Cancer Society. *Breast Cancer Facts & Figures 2011-2012*. American Cancer Society, Inc: Atlanta, p. 1-6.
- [4] Anderson WF, Chu KC, Chang S, Sherman ME. Comparison of age-specific incidence rate patterns for different histopathologic types of breast carcinoma. *Cancer Epidemiology, Biomarkers & Prevention* 2004; 13: 1128-1135.
- [5] Clarke CA, Keegan THM, Yang J, Press DJ, Kurian AW, Patel AH, Lacey JV Jr. Age-specific incidence of breast cancer subtypes: understanding the black-white crossover. *Journal of the National Cancer Institute* 2012; 104(14): 1094-1101.
- [6] Minami Y, Tsubono Y, Nishino Y, Ohuchi N, Shibuya D, Hisamichi S. The increase of female breast cancer incidence in Japan: emergence of birth cohort effect. *Publication of the International*

- Union Against Cancer 2004; 108(6): 901-906.
- [7] Rostgaard K, Væth M, Holst H, Madsen M, Lynge E. Age-period-cohort modelling of breast cancer incidence in the Nordic countries. *Statistics in medicine* 2001; 20(1): 47-61.
- [8] Keyes KM, Li G. Injury research: theories, methods, and approaches. DOI 10.1007/978-1-4614-1599-2_22. Springer Science+Business Media, LLC 2012.
- [9] Statistics Korea. Korean statistical information service. Available at <http://kosis.kr> [accessed on 16 December 2014].
- [10] Shahpar C, Li G. Homicide mortality in the United States, 1935-1994: age, period, and cohort effects. *American Journal of Epidemiology* 1999; 150(11): 1213-22.
- [11] Choi Y, Kim Y, Park SK, Shin HR, Yoo KY. Age-period-cohort analysis of female breast cancer mortality in Korea. *Breast Cancer* 2006; 13(3): 266-271.
- [12] Yang Y, Land KC. Age-period-cohort analysis new models, methods, and empirical applications. CRC Press: New York, 2013, p. 75-123.
- [13] Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971-2006. *Social Science & Medicine* 2010; 70(7): 1100-1108.
- [14] Tukey JW. Exploratory data analysis. Reading, MS: Addison-Wesley Publishing Company: Boston, 1977.
- [15] Keyes KM, Li G. A multi-phase method for estimating cohort effects in age-period contingency table data. *Annals of Epidemiology* 2010; 20(10): 79-785.
- [16] Yang Y, Schulhofer-Wohl S, Fu WJ, Land KC. The intrinsic estimator for age-period-cohort analysis: what it is and how to use it. *American Journal of Sociology* 2008; 113(6): 1697-1736.
- [17] Yang Y, Fu WJ, Land KC. A methodological comparison of age-period-cohort models: intrinsic estimator and conventional generalized linear models. *Sociological Methodology* 2004; 34(1): 75-110.
- [18] Shen YC, Chang CJ, Hsu C, Cheng CC, Chiu CF, Cheng AL. Significant difference in the trends of female breast cancer incidence between taiwanese and caucasian americans: implications from age-period-cohort analysis. *Cancer Epidemiology, Biomarkers & Prevention* 2005; 14(8): 1986-1990.
- [19] Weedon-Fekjær H, Bakken K, Vatten LJ, Tretli S. Understanding recent trends in incidence of invasive breast cancer in Norway: age-period-cohort analysis based on registry data on mammography screening and hormone treatment use. *British Medical Journal* 2012; 344: e299.
- [20] Dhillon PK, Yeole BB, Dikshit R, Kurkure AP, Bray F. Trends in breast, ovarian and cervical cancer incidence in Mumbai, India over a 30-year period, 1976-2005: an age-period-cohort analysis. *British Journal of Cancer* 2011; 105: 723-730.
- [21] Arbeev KG, Ukraintseva SV, Arbeeva LS, Yashin AI. Decline in human cancer incidence rates at old ages: age-period-cohort considerations. *Demographic Research* 2005; 12(11): 273-300.
- [22] Awang H, McNeil D, Chilvers M, Middeldorp J, Taylor R. Age-period-cohort modelling of the incidence of breast and cervical cancer in New South Wales. *The Sciences* 2001; 1(2): 48-51.