

## 빅데이터를 활용한 직업성 질환 코호트 운영(II)

조사기간 | 2020년 2월 ~ 2020년 12월

핵심단어 | 빅데이터, 직업성 질환, 코호트 활용

### 조사배경

- 산업안전보건연구원에서는 2018-2019년 빅데이터를 활용한 직업성 질환코호트 구축 및 운영을 통하여 특수건강검진자료의 사업장, 인적, 유해물질, 검진결과 자료와 국민건강보험공단의 국민건강정보 자료, 고용보험정보 등을 연계하여 코호트를 설계·구축하였다.
- 2003-2005년 고용보험 전체 근로자 및 2013-2017년 특수건강검진단 근로자에 대한 질환 코호트는 단순히 정보의 수집, 저장, 분양하기 위한 고정적 역할을 넘어서 직업 건강분야의 전문가 및 실무자들이 합리적 의사결정을 내릴 수 있도록 과학적 근거를 제시해 주는 역할을 해 줄 것을 목적으로 설계되었으나 아직까지 활용영역이 다양하지 못하다는 한계점이 있다.
- 이번 연구에서는 직업성 질환코호트 의 활용도를 높이고 사전 예방적 역학감시망으로의 역할을 구체화 할 수 있는 방안을 마련하고자 하였다.

### 주요조사내용

#### 조사결과

- 데이터중심의 의사결정 지원기반 마련
  - 산재신청자료, 역학조사 사례보고서, 작업환경실태조사, 화학물질통계조사와 같은 다양한 노동자 건강집단 자료원을 정리하고 기 구축된 직업코호트 분석결과를 세부업종별로 고도화 분석 및 시각화하여 역학조사 우선순위 결정을 위한 체계적 근거를 수집하고 평가한다.
  - 질병부담 및 수요도를 반영한 지표로 최근 5년 산재신청현황을 분석한 결과 노동자 규모 대비 산재신청건수가 많으면서, 질환 발병건수가 큰 대표적인 “업종×질환”을 선정한다.
- 업종에 따른 질환 우선순위 조합에는 대표적으로 “보건 및 사회복지사업×결핵”, “전기·가스·수도업×피부암”, “유리제조업×석면폐증”, “임업×동상”, “수리업×석면폐증”, “고무제품제조업×백혈병”, “임업×피부질환”, “석유정제업×폐암”, “도자기 시멘트업×폐암”, “펄프지류 제조업×위식도암”, “농업×위식도암” 등이 포함된다.

### ● 사업장 기반의 질환감시자료 설계 및 구축

- 개인정보식별 위험에 대한 부담을 줄이고 관리적 측면에서의 사업장 특성에 따른 질환감시기능을 강화 하였다. 단순히 질환 발생 상태에 대한 년도 별 통계 수치가 아닌 시간에 따른 사업장의 근로자 규모 정보와 질환 규모를 현직 및 퇴직자 정보를 포함하였다.

사업장 인력 현황 테이블

파 단 종 류	사업장구분코드	업종	우편번호	년도	연령집단	성 별	인력 현황				질환 A		질환 Z	
							임사자 (명)	퇴직자 (명)	재직자 (명)	누적 근로자 (명)	발생자	누적발생자	현직자	누적 발생자
제조업	XXXXXXXXXX	28519	21417	1995	20-30	M								
	XXXXXXXXXX	28519	21417	1995	30-40	M								
	XXXXXXXXXX	28519	21417	1995	40-50	M								
	XXXXXXXXXX	28519	21417	1995	50-60	M								
	XXXXXXXXXX	28519	21417	1995	>60	M								
	XXXXXXXXXX	28519	21417	1995	20-30	F								
	XXXXXXXXXX	28519	21417	1995	30-40	F								
	XXXXXXXXXX	28519	21417	1995	40-50	F								
	XXXXXXXXXX	28519	21417	1995	50-60	F								
	XXXXXXXXXX	28519	21417	1995	>60	F								
	XXXXXXXXXX	28519	21417	1996	20-30	M								
	...													
	XXXXXXXXXX		21417	2016	50-60	F								
	XXXXXXXXXX		21417	2016	>60	F								
	제조업중공업중공업중공업	47993	15610	1995	20-30	M							...	
	제조업중공업중공업중공업	47993	15610	1995	30-40	M								

사업장 질환 발병 현황 테이블

- 사업장 인력현황 및 질환 발병현황에 대한 두 테이블을 구축하기 위하여 공통으로 적용된 사업장 식별 변수에는, 사업장 관리(등록)번호, 사업장명, 우편번호, 업종코드의 조합을 이용하였다. 추가적으로 성·연령 표준화 지표를 활용할 수 있도록 성별을 증화하여 따로 구축하였다.

### 시사점

- 산재신청자료, 역학조사 사례보고서, 작업환경실태조사, 화학물질통계조사와 같은 다양한 노동자 건강집단 자료원을 정리하고 구축된 직업코호트 분석결과를 세부업종별로 고도화 분석 및 시각화하여 역학조사 우선순위 결정을 위한 체계적 근거를 수집하고 평가하였다.
- 사업장 기반의 질환감시 자료는 사업장별/연도별/성별/연령구간별 근로자 전체의 분포와 질환자 분포로 구성되어 있어, 향후 작업환경 측정자료와의 연계를 통한 작업환경건강영향 역학 연구 설계가 가능할 뿐만 아니라 특정 사업장 혹은 특정 업종 특정 지역 등 사업장 특성에 따라 추출하여 질환에 대한 표준화 발병률을 산출할 수 있다는 장점이 있다.

## 조사활용방안

### 활용방안

#### ● 데이터중심의 의사결정 지원도구

- 빅데이터를 활용하여 역학조사 우선순위 결정과정에서 조사대상에 대한 질병 및 노출부담, 예방가능성, 사회적 근로자 집단 수준의 선호도 등의 객관적 정보를 직업성 질환코호트, 유해물질노출코호트, 작업환경 실태조사자료, 근로자 실태조사자료, 웹 기반 여론 및 언론 정보 등의 데이터로부터 산출하여 제안서와 함께 요약된 정보를 지원함으로써 전문가의 의사결정 지원에 활용 가능하다.

#### ● 사업장 기반의 질환감시 데이터셋 활용

- 시범 분석 결과로 업종별 폐암에 대하여 성, 연령구간으로 층화하여 누적발병률(1,000명당)을 보여주었다. 고형 암의 경우, 실제 발암과정이 시작할 때부터 임상적으로 진단될 때까지의 잠재기간이 상당히 있다는 점에서 발병률이 아닌 누적 발병률을 적용하여 산출한다. 그러나 질환에 따라서 급성기 질환 혹은 손상에 대한 발병률을 감시하고자 활용할 때는 발병률을 산출하여 평가하는 것이 타당할 것이다.
- 감시지표로서 (누적)발병률의 평균을 쓸지, 근로자 및 질환자 수의 합을 통한 1개의 통합 발병률을 쓰는 것이 합당한지에 대한 타당성 평가는 적용할 역학 설계모형에 따라 적합한 값을 산출하여 활용할 필요가 있다.
- 사업장 기반의 질환감시 설계는 퇴직자 일지라도 질환이 발생할 경우 과거 사업장의 누적 근로자 집단 내 발생 건수로 질환자가 집계된다. 비슷한 업종끼리의 비교 및 전체 사업장과의 비교를 통해 위험업종에 대한 평가는 물론 향후, 작업환경 측정 데이터 셋과의 연계가 이루어질 경우 유해요인에 대한 질환 감시 자료로도 활용 가능할 것으로 기대된다.

### 연락처

조사책임자 | 산업안전보건연구원 직업건강연구실 이상길, 연세대학교 윤진하 교수

연락처 | 052-7030-860

e-mail | twincokes@kosha.or.kr