

고용 이력에 따른 암 발생 위험도 추정모형 개발

연구기간 2023년 5월 ~ 2023년 11월

핵심단어 폐암, 백혈병, 위험도

Highlights

개인별 인구통계학적 특성과 고용 이력을 바탕으로 폐암과 백혈병의 위험도를 개인 수준에서 추정할 수 있는 모형을 설계. 향후 흡연상태, 거주지, 가족력 등의 추가 정보를 확대 적용하는 등 모형설계의 고도화 작업을 통해 개인수준에서의 위험감시와 조기개입을 위한 역학적 탐색 도구로 활용될 수 있을 것으로 기대

1

연구배경

○ 직업성 암 예방을 위한 일차 예방의 한계와 암 조기발견의 중요성

- 노출요인을 제거하여 근로자가 건강한 상태로 작업할 수 있도록 만들어주는 1차 예방이 가장 궁극적인 예방책이기는 하나, 직업성 암에서 1차 예방은 쉽지 않음.
- 산업 보건 분야에서 잠복기가 긴 직업성 암은, 발암과정의 복잡성, 원인의 다양성 등으로 그 원인이 되는 노출요인을 찾아내기까지 많은 시간이 소요되며, 노출위험을 발견했다 하더라도 이미 과거에 노출된 경우가 대다수로 개입을 통한 작업환경의 개선이 쉽지 않음.
- 암(특히 고형 암)은 진단 시 병기에 따라 치료 후 재발 및 직업 복귀율이 극적인 차이가 있기 때문에 근로자 집단의 건강관리를 위한 암의 조기발견은 근로자의 사회적 활동을 유지시켜 줄 수 있는 예방 전략임.

표 산업보건 분야에서의 근로자의 암 예방을 위한 단계적 전략

예방 분류	설명	목표	직업성 암과의 관련성
1차 예방	<ul style="list-style-type: none"> 질병이 발생하지 않도록 질환과 관련된 위험요인과의 노출을 차단하거나 제한 	<ul style="list-style-type: none"> 작업장 내 질환관련 위험요인의 식별 및 노출 제어 	<ul style="list-style-type: none"> 암 발생을 원천적으로 차단하기 위해 위험 요인(예: 유해 화학 물질, 방사선 등)을 제거하거나 노출을 감소
2차 예방	<ul style="list-style-type: none"> 증상을 느끼지 못하는 질병이 초기 단계에서 발견되고 완치 될 수 있도록 조기 진단 및 선별검사에 중점을 둠 	<ul style="list-style-type: none"> 이미 위험 요인에 노출된 고위험 군을 대상으로 관련 위험질환의 조기 발견을 위해 주기적인 선별검사를 통한 질환발견 	<ul style="list-style-type: none"> 작업 이력과 노출 데이터를 기반으로 고위험 그룹을 선별하고 조기 진단 및 사전 관리를 유도함
3차 예방	<ul style="list-style-type: none"> 질병이 진행되거나 합병증이 발생하는 것을 방지하고, 이미 발생한 질병으로부터 최적의 삶의 질을 유지하려고 노력함 	<ul style="list-style-type: none"> 질병의 진행을 늦추고, 합병증 및 장애를 최소화하며, 환자의 삶의 질 향상 	<ul style="list-style-type: none"> 직업성 암 환자에 대한 장기 관리 및 재활 프로그램을 통한 삶의 질 향상 지원

○ 포괄적인 발암위험도 추정의 필요성

- 직업성 암은 발암요인 노출뿐만 아니라 다양한 취약 요인(연령, 복합 노출 등)이 암 발병에 영향을 준다는 점에서 암의 조기발견을 위해서는 포괄적이고 통합적인 위험도 평가가 필요함

○ 이에 근로자의 고용이력정보를 활용하여 인구통계학적 특성과 고용이력으로부터 건강 위험, 암의 위험도를 개인 수준에서 추정하는 모형을 설계하고자 하였음.

- 시범적으로 고형암과 혈액암의 대표적인 폐암과 백혈병을 대상으로 위험도 추정 모형을 구축하였음

2

주요연구내용



연구결과

○ 국외 암 위험 예측관련 연구 현황

- 영국 국가검진위원회의 CanPredict(lung) 모형

- 저선량 흉부 CT(low-dose chest CT)를 활용한 폐암 검진의 권고안을 위하여 설계되었음. 영국의 인구 기반 코호트 연구의 내적타당도(internal validation)와 외적타당도(external validation) 검토를 위해 2개의 코호트로부터 자료를 수집. CanPredict(lung) 모형에는 석면 노출, 천식 등의 예측 인자가 추가되었으며, 다른 유사한 모형에 비교하여 예측결과가 더 좋았음
- GLOBOCAN(Global Cancer Observatory)²⁾ 프로젝트
 - 국제암연구소³⁾(International Agency for Research on Cancer; IARC)에서 운영하는 다양한 암 질병 관련 통계를 통합하여 제공하는 플랫폼으로 시점에 따른 암 정보를 제공
 - Cancer Over Time(지난 50년간의 추이), Cancer Today(현재 각 국가의 암 질병의 발생률, 유병률, 사망률), Cancer Tomorrow (2040년까지의 암 질병에 의한 질병 부담(burden of diseases)의 예측모형)가 운영되며, Cancer Tomorrow는 2020년의 Cancer Today를 통하여 수집된 정보를 기반으로 설계된 모형임.

○ 시간가변 콕스 비례위험모형⁴⁾을 활용한 추정모형 설계

- 고용이력에 따른 암 위험 추정모형 설계가정
 - 누적 근로효과: 만일 위험한 업종이 존재한다면, 해당 고위험 업종에 짧게 종사한 근로자 보다는 길게 종사한 근로자의 암 발병 위험이 더 큰 것으로 가정
 - 지연효과(lagged effect): 현재 발생한 암에 대한 발병 위험은 수 년 전의 고용이력(노출)의 효과가 지연되어 나타난 것임. 예를 들어, 폐암의 경우에는 현재의 발병 위험은 10년 전까지의 고용이력과 연관 있고, 10년 이내의 고용이력은 현재의 발병위험과는 관련 없음. 혈액암인 백혈병의 경우 현재의 발병 위험은 5년 전까지의 고용 이력과 관련 있으며 5년 이내의 고용이력과는 관련되지 않음.
 - 정책적 효과의 대리지표로의 입사시기: 나이와 성별이 비슷한 근로자라 하더라도 최초 입사시기가 다르다면 현 시점에서 다른 위험비를 갖고 있을 것이라고 가정하였음. 고용보험정보⁵⁾에서 최초로 감지되는 입사시기 기준으로 95-99년, 00-04년, 05-09년,

2) <https://gco.iarc.fr/tomorrow/en/about>

3) 1965년 세계보건기구가 암 퇴치를 위해 설립한 국제연구기관

4) 특정 요인이 특정 시점에서의 관심사건(폐암 또는 백혈병 발병)의 발생빈도와 맺는 연관성을 분석할 때 이용되는 대표적 모형으로 콕스 비례위험모형은 관찰되는 각 집단(또는 개인)에서의 위험곡선이 서로 비례하며 교차하지 않음을 전제로 함

5) 95년도 고용보험정책 시행 이전의 최초입사시기 정보는 이용불가로 1995년 이전 입사자도 1995년으로 분류

10-15년, 15년 이후 여부에 따라서 발병률 곡선이 다를 것이라 가정

- 시간가변 콕스 비례위험모형의 적용

- 콕스 비례위험모형 하에서 $h(t)$ 는 개인마다 상이하며, 이는 해당 개인마다 서로 다른 요인들 z_1, z_2, \dots, z_p , 그리고 $x_1(t), x_2(t), \dots, x_q(t)$ 에 의하여 결정됨. 이를 수식으로 표현하면 다음과 같으며 각 업종에 대응되는 요인의 회귀계수 β 에 대한 해석은 지수화한 값 즉, $\exp(\beta)$ 는 해당 업종에 1년 더 종사한 사람의 그렇지 아니한 사람에 대한 상대 위험비로 해석

$$h(t) = h_0(t) \times \exp(z_1\alpha_1 + z_2\alpha_2 + \dots + z_p\alpha_p + x_1(t)^T\beta_1 + x_2(t)^T\beta_2 + \dots + x_q(t)^T\beta_q)$$

- t 는 코호트 진입 이후 경과 시간을 의미. 즉 개인마다 $t=0$ 은 (고용보험 상에 등록된) 최초 입사 시점을 의미
- $h(t)$ 는 위험률 함수(hazard function)로 불리는 수량으로, 시점 t 에서의 어떤 사람이 사건을 겪을 위험의 수준을 의미. $h(t)$ 는 시점 t 의 함수로, t 의 값이 바뀔에 따라 (즉 동일한 개인이라 하더라도) 나이 증가에 따라 위험률이 바뀔 수 있음을 뜻함. 즉, $h(t)$ 값이 큰 사람이 해당 시점에서 더 사건 발생의 위험도가 높음.
- $h_0(t)$ 는 기초상태(baseline)의 위험률 함수를 뜻하며, 모든 요인(z 와 $x(t)$)이 0인 가상의 개인이 갖는 위험함수를 의미.
- z_1, z_2, \dots, z_p 는 해당 개인이 갖는 시간-불변 요인⁶⁾을 뜻한다. 옆의 $\alpha_1, \alpha_2, \dots, \alpha_p$ 는 그에 대응되는 회귀계수.
- $x_1(t), x_2(t), \dots, x_q(t)$ 는 해당 개인이 갖는 시간-가변 요인⁷⁾을 뜻한다. 옆의 $\beta_1, \beta_2, \dots, \beta_p$ 는 그에 대응되는 회귀계수

6) 5년 구간 별로 범주한 최초 입사 시기(95-99년도, 00-04년도, 05-09년도, 10-14년도, 15년도 이후)

3

연구활용방안



활용방안

- 개인별로 다양한 인구통계학적 특성 및 다양한 고용이력을 반영하여 개인 수준에서 폐암/백혈병의 위험도를 추정할 수 있는 모형의 개발을 시도하였다는 점에서 의미가 있음.
- 향후 흡연상태, 거주지, 가족력 등의 추가 정보를 확대 적용하는 등 모형설계의 고도화작업을 통해 개인수준의 모형은 특정 개인의 (표준집단 대비) 상대위험도의 추정을 위한 역학적 탐색 도구로 활용될 수 있을 것으로 기대됨.
- (활용 예) 근로자 A (1964년 출생, 남성, 고용보험 기준 1998년 최초입사)에 대하여, 본인의 직업력에 기초하여 참조집단 대비 본인의 폐암발병 위험이 얼마나 높은지 동일한 (5세 연령군, 성별, 입사시기)에 속한 표준집단과 비교하여 근로자 A의 현재 암 발병 위험의 상대위험도를 추정할 수 있음.
- 모형 활용 및 해석 시 고려할 점
 - 모형의 요인의 회귀계수 β 가 인과관계를 의미하지는 않는다는 점에 유의할 필요가 있음
 - 모형 계수는 최소한의 인구통계학적 정보 (5세 연령군, 성별, 입사시기)가 보정된 상태에서의 고용력과 발암위험과의 평균적인 상관도를 나타낸다고 해석하는 것이 바람직함
 - 콕스비례위험모형으로 적합(Fitting)된 위험함숫값은 각 개인과 동일한 직업력과 인구통계학적 정보를 가진 사람의 평균적인 위험함숫값의 추정치로 해석하는 것이 바람직함

✉ 연락처

- 연구책임자: 성균관대학교 최영근 교수
- 연구상대역: 산업안전보건연구원 직업건강연구실
중부권역학조사팀 이경은
- 연락처: 032-510-0753
- e-mail: kyeong85@kosha.or.kr