

최종보고서

근로자 집단 암 발병 예측 모형 시범 개발

최영근·김양우·박성오·이동환·이우주

산업재해예방

안전보건공단

산업안전보건연구원



제 출 문

산업안전보건연구원장 귀하

본 보고서를 “근로자 집단 암 발병 예측 모형 시범 개발”의
최종보고서로 제출합니다.

2022년 11 월

연구진

연구기관 : 숙명여자대학교

연구책임자 : 최영근 (조교수, 숙명여자대학교 통계학과)

연구원 : 김양우 (전임의, 한양대학교구리병원 직업환경의학과)

연구원 : 박성오 (조교수, 성신여자대학교 통계학과)

연구원 : 이동환 (부교수, 이화여자대학교 통계학과)

연구원 : 이우주 (부교수, 서울대학교 보건대학원)

요약문

- 연구기간 2022년 05월 ~ 2022년 11월
- 핵심단어 발병률, 추계, 백혈병, 폐암, 사업장 기반 암 감시 자료
- 연구과제명 근로자 집단 암 발병 예측 모형 시범 개발

1. 연구 배경

그간 산업안전보건연구원에서는 업종별 질환감시를 목적으로 직업성질환 코호트 사업을 운영하여 <사업장 기반 질환 자료>를 구축하였다. 해당 자료는 고용보험에 등록된 근로자들에 대한 개인단위 빅데이터(고용보험DB, 암등록 DB, 통계청 사망DB 및 사망원인DB)를 연계하여, 주요질환 (24종 암과 주요 만성질환)에 대한 사업장 단위의 연간 (누적) 발병률 정보를 수록하고 있다.

질환감시 목적으로 사업장 기반 질환 자료를 활용하는 과정에서 질환발병 추계 모형의 개발이 잠재적으로 기여할 여지가 있는데, 기존에 과소평가/간과 되어 왔던 질환의 발견이나 선제적 조기개입에 도움이 될 수 있다.

다만 질환발병 추계를 위하여는 사업장/업종별로 이질적인 근로자집단의 특성을 반영하면서도 시계열적 흐름을 반영한 추계모형 설계가 과제로 남아 있다.

2. 주요 연구내용

1) 선행연구 및 사례 고찰

- 직업별 암 발생 비교 연구는 주로 직업군을 비롯하여, 국가, 성별, 연령, 기간 그룹별로 정의된 각 집단에서 관찰된 암 발생 건수를 해당 계층별 개인의 연도와 국가인구의 발병률로부터 계산한 예상 건수와 단순 비교 하는 형태로 진행되어 왔다. 그러나, 직업별 암발생 연구는 위험도 평가 등에 주로 집중되어 있고, 시간에 따른 미래 세 추계 연구는 따로 확인되지 않았다.
- 암 발생 추계 모형은 국가별 또는 대규모 집단별 추계 연구에서 간접적으로 수행되어 오거나 특정 암종을 중심으로 제시되어 왔다. 국내외 사례를 종합적 살펴보면, 각 층(예-성별)에서 범주화된 나이와 시간(calendar time)의 모든 조합에서 암 발생 또는 사망 자료가 주어지는 경우 일반화 선형 모형 또는 age-period-cohort(APC) 모형을 이용하는 사례로 이해될 수 있다. 이러한 분석 기법은 나이와 시간에 대한 정보가 다른 형태로 주어지는 경우 직접적으로 적용이 불가능하다는 어려움을 가지고 있다.

2) 데이터셋 검토

- 기 구축된 <사업장 기반 질환 자료>는 각 사업장 수준에서 2개년도 이상 종사한 근로자의 연도별 누적추적인년과 누적질환발병건수가 층화변수별로 (성별×입사시기군×입사당시연령군) 집계되어 있다.
- 현 자료는 근로자의 나이에 대한 정보가 ‘범주화된 입사 시기의 나이’로만 한정되어 있어, 암 발병 추계에 사용된 여러 문헌의 방법론을 그대로 적용할 수 없고 본 자료에 맞는 새로운 모형의 설계가 필요한 것으

로 판단하였다.

3) 질환발병률 추계모형 설계 및 시범 분석

- 기 구축된 사업장 기반 질환 자료에 적용할 수 있으면서, 업종별 특성을 반영한 질환발병률 추계 통계모형을 설계/제안하였다. 제안 모형은 각 질환 대상의 누적발병건수와 누적추적인년의 미래 추세를 추계하려는 목적으로 제안되었으며 다음 특징을 가지고 있다:
 - (1) 각 업종분류수준 (전체 - 대분류 - 중분류)별로 성별(전체/남성/여성) 및 입사시기 (전체/2000년도까지/2000년도이후)의 누적발병건수와 누적추적인년의 추계값을 제공한다.
 - (2) 세부 업종 및 층화별로 발병건수값의 시계열적 특성이 다를 수 있기 때문에, 여러 모형을 이용하여 여러 추계값을 생산한 뒤, 추정오차가 적은 여러 추계값을 합성하였다.
 - (3) 추계 대상의 누적발병건수 및 누적추적인년 시계열은 계층 구조를 가지고 있기 때문에 (하위 세부계층의 관찰값들의 합이 상위 계층의 관찰값과 동일함), 계층시계열 조건을 만족시키는 조정절차를 거쳤다.

4) 개발된 모형의 타당성 검증

- 위 절차를 폐암 및 백혈병에 대한 <사업장 기반 질환 자료>의 2000-2018년도까지의 누적발병건수와 누적추적인년에 적용하여 2022년도의 추계값을 산출하였다.
- 2022년 발병비(전체업종 대비)를 계산하여 발병비 상위 5종의 대분류 업종 및 상위 10종의 중분류 업종을 남녀합계, 남성 또는 여성 집단에서 보고하였다.
- 정량적 타당성 검증 절차로, 관찰값이 가용한 2015-2018년도 업종별 누적발병건수를 예측 대상으로 선정하고, 검증오류(2014년도까지의 자

료로 2015-2018년도를 예측)를 계산하였다. 그 결과, 2000년도까지의 입사자 집단에 대하여는 대부분의 업종에서 모두 오차율 10% 미만의 정확한 추계가 가능하였다. 2000년 이후 입사자 집단에 대하여는 업종별 오차율의 중앙값이 20-25% 정도로 나타났다.

- 정성적 타당성 검증 절차로, 위험군으로 보고된 업종과 기존 백혈병/폐암 발병 연관 업종으로 알려진 업종(국제암연구소 업종별 위험도)을 비교한 결과, 알려진 업종에 포함되나 기존에 알려지지 않은 업종들도 본 연구 결과에 보고되었다.

5) 활용 방안 및 제언

- 본 연구를 통하여, 안정적인 추세가 존재하는 업종에 대하여는 통계적인 기법 기반의 단기추계(3-5년)가 가능함을 확인하였다. 다기관 빅데이터 취득과 전처리에 수 년이 걸림을 감안하면, 본 보고서에서 제안한 방법은 적재된 빅데이터에 기반하여 현 시점에서의 발병률 추정값을 계산하는 데 유용할 수 있다. 다만 암발병의 가장 큰 위험요인인 나이가 반영되지는 않은 결과이다. 본문에는 요약자료에 매 시점 나이를 입력할 수 있는 방안을 제시하였다.
- 암발병 추계의 궁극적인 목적이 알려진 위험요인의 기여도 평가와 업종별 집단 감시라면, 위험기여도 평가 모형과 감시모형은 추계모형과 독립적으로 운용될 필요가 있다. 산업안전보건연구원이 확보가능한 빅데이터로부터 위험기여도 평가모형과 감시 모형을 개발하기 위한 추가적인 연구와 검토가 요구된다.

3. 연락처

- 연구책임자 : 최영근 (숙명여자대학교 통계학과 조교수)
 - ☎ 02) 710-9178
 - E-mail ygchoi@sookmyung.ac.kr

목 차

I. 서 론	1
1. 연구 배경	3
2. 연구 목적	3
II. 연구 내용 및 방법	7
1. 선행연구 및 사례 고찰	9
2. 데이터셋 검토	15
3. 질환발병률 추계모형 설계 및 시범 분석	20
4. 개발 모형의 타당성 평가 방안	29
III. 연구 결과	33
1. 연구결과의 제공 형태, 해석 방법, 활용 방법, 해석상의 주의점	35

2. 백혈병 발병률 추계 결과	41
3. 폐암 발병률 추계 결과	51
IV. 결론 및 제언	61
1. 연구 결과의 활용 방안	63
2. 연구 발전 방안 및 제언	64
참고문헌	69
부록	75
1. 사업장 기반 감시 자료의 기초통계량 요약	75
2. 근로자집단 암 발병률 추계값	82

표 목차

〈표 II-1〉 사업장기반 근로자집단 특성	17
〈표 II-2〉 사업장기반 근로자집단 특성	19
〈표 II-3〉 추계 대상의 계층시계열 수	23
〈표 III-1〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 남/여 합계	44
〈표 III-2〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 남성	45
〈표 III-3〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 여성	46
〈표 III-4〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류), 남/여 합계	47
〈표 III-5〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류), 남성	48
〈표 III-6〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류), 여성	49
〈표 III-7〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 남/여 합계	53
〈표 III-8〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 남성	54
〈표 III-9〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 여성	55
〈표 III-10〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 남/여 합계	56
〈표 III-11〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 남성	57
〈표 III-12〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 여성	58
〈표 부-1〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계	96
〈표 부-2〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계	98
〈표 부-3〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남성	100
〈표 부-4〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남성	102
〈표 부-5〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 여성	104

〈표 부-6〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 여성	106
〈표 부-7〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계	108
〈표 부-8〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계	113
〈표 부-9〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남성	118
〈표 부-10〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남성	123
〈표 부-11〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 여성	128
〈표 부-12〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 여성	133
〈표 부-13〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계	139
〈표 부-14〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계	141
〈표 부-15〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남성	143
〈표 부-16〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남성	145
〈표 부-17〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 여성	147
〈표 부-18〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 여성	149
〈표 부-19〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계	151
〈표 부-20〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계	156
〈표 부-21〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남성	161
〈표 부-22〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남성	166
〈표 부-23〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 여성	171
〈표 부-24〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 여성	176

그림목차

[그림 II-1] 암 발생 감시 설계 예시 (출처: 이경은 외, 2021)	16
[그림 II-2] 사업장 기반 질환 자료를 활용한 질환발병률 추계 통계모형 개요	21
[그림 II-5] 국제암연구소 업종별 위험도, IARC (2022)에서 재인용	31
[그림 III-1] 부록에 수록된 연구결과 제공 형태의 일부: (위) 114쪽의 <표 부-9> 및 (아래) 119쪽의 <표 부-10>	36
[그림 III-6] 백혈병 발병자수에 대한 업종별 추계 오차비율 (추계대상이 '15-'18년도일 때), 50	
[그림 III-7] 폐암 발병자수에 대한 업종별 추계 오차비율 (추계대상이 '15-'18년도일 때), ...	59
[그림 부-1] 대분류(UP1), 중분류(UP2), 소분류(UP3) 업종별 사업장 개수	75
[그림 부-2] 대분류 업종별 연도에 따른 수집/관찰 변수	77
[그림 부-3] 중분류 업종별 연도에 따른 수집/관찰 변수 (상위 15개)	77
[그림 부-4] 코호트별 연도에 따른 수집/관찰 변수	78
[그림 부-5] 코호트별 연도에 따른 수집/관찰 변수	79
[그림 부-6] 대분류 업종별 연도에 따른 수집/관찰 변수	80
[그림 부-7] 중분류 업종별 연도에 따른 수집/관찰 변수	80
[그림 부-8] 코호트별 연도에 따른 수집/관찰 변수	81
[그림 부-9] 사업장(대분류)별 백혈병 누적 발생 건수 추세	83
[그림 부-10] 사업장(대분류)별 남성 근로자 집단 대상 백혈병 누적 발생 건수 추세	84
[그림 부-11] 사업장(대분류)별 여성 근로자 집단 대상 백혈병 누적 발생 건수 추세	85
[그림 부-12] 사업장(중분류)별 백혈병 누적 발생 건수 추세	86
[그림 부-13] 사업장(중분류)별 남성 근로자 집단 대상 백혈병 누적 발생 건수 추세	87

[그림 부-14] 사업장(중분류)별 여성 근로자 집단 대상 백혈병 누적 발생 건수 추세	88
[그림 부-15] 사업장(대분류)별 폐암 누적 발생 건수 추세	89
[그림 부-16] 사업장(대분류)별 남성 근로자 집단 대상 폐암 누적 발생 건수 추세	90
[그림 부-17] 사업장(대분류)별 여성 근로자 집단 대상 폐암 누적 발생 건수 추세	91
[그림 부-18] 사업장(중분류)별 폐암 누적 발생 건수 추세	92
[그림 부-19] 사업장(중분류)별 남성 근로자 집단 대상 폐암 누적 발생 건수 추세	93
[그림 부-20] 사업장(중분류)별 여성 근로자 집단 대상 폐암 누적 발생 건수 추세	94

I. 서론

.....

I. 서론

1. 연구 배경

그간 산업안전보건연구원에서는 업종별 질환감시를 목적으로 직업성질환 코호트 사업을 운영하여 왔다. 직업성질환 코호트 사업의 성과로 사업장 기반 질환 자료가 구축되었다. 해당 자료는 고용보험에 등록된 근로자들에 대하여 개인단위 빅데이터(고용보험DB, 암등록DB, 통계청 사망DB 및 사망원인DB)를 가공하여, 각 사업장 단위에서 질환 (24종 암과 주요 만성질환)에 대한 연간 누적 발병률 정보가 수록되었다.

사업장 기반 질환 자료에는 개인정보가 수록되어 있지 않기 때문에 민감개인정보 노출위험을 피하면서도 사업장 단위가 질환감시를 가능케 하는 이점이 있다. 사업장 기반 질환 자료의 활용방안으로는 (1) 기존에 알려진 업무관련질환 외에도 과소평가되거나 간과되어 왔던 질환의 발견, (2) 선제적인 질환예방활동 (조기개입 등)의 근거자료 등이 제시된 바 있다 (이경은 외, 2021). 동시에, 의도된 활용방안으로 사업장 기반 질환 자료를 활용하기 위하여는 질환발병 추계 모형의 개발이 잠재적으로 기여할 여지가 있는데, 특히 사업장/업종별로 이질적인 근로자집단의 특성을 반영하면서도 시계열적 흐름을 반영한 추계모형 설계가 과제로 남아 있다.

2. 연구 목적

이번 연구에서는 기 구축된 사업장 기반 질환 자료를 활용하는 한편, 고용 규모나 근로자 집단의 특성 등 업종의 특성을 반영하여 발병자수를 추계하는 한편, 업무관련 추가발병률을 추계하고, 그 타당성을 검증하는 것이다.

그러나 비슷한 사례와 테이블이 없기 때문에 (확인 필요), 먼저 목적 달성의 가능성 확인을 위한 시범 모형과 분석 절차가 필요하다. 본 연구에서는 내부 및 산업안전보건연구원과의 협의를 거쳐, 위험업종 정보가 비교적 잘 알려진 폐암 및 백혈병을 대상으로 하여 위의 연구를 실시한다.

세부적인 목표는 다음과 같다.

1) 선행연구 및 사례 고찰

- 암 발병 추계 관련한 국내외 문헌을 수집하고, 사용된 모형과 자료의 형태를 고찰한다.

2) 데이터셋 검토

- 기 구축된 사업장 기반 질환 자료를 탐색하고, 자료의 특성을 요약한다.
- 본 보고서에서 목표하는 업종별 추계 문제에 기존 문헌들의 방법론을 적용할 수 있는지 그 가능성을 검토한다.

3) 질환발병률 추계모형 설계 및 시범 분석

- 기 구축된 사업장 기반 질환 자료에 적용할 수 있으면서, 업종별 특성

을 반영한 질환발병률 추계 통계모형을 설계/제안한다.

- 폐암 및 백혈병에 대하여 제안된 모형을 적합하여, 사업장 기반 질환 자료가 마지막으로 관찰된 시점(2018년) 이후 4년간(2019년-2022년)의 발병자수를 추계하는 한편 업무관련 추가발병률을 산출한다.

4) 개발된 모형의 타당성 검증

- (정량적 타당성 검증) 2019-2022년 4년간의 암 발병 추세가 2015-2018년과 비슷하다는 전제 하에, 제안된 통계 모형을 이용한 2015-2018년에서의 추계 성능을 평가한다.
- (정성적 타당성 검증) 폐암 및 백혈병 발병률 추계에서 검출된 위험업종과 기존 직업역학 문헌의 결과를 비교한다.

5) 근로자 집단의 질환 발병률 추계모형 활용방안 및 개선방안 제언

- 추계된 폐암/백혈병 발병률의 역학조사에서의 활용 가능성을 탐색하여 제언한다.
- 더 나아가, 추계/질환감시 모형을 발전시킬 수 있는 방안들을 폭넓게 검토하여 제언한다.

Ⅱ. 연구 내용 및 방법

.....

II. 연구 내용 및 방법

1. 선행연구 및 사례 고찰

1) 직업별 암 발생 연구

직업별 암 발생 비교 및 추계 연구는 주로 직업군을 비롯하여, 국가, 성별, 연령, 기간 그룹별로 정의된 각 집단에서 관찰된 암 발생 건수를 해당 계층별 개인의 연도와 국가인구의 발생률로부터 계산한 예상 건수와 단순 비교 하는 형태로 진행되어 왔다. Pukkala et al. (2009)은 5개 북유럽 국가 (덴마크, 핀란드, 아이슬란드, 노르웨이, 스웨덴)에서 1960년부터 1990년까지 이뤄진 대규모 인구 조사에서 30세~64세 인구 1,500만 명을 대상으로 280만 건의 암 진단 사례를 가지고 2005년경까지 추적조사를 진행하였으며, 많이 알려진 직업별로 유의한 연관성을 가진 암종을 표준화 발병비 (SIR)을 통해 확인한바 있다.

우리나라에서는 Kim et al. (2010)에서 한국의 직업성 암의 기여 비율 (PAF) 추정치를 제공하였다. 국제암연구소 (IARC)에서 제시한 1군 발암물질 9종과 표적발암물질 및 질병으로 각각 7종을 선정한 후, 발암물질에 노출된 근로자의 유병률을 추정하고 보정계수를 적용하여 그 값이 전체 인구를 대표할 수 있도록 하고, 상대위험도(RR)에 대한 데이터는 IARC 보고서에서 가져와 한국인 근로자에 대한 연구의 RR과 비교하였다.

최근 산업안전보건연구원에서는 2017년부터 지속적으로 국민건강보험공단 빅데이터를 기반으로 직업코호트를 통한 업종별 직업성 암 위험도 평가 연구를 수행하였고, 업종에 따른 직업코호트 및 유해물질 노출에 따른 특수건강진단 노동자 코호트 구축을 통해 우리나라의 직업성질환 코호트를 구축하고 모니터링 하고 있다 (산업안전보건연구원, 2019). 그러나, 직업별 암발생 연구

는 위험도 평가 등에 주로 집중되어 있으며, 시간에 따른 추세 예측 연구는 국가별 또는 대규모 집단별 추세 연구에서 간접적으로 수행되어 오거나 특정 암종을 중심으로 통계적 추세 모형을 제시해왔으며 국내외에서 진행된 집단별 암 발생 추세 방법과 사례들은 다음과 같다.

2) 해외 연구 사례

Yu et al. (2019)는 1988년부터 2018년까지 폐암의 발생률 및 사망률 예측을 한 모형에 대해 체계적 문헌 조사를 시행하였다. 암 관련 추세 예측에 대한 모형 조사 결과 총 101건의 연구가 수집되었으며, 흡연 관련 변수를 사용하였는지에 따라 1차적으로 구분하였다. 대부분의 연구는 흡연 관련 변수를 사용하지 않았으며 이 중에서 Age-Period Cohort (APC) 모형이 44건, 일반화 선형 모형이 35건, 선형 보간법 등의 기타 방법이 12건으로 나타났다. 이들 중, 흡연 관련 변수를 사용한 연구는 16건 정도였으며, 이 중에서 8건은 일반화 선형 모형을 사용하였으며, 2건은 APC 모형, 이 외에는 상용로그 변환을 활용한 선형 모형 등으로 보고되었다. 또한, 각 연구마다 자료의 질도 평가하였다. 5개 대륙의 IARC (International Agency for Research on Cancer)에 포함된 자료이거나, 모집단의 대부분을 자료에 포함하고 있는 경우 자료의 질이 좋다고 평가하였다. Freddie and Møller (2006)에서도 집단별 암 예측을 위한 방법론 비교를 하고 있다. 유용한 통계적 방법들로 APC, Power model, Age만 보정한 간단한 모형, 베이지안 방법들을 소개하고 있다.

Bray and Møller (2006)은 암 발병 추세 예측을 위해서는 성별, 나이, 시간(calendar time)을 구분한 자료가 있어야하며, 특히 APC 모형의 경우 오랜 기간동안 신뢰도가 높은 자료가 수집 되어 코호트 효과를 쉽게 추정할 수 있어야 한다고 보고하였다. 또한, 자료의 신뢰도에 따른 추세 예측을 방법과 다음과 같이 정리하였다. 먼저, 오랜 기간 수집되었으며 신뢰할 수 있는 자료

의 경우 APC모형의 지수 함수 부분을 x^5 로 대체한 power model이 적합하다고 하였으며, 신뢰할 수 있는 자료이지만 짧은 기간 수집된 자료의 경우 코호트 효과를 사용하기 힘들기 때문에 나이 효과 및 기간 효과만 사용한 간단한 모형이 적합하다고 하였다. 또한, 나이 분포의 및 aetiological profile의 변화 등에 영향을 받은 자료, 최근 시점의 발생률 추세에 대한 신뢰도가 높지 않은 경우 GLOBOCAN 등의 여러 국가의 암 추계 자료 등과 연관지어 분석하는 것이 필요하다고 하였다.

Rahib et al. (2021)은 SEER program(Surveillance, Epidemiology and End Results)과 US Census Bureau의 인구 추계자료를 통해 2040년까지의 미국의 암 발생자 수 및 사망자 수를 예측하였다. SEER program에서는 암 등록 자료에서 제공하는 인구학적 특징, 암 정보 및 활력 징후를 수집하였으며, US Census Bureau에서는 성별과 인종을 구분한 2040년까지의 인구추계 자료를 수집하였다. US Census Bureau의 인구 추계자료 기반의 현재 발생자 수 (연령별 발생률과 인구 추계의 가중 평균)와 SEER program을 통해 암 발생에 대한 연간 %변화율 평균(Average Annual Percent Change, AAPC)을 추정하고 지수함수 형태로 변환한 뒤 곱하여 발생자 수를 예측하였다. 미래 사망자 수의 경우 실제 사망자 수에 지수 함수화된 연간 %변화율 평균과 발생자 수를 곱하여 예측하였다. 이와 같은 분석 방법은 기존 연구(Rahib et al., 2014; Quante et al., 2016)에도 사용되어진 것으로 확인되었다. Weir et al. (2021)에서도 SEER program과 US Census Bureau 자료를 통해 2050년까지의 미국의 암 발생자 수를 예측하였다. 5년 단위 (1996-2000, 2001-2005, 2006-2010, 2011-2015) 및 5세 간격의 나이 (15-19세부터 85세 이상까지 총 15개)를 고려한 APC 회귀 모형 (age-period-cohort regression model)을 사용하였으며, 성별과 암 유형에 따라 개별적으로 모형을 적합하여 미래의 연령 표준화 발생률 (age-standardized incidence rate)를 예측하였다. 또한, 발생률의 지수적 증감을 고려하기 위해 5제곱($g(x) = x^5$) 연결 함수(link function)를 고려하였

다. Petrick et al. (2016)은 2000년부터 2012년까지 35세에서 84세 사이 단일 연도 HCC(간세포 암종) 사례 및 인구 데이터를 SEER Registry Database에서 수집한 후, APC모형을 이용하고 성별, 인종, 민족 및 연령별로 계층화하여 2030년까지 HCC 발생률을 예측하였다. 기존 출생 코호트에서 추정된 종단 연령 발생률 곡선에 특정 출생 코호트와 기존 코호트 간의 비율을 곱하여 연령별 미래출생 코호트별 발생률과 95% 신뢰구간을 추정하였다. 이때 미래 연도에 새로운 출생 코호트가 입력되어 예측에 반영되어야 하므로 새로운 코호트에 대한 비율은 관찰된 코호트의 비율을 사용하여 곡선 기울기의 변화를 설명할 수 있는 join point piecewise log-linear regression model에서 외삽하여 얻은 값을 이용하였다.

Smittenaar et al. (2016)은 NCRAS (National cancer Registration and Analysis Service)에서 수집한 자료 중 1979년부터 2014년까지의 암 발생 및 암으로 인한 사망 정보를 통해 2035년까지 영국의 암 발생률 및 사망률을 예측하였다. 5세 간격의 나이(15-19세부터 90세 이상 까지)와 성별을 고려한 APC 회귀 모형을 사용하였으며, 생물학적으로 타당한 비전염성 질환 자료 분석을 위해 각 성분(age, period, cohort)에 3차 자연 스플라인(natural cubic splines)의 함수로 설정하여 모형을 적합하였다. 연결 함수로는 5제곱 함수 및 로그 함수를 둘 다 사용하여 민감도 분석을 시행하였다.

Mistry et al. (2011)는 1975년부터 2007년까지 영국에서 암 발병 데이터에 APC(Age- Period-Cohort) smooth function modeling을 이용하여 2030년 영국 인구에 적용하였다. 5세 연령 그룹 및 연도별, 성별, 1995-2007년 28개 암의 발생률 데이터는 Cancer Research UK에서, 5세 연령 그룹 및 연도별 2008-2030년 국가인구 예측은 통계청에서 수집하였다. 암발생률은 APC모형을 사용하여 모델링 되었으며, link function으로 지수함수(exponential)와 5 제곱 함수(power-5) function을 이용하였고 각 성분(age, period, cohort)에는 5년 간격의 step functions과 cubic spline을 적용하였다.

Jakobsen et al. (2021)은 CEDAR (Cancer Impact in Denmark Study) 연구의 일환으로 2006년부터 2015년까지의 의료 기록 정보인 Danish Cancer Registry 자료 및 NORDCAN (Cancer Statistics for the Nordic Countries) 자료를 통해 2030년까지의 폐암의 발생률, 사망률, 유병률을 성별 및 나이를 구분하여 예측하였다. Danish Cancer Registry에서 제공하는 연도, 성별, 나이(40-59세, 60-69세, 70-79세, 80세 이상) 정보를 통해 폐암의 발생률과 사망률을 포아송 일반화 선형 모형(Poisson generalized linear model)으로 예측하였으며, 유병률은 기존 유병 건수에 예측된 발생자 수와 사망자 수의 차이를 더한 뒤 인구수로 나누어 예측하였다.

3) 국내 연구 사례

Jung et al. (2021)은 2021년 한국의 예상 암 발생률 및 사망률을 보고하여 암 부담을 추정하였다. 1999년부터 2018년까지 암 발생률 데이터는 Korea National Cancer Incidence Database에서, 1993년에서 2019년까지 암 사망률 데이터는 통계청에서 수집되었고, 해당 데이터를 이용하여 암 발생률과 사망률을 예측하였다. join point regression model을 사용하여 선형 추세가 크게 변경된 연도를 찾은 후, 최신 트렌드를 보이는 데이터만 이용하여 연령별 5세 집단별 발생률에 선형회귀모형을 적합한 후, 2021년 예상 연령별 발생률에 예상 연령별 암 발생률을 곱하여 예측되었다.

Shim and Choi (2014)는 1999년부터 2011년까지 작성된 국가 암 등록 통계사업에서 얻은 자료를 기반으로 일반화선형모형을 이용하여 암의 연령별 발생률을 추계하였다. 본 논문에서는 사회구조적 변화를 반영하기 위해 암발생률의 로그값이 경과기간에 따라 선형적으로 변화하는 일반화 선형모형을 가정하였다. 이때 모형의 매개변수들은 1999~2011년까지 12년간 암발생률 자료를 이용하여 주성분분석 방법으로 적합하였다. 2012년부터 10년간 암발

생률 지표를 시계열로 예측해본 결과, 연령에 상관없이 암발생률의 상승 패턴이 관측되었으며, 암발생률의 증가 속도는 30-34세가 가장 높은 것으로 나타났다.

Shin et al.(2002)은 국가 사망률 데이터와 강화, 서울, 부산, 대구 4개 지역 암 등록기관의 암 발생률 데이터를 이용하여 암의 발생률을 추정하였다. 7개의 연령 그룹(0~44, 45~49... 70~75+)이 이용되었고 0-44세 그룹을 암 등록부에서 제공한 비율을 사용하여 0-14, 15-44세로 나누어서 이용하였다. 특정 암 발생률의 로그가 각 성별 그룹의 사망률 로그의 선형 함수로 표현할 수 있다고 가정하였고 25개의 암에 따라 별도의 모델을 fitting하였다.

Bae et al.(2002)는 모집단을 대상으로 하는 지역 암 등록 자료를 이용하여 국가 전체를 대표하는 암 발생률을 추계 하기 위한 방법들을 정리하였다. 암 발생자료가 전혀 없고, 전체 암에 대한 사망률 자료만 있는 경우 모든 암에 대한 연령/성별 발생률과 암종에 대한 상대빈도를 이용하여 단순 추정하는 방법이 있다. 또한, 질 높은 사망률 자료가 연령별, 성별, 암종별로 있는 경우 사망 교정계수를 이용하거나 사망률과 성별, 연령 그룹변수를 설명변수로 하는 로그 선형 모형을 적합하여 이용하기도 한다. Bae et al.(2002)이 정리한 시점 기준으로, 2000년대 초반의 미국에서는 SEER 프로그램을 중심으로, 각 SEER 암 등록소의 발생률을 합한 발생률 (aggregate incidence)에 연령 및 성별 인구집단에 곱하고, 각 암의 발생률:사망률 비를 사용하여, 각 지역의 추정 발생률을 유도해낸다. 미래에 대한 예측은 AR 모형을 사용한바 있다.

2. 데이터셋 검토

본 연구에서 다루는 “사업장기반 질환 감시자료”는 윤진하 외 (2017)에서 처음 제안되었으며, 후속 연구(이상길 외 (2018, 2020), 이경은 외 (2021))에서 수정/보완된 것이다. 특히 이경은 외 (2021)에서는 사업장 기반 질환 감시 데이터베이스를 구축하였고 이전 조사에서 다루지 않은 직업성 질환(e.g. 백혈병)까지 포함하여 업종별로 감시하는 연구를 수행하였다. 본 절에서는 본 연구진이 확보한 사업장 기반 감시 자료의 기본적인 특성을 서술한다. 보다 구체적인 사항은 사업장 기반 감시 자료 관련 가장 최신 문헌인 이경은 외 (2021)에 자세히 소개되어 있으나, 본 연구진이 확보한 데이터셋에는 일부 차이가 있어, 본 연구에서 직접적으로 이용한 부분에 대하여는 재소개가 필요하다고 판단하였다.

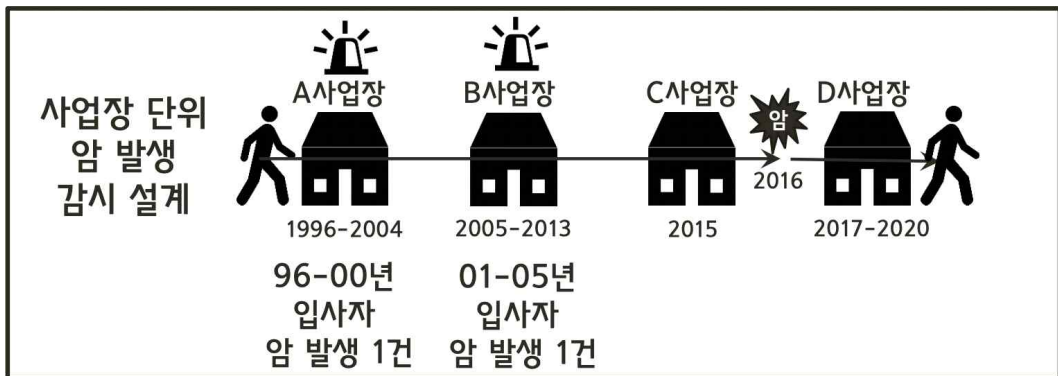
사업장기반 질환 감시자료는 세 가지 원시자료(고용노동부의 고용보험자료, 국립암센터의 암등록자료, 통계청의 원인별사망자료)로 구성된다. 사업장기반 질환 감시자료는 관찰단위가 사업장 및 하위 층화하였기 때문에 개인정보식별의 위험성이 줄어드는 장점이 있는 반면, 앞서 살펴본 암 추계 방법론의 적용이 본 데이터셋 대상으로 가능할지는 검토가 필요한데, 이는 자료 소개 이후에 언급토록 한다.

1) 관찰 대상

관찰 대상은 연도 기준으로 2년 이상 종사한 근로자이다. 이는 한 사업장에 만 1년 이상 종사한 근로자만을 추출하는 의미를 갖고 있다. 예를 들어, 연도 기준으로 1년을 기준으로 할 경우, 2010년 12월에 입사 후 2011년 1월에 퇴직한 경우도 포함이 된다 (데이터베이스 연도에서는 2010년에 1년 근무한 것으로 계산됨). 이와 같은 경우에 사업장의 위험 노출이 해당 근로자의 발병

에 영향을 미쳤다고 보기 어렵기 때문에 “연도 기준으로 2년 이상” 조건이 고려되어 있다.

암이나 심혈관질환 등과 같은 만성질환의 경우 최소 몇 년 동안의 잠복기를 거쳐 발생하는 것이 일반적이기 때문에 퇴직 이전의 사업장 요인이 질환 발생에 영향을 줄 수 있는 점에 착안하여, 사업장 기반 감시자료에는 사업장 별 퇴사자를 포함한 누적 근로자 수와 누적 발생자 수에 대한 정보가 집계되어 개별 변수로 구성되어 있다. 누적 변수의 관찰 대상은 암이 노출 시점으로부터 최소 5년 이상의 잠복기를 가진다고 가정하고, 입사 후 5년 경과 이전 발생한 암 사례는 누적 대상에서 제외되어 있다. 예를 들어, [그림 I-1]에서 2016년에 근로자에게 발생한 암은 A, B 사업장에는 발생 사건으로 고려되지만 C 사업장(2015년에 근무)에는 집계에 포함되지 않는다.



[그림 II-1] 암 발생 감시 설계 예시 (출처: 이경은 외, 2021)

2) 자료 구조

감시자료의 각 행은 사업장과 그 하위의 근로자집단 층화변수 (성별, 입사 시기군, 입사시연령군) 및 연도(2000-2018)가 단위이며, 열을 구성하는 변수는 상기 언급된 집계단위 5개 (사업장 고유번호, 입사시기군, 입사시연령군, 연도)와 수집/관찰 변수 (백혈병 누적발병건수, 백혈병 누적추적인년합계, 폐

암 누적발병건수, 폐암 누적추적인년합계)로 구성되어 있다. 근로자집단 특성을 정의하는 세 층화변수인 성별, 입사시기군, 입사시연령군의 값은 아래 <표 I-2>에 소개되어 있다. 세 층화변수는 총 48개(=성별2 × 입사시기6 × 입사시연령4) 조합을 생성한다. 이전 보고서들에서는 각 조합에 해당하는 근로자 집단을 코호트라고 구분하여 명명하고 있다.

위의 집계정보와는 별도로, 각 사업장을 설명하는 정보로서 각 사업장이 속한 대분류/중분류/소분류 업종 정보가 수록되어 있다. 대분류 업종코드(UP1)는 A부터 U까지의 알파벳 대문자 21개, 중분류 업종코드(UP2)는 1부터 99 사이의 자연수 77개, 소분류 업종코드(UP3)는 11부터 990 사이의 자연수 232개로 이루어졌다. 대분류, 중분류, 소분류 업종을 알 수 없는 사업장은 NA로 기재되어 있는데, 본 분석에서는 제외하였다.

〈표 II-1〉 사업장기반 근로자집단 특성

변수명	변수설명	코딩	코딩설명
SEX	성별	남	남성
		여	여성
CAL	입사시기	0	95년 및 그 이전
		1	96-00년
		2	01-05년
		3	06-10년
		4	11-15년
		5	16-21년
SEX	입사시 연령	20	20-29세
		30	30-39세
		40	40-49세
		50	50-59세

수집/관찰 변수는 질병(폐암, 백혈병) 별 누적추적인년 합계 및 누적발병건수로 이루어져 있어, 근로자 집단마다 연도별로 누적발병률을 계산할 수 있게 한다. 보다 자세히 알아보면, 누적추적인년합은 해당 관찰시기에서 관찰한 해당 사업장 내 해당 코호트의 누적인년수치이다. 누적발병건수(LEUKEMIA_BZ)는 해당 관찰시기에서 관찰한 해당 사업장 내 해당 코호트의 퇴사자를 포함한 누적 질병 발병건수이다. 따라서 동일 사업장 내 동일 코호트의 경우 연도가 지날수록 값이 같거나 커진다.

해당 자료에 집계되지 않은 코호트 혹은 관찰연도는 해당하는 근로자가 없다는 것으로 해석할 수 있다. 예를 들어 사업장 A를 구성하는 코호트가 (SEX / CAL / ECNY_AGE)=(남 / 4 / 20), (남 / 5 / 20), (여 / 4 / 20), (여 / 5 / 20)인 경우, 사업장 A에 2010년 이전에 입사한 근로자는 없는 것으로 해석하면 된다. 또한 사업장에서 특정 코호트(예를 들어, (남 / 4 / 20)의 경우)가 관찰되어도 관찰연도(YEAR)의 시작값이 2013이라면, 20대에 입사한 남자 근로자 중 2011년 이후에 입사한 자가 없거나, 2011년 이후에 입사했지만 2년 미만 근무한 자로서 집계되지 않았을 수 있다.

〈표 I-2〉는 특정 사업장에 내에서 추출한 특정 코호트의 테이블 예시이다. 해당 코호트는 성별이 남성이고 입사시기가 95-99년 사이, 입사시연령 50-59세를 대상으로 하고 있다. 연도에 따라 질환(백혈병 및 폐암)의 누적추적인년합계와 누적발병건수를 추적할 수 있다.

〈표 II-2〉 사업장기반 근로자집단 특성

사업장 식별번호	성별	입사시 기	입사시 연령	관찰연 도	백혈병 -누적 추적인 년합계	백혈병 -누적 발병건 수	폐암- 누적추 적인년 합계	폐암- 누적발 병건수
103*****	남	1	50	2000	6352	0	...	...
103*****	남	1	50	2001	8466	0	...	...
103*****	남	1	50	2002	10557	0	...	...
103*****	남	1	50	2003	12634	1		
103*****	남	1	50	2004	14697	1		
103*****	남	1	50	2005	16738	1		
103*****	남	1	50	2006	18754	1		
103*****	남	1	50	2007	20750	1		
103*****	남	1	50	2008	22719	3		
103*****	남	1	50	2009	24662	3		
103*****	남	1	50	2010	26575	3		
103*****	남	1	50	2011	28452	3		
103*****	남	1	50	2012	30303	4		
103*****	남	1	50	2013	32123	4		
103*****	남	1	50	2014	33910	4		
103*****	남	1	50	2015	35664	5		
103*****	남	1	50	2016	37384	6		
103*****	남	1	50	2017	39058	6		
103*****	남	1	50	2018	40694	7		

* 1996-2000년 입사 당시 50-59세 남자 근로자 예시

3. 질환발병률 추계모형 설계 및 시범 분석

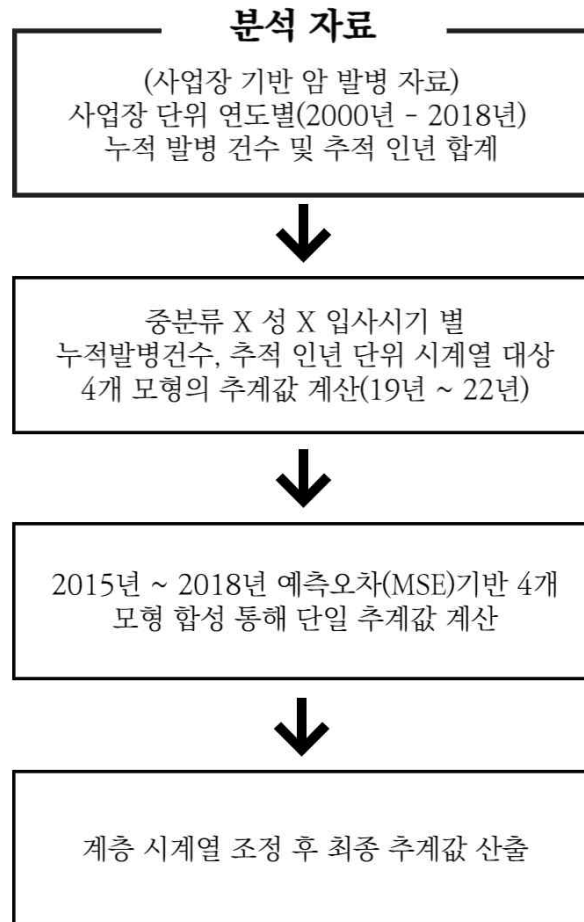
국내외 사례를 종합적 살펴보았을 때, 각 층(예-성별)에서 범주화된 현재나 이와 시간(calendar time)의 모든 조합에서 암 발생 또는 사망 자료가 주어지는 경우 일반화 선형 모형 또는 APC 모형을 이용하는 사례로 이해될 수 있다. 이러한 분석 기법은 나이와 시간에 대한 정보가 다른 형태로 주어지는 경우 직접적으로 적용이 불가능하다는 어려움을 가지고 있다. 예를 들어 현시점에서 가용자료인 사업장 질환 감시자료에서처럼 근로자의 나이에 대한 정보가 '입사시기의 범주화된 나이'로 한정되는 경우 해외 사례의 방법론은 그대로 적용하기 어려운 것으로 판단하였다. 이에 사업장 질환 감시자료의 자료 형태에 맞는 새로운 모형을 고려하게 되었다.

1) 추계모형 설계 개요

본 연구진은 확보된 자료의 특징과 자료를 이용한 암 발병률 추계에서의 도전점을 다음과 같이 정리하였다.

- (1) 이상적으로는 모든 가용 세부층화(업종소분류, 성별, 입사시기군, 입사시연령군)의 임의의 조합에서의 암발병 추계가 가장 많은 정보를 줄 수 있겠으나, 계층을 세분화할 경우 발병 건수 및 추적대상의 표본수가 부족해지는 조합이 존재하여 추계값의 품질을 보장하기 어려움.
- (2) 세부층화별 연간 암발병 관련 관찰값들은 시계열적 특성이 각각 다르므로, 어느 단일한 모형이 모든 세부층화별 암관련 지표를 잘 적합하기보다는 적합한 모형이 각기 다를 수 있음.
- (3) 추계 대상의 시계열은 계층 구조를 가지고 있기 때문에 (하위 세부계층의 누적발병건수의 합이 상위 계층의 누적발병건수와 동일함), 계층시계열 조건을 만족시키는 암발병건수 추계가 필요함.

위 사항들에 대하여, 아래와 같은 접근법을 취하기로 결정하였다. 전체적인 흐름은 아래 그림과 같으며, 단계별 핵심은 아래와 같다. 각 단계에 대한 통계모형의 세부사항은 후술한다.



[그림 II-2] 사업장 기반 질환 자료를 활용한 질환발병률 추계 통계모형 개요

- (1) 일부 세부층화 (중분류 업종, 성별, 일부 병합된 입사시기)를 선택하여, 해당 세부층화 변수의 모든 조합 대하여 폐암, 백혈병의 누적 발생 빈도와 누적추적인년에 추계 모형을 적합였다. 여기서 업종은 중분류별로, 성별은 남성 및 여성, 입사시기는 입사연도가 2000년도까지 및 2000년도 이후의 값을 갖는다.¹⁾

- (2) 암종과 세부층화의 누적추적인년과 누적발병건수의 각 시계열에 대하여, 독립적으로, 여러 모형을 적합하였다. 또한 각 시계열에 대하여 단일한 모형을 선택하지 않고, 예측 오차에 따라 가중치를 달리하여 방법론별 추계값을 합성하여 합성 추계값을 계산하였다.
- (3) 2번에서 계산된 시계열별 누적 예측값에 대하여, 계층 조건을 만족시켜주는 사후조정 방법을 적용하였다.

2) 추계 대상 시계열 선정

추계가 필요한 시계열을 다음과 같이 정리하였다. 본 모형 적용의 우선순위를 협의하였을 때, 성별 차이를 반영하고 제도변경의 효과를 추적하면서도 통계적인 안정성을 확보하기 위하여 성별(남/여) 및 입사년도 (2000년도까지/2000년도 이후)의 층화를 유지하되 다른 층화들은 병합하는 것으로 결정하였다. 입사년도의 경우 최근 시점으로 올수록 추적인년이 짧기 때문에, 충분한 표본을 확보하기 위하여 기존 자료의 분류('95년도까지, '96-'00년, '01년-'05년, '06-'10년, '11년-'15년)을 2000년도('95년도까지 및 '96-'00년)와 2000년도 이후('01년-'05년, '06-'10년, '11년-'15년)으로 통합하였다. 결과적으로, 각 846개의 계층시계열은 연도(2000년-2018년)별로 폐암 및 백혈병의 발병 빈도값 19개씩을 갖는다.

1) 따라서 원자료에 대하여는 성-연령 표준화 발병비의 계산이 가능하나, 본 연구에서 추계된 값들에 대하여는 성-연령 표준화 발병비의 계산은 불가능하게 된다. 대신 성-표준화 발병비는 계산이 가능하다.

〈표 II-3〉 추계 대상의 계층시계열 수

(단위 : 단위시계열의 개수)

특성구분 업종분류 수준(Level)	전체	성별(남성/ 여성)	입사시기(2 000년도까 지/2000년 도 이후)	성별 × 입사시기	합계
전체	1	2	2	4	9
대분류	19	38	38	76	171
중분류	74	148	148	296	666
합계	94	182	182	364	846

3) 단위시계열 추계 방법론들

각 846개의 시계열에 대하여, 아래의 모형들을 이용한 추계값을 산출하였다. 즉 하나의 단위 시계열마다 4개의 추계값이 존재하게 된다. 이들의 합성 방안은 후술하였다. 추계의 범위는 2019-2022년 (4개년도)으로 설정하였다. 이론적으로 각 모형들은 2023년도 이후의 추계값들도 모두 계산할 수는 있으나, 데이터 기반의 추계 및 추계모형들을 사용하는 경험적 관례에 비추어 보았을 때, 모형에 인입된 관찰기간 (19개년도)의 20% 정도인 4개년도가 적당하다고 판단하였다.

추계에 사용된 통계모형들은 아래와 같다.

- (1) 단순 선형 모형: $E(y_t) = \beta_0 + \beta_1 \times t$
- (2) 이차식 모형: $E(y_t) = \beta_0 + \beta_1 \times t + \beta_2 \times t^2$
- (3) 선형 스플라인 모형:

$$E(y_t) = \beta_0 + \beta_1 \times (t - 2005)_+ + \beta_2 \times (t - 2010)_+ + \beta_3 \times t$$
- (4) 푸아송 모형: $E(y_t) = \exp(\beta_0 + \beta_1 \times t)$

여기서 $t = 2000, 2001, \dots, 2019$ 는 연도를 뜻하고, y_t 는 t 년도에서의 누적건수 관찰값을 뜻한다. 하나의 시계열의 관찰값은 y_{2000} 부터 y_{2019} 까지 존재하는데, 위 모형들을 적합한 뒤의 $t = 2019, 2020, 2021, 2022$ 에 대한 적합값 $\hat{y}_{2019}, \hat{y}_{2020}, \hat{y}_{2021}, \hat{y}_{2022}$ 를 추계값으로 사용하게 된다.

(1) 단순 선형 모형은 간단한 해석을 가능케 하는 이점이 있다. 또한 단순 선형 모형 기반 추계는 신규 발병건수가 해마다 동일한 경우에는 타당한 추계값을 제공할 수 있다. 그러나 신규 발병건수가 증가하거나 감소하는 업종에 대하여는 적합하지 않을 수 있다. (2) 선형 스플라인 모형은 입사시기가 5년 단위로 구분되어 있는 데이터의 특성을 반영하기 위해 설계한 모형으로, 예를 들어 특정 업종이 2000년도 이후 급히 팽창하여 사업장이 늘어난 업종이라 2005, 2010년에 누적발병건수에 꺾임 경향이 있는 경우 모형적합이 잘 될 것으로 기대된다. (3) 이차식 모형은 자료탐색 과정에서 누적발병건수의 증가추세가 대체로 선형추세보다 빠른 점에서 착안하여 추가한 모형으로, 신규 발병건수가 꾸준히 증가하거나 꾸준히 감소하는 업종에 대하여 타당한 추계값을 제공할 수 있다. (4) 푸아송 모형도 이차식 모형과 마찬가지로 동기에서 고려된 모형이다. 다만 푸아송 모형의 경우 신규발병건수가 직전해 누적발병건수에 비례함을 가정하고 있어, 전염 관련 누적발병건수 예측에는 적합할 수도 있으나 암 발생처럼 전염성이 적고 환경적/유전적 요인이 많은 질환에 대하여는 적합하지 않을 수 있다.

4) 추계값 합성

앞서 살펴보았다시피, 여러 세부계층의 시계열들은 서로 기저의 현상과 가정이 다르므로, 하나의 모형만으로 추계하기에는 한계가 있고 오히려 다양한 모형을 합성하였을 때 예측력을 개선할 만한 여지가 있다. 다시 말하여, 하나의 모형만 선택하는 것보다는, 예측력이 좋은 모형들이 여러 개를 합성하는 앙상블(ensemble) 또는 모델 평균화(model averaging) 기법을 이용하면 여

러 가정에 더 강건한 추계값을 얻을 수 있다.

양상블 방법을 상술하면 다음과 같다. 예를 들어 단위 시계열에 대하여 추계 방법론이 M 개가 있고 m 번째 추계방법에 대한 t 시점에서의 추계값을 $\hat{y}_t^{(m)}$ 이라고 하면, 다음과 같은 가중-합성추계값 $\hat{y}_t$ 를 생각할 수 있다.

$$\hat{y}_t = w_1 \hat{y}_t^{(1)} + w_2 \hat{y}_t^{(2)} + \dots + w_M \hat{y}_t^{(M)}$$

단, 여기서 w_m ($w_m \geq 0$, $m = 1, \dots, M$)들은 가중치이며, $w_1 + w_2 + \dots + w_M = 1$. 가중치를 계산하는 여러 방법이 있으나, 대개의 경우 예측오차가 커질수록 가중치가 지수적으로 감소하는 방법을 많이 사용한다. 여기서는 추계모형별로 계산한 평균제곱오차(MSE)를 이용하여 가중치를 계산하였다. 즉,

$$w_m = \frac{\exp(-MSE_m)}{\sum_{m=1}^M \exp(-MSE_m)}, \quad m = 1, \dots, M$$

여기서 MSE_m 은 m 번째 모형의 최근 4개년(2015-2018) 예측오차제곱의 평균, 즉

$$MSE_m = \frac{1}{4} \sum_{t=2015}^{2018} (\hat{y}_t^{(m)} - y_t)^2$$

이다. 예를 들어, A업종 백혈병 누적발병건수의 2019년 값에 대하여 단순 선형법 추계치가 10명이고 이차식법 추계치가 12명인 상황에서 두 개의 방법론만을 합성하는 상황을 고려하자. 만약 단순선형법의 MSE가 1, 이차식 추계의 MSE가 0.5이면 합성된 추계치는

$$\frac{e^{-1.0}}{e^{-(1.0+0.5)}} \times 10 + \frac{e^{-0.5}}{e^{-(1.0+0.5)}} \times 12 = 11.24 \text{가 된다.}$$

위 앙상블 방법 사용에 있어 유의점은 다음과 같다. 첫째, 1999-2018년 19개년 자료를 모두 사용하여 위의 추정오차를 구할 경우 늘 복잡한 모형의 오차가 더 적고 (과적합) 미래의 추계값의 추정오차를 늘 과소추정하게 되기 때문에, 이를 방지하기 위하여, 위의 MSE_m 를 계산할 때는 모든 모형에 대하여 1999-2014년의 15개년 관찰값만 이용하였다. 둘째, 자료의 특성상 단위 시계열마다 스케일이 다르며, MSE의 단위가 수천을 넘는 단위시계열들은 컴퓨터를 이용한 $\exp(-MSE)$ 의 수치계산이 컴퓨터의 부동소숫점 표현 범위를 넘을 가능성이 크다. 따라서 위의 MSE_m 를 계산할 때는 단위시계열의 스케일을 통일하였는데, 구체적으로는 2014년의 누적발병건수가 1이 되도록 상수 배하였다. 셋째, 2014년까지 누적발병건수가 0이어서 모형 합성이 의미없는 경우, 본 과제의 추계가 의미없는 경우라고 판단하여 가장 기본적인 단순선형법의 가중치를 1로 두고 다른 모형은 이용하지 않았다.

5) 계층 시계열 조정

앞서 서술했듯, 추계 대상의 846개의 누적발병건수 시계열은 계층시계열 구조를 가지고 있다. 다시 말해서, 전체집단에 해당하는 폐암 누적발병자수의 시계열의 연도별 값은 가장 하위 수준인 중분류×성별×입사시기의 296개의 시계열들의 폐암 누적발병자수들의 합계와 동일하다. 그렇지만 B, C에서 서술된 추계 방법은 계층시계열 조건에 상관없이 적용된 것이므로, B-C 절차를 적용한 2019-2022년도의 추계값들은 계층시계열 조건을 만족시키지 않는다. 예를 들어 C에서 소개한 예시로 돌아가면, A업종 백혈병 발병건수의 2019년 추계치가 11.24로 산출되었을 때, A업종 남성의 백혈병 발병건수 추계치가 5.1이고 여성의 발병건수 추계치가 7.2으로 둘을 더한 값 $5.1+7.2=12.3$ 이 11.24와 같지 않게 된다.

그러나 계층시계열 조건을 만족시켜야 후속절차 (이를테면 표준화 발병률비 계산 등)가 가능해지기 때문에, 추계값이 계층시계열 조건을 만족시키는

것은 필수불가결하다.

계층시계열에서 하위계층 시계열의 추계값의 합이 상위계층 시계열의 추계값과 일치하지 않는 문제를 해결하기 위한 연구는 다양하게 진행되어 왔는데, 크게 (1) top-down 방법, (2) bottom-up 방법 및 (3) optimal combination 방법이 있다. (1) Top-down 방법은 최상위계층 시계열(들)을 먼저 추계하고, 하위 시계열들에 대하여는 상위 시계열에서의 구성비율을 추계한 뒤 상위시계열 추계값에 곱하는 방법이다. Top-down 방법의 경우 상위시계열 추계값에 편향이 없다 하더라도 하위시계열 추계값에 편향이 생길 수 있음이 Hydeman et al. (2011)에 의해 지적되었다. (2) Bottom-up 방법은 최하위계층 시계열들을 먼저 추계한 뒤 이들을 더하여 상위계층 시계열을 추계하는 방법이다. Bottom-up 방법의 경우 하위 시계열 추계값은 대개 불안정하고 변동성이 클뿐더러, 그 시계열들이 독립적이라고 가정하고 상위계층 시계열을 구성하므로, 상위계층 시계열 추계값들이 top-down 방식에 비하여 부정확하고 변동성이 큰 단점이 있다. (3) Optimal combination 방법은 각 단위 계층시계열에 대한 추계값을 먼저 계산한 뒤 (기본추계값), 최종 추계값은 기본추계값들의 가중결합으로 계산하는 방법이다. 가중결합의 가중치 결정에 대하여 다양한 방법이 제안되었는데, 특히 Wickramasuriya et al. (2019)의 minimun trace (MinT) 방법의 경우 가중치는 최종추계값의 분산을 최소화하면서 최종추계값이 계층시계열 조건을 만족시키도록 하는 수치최적화(numerical optimization) 문제의 해로 주어지게 된다. 한편, 수치최적화의 해를 이용하면 최종추계값의 분산을 추정할 수 있으므로, 이를 이용하면 최종추계값의 기댓값에 대한 신뢰구간을 건설하였다. 자세한 수식은 Panagiotelis et al. (2022+)에 수록되어 있다.

본 보고서에서는 계층시계열 조정을 위하여 MinT 방법을 이용하였는데, 가장 큰 이유는 MinT 방법이 상위시계열은 물론 하위시계열 추계에서 MinT 방법이 top-down 및 bottom-up보다 더 좋은 예측력을 보여줌이 여러 전

산실험으로 확인된 바 있으며 (Wickramasuriya et al., 2019), 본 연구문제의 특성상 하위시계열 뿐만 아니라 상위시계열의 추계 또한 중요하기 때문이다. 또한 MinT방법은 통계청의 가계동향조사 시계열조사 단절로 인한 연계 연구문제에서도, 가계동향조사 항목의 계층구조 조정에 성공적으로 적용된 바 있다 (여인권 외, 2022)

6) 추가 발병률 계산

상기 절차들을 아래 네 개의 누적값들에 적용한다:

- (1) 폐암, 연간 누적 발병건수
- (2) 폐암, 연간 누적 추적인년
- (3) 백혈병, 연간 누적 발병건수
- (4) 백혈병, 연간 누적 추적인년

그 결과로, (1)-(4)의 각 누적값에 대하여 하위계층(분류, 성별, 입사년도)을 포함한 846개의 전 계층의 2019-2022년의 추계값을 확보하게 된다. 이를 이용하여 각 계층별 폐암/백혈병의 누적 발병건수에 더하여, 전체집단을 표준 집단으로 상정하였을 때의 각 업종의 발병비를 계산하였다.

4. 개발 모형의 타당성 평가 방안

1) 정량적 평가 방안

가용자료는 2018년까지의 관찰값만 존재하므로, 2019-2022년까지의 추계값을 실측값에 기반하여 정량적으로 평가할 방법은 보고서 작성 시점에서 존재하지 않았다. 일반적으로 실측값이 없는 추계값의 정량적 성능평가가 필요한 경우에는, (현재의 추세가 미래에도 지속된다는 전제 하에) 실측값이 가용한 가장 최근시점들에서 예측값과 실제값을 비교하는 평가 방법이 흔히 고려된다.

본 연구에서는 2015-2018년을 이용하여 정량적인 평가를 실시하는 방안을 고려하였다. 아래 두 가지 지표를 모두 고려하였는데, 둘의 가장 큰 차이는 통계적 방법을 적용하는 자료 범위가 다르다는 점이다.

- **훈련오류 (training error, in-sample error):** 2015-2018년 발병자수를 전제(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 오차
- **검증오류 (validation error, out-sample error):** 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 오차
- 두 오류에 대하여 오차의 지표는 제곱근평균오차비율(root-mean-squared error ratio, RMSE ratio) 이용

$$\text{제곱근평균오차비율} = \frac{\sqrt{\frac{1}{4} \sum_{t=2015}^{2018} (\tilde{y}_t - y_t)^2}}{\sqrt{\frac{1}{4} \sum_{t=2015}^{2018} y_t^2}}$$

y_t : t 년도에서의 누적 발병건수 관찰값

$\tilde{y}_t$: 제안된 방법 (모형적합, 앙상블, 계층시계열 조정)을 거친 y_t 의 예측값

직관적으로, 훈련오류는 현재의 정보를 모두 아는 상태에서 실제값과 적합값을 비교하는 것으로, 순수하게 모형의 적합 성능을 평가하는데 주로 사용된다. 검증오류는 현재의 정보를 모르는 상태에서 실제값과 적합값을 비교한다.

예를 들어 업종 A의 (남성, 2000년 이후 입사자) 집단의 폐암 누적 발병건수에 대한 훈련오류가 0.10이고 검증오류가 0.25라 하자. 먼저 훈련오류=0.1은 살펴보면 모든 가용 정보가 주어진 상태에서도 관찰값 대비 10%의 오차가 있다는 뜻으로, 가용정보를 모형이 설명하지 못하는 임의의 변동이 10%의 오차를 야기하는 것으로 해석할 수 있다. 다음으로 검증오류=0.25를 살펴보면, 과거의 정보만 이용가능한 상태에서 관찰값 대비 25%의 오차가 있다는 뜻으로, 모형이 설명하지 못한 임의의 변동에 더하여 모형이 잡아내지 못한 추세에서의 오류까지 더하여 25%의 오차가 발생한 것으로 해석할 수 있다. 직관적으로 검증오류가 훈련오류보다 큰 편이다.

2) 정성적 평가 방안

만일 추계값 기준으로 참조집단 대비 발병비율을 계산하였을 때의 위험업종이 기존에 알려진 암종별 위험업종과 중첩되는 바가 있다면, 제안된 방법이 어느 정도의 타당성이 있다고 판단하였다. 암종별 위험업종 자료로서 국제암연구소의 암종별 위험업종 보고를 참조하였다(cite). 폐암의 경우 충분한 증거의 연관성이 있는 업종으로 고무, 페인트, 철강 관련 제조업이 지적되었고, 한정적인 증거의 연관업종으로 인쇄업이 지적된 바 있다. 백혈병의 경우 충분한 증거의 연관업종으로 고무 제조업이 지적되었고, 한정적 증거의 연관업종으로 석유 정제업이 지적된 바 있다.

통계적인 관점에서는 산출된 추계값의 타당성을 평가하기 위하여 추계대상

(2019-2022년도)의 실측값을 확보한 뒤 추계값과 실측값의 오차를 계산하는 것으로, 이 작업은 추후 사업장 기반 질환 자료가 업데이트되면 가능할 것으로 보인다.

	Carcinogenic agents with sufficient evidence in humans	Agents with limited evidence in humans
Oesophagus		Dry cleaning
Nasal cavity and paranasal sinus		Rubber manufacturing industry
Larynx		Textile manufacturing industry (work in)
	Rubber manufacturing industry	Rubber manufacturing industry
Lung	Painter (occupational exposure as a)	Printing processes (occupational exposures in)
	Iron and steel founding (occupational exposure during)	
TESTIS		Firefighter (occupational exposure as a)
Prostate		Rubber manufacturing industry
SKIN(MELANOMA)		Firefighter (occupational exposure as a)
SKIN(NON-MELANOMA)		Petroleum refining (occupational exposures in)
	Painter (occupational exposure as a)	Petroleum refining (occupational exposures in)
Urinary bladder	Rubber manufacturing industry	Dry cleaning (occupational exposures in)
		Hairdresser or barber (occupational exposure as a)
		Printing processes (occupational exposures in)
		Textile manufacturing industry (work in)
Leukaemia: all combined	Rubber manufacturing industry	Petroleum refining (occupational exposures in)
Non-Hodgkin lymphoma: all combined		Firefighter (occupational exposure as a)
Lymphoma: all combined	Rubber manufacturing industry	
Mesothelium (pleura, peritoneum, and other)	Painter (occupational exposure as a)	

**[그림 II-5] 국제암연구소 업종별 위험도,
IARC (2022)에서 재인용**

Ⅲ. 연구 결과

.....

Ⅲ. 연구 결과

1. 연구결과의 제공 형태, 해석 방법, 활용 방법, 해석상의 주의점

1) 연구결과의 제공 형태

부록 2의 96쪽부터 176쪽에 걸쳐 제시된 <표 부-1>-<표 부-24>는 모든 업종(대분류별/중분류별)에 대하여 하위그룹의 모든 조합({남녀합계 집단, 남성 집단, 여성 집단} × {2000년도까지의 입사자, 2000년도 이후 입사자})별로 발병자수의 2018년 실제값, 2022년 추계값 및 예측 오류를 보고하였다. 추계값을 보고할 때는 표준집단의 기대발병자수에 비교하여 초과발병자 수, 표준집단 대비 발병비 또한 계산하였다. 마지막으로 혼련오류 및 검증오류 지표를 수록하였다. 예를 들어 표의 형태는 [그림 III-1]과 같다. 백혈병에 대한 남성집단의 중분류 업종별 추계결과를 2000년도까지의 남성 입사자 집단(118쪽의 <표 부-9>) 및 2000년도 이후 입사자 집단(123쪽의 <표 부-10>)에 대하여 일부만 캡처한 것이다. 각 연구자가 추계 결과를 추가적으로 가공하여 원하는 인사이트를 얻기 용이하도록, 동일한 표와 추가적인 정보들을 Microsoft Excel 스프레드시트 파일로 별첨하였다.

추가로, 위의 모든 조합에 대하여 백혈병, 폐암 발병자수의 2000년-2022년 추세 및 표준집단과의 비교 그림은 부록 2의 83쪽부터 94쪽에 걸쳐 [그림 부-9]-[그림 부-20]에 수록하였다.

본 절에서는 제공되는 결과표의 해석 방법과 예를 논하고, 백혈병, 폐암에 대하여 주목할 만한 관찰은 다음 2절과 3절에서 다룬다.

중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남성 (코드 32-51)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발생비	발병자수	초과분	발생비	훈련 [†]	검증 [‡]
32	가구 제조업	28	3	1.13	40	7	1.2	0.06	0.21
33	기타 제품 제조업	31	-3	0.92	46	0	1	0.09	0.09
34	산업용 기계 및 장비 수리업	6	-1	0.9	8	-1	0.9	0.04	0.06
35	전기, 가스, 증기 및 공기 조절 공급업	100	-6	0.94	145	2	1.01	0.06	0.16
36	수도업	5	1	1.25	8	2	1.41	0.13	0.18
37	하수, 폐수 및 분뇨 처리업	6	1	1.13	8	1	1.13	0.04	0.08
38	폐기물 수집, 운반, 처리 및 원료 재생업	44	19	1.77	56	22	1.65	0.02	0.02
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	418	36	1.1	557	36	1.07	0.01	0.01
42	전문직별 공사업	20	-5	0.81	24	-10	0.7	0.03	0.03
45	자동차 및 부품 판매업	29	-13	0.7	36	-21	0.64	0.07	0.1
46	도매 및 상품 중개업	321	9	1.03	430	2	1	0.01	0.01
47	소매업; 자동차 제외	90	-30	0.75	120	-45	0.73	0.02	0.07
49	육상 운송 및 파이프라인 운송업	419	27	1.07	544	16	1.03	0.01	0.02
50	수상 운송업	17	-2	0.9	24	-2	0.94	0.04	0.08
51	항공 운송업	17	-1	0.97	21	-2	0.9	0.03	0.03

중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남성 (코드 32-51)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발생비	발병자수	초과분	발생비	훈련 [†]	검증 [‡]
32	가구 제조업	22	8	1.59	34	14	1.65	0.08	0.12
33	기타 제품 제조업	18	-3	0.84	28	-5	0.85	0.07	0.08
34	산업용 기계 및 장비 수리업	8	-1	0.93	11	-3	0.79	0.17	0.77
35	전기, 가스, 증기 및 공기 조절 공급업	32	8	1.32	45	10	1.29	0.08	0.38
36	수도업	2	1	1.49	3	1	1.33	0.42	1
37	하수, 폐수 및 분뇨 처리업	11	7	2.45	19	12	2.79	0.07	0.18
38	폐기물 수집, 운반, 처리 및 원료 재생업	22	2	1.08	29	-2	0.94	0.04	0.1
39	환경 정화 및 복원업	0	0	0	0	-1	0		
41	종합 건설업	329	-14	0.96	530	11	1.02	0.03	0.07
42	전문직별 공사업	7	-3	0.67	9	-7	0.56	0.19	0.33
45	자동차 및 부품 판매업	18	1	1.05	28	2	1.08	0.05	0.12
46	도매 및 상품 중개업	282	1	1	447	17	1.04	0.02	0.15
47	소매업; 자동차 제외	118	-1	0.99	172	-8	0.96	0.08	0.11
49	육상 운송 및 파이프라인 운송업	217	60	1.38	376	145	1.63	0.02	0.06
50	수상 운송업	6	-1	0.83	8	-3	0.7	0.06	0.13
51	항공 운송업	3	0	0.89	5	-1	0.85	0.25	0.56

[그림 Ⅲ-1] 부록에 수록된 연구결과 제공 형태의 일부: (위) 114쪽의 <표 부-9> 및 (아래) 119쪽의 <표 부-10>

2) 연구결과표의 해석 방법

앞 장에서 각 수치의 정의와 간단한 설명을 살펴 보았으나, 뒤이은 절에서는 백혈병, 폐암에 대하여 주목할 만한 관찰을 기록하고, 여기서는 각 수치의 해석 방법과 예를 논한다.

먼저 2018년의 발병자수는 실제 관찰값이고, 2022년의 발병자수는 방법론이 적용된 추계값이다. 초과발병건수(“초과분”)은 해당업종 발병자수에서 표준집단의 기대도수를 뺀 값이고, 발병비는 해당업종 발병자수에서 표준집단의 기대도수를 나눈 값이다. 여기서 표준집단의 기대도수란, 해당 업종의 질환발병률이 전체집단에서의 발병률과 동일하다면 몇 명 정도가 발병하였을지를 나타내는 지표이다. 예를 들어 [그림 III-1]의 특정 업종에 대하여 아래와 같은 해석이 가능하다.

〈육상 운송업 및 파이프라인 운송업 (49)〉, 입사시기가 2001년 이후인 남성

- 2018년: 백혈병 발병자수가 217명, 초과분이 60, 발병비가 1.38
 - ✓ 해당업종에 한번이라도 종사한 적 있고 입사시기가 2001년 이후인 남성집단의 2018년까지의 백혈병 누적발병건수가 217명
 - ✓ 만약 해당 업종에서의 백혈병 발병률이 표준군(전체집단)에서의 발병률과 동일하다면, 그 기대도수는 157명이고, 기대도수에 비교하여 60(=217-157)명이 초과발생 또는 1.38배(=217/157) 초과발생
- 2022년 (추계): 백혈병 발병자수가 376명, 초과분이 145, 발병비가 1.63
 - ✓ 해당업종에 한번이라도 종사한 적 있고 입사시기가 2001년 이후인 남성집단의 2022년까지의 백혈병 누적발병건수 전망치가 376명
 - ✓ 만약 해당 업종에서의 백혈병 발병률이 표준군(전체집단)에서의 발병률과 동일하다면, 그 기대도수 전망치는 231명이고, 기대도수에 비교하여 145(=376-231)명이 초과발생 또는 1.63배(=376/231) 초과발생 전망

남녀합계 결과처럼 누적인년에 서로 다른 성별이 혼합된 경우, 성별로 간접표준화(전체집단의 남녀비율을 해당업종의 남녀비율로 재계산)한 후 기대도수를 계산된다.

다음으로 표 오른쪽의 훈련 오류와 검증 오류를 살펴 보자. 2장 4절에서 그 정의와 의미를 살펴보았다. 표의 특정 업종에 대하여는 아래와 같이 해석이 가능하다.

〈육상 운송업 및 파이프라인 운송업 (49)〉, 입사시기가 2001년 이후인 남성	
• 훈련 오류 0.02	✓ 2018년까지의 모든 자료를 사용하여 추세선을 그었을 때, 2015-2018년 관찰값과 추세선 사이에 2%의 오차가 있음. ✓ 표에 의하면 2018년 백혈병 발병자수 관찰값이 217명 정도이므로, 대략 $217 \times 2\% \approx 4$명 정도의 오차로 추세선이 적합되었음 ✓ 본 오차는 추세선 (일차식, 이차식, 스플라인의 혼합)으로 설명되지 못하는, 관찰값 자체의 변동에 의한 오차임
• 검증 오류 0.06	✓ 과거의 정보(-2014년)만 이용하여 이후 추세를 외삽(extrapolation)하였을 때, 2015-2018년 관찰값과 추세선 사이에 6%의 오차가 있음 ✓ 표에 의하면 2018년 백혈병 발병자수 관찰값이 217명 정도이므로, 대략 $217 \times 6\% \approx 13$명 정도의 오차로 추세선이 적합되었음. ✓ 본 오차는 추세선이 설명하지 못하는 관찰값 자체의 변동에 더하여 추세선이 예측하지 못한 미래의 추세변화가 부가된 오차임 ✓ 만약 2018년 자료로 2022년을 추계하였을 때도 변동과 추세변화가 비슷한 정도로 존재한다고 가정하면, 추계값에는 $376 \times 6\% \approx 23$명 정도의 오차가 존재할 것으로 예상할 수 있음

만일 검증오류와 훈련오류가 모두 준수한 수준이라면 해당 업종의 질환 누적 발병건수는 안정적인 추계가 가능할 것으로 판단할 수 있다. 직관적으로, 관찰 기간이 길고 누적발병건수가 많은 집단일수록 추세가 안정적이고 임의의 변동이 적기 때문에 검증오류와 훈련오류 모두 적을 것으로 예상할 수 있다. 반대로, 관찰된 발병건수가 적은 집단일수록 훈련오류와 검증오류 모두 커질 것으로 예상할 수 있다. 왜냐하면 이러한 집단은 임의의 변동이 크고, 추세가 안정적이지 않을뿐더러 오류의 분모(2015-2018 누적발병건수)도 작아지기 때문이다.

3) 연구결과표의 활용 방법

먼저 역학조사자의 실제적인 필요에 따라, 산출 수치들을 정렬하여 필요에 따라 위험한 업종을 선별해볼 수 있다. 이를테면 다음 지표 중 하나를 기준으로 내림차순으로 정렬하여 상위의 업종들을 확인하는 작업들을 수행해볼 수 있고, 각 조사자가 조사 맥락 속에서 더 중요하다고 생각하는 지표를 설정해볼 수도 있다.

- 표준집단대비 추가발병자수 또는 발병비 (2018년 또는 2022년 기준)
- 표준집단대비 추가발병자수의 증가량 (2018년 → 2022년)
- 표준집단대비 발병비의 증가량 (2018년 → 2022년)

뒤이은 절에서는 표준집단 대비 발병비(2022년)를 기준으로 선정 후, 가능한 모든 조합 ({소분류별, 중분류별} × {남여합계, 남성, 여성} × {2000년도까지 입사, 2000년도 이후 입사¹⁾})에 대하여 발병비 상위 업종을 정렬하고 있다.

정렬한 이후에는 각 연구자가 결과에 기반하여 다양한 가설을 고려해볼 수 있다. 작업환경노출 측면에서 2000년을 전후로 산업보건정책적인 중요한 변화와 함께 발암관련 유해인자의 노출수준에 큰 감소가 있었을 것이라고 추정된다. 이러한 점에서 2000년까지의 입사자 집단과 2000년 이후 입사자 집단에서 발병률의 유형변화가 큰 업종일수록 해당시기의 질환관련 발암인자에 대한 여러 정책적 결정과 작업환경변화의 영향을 받았다는 가설을 세워볼 수 있다. 특정 업종을 예를 들면 아래와 같은 해석을 고려해볼 수 있다.

1) II장의 모형 설계에서도 언급되었듯, 백혈병 및 폐암의 발생률을 고려하여 층화계층별 자료의 규모가 너무 작을 경우 발병률에 대한 값의 신뢰도가 떨어지기 때문에 사업장 기반 질환자료의 층화구간을 합쳐 재가공하였다. 입사시기의 경우 6개 구간으로 층화된 입사시기별 산출값을 2개의 구간(2000년도까지 입사, 2000년도 이후 입사)으로 합산하였다.

〈육상 운송업 및 파이프라인 운송업 (49)〉, 남성

- 2000년도까지의 입사자 집단:
2018년: 누적발생자수 419명, 표준집단대비 초과분 27명 및 발병비 1.07
2022년: 누적발생자수 544명, 표준집단대비 초과분 16명 및 발병비 1.03
※ 2022년 표준집단대비 발병비 기준 상위 10업종에 포함되지 않음
- 2000년도 이후 입사자 집단:
2018년: 누적발생자수 217명, 표준집단대비 초과분 60명 및 발병비 1.38
2022년: 누적발생자수 376명, 표준집단대비 초과분 145명 및 발병비 1.63
※ 2022년 표준집단대비 발병비 기준 상위 10업종에 포함됨
- 표준집단대비 발병비 기준의 해석: 2000년도까지의 입사자 집단에 비하여 2000년도 이후 입사자 집단에서 발병비 및 초과분이 크고, 2000년도 이후 입사자 집단의 다른 중분류에 비하여도 2022년 발병비가 높은 편
✓ 타 요인 (나이, 또는 흡연여부, 가족력 등의 환경적 요인)의 영향을 추후 따로 평가할 필요는 있으나, 이 요인들을 포함하여 2000년 전후의 정책적 변화들이 2000년도 이후 입사자 집단의 높은 초과발병과 연관되어 있을 가능성이 있음.
- 표준집단대비 발병비 증가분 기준의 해석: 2000년도 이후 입사자 집단에서 2018년 발병비(1.38)에 비하여 2022년 발병비(1.63)가 증가하고 있음
✓ 해당 업종과 관련하여, 미확인 요인으로 인한 위험이 해소되지 않고 있는 것으로 추측됨

4) 연구결과표 해석상의 주의점

첫째, 표준집단대비 추가발병자수 및 발병비의 과소(과대) 추정 가능성이 존재한다. 현재의 표준집단 기대도수는 성별에 대하여만 층화 또는 표준화가 되어 있기 때문이다. 암의 다른 주요한 요인 (나이, 흡연여부, 가족력 등)이 더 반영된다면 업종마다 표준집단 기대도수가 달라질 수도 있다. 예를 들어 특정 업종의 코호트에 고령자, 흡연자 등 암 위험군의 비율이 높다면, 이들을 추가적으로 보정하였을 경우 표준집단 기대도수가 높아져 결론적으로 표준집단대비 초과분과 발병비가 현재 연구결과보다 낮아질 가능성이 있다. 반대로,

특정 업종에서 표준집단대비 초과분과 발병비가 작게 나왔다면 (음수의 초과분, 1보다 작은 발병비), 낮은 위험도가 업종의 특성일 가능성도 물론 있으나 해당 업종의 인구구성이 암 저위험군으로 구성되어 있을 가능성 또한 배제할 수 없다. 본 설명을 특정 업종을 예를 들어 설명하면 아래와 같다.

〈육상 운송업 및 파이프라인 운송업 (49)〉, 2000년도 이후 입사자 집단, 남성

- 2018년: 누적발생자수 217명, 표준집단대비 초과분 60명 및 발병비 1.38
2022년: 누적발생자수 376명, 표준집단대비 초과분 145명 및 발병비 1.63
※ 2022년 표준집단대비 발병비 기준 상위 10업종에 포함됨
- 만약 코호트에서 고령자의 비율과 흡연자의 비율표준집단에서의 고령자, 흡연자 비율보다 높다면, 표준집단의 기대도수 (2018년 157명; 2022년 231명)가 과소추정되어있을 가능성이 있고, 따라서 표준집단대비 초과분과 발병비가 과대추정되어있을 가능성이 있음

해석이 어려운 업종들에 대한 추가적인 연구를 위하여는 먼저 제공되는 초과분과 발병비의 추정 성능을 고도화할 필요가 있다. 예를 들어 위에서 살펴본듯 각 코호트의 연도별 연령대 분포, 흡연자의 분포, 가족력 등의 정보를 반영하기 위한 추가적인 연구와 자료원이 필요하다.

둘째, 추계에서 훈련, 검증 오차는 2015-2018년 관찰값에 대해 상대적으로 계산된 비율이기 때문에, 2015-2018년 누적발병자수가 희소한 업종의 경우에는 절대적 규모에서는 작은 오차가 비율에서는 증폭되어 보이는 현상이 있을 수 있다. 가령 2018년 실제값이 2명이고 적합값이 3.5명인 경우, 오류 계산값으로는 $(3.5-2)/2 = 75\%$ 가 발생할 수 있다. 따라서 오차율을 확인할 때는 2018년도의 발병규모를 같이 확인할 필요가 있다.

2. 백혈병 발병률 추계 결과

2022년 추계값을 이용하여, 2022년의 참조집단(전체업종 합계) 발병률과

비교하여 발병비율이 가장 높은 대분류 업종 5종을 위험집단으로 삼아, 남녀 합계 집단, 남성 집단, 여성 집단에 대하여 아래 <표 III-1>부터 <표 III-3>에 정리하였다. 단순히 발병자수의 추계뿐만 아니라 참조집단 대비 발병비율(해당업종 발병자수를 참조집단 발병자수로 나눈 값) 및 초과발병자 수(해당업종 발병자수에서 참조집단 발병자수를 뺀 값)을 함께 수록하여 비교가 용이하도록 하였다. 추가적으로, 남녀합계 집단에서 참조집단의 발병자수를 계산할 때는 업종별 2022년 추계인년 기준 성비를 이용한 표준화된 발병자수를 계산하였다. 마찬가지로, 참조집단과 비교하여 발병비율이 가장 높은 중분류 업종 10종을 남녀합계 집단, 남성 집단, 여성 집단에 대하여 아래 <표 III-4>부터 <표 III-6>에 정리하였다. 단, 추계값의 통계적인 안정성을 고려하여, 2018년 기준으로 발병자수가 2명 이하인 경우는 생략하였다.

본 연구에서 위험집단으로 검출된 업종 중에는 기존에 백혈병 발병 연관 업종으로 알려진 업종 (고무/석유정제 관련 제조업)과 연관있어 보이는 업종들이 검출되었다. 먼저 대분류를 살펴보면 2000년도까지/2000년도이후 입사자 집단에서 <광업(05~08)>과 <수도, 하수 및 폐기물 처리, 원료재생업(36~39)>이 공통적으로 검출되어 왔다. 한편, 중분류 결과를 살펴보면 <석탄, 석유 및 천연가스 광업 (5)>, <비금속광물 광업: 연료용 제외 (7)>, <담배 제조업(12)>, <목재 및 나무제품 제조업; 가구 제외 (16)>, <하수, 폐수 및 분뇨 처리업 (37)>, <폐기물 수집, 운반, 처리 및 원료 재생업 (38)>, <환경 정화 및 복원업 (9)> 등이 검출됨이 확인된다. 본 연구 결과는 업종의 연령대 구성은 고려되지 않은 것이기 때문에, 본 결과들을 활용하는 과정에서는 업종의 연령대 구성의 영향도 함께 살펴보는 것이 요구된다는 점에 유의할 필요가 있다.

한편, 정량적 타당성 검증을 위한 훈련 및 검증오류는 (대분류, 중분류) × (남녀합계 집단, 남성 집단, 여성 집단) × (2000년도까지의 입사자, 2000년도 이후 입사자) 모든 조합에 대하여 <그림 III-6>에 상자그림으로 요약하였다. 상자그림에서 상자의 분포를 구성하는 점들은 각 업종을 의미한다.

〈그림 III-6〉에 의하면 전반적으로 2000년 이후 입사 집단에서 검증 오류가 2000년도 이전 입사 집단보다 대체로 큰 편이다 (RMSE ratio의 중앙값이 약 0.2). 이는 2000년도 이후 입사자 집단에 대하여는 2000-2014 자료만을 이용하여는 2015-18년의 발병자수를 추계하는 데에 어려움이 있음을 뜻한다. 이 관찰은 자연스러운데, 2000년도 이후 입사자의 경우 관찰된 발병건수가 충분히 누적되지 않았고, 때문에 II장에서 언급한 것처럼 집단은 임의의 변동이 크고, 추세가 안정적이지 않을뿐더러 RMSE ratio의 분모도 작아졌을 수 있다. 중분류의 경우 훈련/검증오차가 모두 큰 업종이 소수 존재하는데, 이들도 관찰된 발병건수가 충분히 누적되지 않은 경우였다.

본 장에서 보고된 결과들의 세부사항으로써, 부록에서는 모든 업종(대분류별/중분류별)에 대한 모든 조합에 대하여 (발병률, 초과발병자 수, 발병비)의 2022년 추계값 및, 훈련오류, 검증오류 지표를 수록하였고, 추계값의 추이를 시각적으로 요약하는 추계값 그래프를 수록하였다.

〈표 Ⅲ-1〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 남/여
합계

(단위 : 명)

	코 드	분류명	발병자수 (성-표준화 집단 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	B	광업(05~08)	33 (12, 1.58)	51 (23, 1.83)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	52 (19, 1.56)	67 (22, 1.48)
	L	부동산업(68)	292 (95, 1.48)	389 (123, 1.46)
	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	409 (125, 1.44)	552 (170, 1.44)
	R	예술, 스포츠 및 여가관련 서비스업(90~91)	59 (8, 1.17)	79 (10, 1.15)
2000 년도 이후 입사자	L	부동산업(68)	219 (75, 1.52)	368 (151, 1.69)
	B	광업(05~08)	8 (2, 1.43)	13 (5, 1.67)
	O	공공 행정, 국방 및 사회보장 행정(84)	101 (30, 1.41)	158 (50, 1.46)
	H	운수 및 창고업(49~52)	363 (84, 1.3)	586 (168, 1.4)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	37 (8, 1.27)	55 (11, 1.24)

* 표준군 = 전체업종 합계

〈표 III-2〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 남성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	B	광업(05~08)	31 (11, 1.57)	48 (21, 1.8)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	50 (20, 1.64)	64 (22, 1.54)
	L	부동산업(68)	254 (87, 1.52)	339 (113, 1.5)
	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	258 (61, 1.31)	360 (93, 1.35)
	R	예술, 스포츠 및 여가관련 서비스업(90~91)	42 (5, 1.13)	58 (7, 1.14)
2000 년도 이후 입사자	B	광업(05~08)	8 (3, 1.57)	13 (6, 1.83)
	O	공공 행정, 국방 및 사회보장 행정(84)	75 (25, 1.5)	121 (49, 1.69)
	L	부동산업(68)	157 (49, 1.45)	268 (105, 1.64)
	H	운수 및 창고업(49~52)	319 (78, 1.32)	522 (161, 1.45)
	D	전기, 가스, 증기 및 공기 조절 공급업(35)	34 (8, 1.32)	48 (11, 1.3)

* 표준군 = 전체업종 합계

〈표 Ⅲ-3〉 2022년 추계값 기준 백혈병 표준화발병비 상위 5개 업종(대분류), 여성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	B	광업(05~08)	2 (1, 1.63)	4 (2, 2.24) **
	A	농업, 임업 및 어업(01~03)	7 (4, 2.2)	8 (4, 1.96)
	N	사업시설 관리, 사업 자원 및 임대 서비스업(74~76)	151 (64, 1.73)	192 (77, 1.66)
	L	부동산업(68)	38 (8, 1.28)	50 (10, 1.26)
	R	예술, 스포츠 및 여가관련 서비스업(90~91)	17 (4, 1.27)	21 (3, 1.16)
2000 년도 이후 입사자	L	부동산업(68)	62 (27, 1.75)	101 (46, 1.84)
	S	협회 및 단체, 수리 및 기타 개인 서비스업(94~96)	46 (10, 1.27)	73 (17, 1.32)
	F	건설업(41~42)	59 (6, 1.11)	101 (20, 1.24)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	4 (0, 1.02)	7 (1, 1.14)
	H	운수 및 창고업(49~52)	44 (6, 1.17)	65 (7, 1.12)

* 표준군 = 전체업종 합계

** 2.5%분위 추계값 3 (1, 1.87), 97.5%분위 추계값 (3, 2.85) [신뢰구간 기준]

〈표 III-4〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류),
남/여 합계

(단위 : 명)

	코드	분류명	발병자수 (성-표준화 집단 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	12	담배 제조업	17 (9, 2.04)	22 (11, 2)
	76	임대업; 부동산 제외	13 (6, 1.77)	20 (10, 1.95)
	05	석탄, 원유 및 천연가스 광업	19 (7, 1.64)	29 (14, 1.89)
	07	비금속광물 광업; 연료용 제외	13 (4, 1.5)	20 (8, 1.68)
	74	사업시설 관리 및 조경 서비스업	168 (72, 1.75)	217 (88, 1.68)
	90	창작, 예술 및 여가관련 서비스업	19 (8, 1.77)	24 (10, 1.67)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	46 (19, 1.69)	59 (22, 1.6)
	16	목재 및 나무제품 제조업; 가구 제외	35 (7, 1.25)	56 (18, 1.48)
	68	부동산업	292 (95, 1.48)	389 (123, 1.46)
	15	가죽, 가방 및 신발 제조업	70 (20, 1.39)	92 (24, 1.36)
2000 년도 이후 입사자	37	하수, 폐수 및 분뇨 처리업	11 (6, 2.17)	19 (12, 2.5)
	68	부동산업	219 (75, 1.52)	368 (151, 1.69)
	49	육상 운송 및 파이프라인 운송업	234 (68, 1.41)	400 (156, 1.64)
	74	사업시설 관리 및 조경 서비스업	166 (51, 1.45)	271 (94, 1.53)
	16	목재 및 나무제품 제조업; 가구 제외	26 (9, 1.52)	39 (14, 1.52)
	32	가구 제조업	24 (7, 1.43)	37 (12, 1.48)
	07	비금속광물 광업; 연료용 제외	6 (2, 1.36)	9 (3, 1.47)
	84	공공 행정, 국방 및 사회보장 행정	101 (30, 1.41)	158 (50, 1.46)
	60	방송업	14 (4, 1.42)	19 (5, 1.32)
	23	비금속 광물제품 제조업	66 (15, 1.3)	97 (20, 1.27)

* 표준군 = 전체업종 합계

* 발병자수 2명 이하(2018년 기준)의 중분류는 제외하였음

〈표 III-5〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류),
남성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	76	임대업; 부동산 제외	13 (7, 2.01)	20 (11, 2.22)
	90	창작, 예술 및 여가관련 서비스업	16 (8, 2.11)	20 (9, 1.91)
	05	석탄, 원유 및 천연가스 광업	18 (7, 1.62)	27 (12, 1.84)
	12	담배 제조업	14 (7, 1.88)	18 (8, 1.79)
	07	비금속광물 광업; 연료용 제외	12 (4, 1.5)	18 (7, 1.69)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	44 (19, 1.77)	56 (22, 1.65)
	16	목재 및 나무제품 제조업; 가구 제외	31 (8, 1.32)	51 (19, 1.59)
	68	부동산업	254 (87, 1.52)	339 (113, 1.5)
	96	기타 개인 서비스업	17 (4, 1.32)	25 (7, 1.42)
	36	수도업	5 (1, 1.25)	8 (2, 1.41)
2000 년도 이후 입사자	37	하수, 폐수 및 분뇨 처리업	11 (7, 2.45)	19 (12, 2.79)
	84	공공 행정, 국방 및 사회보장 행정	75 (25, 1.5)	121 (49, 1.69)
	32	가구 제조업	22 (8, 1.59)	34 (14, 1.65)
	68	부동산업	157 (49, 1.45)	268 (105, 1.64)
	49	육상 운송 및 파이프라인 운송업	217 (60, 1.38)	376 (145, 1.63)
	07	비금속광물 광업; 연료용 제외	6 (2, 1.49)	9 (4, 1.61)
	16	목재 및 나무제품 제조업; 가구 제외	23 (9, 1.63)	34 (12, 1.58)
	96	기타 개인 서비스업	19 (5, 1.35)	32 (10, 1.48)
	74	사업시설 관리 및 조경 서비스업	99 (26, 1.35)	163 (50, 1.45)
	60	방송업	11 (4, 1.47)	15 (5, 1.41)

* 표준군 = 전체업종 합계

* 발병자수 2명 이하(2018년 기준)의 중분류는 제외하였음

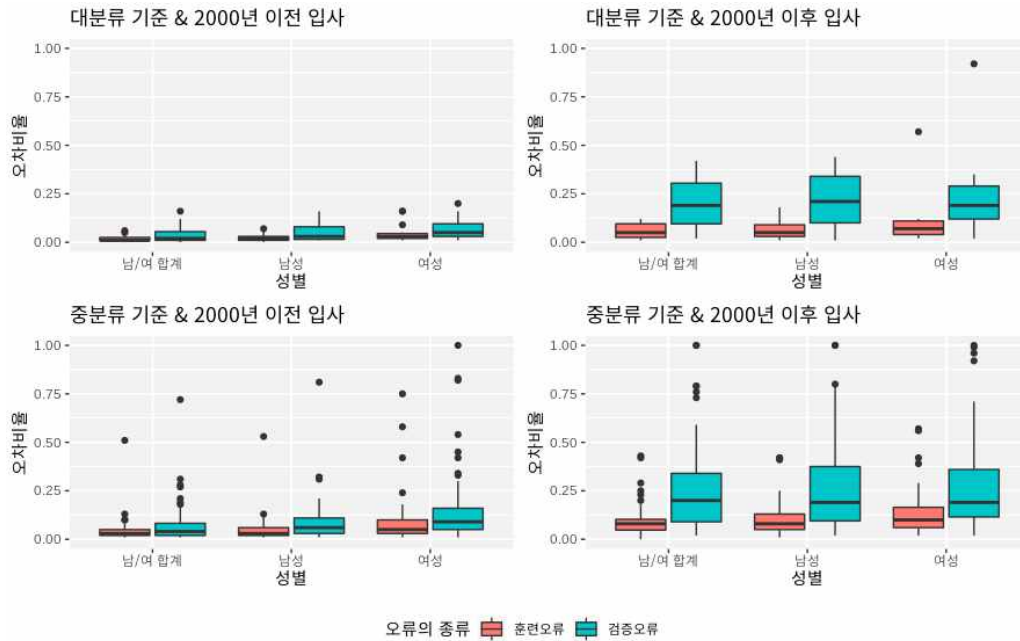
〈표 III-6〉 2022년 추계값 기준 백혈병 표준화발병비 상위 10개 업종(중분류),
여성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	12	담배 제조업	17 (9, 2.04)	22 (11, 2)
	76	임대업; 부동산 제외	13 (6, 1.77)	20 (10, 1.95)
	05	석탄, 원유 및 천연가스 광업	19 (7, 1.64)	29 (14, 1.89)
	07	비금속광물 광업; 연료용 제외	13 (4, 1.5)	20 (8, 1.68)
	74	사업시설 관리 및 조경 서비스업	168 (72, 1.75)	217 (88, 1.68)
	90	창작, 예술 및 여가관련 서비스업	19 (8, 1.77)	24 (10, 1.67)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	46 (19, 1.69)	59 (22, 1.6)
	16	목재 및 나무제품 제조업; 가구 제외	35 (7, 1.25)	56 (18, 1.48)
	68	부동산업	292 (95, 1.48)	389 (123, 1.46)
	15	가죽, 가방 및 신발 제조업	70 (20, 1.39)	92 (24, 1.36)
2000 년도 이후 입사자	37	하수, 폐수 및 분뇨 처리업	11 (6, 2.17)	19 (12, 2.5)
	68	부동산업	219 (75, 1.52)	368 (151, 1.69)
	49	육상 운송 및 파이프라인 운송업	234 (68, 1.41)	400 (156, 1.64)
	74	사업시설 관리 및 조경 서비스업	166 (51, 1.45)	271 (94, 1.53)
	16	목재 및 나무제품 제조업; 가구 제외	26 (9, 1.52)	39 (14, 1.52)
	32	가구 제조업	24 (7, 1.43)	37 (12, 1.48)
	07	비금속광물 광업; 연료용 제외	6 (2, 1.36)	9 (3, 1.47)
	84	공공 행정, 국방 및 사회보장 행정	101 (30, 1.41)	158 (50, 1.46)
	60	방송업	14 (4, 1.42)	19 (5, 1.32)
	23	비금속 광물제품 제조업	66 (15, 1.3)	97 (20, 1.27)

* 표준군 = 전체업종 합계

* 발병자수 2명 이하(2018년 기준)의 중분류는 제외하였음



**[그림 III-6] 백혈병 발병자수에 대한 업종별 추계 오차비율
(추계대상이 '15-'18년도일 때),**

* 훈련오류: 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

* 검증오류: 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

3. 폐암 발병률 추계 결과

백혈병과 마찬가지로, 2022년 추계값을 이용하여, 2022년의 참조집단(전 체업종 합계) 발병률과 비교하여 발병비율이 가장 높은 대분류 업종 5종 및 중분류 10종을 남녀합계 집단, 남성 집단, 여성 집단에 대하여 <표 III-7>부터 <표 III-12>에 정리하고, (성-표준화된) 참조집단의 누적발병자수와 비교하였을 때 초과발병자수(차이) 및 발병비(비율)을 보고하였다. 여기서도 추계값의 통계적인 안정성을 고려하여, 2018년 누적발병자수가 10명 이하인 업종은 본 장의 보고에서는 생략하되 부록에서는 보고하였다. 백혈병보다 5배 많은 10명을 생략 기준으로 삼은 이유는 폐암의 발생빈도가 백혈병보다 5배 정도 많은 것으로 알려져 있기 때문이었다 (2019년 국가암등록통계 참고자료에 의하면 2019년 현재 표준화발병비: 백혈병 5.5명/10만명, 폐암 28.2명/10만명).

먼저 위험집단으로 검출된 업종 중, 기존에 폐암 발병과 관련된 연관 업종으로 알려진 업종 (고무, 페인트, 철강, 인쇄 관련 제조업)과 연관있어 보이는 업종들을 살펴보자. 대분류를 살펴보면 백혈병에서의 결과와 마찬가지로 2000년도까지/2000년도이후 입사자 집단에서 <광업(05~08)>과 <수도, 하수 및 폐기물 처리, 원료재생업(36~39)>이 공통적으로 검출되어 왔다. 한편, 중분류 결과를 살펴보면, 남녀합계 집단 및 남성 집단에서는 <석탄, 석유 및 천연가스 광업 (05)>, <금속 광업(06)>, <비금속광물 광업: 연료용 제외 (07)>, <담배 제조업(12)>, <목재 및 나무제품 제조업; 가구 제외 (16)>, <하수, 폐수 및 분뇨 처리업 (37)>, <폐기물 수집, 운반, 처리 및 원료 재생업 (38)>, <환경 정화 및 복원업 (39)> 등이 검출됨이 확인된다. 여성 집단을 살펴 보았을 때는 추가적으로 <섬유제품 제조업: 의복 제외 (13)>, <가죽, 가방 및 신발 제조업 (15)>, <비금속 광물제품 제조업 (23)>, <가구 제조업(32)> 및 <기타 제조업 (33)>이 검출되고 있다. 백혈병에서와 마찬가지로, 본 연구 결과는 업종

의 연령대 구성은 고려되지 않았으므로, 업종의 연령대 구성의 영향은 추후 평가가 필요하다.

한편, 정량적 타당성 검증을 위한 훈련 및 검증오류 또한 백혈병 결과와 마찬가지로 (대분류, 중분류) × (남녀합계 집단, 남성 집단, 여성 집단) × (2000년도까지의 입사자, 2000년도 이후 입사자) 모든 조합에 대하여 <그림 III-7>에 상자그림으로 요약하였다. 각 상자의 분포를 구성하는 단위자료는 각 업종에서의 오류이다.

<그림 III-7>에서도 백혈병에서와 비슷한 관찰을 얻을 수 있는데, 2000년 이후 입사 집단의 검증 오류가 2000년도 이전 입사 집단보다 대체로 큰 편이다 (RMSE ratio의 중앙값이 약 0.25). 또한 중분류의 경우 훈련/검증오차가 모두 큰 업종이 소수 존재하나 백혈병에서보다는 그 수가 적었다. 역시 2000년도 이후 입사자의 경우 관찰된 발병건수가 충분히 누적되지 않았고, 때문에 II장에서 언급한 것처럼 집단은 임의의 변동이 크고, 추세가 안정적이지 않을 뿐더러 RMSE ratio의 분모도 작아졌을 수 있다.

세부적인 보고 결과로, 부록에서는 모든 업종(대분류별/중분류별)에 대한 모든 조합에 대하여 (발병률, 초과발병자 수, 발병비)의 2022년 추계값 및, 훈련오류, 검증오류 지표를 수록하였고, 추계값의 추이를 시각적으로 요약하는 추계값 그래프를 수록하고 있다.

〈표 III-7〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 남/여
합계

(단위 : 명)

	코드	분류명	발병자수 (성-표준화 집단 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	L	부동산업(68)	4087 (2771, 3.11)	5537 (3716, 3.04)
	B	광업(05~08)	407 (260, 2.77)	550 (349, 2.73)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	561 (330, 2.43)	763 (443, 2.38)
	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	3656 (1926, 2.11)	5084 (2675, 2.11)
	O	공공 행정, 국방 및 사회보장 행정(84)	889 (430, 1.94)	1248 (605, 1.94)
2000 년도 이후 입사자	L	부동산업(68)	2173 (1489, 3.18)	3539 (2539, 3.54)
	B	광업(05~08)	63 (35, 2.21)	92 (52, 2.31)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	328 (183, 2.26)	466 (251, 2.17)
	A	농업, 임업 및 어업(01~03)	142 (74, 2.08)	212 (110, 2.09)
	H	운수 및 창고업(49~52)	2745 (1354, 1.97)	3958 (1931, 1.95)

* 표준군 = 전체업종 합계

〈표 Ⅲ-8〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 남성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	L	부동산업(68)	3913 (2701, 3.23)	5290 (3617, 3.16)
	B	광업(05~08)	402 (259, 2.82)	542 (346, 2.77)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	550 (330, 2.5)	749 (444, 2.45)
	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	3010 (1586, 2.11)	4144 (2165, 2.09)
	O	공공 행정, 국방 및 사회보장 행정(84)	844 (432, 2.05)	1185 (610, 2.06)
2000 년도 이후 입사자	L	부동산업(68)	1939 (1370, 3.41)	3215 (2379, 3.85)
	B	광업(05~08)	61 (34, 2.27)	89 (51, 2.37)
	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	303 (171, 2.29)	428 (231, 2.18)
	A	농업, 임업 및 어업(01~03)	126 (68, 2.16)	190 (102, 2.17)
	H	운수 및 창고업(49~52)	2627 (1358, 2.07)	3782 (1931, 2.04)

* 표준군 = 전체업종 합계

〈표 III-9〉 2022년 추계값 기준 폐암 표준화발병비 상위 5개 업종(대분류), 여성

(단위 : 명)

	코 드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	646 (341, 2.12)	940 (510, 2.18)
	I	숙박 및 음식점업(55~56)	378 (172, 1.83)	550 (259, 1.89)
	L	부동산업(68)	174 (70, 1.67)	247 (99, 1.68)
	A	농업, 임업 및 어업(01~03)	16 (5, 1.43)	22 (6, 1.37)
	B	광업(05~08)	5 (1, 1.16)	8 (2, 1.35)
2000 년도 이후 입사자	E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	25 (12, 1.96)	38 (20, 2.11)
	L	부동산업(68)	234 (119, 2.03)	324 (160, 1.97)
	A	농업, 임업 및 어업(01~03)	16 (6, 1.61)	22 (8, 1.56)
	I	숙박 및 음식점업(55~56)	357 (117, 1.49)	506 (182, 1.56)
	N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	997 (367, 1.58)	1388 (485, 1.54)

* 표준군 = 전체업종 합계

〈표 III-10〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 남/여
합계

(단위 : 명)

	코 드	분류명	발병자수 (성-표준화 집단 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지 의 입사 자	06	금속 광업	24 (20, 5.47)	31 (25, 5.19)
	68	부동산업	4087 (2771, 3.11)	5537 (3716, 3.04)
	05	석탄, 원유 및 천연가스 광업	250 (168, 3.04)	332 (219, 2.96)
	74	사업시설 관리 및 조경 서비스업	1660 (1085, 2.89)	2276 (1477, 2.85)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	466 (278, 2.48)	637 (376, 2.44)
	07	비금속광물 광업; 연료용 제외	133 (73, 2.21)	188 (104, 2.25)
	12	담배 제조업	128 (71, 2.25)	173 (95, 2.21)
	37	하수, 폐수 및 분뇨 처리업	86 (46, 2.13)	116 (59, 2.05)
	84	공공 행정, 국방 및 사회보장 행정	889 (430, 1.94)	1248 (605, 1.94)
	75	사업 지원 서비스업	1940 (835, 1.76)	2729 (1188, 1.77)
2000 년도 이후 입사 자	03	어업	58 (39, 3.12)	95 (68, 3.54)
	68	부동산업	2173 (1489, 3.18)	3539 (2539, 3.54)
	74	사업시설 관리 및 조경 서비스업	1536 (1016, 2.96)	2299 (1526, 2.97)
	49	육상 운송 및 파이프라인 운송업	2138 (1284, 2.5)	3043 (1818, 2.49)
	07	비금속광물 광업; 연료용 제외	51 (29, 2.28)	75 (43, 2.37)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	272 (155, 2.32)	387 (214, 2.23)
	16	목재 및 나무제품 제조업; 가구 제외	172 (88, 2.05)	249 (125, 2.02)
	37	하수, 폐수 및 분뇨 처리업	54 (28, 2.12)	76 (38, 2.01)
	02	임업	28 (16, 2.33)	37 (18, 2)
	50	수상 운송업	73 (32, 1.76)	114 (53, 1.86)

* 표준군 = 전체업종 합계, 발병자수 10명 이하(2018년 기준)의 중분류는 제외하였음

〈표 Ⅲ-11〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 남성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	06	금속 광업	23 (19, 5.54)	30 (24, 5.26)
	68	부동산업	3913 (2701, 3.23)	5290 (3617, 3.16)
	05	석탄, 원유 및 천연가스 광업	248 (168, 3.08)	328 (218, 2.99)
	74	사업시설 관리 및 조경 서비스업	1255 (794, 2.72)	1688 (1048, 2.64)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	457 (277, 2.54)	625 (376, 2.51)
	07	비금속광물 광업; 연료용 제외	131 (73, 2.26)	184 (104, 2.3)
	12	담배 제조업	121 (67, 2.25)	163 (90, 2.21)
	37	하수, 폐수 및 분뇨 처리업	84 (46, 2.19)	113 (60, 2.12)
	84	공공 행정, 국방 및 사회보장 행정	844 (432, 2.05)	1185 (610, 2.06)
	75	사업 지원 서비스업	1699 (783, 1.85)	2377 (1102, 1.87)
2000 년도 이후 입사자	68	부동산업	1939 (1370, 3.41)	3215 (2379, 3.85)
	03	어업	57 (40, 3.31)	93 (68, 3.75)
	74	사업시설 관리 및 조경 서비스업	1081 (697, 2.81)	1656 (1077, 2.86)
	49	육상 운송 및 파이프라인 운송업	2097 (1272, 2.54)	2975 (1791, 2.51)
	07	비금속광물 광업; 연료용 제외	50 (29, 2.37)	73 (43, 2.45)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	250 (143, 2.34)	355 (196, 2.24)
	02	임업	28 (17, 2.61)	37 (20, 2.22)
	16	목재 및 나무제품 제조업; 가구 제외	157 (83, 2.11)	229 (119, 2.09)
	37	하수, 폐수 및 분뇨 처리업	51 (27, 2.16)	71 (36, 2.01)
	50	수상 운송업	73 (35, 1.91)	114 (57, 2.01)

* 표준군 = 전체업종 합계

* 발병자수 10명 이하(2018년 기준)의 중분류는 제외하였음

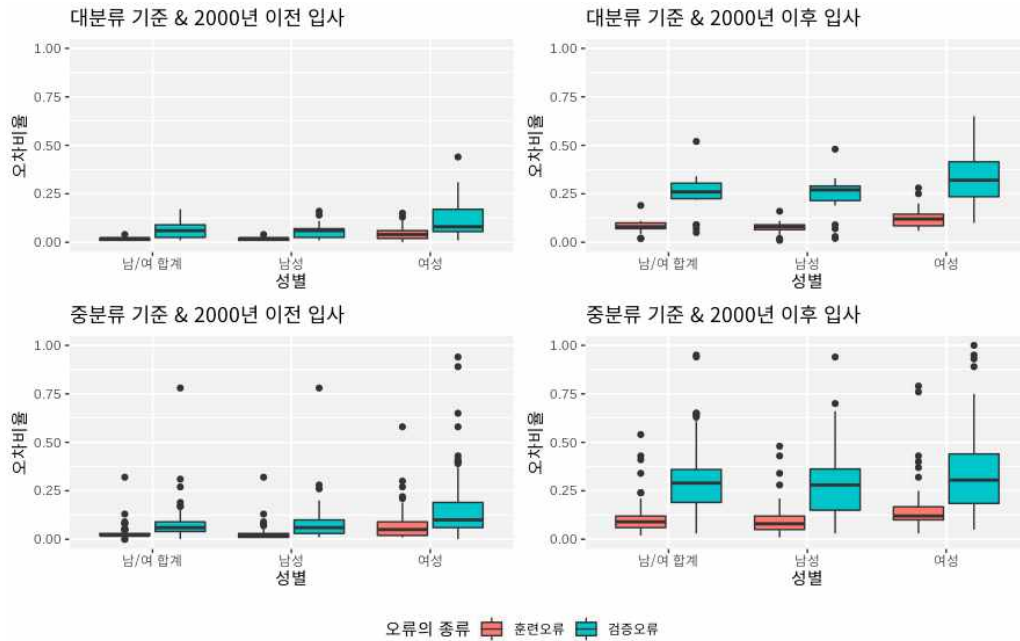
〈표 III-12〉 2022년 추계값 기준 폐암 표준화발병비 상위 10개 업종(중분류), 여성

(단위 : 명)

	코드	분류명	발병자수 (표준군 대비 추가발병자수, 발병비)	
			2018년	2022년 추계
2000 년도 까지의 입사자	74	사업시설 관리 및 조경 서비스업	405 (291, 3.56)	588 (428, 3.69)
	32	가구 제조업	46 (24, 2.13)	63 (33, 2.09)
	56	음식점 및 주점업	264 (130, 1.97)	382 (193, 2.01)
	15	가죽, 가방 및 신발 제조업	123 (63, 2.05)	159 (75, 1.9)
	23	비금속 광물제품 제조업	94 (37, 1.65)	137 (58, 1.73)
	68	부동산업	174 (70, 1.67)	247 (99, 1.68)
	55	숙박업	114 (42, 1.57)	168 (67, 1.66)
	33	기타 제품 제조업	78 (30, 1.63)	109 (42, 1.62)
	25	금속 가공제품 제조업; 기계 및 가구 제외	139 (55, 1.66)	189 (72, 1.61)
	24	1차 금속 제조업	53 (21, 1.66)	69 (24, 1.54)
2000 년도 이후 입사자	74	사업시설 관리 및 조경 서비스업	455 (320, 3.36)	643 (449, 3.31)
	38	폐기물 수집, 운반, 처리 및 원료 재생업	22 (11, 2.09)	32 (18, 2.18)
	32	가구 제조업	19 (9, 1.97)	27 (14, 2.05)
	68	부동산업	234 (119, 2.03)	324 (160, 1.97)
	01	농업	15 (8, 2.07)	20 (10, 1.95)
	13	섬유제품 제조업; 의복 제외	131 (57, 1.76)	191 (89, 1.87)
	49	육상 운송 및 파이프라인 운송업	41 (12, 1.43)	68 (28, 1.7)
	33	기타 제품 제조업	46 (18, 1.66)	66 (27, 1.69)
	56	음식점 및 주점업	322 (116, 1.56)	452 (175, 1.63)
	10	식품 제조업	225 (87, 1.63)	310 (111, 1.56)

* 표준군 = 전체업종 합계

* 발병자수 10명 이하(2018년 기준)의 중분류는 제외하였음



[그림 Ⅲ-7] 폐암 발병자수에 대한 업종별 추계 오차비율
(추계대상이 '15-'18년도일 때),

* 훈련오류: 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

* 검증오류: 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

IV. 결론 및 제언

.....

IV. 결론 및 제언

1. 연구 결과의 활용 방안

본 연구에 사용된 자료는 개인정보의 노출을 최소화하기 위하여 기업별로 집계된 통계량으로 구성되어 있었다. 이를 이용하여 입사시 사업장의 업종별 코호트를 성별(남/여), 입사시기(2000년도까지/2000년도 이후)로 층화하여 추계값을 살펴보았다.

본 연구의 의의는, 일정 수의 발생건수를 보인 업종에 대하여는 통계적인 시계열추계 기법 기반의 단기추계(3-5년)값은 안정적으로 얻을 수 있음을 확인한 것이다. 이를 활용하는 방법으로는, 이를테면 개인단위 빅데이터가 원시 출처로부터 수집되어 기관 내 데이터베이스에 접근가능한 형태가 되는데 평균 3-4년의 시간이 걸림을 감안하면 (본 보고서가 작성된 2022년에는 2018년까지의 자료만 이용가능함), 본 보고서에서 제안한 방법은 적재된 빅데이터 이용하여 현 시점에서의 발병률 추정값을 계산하는 데 유용할 수 있다. 다만 암발병의 가장 큰 위험요인인 나이가 반영되지 않았는데, 이는 하기 서술된 데이터셋의 개선 방안을 고려한다면 추후 비슷한 연구 기회에서 나이군을 반영한 모형 설계가 가능할 것으로 예상된다.

다만 통계적인 추계모형이 장기적인 추계에 사용가능한지에 대하여는 충분한 검증과 검토가 필요하다. 문헌고찰에서 살펴보았듯, 장기추계는 대규모 집단 (국가수준)에서 총인구수와 발병률에 대한 합리적인 가정 하에 암발병자 추계가 다루어진 바 있으나, 그러나 소규모 집단에 대한 장기추계는 문헌에 존재하지 않는다. 대개 장기적인 추계를 위하여는 관찰치가 나온 동역학적인 구조의 방정식이 필요하다. 대표적으로 장기 기상예보를 위한 기후모형, 장기 인구추계를 위한 유입-유출 모형 모두 현상에 대한 합리적인 가정을 미분방

정식으로 서술하고, 이를 반복적으로 품으로써 추계값을 얻는다. 업종별 감시에 이와 같은 동역학적 모형이 제안가능하고 실증적으로 유용한지를 검토하는 것은 완전히 새로운 연구를 필요로 할 것이다.

2. 연구 발전 방안 및 제언

연구의 발전방향을 논하기 위하여 먼저 현 연구의 제한점을 짚어보자. 본 연구는 사업장별 집계 데이터를 이용하여 대분류/중분류 업종까지에 단기추계값을 계산하고 있다. 직업성 암 연구의 목표를 달성하기 위하여는 다른 주요한 요인(나이, 가족력, 흡연력, 암 잠복기 등) 및 암 발생기전의 직업환경적 기여요인들(위험물질별 노출량)에 대한 고려가 필요하다. 이 요인들을 고려하기 위하여는 해당 정보를 담고 있는 자료원들의 확보와 개인수준의 통계모형을 활용한 접근법이 필요하다. 그러나 자료원 확보에 많은 시간과 자원이 필요하며 개인정보보안 문제 또한 고려해야 하는 난점이 있으므로, 후속 연구의 모형과 데이터는 점진적으로 발전되는 형태로 계획되는 것이 바람직하다 판단된다.

1) (단기)추계 모형의 개선 방안

본 연구에서 사용된 데이터셋의 전처리법을 조금 변형하면, 성별에 나이 변인을 추가하여 국가 암 추계에 표준적으로 사용되는 Age-Period-Cohort 기반의 추계 모형을 소분류 수준의 단기 추계에 적용하는 방안을 고려해 볼 수 있다.

데이터 구성 방식 제안 방향으로 다음의 요약 데이터 형태를 고려해볼 수 있다. 소분류 수준에서, 1년 단위로 만들어진 나이(age) 구간을 행(row)으로 하고, 1년 단위로 만들어진 시기(period) 구간을 열(column)로 하는 Age-Period 행렬을 생각해 볼 수 있다 (1년이 너무 세밀한 경우, 2년-3년 정도의 구간도 가능하다). 유사한 방식으로 1년 단위로 만들어진 시기

(period) 구간을 행으로 하고 1년 단위로 만들어진 출생년도 구간을 열로 하는 Period-Cohort 행렬을 생각해 볼 수 있다.

만약 Age-Period 행렬 또는 Period-Cohort 행렬의 각 셀 안에는 해당 나이-시기/시기-출생년도를 가지면서 소분류에 속하는 회사에 재직했던 사람들을 대상으로 셀 별 재직기간의 총합과 관심있는 암의 발병자의 총수가 있다면, 상기 언급된 Age-Period-Cohort 기반의 추계 모형을 소분류 수준에서 적용해 볼 수 있다.

다만, 소분류 수준과 국가 수준은 양적인 측면에서 매우 다르므로 소분류 수준에서 사용되는 모형의 추계 모형에는 예측의 안정성을 더하기 위해 적절한 모수규제법(penalty 혹은 regularization)을 고려하는 것이 필요할 것으로 예상된다.

2) (추계 외) 감시목적의 후속연구를 위한 제언

위에서 논의된 추계 모형을 통해서는 해당 추계값에 대한 주요 (알려진) 위험요인(risk factor)의 기여분을 파악하기는 어렵다는 점을 먼저 주지할 필요가 있다. 성/연령 외의 위험요인은 그 요인효과의 암발병여부에 대한 효과의 유의성을 떠나 효과크기와 설명력의 크기가 약하고 측정 오차도 큰 편이다. 이런 특성의 설명변수들을 인입할수록 오히려 추계의 정확도는 떨어지는 것은 통계실무적으로 잘 알려져 있다.

따라서 여러 알려진 위험 요인의 효과크기를 평가하거나, 아니면 알려진 위험 요인을 모두 보정한 이후 (요인들로 설명되지 않는) 이상징후(abnormality)를 검출하는 것이 목적이라면 추계목적의 모형과는 별도로, 위험요인 평가와 이상징후 검출을 목적으로 하는 모형들이 연구개발될 필요가 있다.

(1) 알려진 위험 요인의 기여도 측정 관련

알려진 위험요인의 기여도를 측정하기 위하여는 사업장 별 특징(feature)를 반영하는 것이 중요하다. 예를 들어, (직무 노출 매트릭스 등을 이용한) 특정 물질의 노출정도 대한 노출정도 등, 추계모형에서 사용된 정보보다 더 높은 수준의 정보가 필요할 것이다. 만약 개인별 수준의 데이터가 이용이 가능하다면, 개인별 사업장 종사의 여정(history)에 기반하여 개인마다 노출물질별 누적노출량을 설명변수로 인입시키고 암발병 관련 표지자를 반응변수로 함으로써 노출물질별 영향을 평가하는 방안을 생각할 수 있다. 물론 이러한 작업을 위한 데이터셋이 개인정보 노출이 가능한 수준에서 만들어질 수 있는 데이터셋인지에 대해 검토가 필요하다. 만약 집단 수준에서의 데이터만 제공이 가능한 경우 작업장/소분류 수준에서의 평균적인 노출량과 암 발생 정보가 필요한데, 이 때는 집단 수준에서의 변수들 간의 관계가 개인 수준에서의 변수들 사이의 관계로 해석될 수 있는지에 대한 해석의 어려움이 생기므로 위험 요인을 정확히 판단하는데 본질적인 한계가 있을 수 있음에 주의해야 한다.

한편, 국가 규모의 직업별/산업별 위험요인(예-발암물질)에 대한 노출평가 자료를 바탕으로 위험요인 노출인구를 평가하고, 인구집단에서의 귀속위험도(Attributable Fraction)를 계산하는 방안을 고려해볼 수 있다. 관련하여 박소희 외 (2009) 및 손미아 외 (2010)에서는 문헌 메타분석을 통한 주요 직업 노출물질의 귀속위험도를 측정하고, 국가통계의 주요 집계자료들에 기반한 업종별 종사자수를 곱하여 노출추정인구와 암부담을 계산한 바 있다. 앞 문단에서 살펴보았듯, 산보연 내에서 개인수준으로 이용가능한 자료를 이용하여 귀속위험도를 더 정교하게 추정할 수 있는지 알아보기 위한 기법 및 가용자료들을 검토하는 것도 향후의 관심사가 될 수 있을 것이다.

(2) 추계를 이용하지 않은 감시 모형

한편, 위험요인에 대한 평가 외에 소분류 및 여러 분류 수준에서 관심 질병의 최근 발병 상황에 이상(anomaly) 현상이 나타나고 있는지 아닌지를 살펴보는 모니터링 또는 감시(surveillance) 방안에 대해 검토하는 것이 산업보건의 관점에서 매우 중요하다고 판단된다. 통계적으로 이상현상 감지를 위하여는 (1) 귀무가설을 “현재의 관찰값이 이전의 추세로 설명가능하다”로 설정한 뒤, 귀무가설을 기각하는 경우를 이상현상으로 선언하는 접근법 혹은 (2) 이전의 추세로 추계한 현시점의 예측값과 실제값의 잔차를 비교하는 방법 등이 사용되며, 실제로 미국에서는 긴급승인된 코로나19 백신의 이상반응을 사후 모니터링하기 위하여 위 방법론들이 사용된다 사업장별 혹은 집단수준에 대한 추적인년을 정교하게 정의할 수 있는 방법이 있다면, 이를 바탕으로 감시에 필요한 위험 평가를 실시간으로 살펴볼 수 있을 것이다.

3) 구득자료의 개선 방안

현재 연구팀에게 공개된 빅데이터 코호트는 개인 정보의 노출을 최소화하기 위해 기업별로 집계된 통계량으로 구성되어 있다. 또한, 입사자에 대해서만 집계가 되어 시간에 따라 근로자 수가 누적되는 형태로 자료가 생성되고 있다. 본 연구에 주어진 데이터는 이전의 보고서 (이상길 외 2018, 2020, 2021)에서 꾸준히 개발된 것으로 사업장별로 입사, 퇴사자가 매년 기록되고 있으나, 입사와 재입사가 구분되지 않고 재직기간에 상관없이 특정 사업장에 근무경험만 있으면 (질병 발병 전까지는) 추적인년이 매해 균등하게 누적되고 있다. 이 점은 추적 인년의 추계 예측 결과에 대해 해석적으로 오해를 불러일으킬 여지가 존재하여, 추후 질병의 역학적인 분석에서도 어떤 형태로든 편향된 결과가 도출될 수 있을 것으로 판단된다. 만약 업종이나 노출물질의 위험 기여를 더 정확하게 평가하고 감시모형을 고도화하기 위하여는, 집계 자료에서도 개인별로 노출기간이 다른 점을 반영할 수 있는 방안에 대한 연구가 필요할 것으로 예상된다.

참고문헌

- 산업안전보건연구원 (2019). 빅데이터를 활용한 직업성 질환 코호트 운영.
- 박민정, 김정연 (2015). 재현자료 작성 방법론 검토. 통계개발원.
- 박민정, 이용희, 권성훈 (2019). 차등정보보호에 관한 연구. 통계개발원.
- 박민정, 한정임, 박노성(2021), 통계데이터센터 DB를 활용한 재현자료 생성 방법 연구. 통계개발원.
- 박소희 외 16인 (2009). 국내 근거 암위험요인의 기여위험도 추정 및 암발생, 사망의 장기적 경향 추계. 국립암센터.
- 손미아, 이원진, 진영우, 문은경, 유병옥, 박소희, 피영규, 박미진, 김은아, 조삼래, 안연순, 백도명 (2010). 우리나라의 직업성 암 부담 연구. 보건복지부.
- 윤진하, 고상백, 안연순, 김환철, 강모열, 박철용, 김지현, 임성실, 김민석, 육지후, 주진한, 김은미 (2017). 빅데이터 기반 직업코호트 구축을 통한 질병 발생 연구. 산업안전보건연구원.
- 이상길, 김은아, 이경은, 엄희수, 신무영, 윤민주, 전교연 (2018). 빅데이터를 활용한 직업성 질환 코호트 운영(I). 산업안전보건연구원.
- 이상길, 이경은, 예신희, 윤민주, 전교연, 신무영 (2020). 빅데이터를 활용한 직업성 질환 코호트 운영(II). 산업안전보건연구원.
- 이경은, 이상길, 예신희, 전교연, 윤민주, 성정민 (2021). 빅데이터를 활용한

직업성 질환 코호트 운영 - 사업장기반 질환 감시자료 (1). 산업안전보건
연구원.

중앙암등록본부 (2021), 2019년 국가암등록통계 참고자료.

Bae, J. M., Lee, C. W., Lee, D. H. and Ahn, Y.O. (2002). Methods for Estimation of Nationwide Cancer Incidence from Population-based Cancer Registries. Korean Journal of Epidemiology 24(1), 1-6.

<https://www.e-epih.org/upload/pdf/kje-24-1-1.pdf>

Bray, F., & Moller, B. (2006). Predicting the future burden of cancer. Nature Reviews Cancer, 6(1), 63-74.

<https://doi.org/10.1038/nrc1781>

Freddie, B. and Møller, B. (2006). Predicting the future burden of cancer. Nature Reviews Cancer 6(1), 63-74.

International Agency of Research on Cancer (IARC) (2022), List of classifications by cancer sites with sufficient or limited evidence in humans, IARC Monographs Volumes 1-132,

https://monographs.iarc.who.int/wp-content/uploads/2019/07/Classifications_by_cancer_site.pdf

Jakobsen, E., Olsen, K. E., Bliddal, M., Hornbak, M., Persson, G. F., and Green, A. (2021). Forecasting lung cancer incidence, mortality, and prevalence to year 2030. BMC Cancer, 21(1).

<https://doi.org/10.1186/s12885-021-08696-6>

Jung, K. W., Won, Y. J., Hong, S., Kong, H. J., Im, J. S., and Seo,

- H. G. (2021). Prediction of Cancer Incidence and Mortality in Korea, 2021. *Cancer research and treatment*, 53(2), 316-322.
<https://doi.org/10.4143/crt.2021.290>
- Kim, E.-A., Lee, H.-E., and Kang, S.-K. Occupational burden of cancer in Korea. *Safety and Health at Work*, 1, 61-68.
- MIT Technology Review (2022), 합성 데이터의 시대가 오고 있다, 2022년 10월 24일 수정, 2022년 10월 31일 접속.
- Mistry, M., Parkin, D. M., Ahmad, A. S., and Sasieni, P. (2011). Cancer incidence in the United Kingdom: projections to the year 2030. *British journal of cancer*, 105(11), 1795-1803.
<https://doi.org/10.1038/bjc.2011.430>
- Panagiotelis, A., Gamakumara, P., Athanasopoulos, G., and Hyndman, R. J. (2022+). Probabilistic forecast reconciliation: Properties, evaluation and score optimisation. *European Journal of Operational Research*, to appear.
- Petrack, J. L., Kelly, S. P., Altekruse, S. F., McGlynn, K. A., and Rosenberg, P. S. (2016). Future of hepatocellular carcinoma incidence in the United States forecast through 2030. *Journal of Clinical Oncology*, 34(15), 178
<https://doi.org/10.1200/JCO.2015.64.7412>
- Pukkala, E., Martinsen, J. I., Lynge, E., Gunnarsdottir, H. K., Sparén, P., Tryggvadottir, L., Weiderpass, E., and Kjaerheim, K. (2009). Occupation and cancer – follow-up of 15 million people in five Nordic countries, *Acta Oncologica*, 48 (5),

646-790, DOI: 10.1080/02841860902913546

- Quante, A. S., Ming, C., Rottmann, M., Engel, J., Boeck, S., Heinemann, V., Westphalen, C. B., and Strauch, K. (2016). Projections of cancer incidence and cancer-related deaths in Germany by 2020 and 2030. *Cancer Medicine*, 5(9), 2649-2656. <https://doi.org/10.1002/cam4.767>
- Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States (vol 74, pg 2913, 2014). *Cancer Research*, 74(14), 4006-4006. <https://doi.org/10.1158/0008-5472.Can-14-1642>
- Rahib, L., Wehner, M. R., Matrisian, L. M., and Nead, K. T. (2021). Estimated Projection of US Cancer Incidence and Death to 2040. *Jama Network Open*, 4(4). <https://doi.org/10.1001/jamanetworkopen.2021.4708>
- Shim, H. and Choi, Y. H. (2014). Modeling and Forecasting of Cancer Incidence Rate of Korea, *Journal of The Korean Data Analysis Society* 16.6 (2014): 3107-3120. <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArticleView.kci?sereArticleSearchBean.artilId=ART001941346>
- Shin, H. R. et al. (2002). Cancer incidence in Korea. *Cancer Research Treatment*, 34(6), 405-408. <https://ir.ymlib.yonsei.ac.kr/handle/22282913/144200>

- Smittenaar, C. R., Petersen, K. A., Stewart, K., & Moitt, N. (2016). Cancer incidence and mortality projections in the UK until 2035. *Br J Cancer*, 115(9), 1147-1155. <https://doi.org/10.1038/bjc.2016.304>
- Yu, X. Q., Luo, Q. W., Hughes, S., Wade, S., Caruana, M., Canfell, K., and O'Connell, D. L. (2019). Statistical projection methods for lung cancer incidence and mortality: a systematic review. *Bmj Open*, 9(8). <https://doi.org/10.1136/bmjopen-2018-028497>
- Weir, H. K., Thompson, T. D., Stewart, S. L., and White, M. C. (2021). Cancer Incidence Projections in the United States Between 2015 and 2050. *Preventing Chronic Disease*, 18. <https://doi.org/10.5888/pcd18.210006>

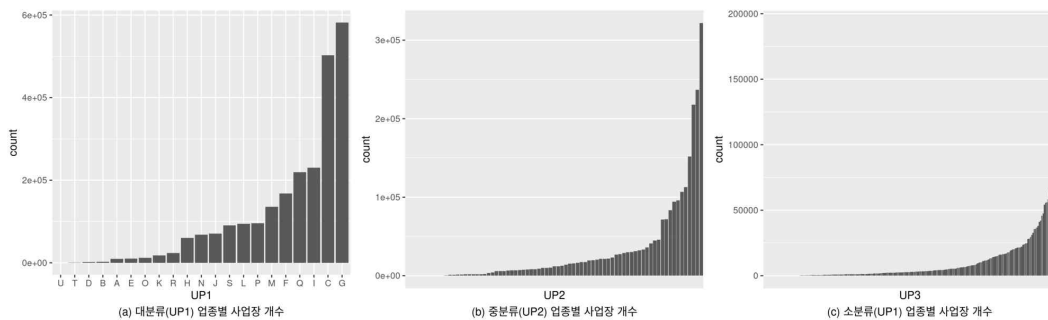
부록

1. 사업장 기반 감시 자료의 기초통계량 요약

(1) 사업장기반 백혈병 감시자료: ref_leukemia_bz_v4.csv

가) 대분류, 중분류, 소분류 업종별 사업장 개수

[그림 I-2]은 (a) 대분류, (b) 중분류, (c) 소분류 별 각 업종에 속하는 사업장의 개수를 나타낸 빈도막대그래프이다. 대분류(UP1) 업종마다 소속된 사업장의 개수는 최소 85개(U;국제 및 외국기관)에서 최대 581830개(G;도매 및 소매업)로 집계된다. 중분류(UP2)에서는 최소 8개(98;달리 분류되지 않은 자가 소비를 위한 가구의 재화 및 서비스 생산활동)에서 최대 321677개(46;도매 및 상품 중개업)로 집계되며, 소분류(UP3)에서는 최소 1개(425;시설물 유지관리 공사업, 981;자가 소비를 위한 가사 생산 활동)에서 최대 192814개(561;음식점업)로 집계된다.



[그림 부-1] 대분류(UP1), 중분류(UP2), 소분류(UP3) 업종별 사업장 개수

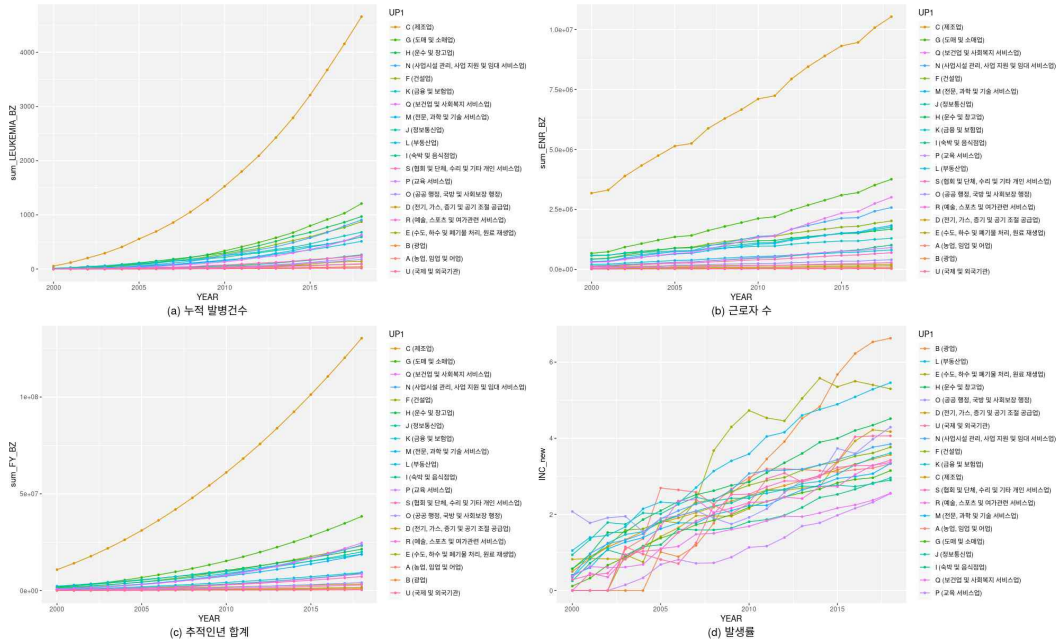
나) 대분류, 중분류 별 수집/관찰 변수

우선 대분류(UP1), 중분류(UP2), 소분류(UP3) 업종이 모두 NA인 관측치

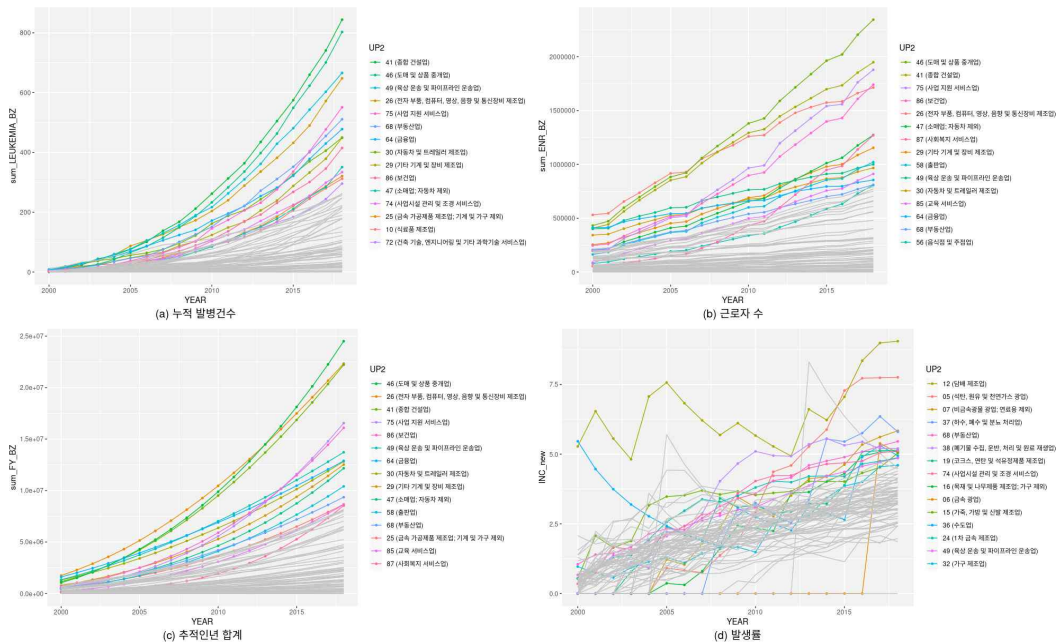
3325개를 제거하였다. 대분류(UP1) 업종이 T(가구 내 고용활동 및 달리 분류되지 않은 자가 소비 생산활동)인 사업장의 경우, 추적인년 합계(FY_BZ) 평균이 14.82이므로 집단의 대표성을 반영하지 못하며 업종 자체의 특징이 명확하지 않으므로 해당 업종은 분석에서 제외하였다.

[그림 I-3]는 대분류(UP1) 업종별 수집/관찰 변수의 연도(YEAR)에 따른 추세를 보여준다. (a)는 같은 대분류(UP1) 업종에 속하는 사업장의 관찰 연도(YEAR)별 누적 발병건수(LEUKEMIA_BZ)의 합(sum_LEUKEMIA_BZ)을 보여준다. 이와 비슷하게 (b),(c)는 근로자 수(ENR_BZ), 추적인년 합계(FY_BZ)의 합(sum_ENR_BZ, sum_FY_BZ)을 보여준다. ENR_BZ는 누적 값이 아님에도 연도가 증가할수록 선형적으로 증가하고 있는 것이 주목할 만한데, 2018년에 가까워질수록 관측하는 사업장과 코호트의 수가 증가하기 때문이다. (d)는 관찰 연도(YEAR)마다 발병률($INC_new = \frac{sum_LEUKEMIA_BZ}{sum_FY_BZ} \times 100000$)을 나타낸 것이다.

[그림 I-4]은 중분류(UP2)에 대하여 [그림 I-3]와 같은 그림을 제공한다. 77개의 중분류(UP2)를 비교하기가 쉽지 않으므로 2018년 Y값을 기준으로 상위 15개만 범례로써 나타내었다.



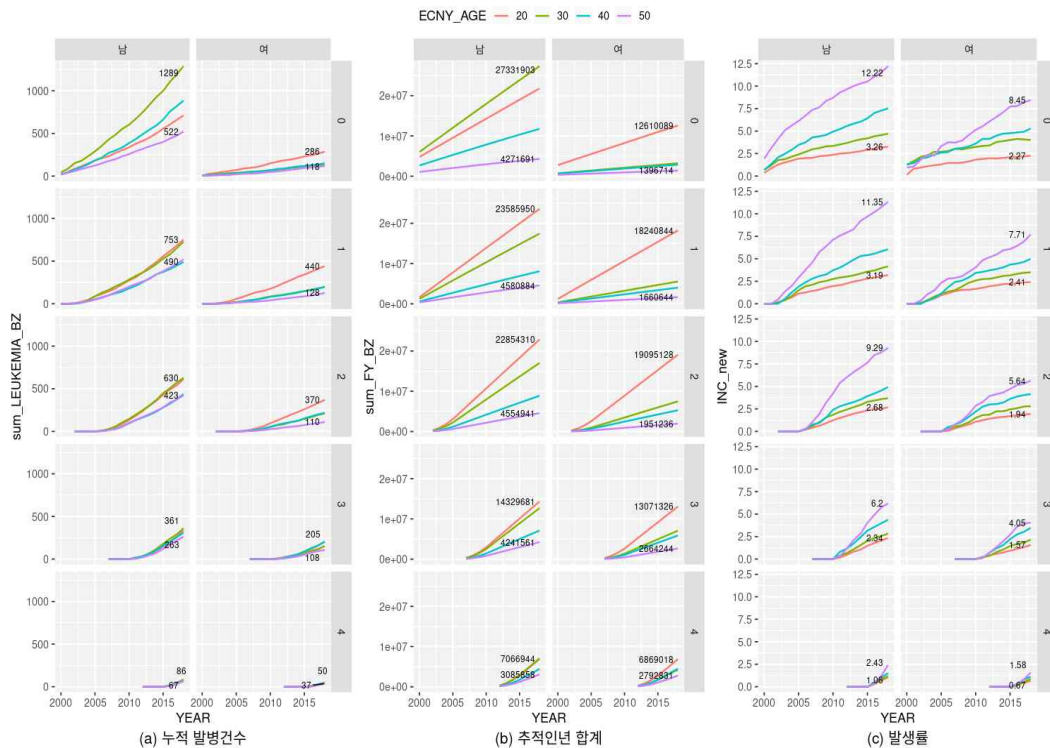
[그림 부-2] 대분류 업종별 연도에 따른 수집/관찰 변수



[그림 부-3] 중분류 업종별 연도에 따른 수집/관찰 변수 (상위 15개)

다) 코호트별 수집/관찰 변수

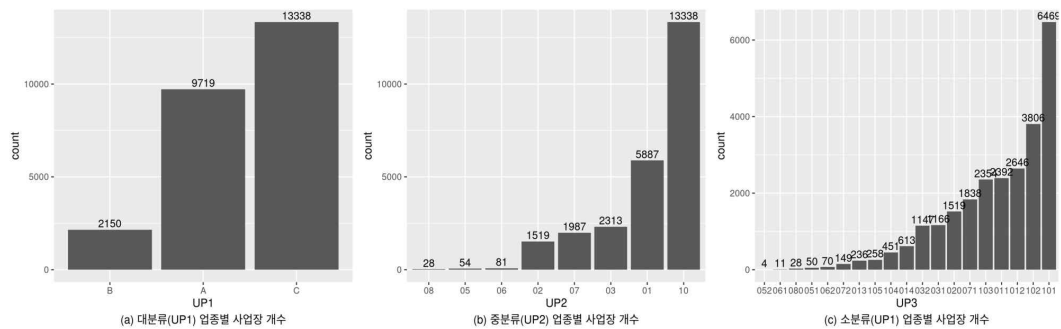
[그림 I-5]는 코호트별 수집/관찰 변수의 연도(YEAR)에 따른 추세를 보여 준다. (a)의 X축은 연도(YEAR)이고, Y축은 같은 코호트에 속하는 사업장의 연도(YEAR)별 누적 발병건수(LEUKEMIA_BZ)의 합(sum_LEUKEMIA_BZ)을 나타낸다. 각 코호트는 그래프의 위치와 선의 색깔로 구분한다. 그래프 위치는 열 방향으로 성별(SEX), 행 방향으로 입사시기(CAL)에 따라 구분하고 선의 색깔로 입사시연령(ECNY_AGE)을 구분한다. (b), (c)는 이와 동일하게 구성되며 Y축은 각각 추적인년 합계의 합(sum_FY_BZ), 발병률(INC_new)을 나타낸다. 참고로 CAL=5인 경우는 누적발병건수(LEUKEMIA_BZ)가 0으로 집계되었기 때문에 그림에 추가하지 않았다.



[그림 부-4] 코호트별 연도에 따른 수집/관찰 변수

(2) 사업장기반 폐암 감시자료: ref_lung_bz_v4.csv

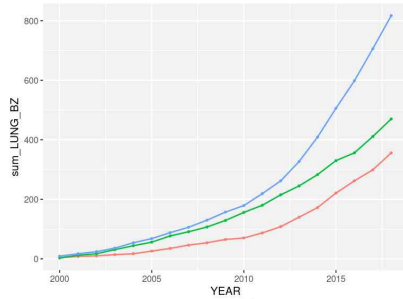
가) 대분류, 중분류, 소분류 업종별 사업장 개수 - [그림 I-6]



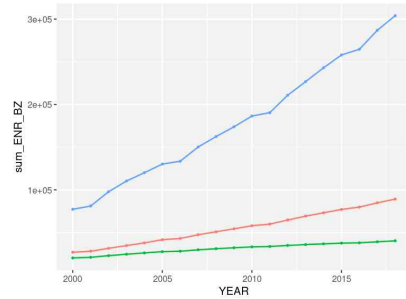
[그림 부-5] 코호트별 연도에 따른 수집/관찰 변수

나) 대분류, 중분류 별 수집/관찰 변수 - [그림 I-7], [그림 I-8]

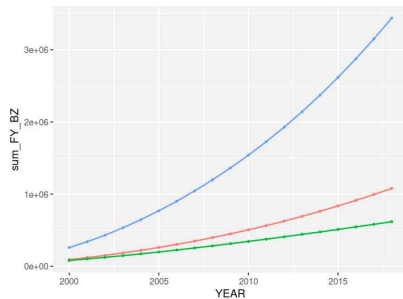
대분류, 중분류, 소분류 업종이 모두 NA인 관측치 3325개를 제거한 후 코호트별 수집/관찰 변수의 통계량을 계산하였다.



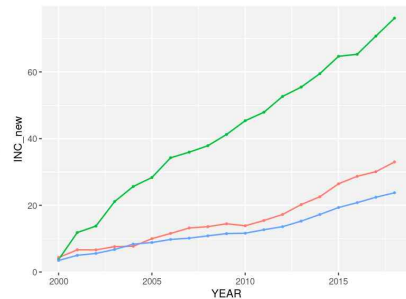
(a) 누적 발병건수



(b) 근로자 수

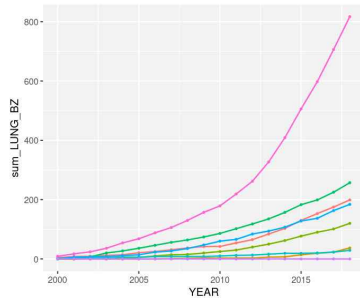


(c) 추적인년 합계

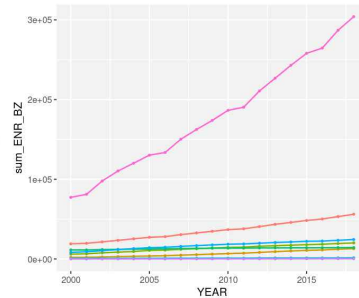


(d) 발생률

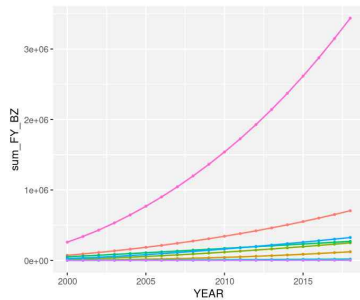
[그림 부-6] 대분류 업종별 연도에 따른 수집/관찰 변수



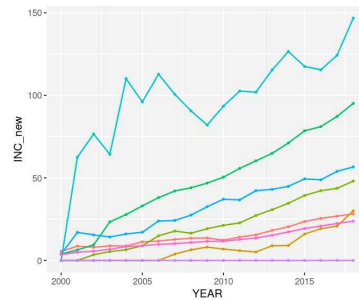
(a) 누적 발병건수



(b) 근로자 수



(c) 추적인년 합계

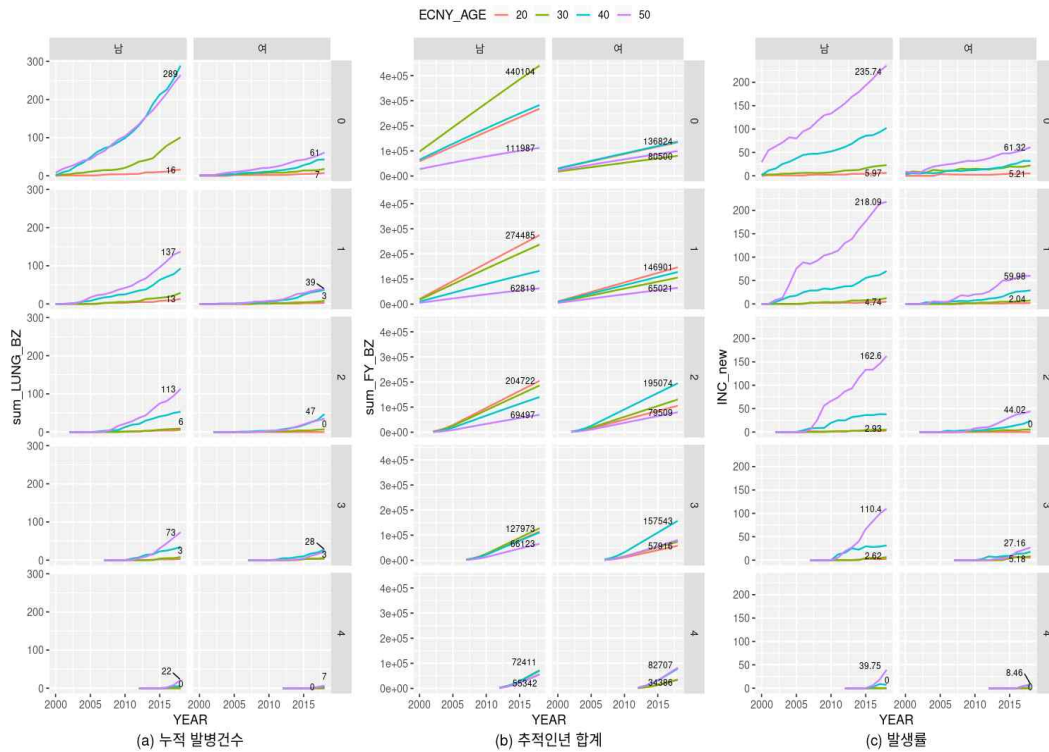


(d) 발생률

[그림 부-7] 중분류 업종별 연도에 따른 수집/관찰 변수

다) 코호트별 수집/관찰 변수 - [그림 I-9]

CAL=5인 경우는 누적발병건수(LUNG_BZ)가 0으로 집계되었기 때문에 그림에 추가하지 않았다.

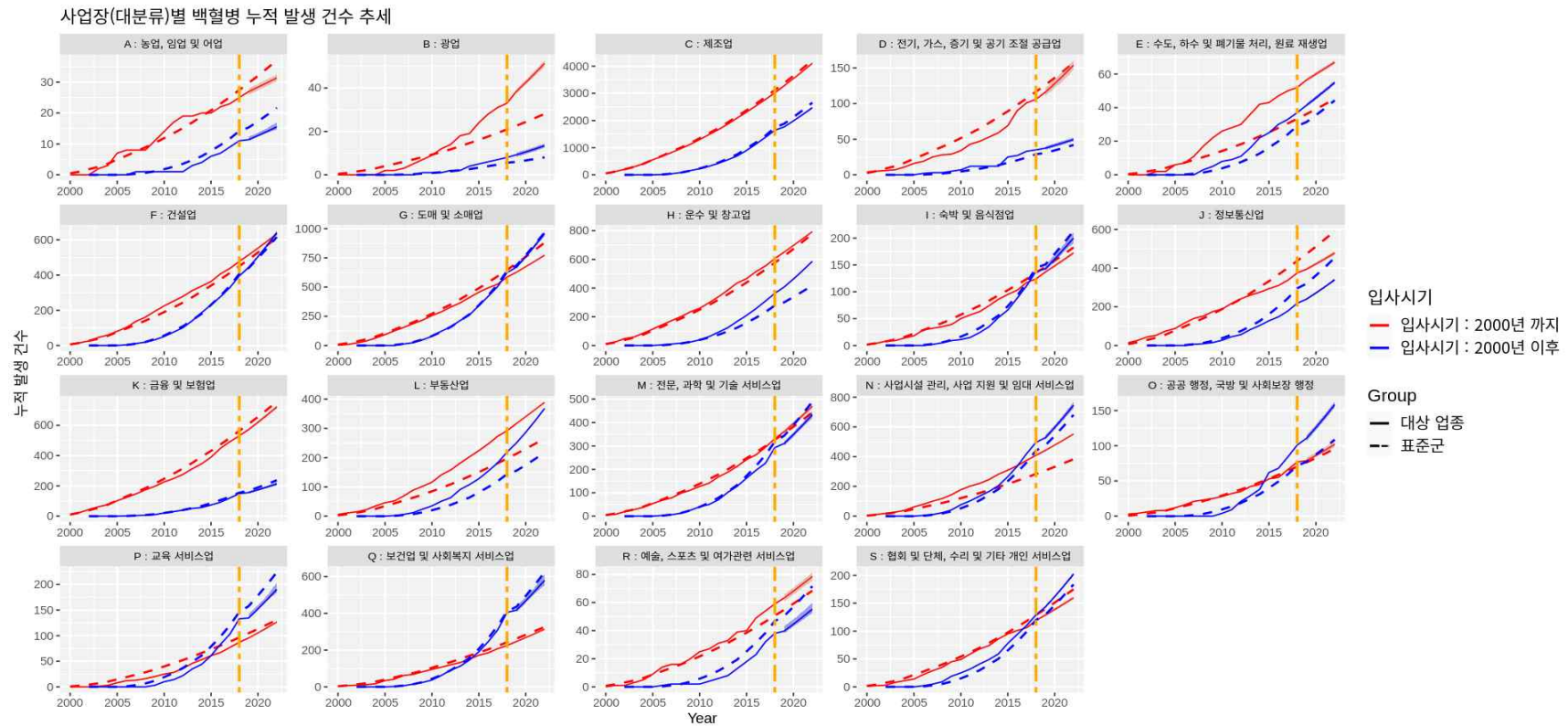


[그림 부-8] 코호트별 연도에 따른 수집/관찰 변수

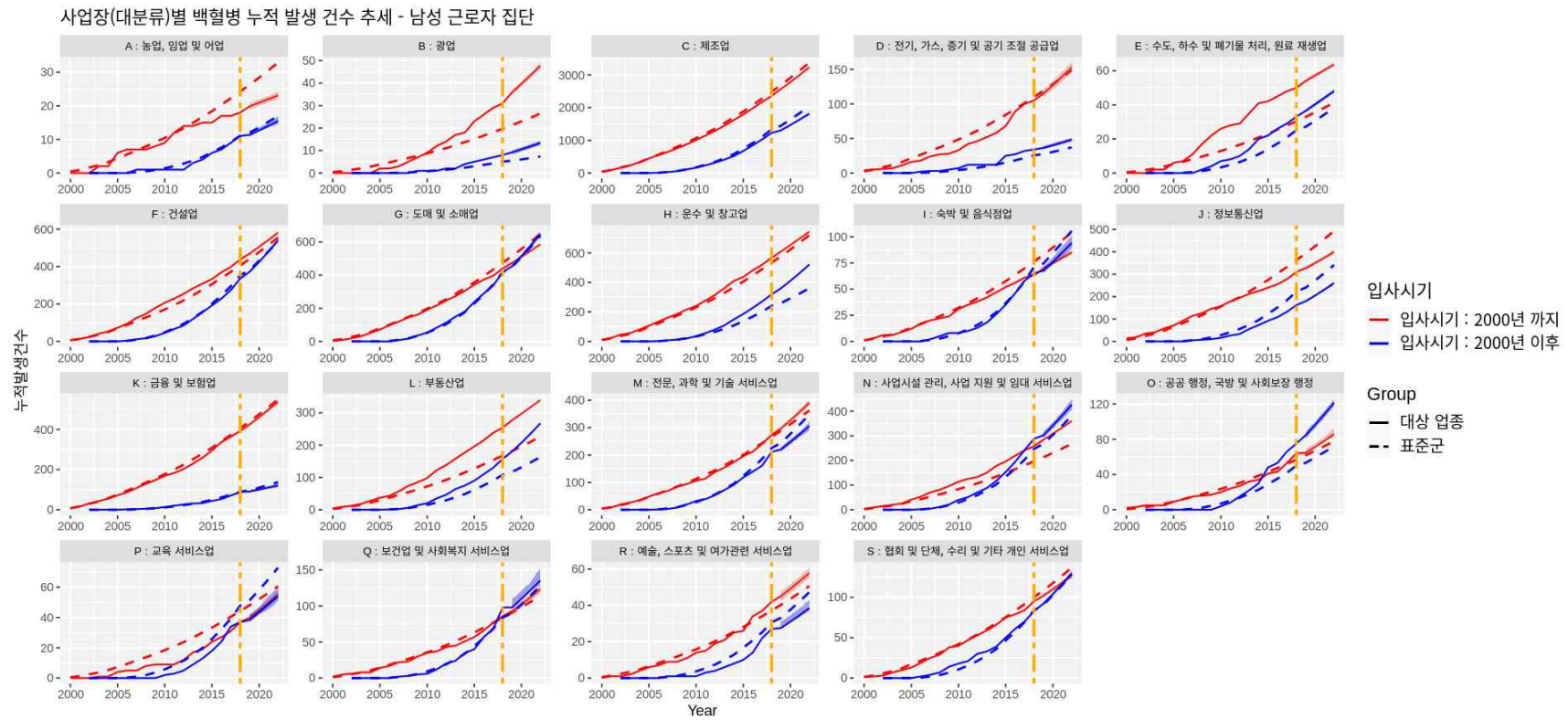
2. 근로자집단 암 발병률 추계값

여기서는 모든 업종(대분류별/중분류별)에 대하여 2000년도까지 입사 또는 2000년도 이후 입사한 남성, 여성 근로자집단 각각의 발병률과 참조집단 발병률의 추계값 그래프를 수록하였다.

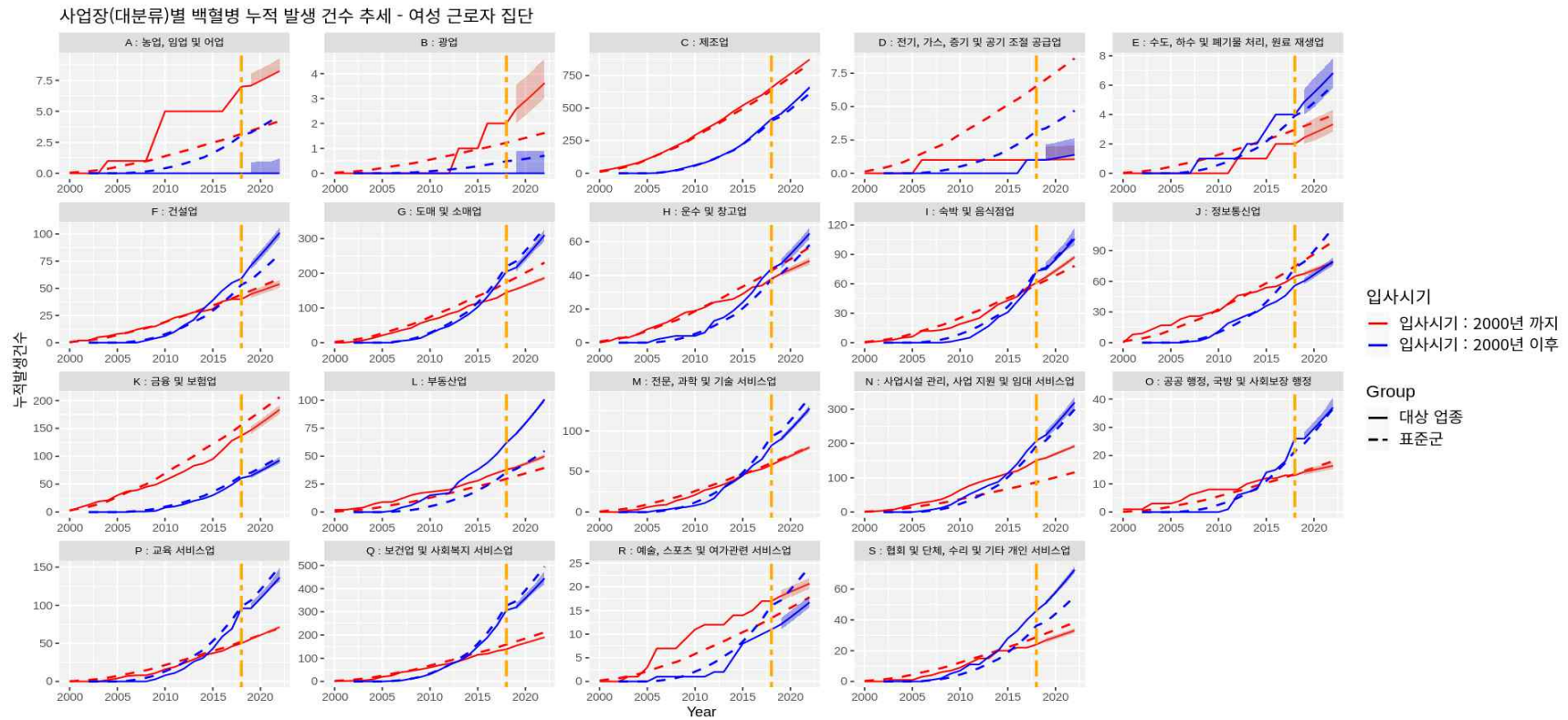
순서는 백혈병 - 폐암 순이며, 각 질환에 대하여 대분류(남여합계, 남성, 여성) 및 중분류(남여합계, 남성, 여성)의 그래프를 업종별로 정리하였다.



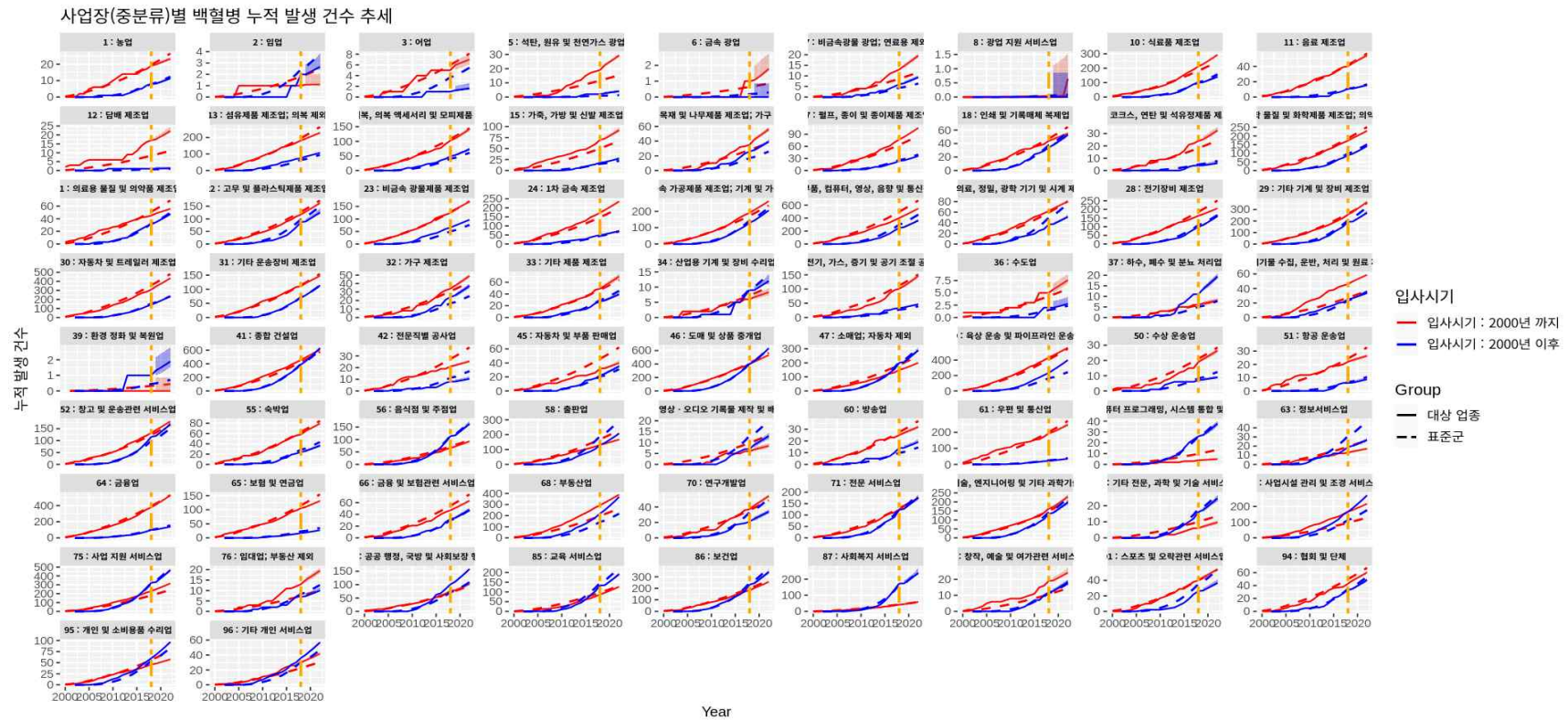
[그림 부-9] 사업장(대분류)별 백혈병 누적 발생 건수 추세



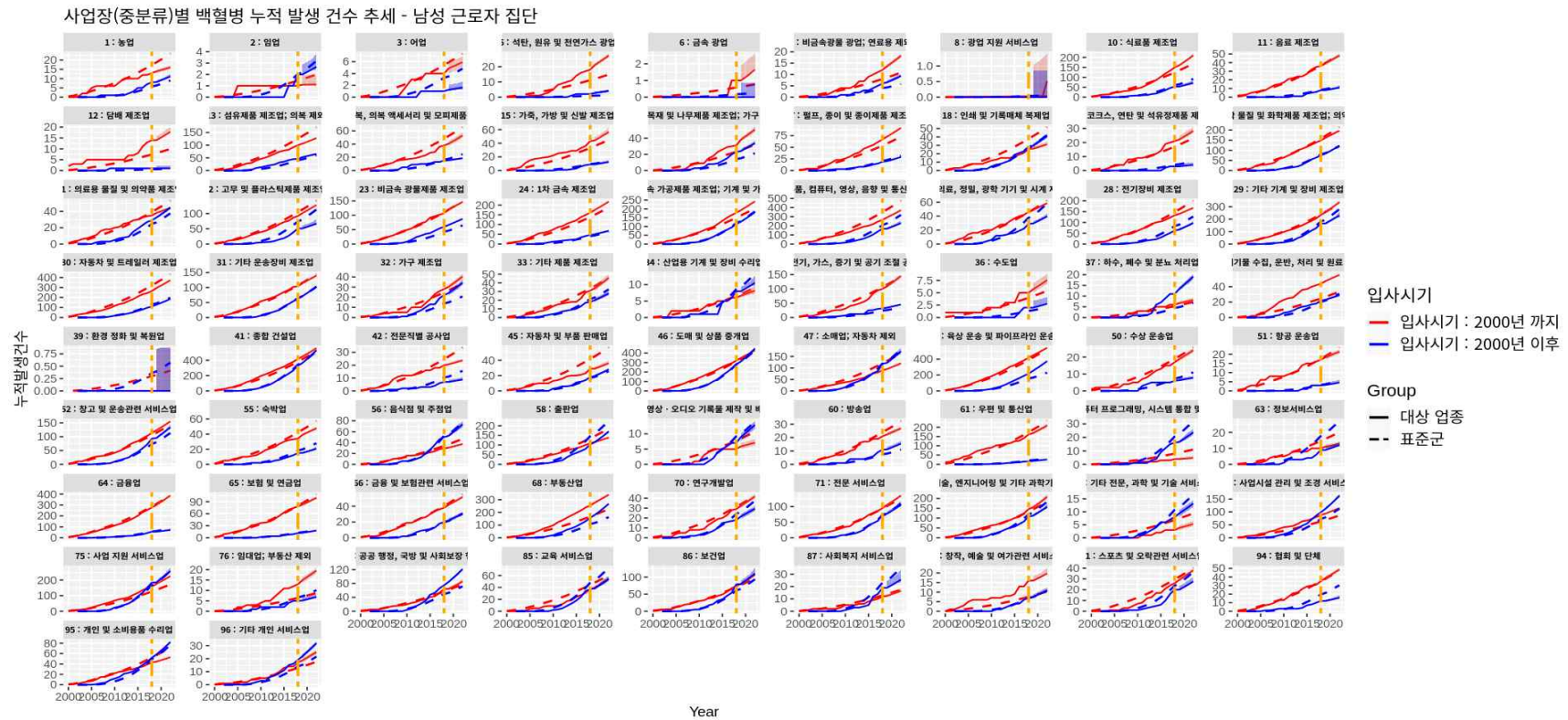
[그림 부-10] 사업장(대분류)별 남성 근로자 집단 대상 백혈병 누적 발생 건수 추세



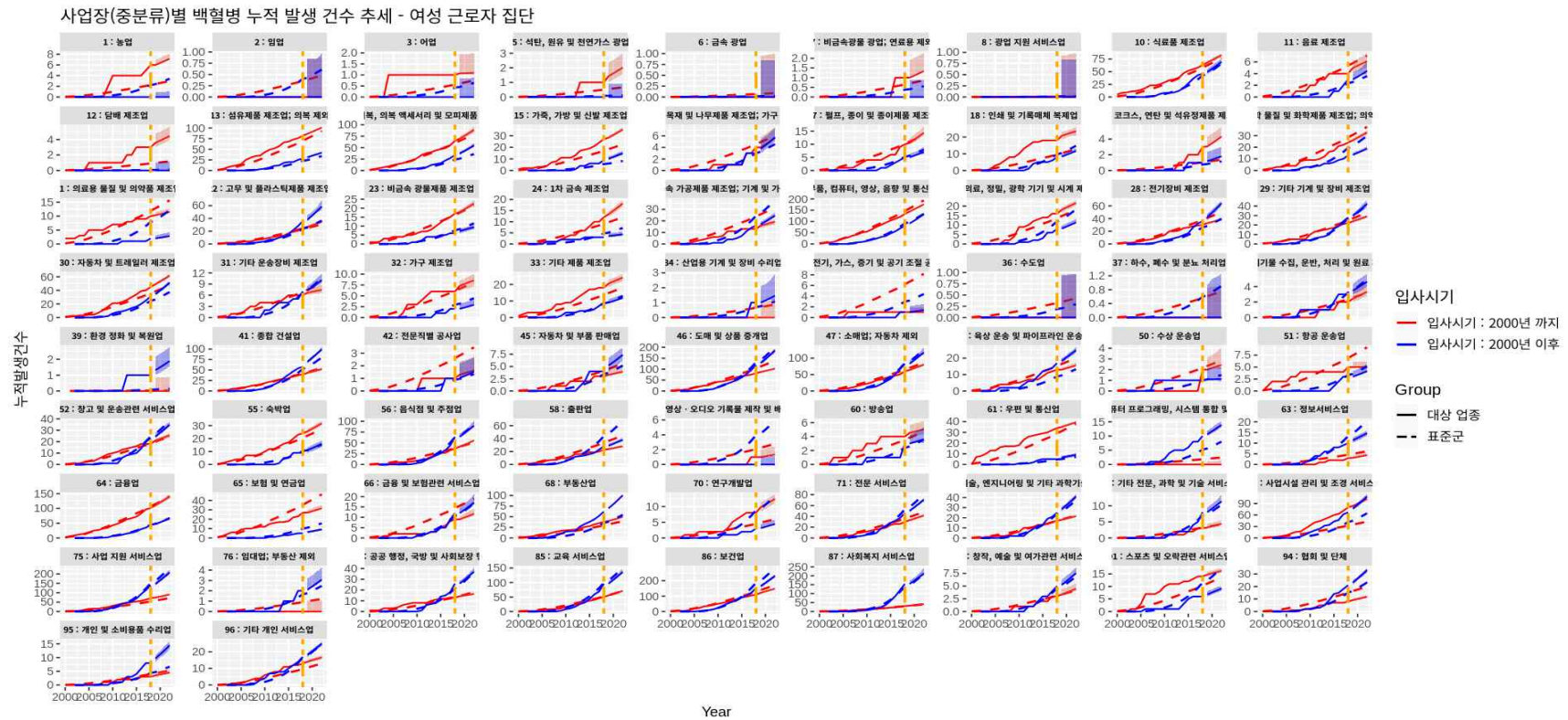
[그림 부-11] 사업장(대분류)별 여성 근로자 집단 대상 백혈병 누적 발생 건수 추세



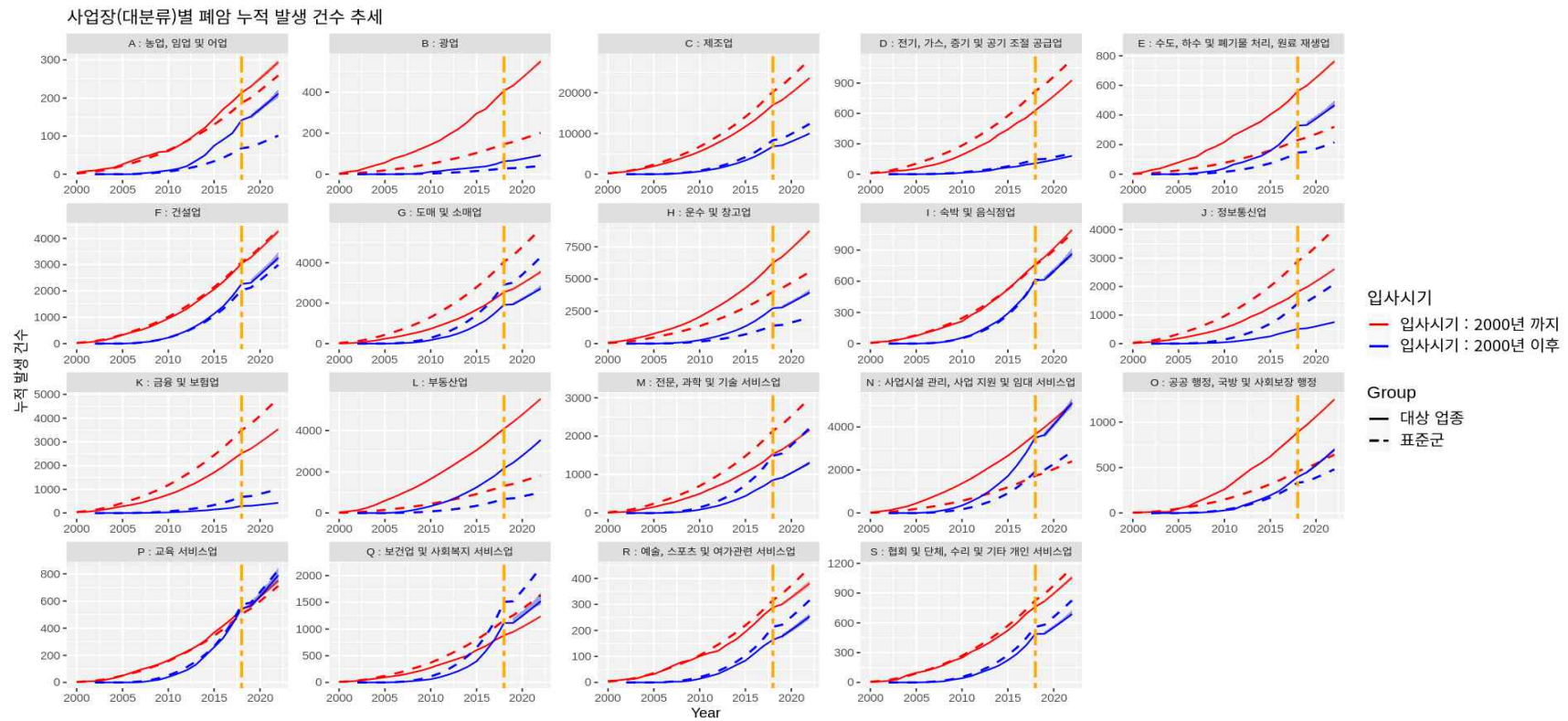
[그림 부-12] 사업장(중분류)별 백혈병 누적 발생 건수 추세



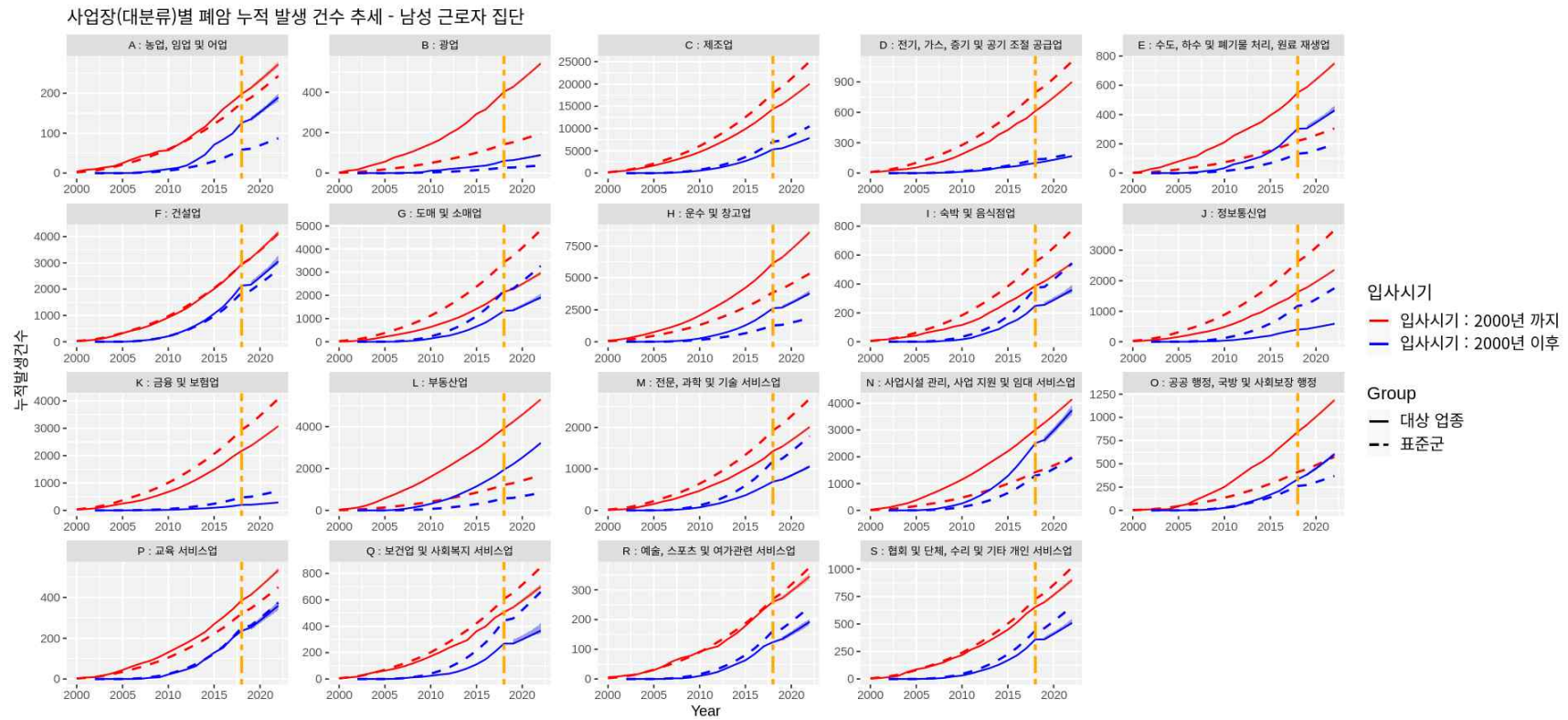
[그림 부-13] 사업장(중분류)별 남성 근로자 집단 대상 백혈병 누적 발생 건수 추세



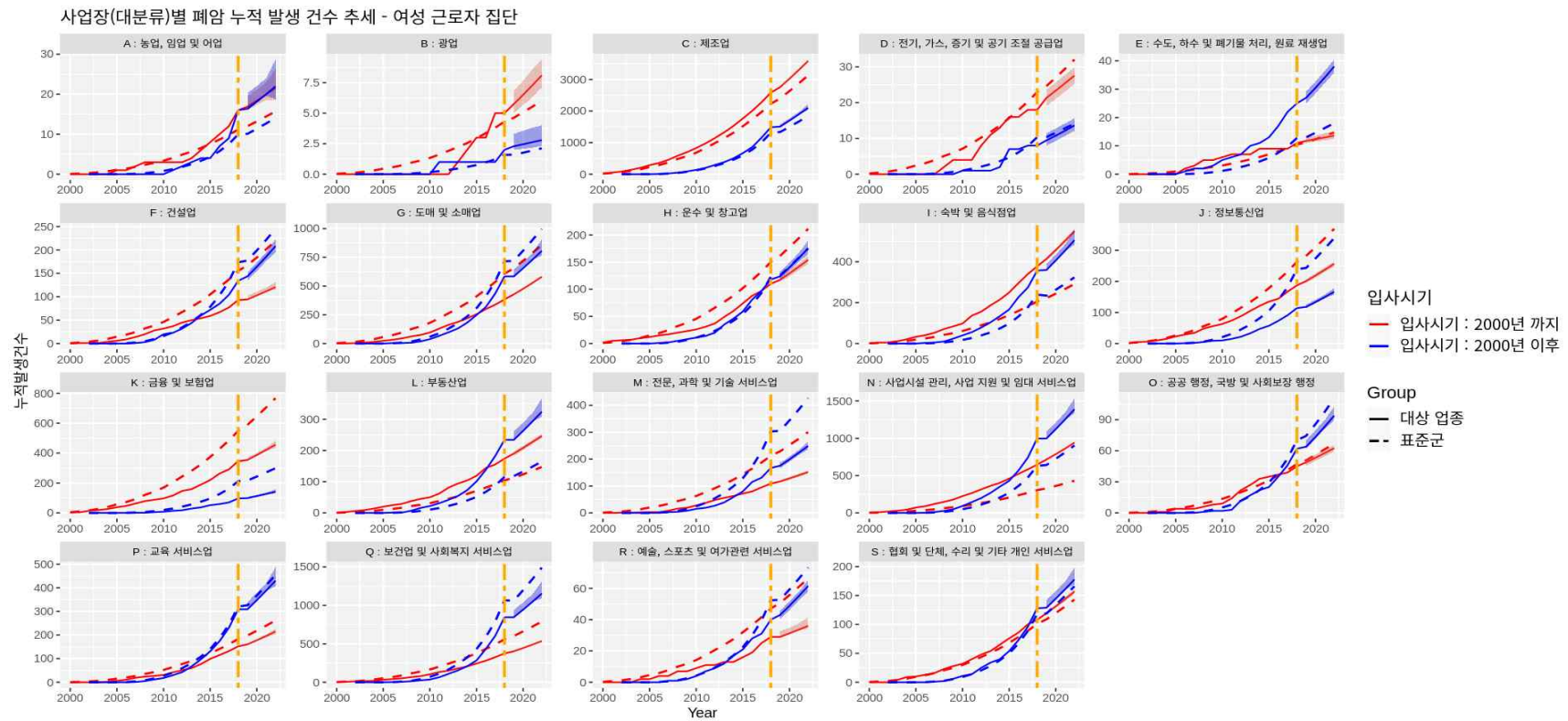
[그림 부-14] 사업장(중분류)별 여성 근로자 집단 대상 백혈병 누적 발생 건수 추세



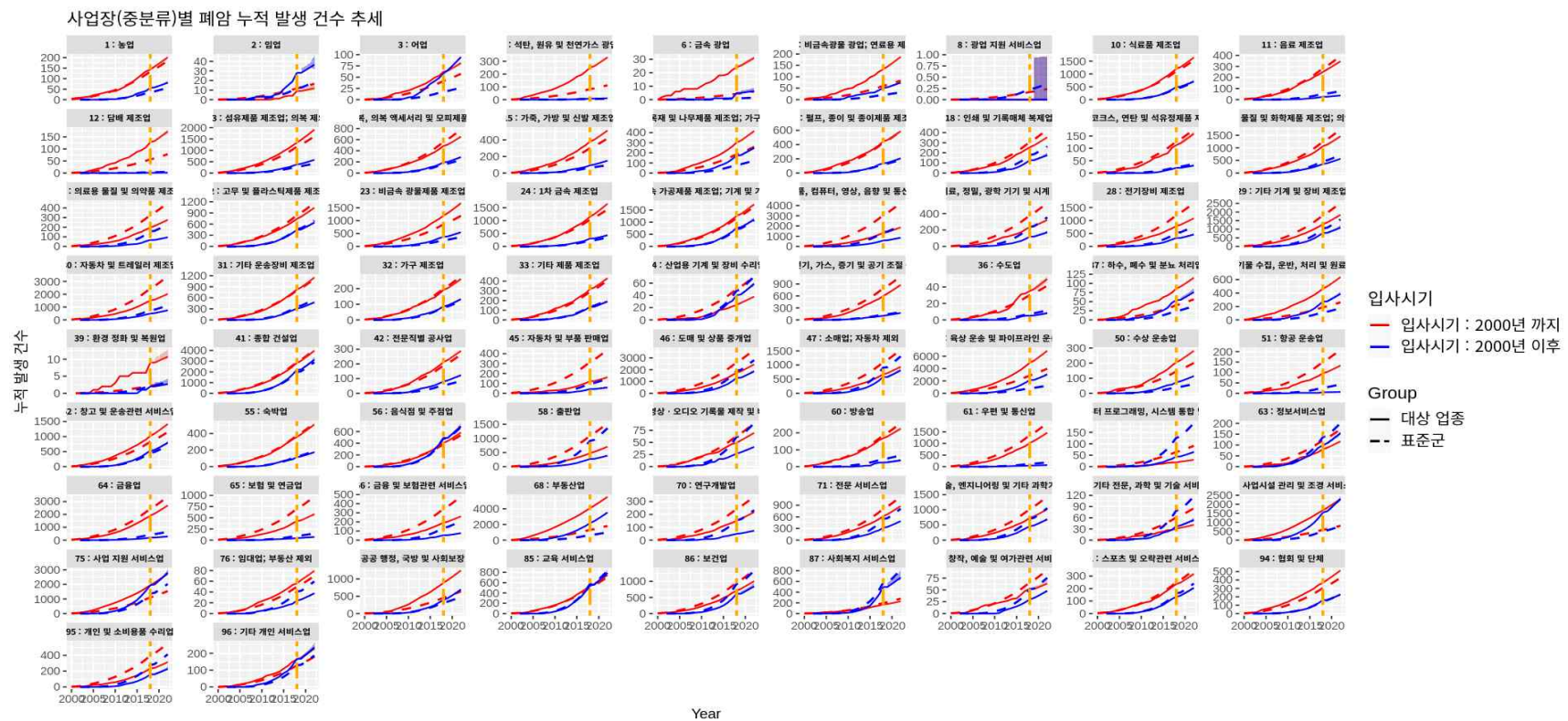
[그림 부-15] 사업장(대분류)별 폐암 누적 발생 건수 추세



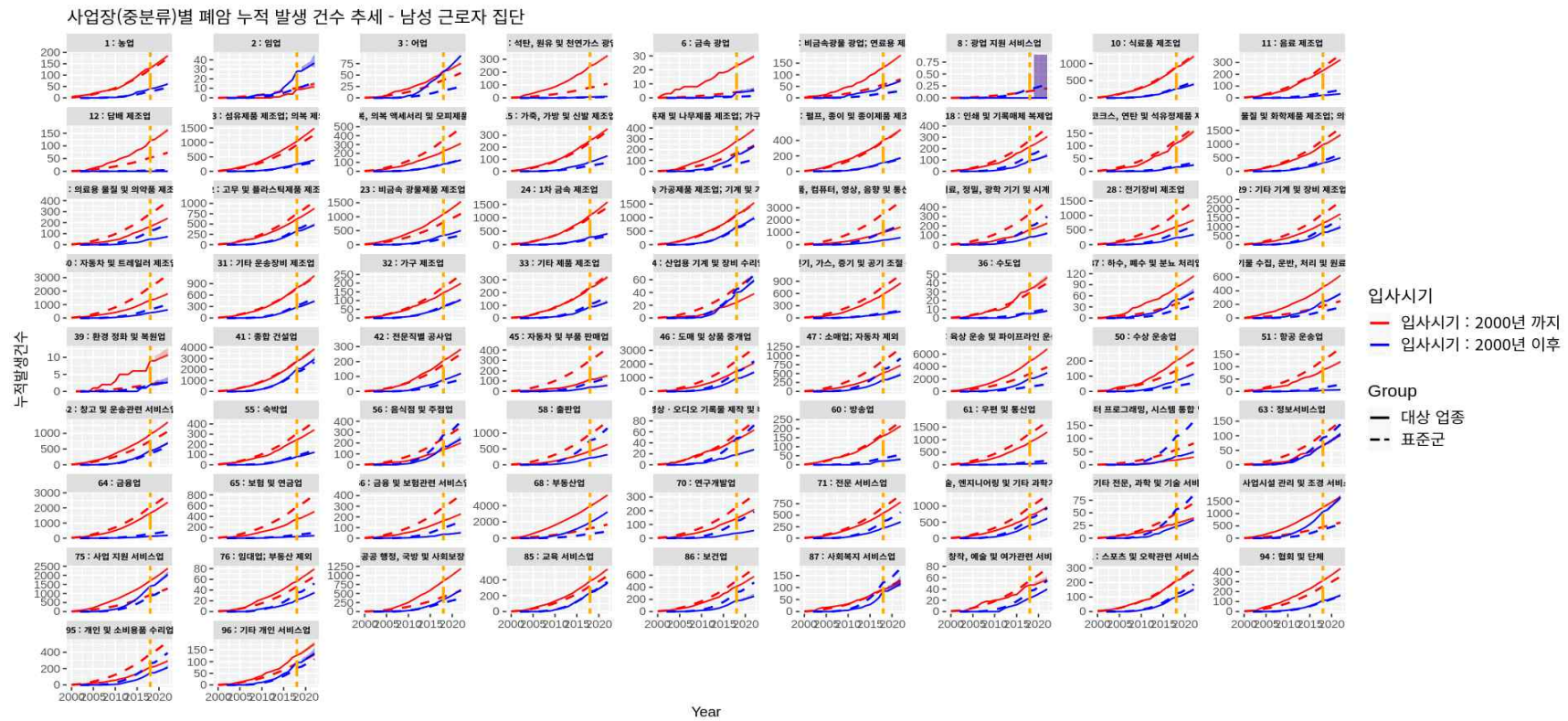
[그림 부-16] 사업장(대분류)별 남성 근로자 집단 대상 폐암 누적 발생 건수 추세



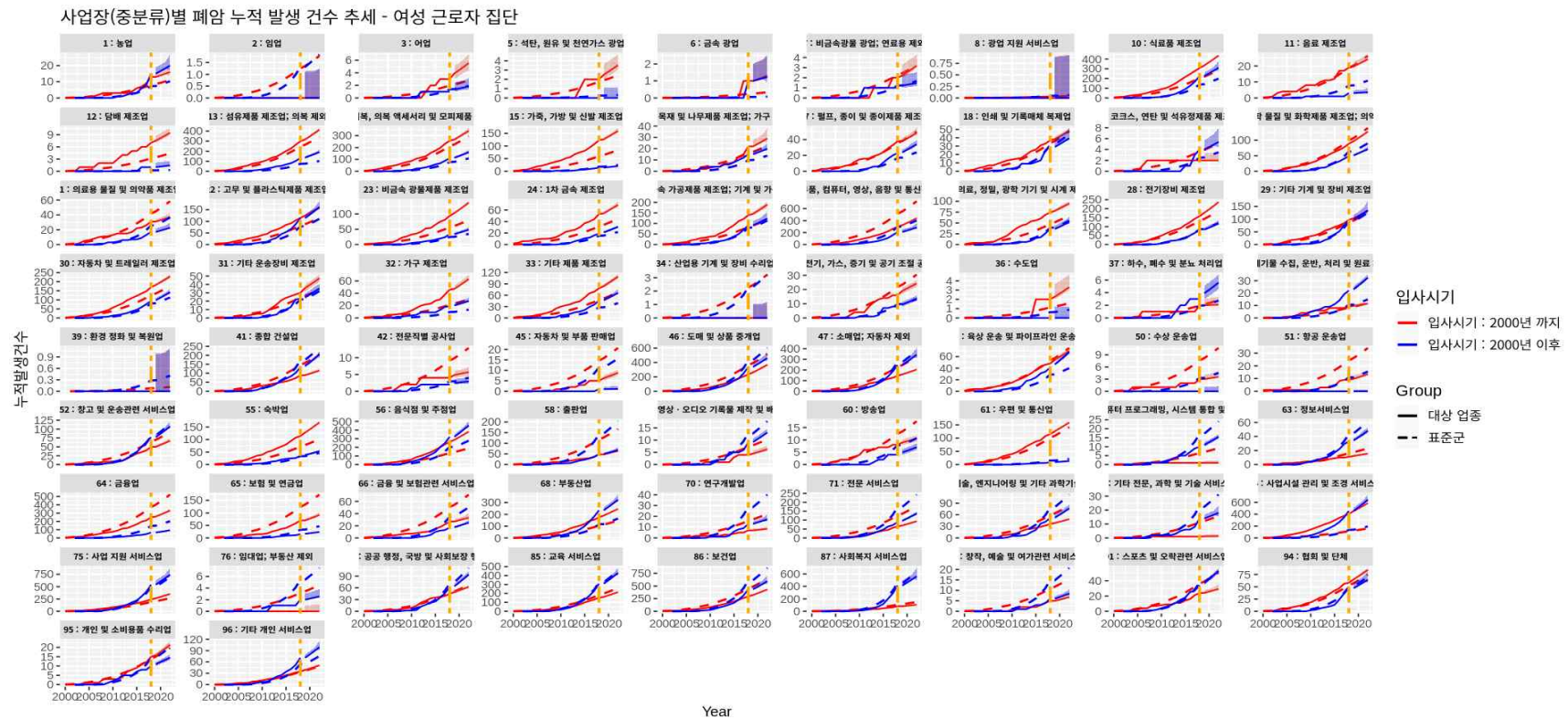
[그림 부-17] 사업장(대분류)별 여성 근로자 집단 대상 폐암 누적 발생 건수 추세



[그림 부-18] 사업장(중분류)별 폐암 누적 발생 건수 추세



[그림 부-19] 사업장(중분류)별 남성 근로자 집단 대상 폐암 누적 발생 건수 추세



[그림 부-20] 사업장(중분류)별 여성 근로자 집단 대상 폐암 누적 발생 건수 추세

〈표 부-1〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	25	-2	0.91	31	-6	0.85	0.03	0.12
B	광업(05~08)	33	12	1.58	51	23	1.83	0.03	0.04
C	제조업(10~34)	3016	-106	0.97	4107	-111	0.97	0	0
D	전기, 가스, 증기 및 공기 조절 공급업(35)	106	-11	0.91	154	-3	0.98	0.06	0.16
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	52	19	1.56	67	22	1.48	0.02	0.02
F	건설업(41~42)	478	27	1.06	635	21	1.03	0.01	0.01
G	도매 및 소매업(45~47)	585	-62	0.9	773	-108	0.88	0.01	0.01
H	운수 및 창고업(49~52)	605	28	1.05	794	16	1.02	0.01	0.01
I	숙박 및 음식점업(55~56)	124	-11	0.92	172	-10	0.94	0.01	0.02
J	정보통신업(58~63)	375	-63	0.86	478	-114	0.81	0.02	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	530	-33	0.94	722	-35	0.95	0.02	0.07
L	부동산업(68)	292	95	1.48	389	123	1.46	0.01	0.01
M	전문, 과학 및 기술 서비스업(70~73)	331	4	1.01	470	27	1.06	0.01	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	409	125	1.44	552	170	1.44	0	0.01
O	공공 행정, 국방 및 사회보장 행정(84)	77	7	1.09	102	6	1.07	0.05	0.11
P	교육 서비스업(85)	87	-10	0.9	127	-4	0.97	0.01	0.02
Q	보건업 및 사회복지 서비스업(86~87)	225	-18	0.93	314	-12	0.96	0.01	0.01
R	예술, 스포츠 및 여가관련 서비스업(90~91)	59	8	1.17	79	10	1.15	0.04	0.07

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-2〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	11	-3	0.78	16	-6	0.72	0.1	0.35
B	광업(05~08)	8	2	1.43	13	5	1.67	0.03	0.1
C	제조업(10~34)	1640	-98	0.94	2473	-183	0.93	0.05	0.19
D	전기, 가스, 증기 및 공기 조절 공급업(35)	35	6	1.21	50	8	1.18	0.08	0.42
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	37	8	1.27	55	11	1.24	0.02	0.02
F	건설업(41~42)	395	-11	0.97	640	24	1.04	0.02	0.07
G	도매 및 소매업(45~47)	623	-14	0.98	958	-10	0.99	0.04	0.15
H	운수 및 창고업(49~52)	363	84	1.3	586	168	1.4	0.02	0.12
I	숙박 및 음식점업(55~56)	138	-7	0.95	200	-13	0.94	0.09	0.32
J	정보통신업(58~63)	218	-77	0.74	340	-114	0.75	0.03	0.06

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	149	-6	0.96	213	-24	0.9	0.1	0.24
L	부동산업(68)	219	75	1.52	368	151	1.69	0.01	0.02
M	전문, 과학 및 기술 서비스업(70~73)	293	-25	0.92	433	-56	0.89	0.06	0.16
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	496	56	1.13	746	65	1.1	0.06	0.2
O	공공 행정, 국방 및 사회보장 행정(84)	101	30	1.41	158	50	1.46	0.05	0.24
P	교육 서비스업(85)	133	-13	0.91	190	-33	0.85	0.1	0.34
Q	보건업 및 사회복지 서비스업(86~87)	405	-6	0.99	580	-44	0.93	0.11	0.29
R	예술, 스포츠 및 여가관련 서비스업(90~91)	38	-9	0.81	55	-16	0.77	0.12	0.4

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-3〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	18	-6	0.74	23	-10	0.7	0.04	0.13
B	광업(05~08)	31	11	1.57	48	21	1.8	0.03	0.03
C	제조업(10~34)	2359	-127	0.95	3234	-142	0.96	0	0.01
D	전기, 가스, 증기 및 공기 조절 공급업(35)	105	-5	0.95	153	4	1.03	0.06	0.16
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	50	20	1.64	64	22	1.54	0.01	0.02
F	건설업(41~42)	438	32	1.08	581	26	1.05	0.01	0.01
G	도매 및 소매업(45~47)	440	-34	0.93	586	-64	0.9	0.01	0.01
H	운수 및 창고업(49~52)	567	33	1.06	745	24	1.03	0	0.01
I	숙박 및 음식점업(55~56)	63	-13	0.82	85	-19	0.82	0.02	0.03
J	정보통신업(58~63)	310	-53	0.85	399	-94	0.81	0.02	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	393	-14	0.97	538	-12	0.98	0.02	0.06
L	부동산업(68)	254	87	1.52	339	113	1.5	0.01	0.01
M	전문, 과학 및 기술 서비스업(70~73)	273	7	1.03	391	28	1.08	0.02	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	258	61	1.31	360	93	1.35	0.01	0.02
O	공공 행정, 국방 및 사회보장 행정(84)	64	7	1.12	86	8	1.1	0.07	0.13
P	교육 서비스업(85)	37	-7	0.84	55	-5	0.91	0.03	0.14
Q	보건업 및 사회복지 서비스업(86~87)	86	2	1.03	123	9	1.08	0.03	0.04
R	예술, 스포츠 및 여가관련 서비스업(90~91)	42	5	1.13	58	7	1.14	0.06	0.1

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-4〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
A	농업, 임업 및 어업(01~03)	11	0	0.99	15	-2	0.91	0.1	0.35
B	광업(05~08)	8	3	1.57	13	6	1.83	0.03	0.11
C	제조업(10~34)	1224	-111	0.92	1814	-230	0.89	0.05	0.21
D	전기, 가스, 증기 및 공기 조절 공급업(35)	34	8	1.32	48	11	1.3	0.08	0.41
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	33	8	1.31	48	10	1.26	0.01	0.01
F	건설업(41~42)	336	-17	0.95	539	4	1.01	0.03	0.08
G	도매 및 소매업(45~47)	418	0	1	648	11	1.02	0.03	0.13
H	운수 및 창고업(49~52)	319	78	1.32	522	161	1.45	0.01	0.11
I	숙박 및 음식점업(55~56)	65	-6	0.91	94	-11	0.89	0.08	0.35
J	정보통신업(58~63)	162	-60	0.73	260	-81	0.76	0.03	0.09

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	88	-2	0.98	120	-17	0.87	0.11	0.28
L	부동산업(68)	157	49	1.45	268	105	1.64	0.01	0.03
M	전문, 과학 및 기술 서비스업(70~73)	211	-14	0.94	306	-42	0.88	0.07	0.17
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	288	41	1.17	426	45	1.12	0.06	0.25
O	공공 행정, 국방 및 사회보장 행정(84)	75	25	1.5	121	49	1.69	0.05	0.23
P	교육 서비스업(85)	37	-11	0.77	54	-19	0.74	0.11	0.4
Q	보건업 및 사회복지 서비스업(86~87)	98	14	1.17	135	7	1.06	0.14	0.33
R	예술, 스포츠 및 여가관련 서비스업(90~91)	27	-4	0.88	39	-9	0.82	0.18	0.44

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-5〉 대분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	7	4	2.2	8	4	1.96	0.09	0.14
B	광업(05~08)	2	1	1.63	4	2	2.24	0.16	0.2
C	제조업(10~34)	657	21	1.03	873	32	1.04	0.01	0.01
D	전기, 가스, 증기 및 공기 조절 공급업(35)	1	-6	0.15	1	-8	0.12	0.05	0.14
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	2	-1	0.67	3	-1	0.84	0.16	0.16
F	건설업(41~42)	40	-4	0.9	54	-5	0.91	0.05	0.05
G	도매 및 소매업(45~47)	145	-28	0.84	187	-44	0.81	0.02	0.03
H	운수 및 창고업(49~52)	38	-5	0.89	49	-8	0.86	0.03	0.05
I	숙박 및 음식점업(55~56)	61	2	1.04	87	9	1.11	0.02	0.04
J	정보통신업(58~63)	65	-10	0.87	79	-21	0.79	0.02	0.02

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	137	-19	0.88	184	-22	0.89	0.04	0.08
L	부동산업(68)	38	8	1.28	50	10	1.26	0.02	0.06
M	전문, 과학 및 기술 서비스업(70~73)	58	-3	0.96	80	-1	0.99	0.02	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	151	64	1.73	192	77	1.66	0.03	0.04
O	공공 행정, 국방 및 사회보장 행정(84)	13	0	0.97	16	-2	0.91	0.03	0.03
P	교육 서비스업(85)	50	-2	0.96	71	1	1.02	0.02	0.07
Q	보건업 및 사회복지 서비스업(86~87)	139	-20	0.87	191	-21	0.9	0.03	0.03
R	예술, 스포츠 및 여가관련 서비스업(90~91)	17	4	1.27	21	3	1.16	0.03	0.04

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-6〉 대분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	0	-3	0	0	-5	0		
B	광업(05~08)	0	0	0	0	-1	0		
C	제조업(10~34)	416	13	1.03	659	47	1.08	0.04	0.12
D	전기, 가스, 증기 및 공기 조절 공급업(35)	1	-2	0.32	1	-3	0.3	0.57	0.92
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	4	0	1.02	7	1	1.14	0.11	0.23
F	건설업(41~42)	59	6	1.11	101	20	1.24	0.05	0.08
G	도매 및 소매업(45~47)	205	-15	0.93	310	-21	0.94	0.07	0.2
H	운수 및 창고업(49~52)	44	6	1.17	65	7	1.12	0.06	0.19
I	숙박 및 음식점업(55~56)	73	-1	0.99	106	-2	0.98	0.12	0.29
J	정보통신업(58~63)	56	-17	0.76	79	-33	0.71	0.04	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	61	-4	0.94	93	-7	0.93	0.07	0.19
L	부동산업(68)	62	27	1.75	101	46	1.84	0.02	0.02
M	전문, 과학 및 기술 서비스업(70~73)	82	-11	0.88	128	-14	0.9	0.04	0.12
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	208	15	1.08	320	20	1.07	0.05	0.12
O	공공 행정, 국방 및 사회보장 행정(84)	26	4	1.21	37	0	1.01	0.11	0.3
P	교육 서비스업(85)	96	-2	0.98	136	-14	0.9	0.11	0.32
Q	보건업 및 사회복지 서비스업(86~87)	307	-20	0.94	444	-51	0.9	0.1	0.29
R	예술, 스포츠 및 여가관련 서비스업(90~91)	11	-5	0.68	17	-8	0.69	0.1	0.35

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-7〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	19	-1	0.97	23	-3	0.88	0.05	0.11
02	임업	1	-1	0.55	1	-1	0.45	0.06	0.31
03	어업	5	-1	0.83	7	-1	0.86	0.08	0.17
05	석탄, 원유 및 천연가스 광업	19	7	1.64	29	14	1.89	0.05	0.08
06	금속 광업	1	0	1.56	2	1	2.08	0.51	0.72
07	비금속광물 광업; 연료용 제외	13	4	1.5	20	8	1.68	0.03	0.12
08	광업 지원 서비스업	0	0	0	1	1	17.37		
10	식료품 제조업	213	30	1.17	293	47	1.19	0.02	0.06
11	음료 제조업	40	-2	0.95	54	-2	0.96	0.03	0.03
12	담배 제조업	17	9	2.04	22	11	2	0.08	0.28
13	섬유제품 제조업; 의복 제외	182	-14	0.93	229	-35	0.87	0.01	0.02
14	의복, 의복 액세서리 및 모피제품 제조업	103	-5	0.95	142	-3	0.98	0.04	0.05
15	가죽, 가방 및 신발 제조업	70	20	1.39	92	24	1.36	0.05	0.05

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	35	7	1.25	56	18	1.48	0.06	0.09
17	펄프, 종이 및 종이제품 제조업	77	13	1.2	106	20	1.23	0.02	0.02
18	인쇄 및 기록매체 복제업	44	-3	0.93	55	-10	0.85	0.02	0.03
19	코크스, 연탄 및 석유정제품 제조업	24	6	1.32	33	8	1.33	0.07	0.21
20	화학 물질 및 화학제품 제조업; 의약품 제외	173	-14	0.93	230	-22	0.91	0.01	0.03
21	의료용 물질 및 의약품 제조업	45	-6	0.88	56	-13	0.81	0.02	0.03
22	고무 및 플라스틱제품 제조업	117	-11	0.92	163	-10	0.94	0.01	0.03
23	비금속 광물제품 제조업	124	-3	0.98	169	-2	0.99	0.03	0.04
24	1차 금속 제조업	173	25	1.17	236	36	1.18	0.01	0.02
25	금속 가공제품 제조업; 기계 및 가구 제외	191	17	1.1	262	26	1.11	0.02	0.06
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	399	-98	0.8	547	-124	0.82	0.01	0.03
27	의료, 정밀, 광학 기기 및 시계 제조업	62	3	1.04	80	0	1	0.01	0.02
28	전기장비 제조업	163	-23	0.88	206	-46	0.82	0.01	0.02
29	기타 기계 및 장비 제조업	256	-7	0.97	366	8	1.02	0.01	0.02
30	자동차 및 트레일러 제조업	309	-48	0.87	435	-48	0.9	0.02	0.03
31	기타 운송장비 제조업	110	-4	0.96	148	-7	0.96	0.02	0.02

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	34	3	1.1	49	7	1.18	0.04	0.13
33	기타 제품 제조업	49	2	1.04	70	6	1.1	0.04	0.04
34	산업용 기계 및 장비 수리업	6	-1	0.82	8	-2	0.83	0.04	0.07
35	전기, 가스, 증기 및 공기 조절 공급업	101	-11	0.9	146	-5	0.97	0.06	0.16
36	수도업	5	1	1.16	8	2	1.31	0.13	0.18
37	하수, 폐수 및 분뇨 처리업	6	0	1.02	8	0	1.02	0.04	0.08
38	폐기물 수집, 운반, 처리 및 원료 재생업	46	19	1.69	59	22	1.6	0.02	0.02
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	457	34	1.08	609	33	1.06	0.01	0.01
42	전문직별 공사업	21	-6	0.77	25	-12	0.68	0.03	0.04
45	자동차 및 부품 판매업	32	-14	0.7	40	-22	0.64	0.06	0.08
46	도매 및 상품 중개업	403	-13	0.97	532	-35	0.94	0.01	0.01
47	소매업; 자동차 제외	150	-35	0.81	200	-51	0.8	0.02	0.03
49	육상 운송 및 파이프라인 운송업	432	25	1.06	560	11	1.02	0.01	0.01
50	수상 운송업	19	-2	0.9	26	-2	0.93	0.05	0.07
51	항공 운송업	22	-2	0.9	26	-6	0.81	0.03	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	132	8	1.07	181	13	1.08	0.01	0.01
55	숙박업	58	-5	0.92	79	-5	0.94	0.02	0.02
56	음식점 및 주점업	66	-6	0.91	93	-5	0.95	0.03	0.03
58	출판업	132	-37	0.78	167	-63	0.73	0.01	0.02
59	영상·오디오 기록물 제작 및 배급업	6	-4	0.61	8	-5	0.63	0.05	0.08
60	방송업	25	-3	0.91	32	-5	0.86	0.03	0.04
61	우편 및 통신업	195	-7	0.96	249	-23	0.92	0.03	0.04
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	4	-6	0.41	5	-9	0.36	0.1	0.19
63	정보서비스업	13	-6	0.67	17	-9	0.65	0.04	0.08
64	금융업	378	-15	0.96	527	-1	1	0.02	0.05
65	보험 및 연금업	105	-10	0.91	131	-23	0.85	0.03	0.12
66	금융 및 보험관련 서비스업	47	-8	0.86	63	-11	0.85	0.03	0.1
68	부동산업	292	95	1.48	389	123	1.46	0.01	0.01
70	연구개발업	37	1	1.03	54	5	1.11	0.03	0.03
71	전문 서비스업	125	-5	0.96	177	1	1.01	0.03	0.03
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	163	12	1.08	230	24	1.12	0.04	0.05

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	6	-4	0.6	10	-4	0.71	0.06	0.16
74	사업시설 관리 및 조경 서비스업	168	72	1.75	217	88	1.68	0.02	0.02
75	사업 지원 서비스업	228	48	1.27	316	72	1.3	0.01	0.01
76	임대업; 부동산 제외	13	6	1.77	20	10	1.95	0.06	0.07
84	공공 행정, 국방 및 사회보장 행정	77	7	1.09	102	6	1.07	0.05	0.11
85	교육 서비스업	87	-10	0.9	127	-4	0.97	0.01	0.02
86	보건업	183	-17	0.91	256	-13	0.95	0.01	0.01
87	사회복지 서비스업	42	-1	0.98	58	1	1.01	0.02	0.03
90	창작, 예술 및 여가관련 서비스업	19	8	1.77	24	10	1.67	0.1	0.27
91	스포츠 및 오락관련 서비스업	40	0	1	54	0	1.01	0.03	0.05
94	협회 및 단체	43	-7	0.86	60	-7	0.89	0.05	0.1
95	개인 및 소비용품 수리업	46	-10	0.82	58	-19	0.75	0.02	0.02
96	기타 개인 서비스업	30	7	1.33	42	11	1.36	0.05	0.06

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-8〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	8	0	1	11	-1	0.91	0.1	0.37
02	임업	2	0	0.82	3	-1	0.7	0.42	0.79
03	어업	1	-3	0.27	2	-4	0.3	0.14	0.27
05	석탄, 원유 및 천연가스 광업	2	1	2.09	4	3	3.11	0.23	0.36
06	금속 광업	0	0	0	0	0	0		
07	비금속광물 광업; 연료용 제외	6	2	1.36	9	3	1.47	0.07	0.07
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	99	-4	0.96	141	-17	0.89	0.1	0.44
11	음료 제조업	11	0	1	15	-1	0.94	0.17	0.73
12	담배 제조업	1	0	1.54	1	0	1.26	0	0.07
13	섬유제품 제조업; 의복 제외	74	10	1.16	107	12	1.13	0.02	0.07
14	의복, 의복 액세서리 및 모피제품 제조업	49	8	1.21	73	13	1.21	0.05	0.06
15	가죽, 가방 및 신발 제조업	18	3	1.2	27	5	1.23	0.07	0.31

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	26	9	1.52	39	14	1.52	0.08	0.21
17	펄프, 종이 및 종이제품 제조업	24	-2	0.92	36	-4	0.91	0.07	0.2
18	인쇄 및 기록매체 복제업	34	-4	0.9	53	-4	0.92	0.05	0.11
19	코크스, 연탄 및 석유정제품 제조업	4	-1	0.81	6	-2	0.75	0.29	1
20	화학 물질 및 화학제품 제조업; 의약품 제외	88	-12	0.88	142	-10	0.94	0.04	0.25
21	의료용 물질 및 의약품 제조업	31	-2	0.95	47	-2	0.95	0.04	0.2
22	고무 및 플라스틱제품 제조업	87	-11	0.89	126	-24	0.84	0.15	0.39
23	비금속 광물제품 제조업	66	15	1.3	97	20	1.27	0.02	0.09
24	1차 금속 제조업	50	5	1.1	72	2	1.03	0.08	0.11
25	금속 가공제품 제조업; 기계 및 가구 제외	130	-14	0.9	205	-17	0.92	0.02	0.11
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	249	-50	0.83	361	-92	0.8	0.08	0.28
27	의료, 정밀, 광학 기기 및 시계 제조업	37	-13	0.75	51	-25	0.67	0.09	0.09
28	전기장비 제조업	98	-11	0.9	161	-6	0.97	0.06	0.11
29	기타 기계 및 장비 제조업	193	-15	0.93	272	-50	0.85	0.08	0.34
30	자동차 및 트레일러 제조업	141	-14	0.91	237	2	1.01	0.01	0.04
31	기타 운송장비 제조업	70	-2	0.98	113	0	1	0.02	0.14

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	24	7	1.43	37	12	1.48	0.08	0.13
33	기타 제품 제조업	27	-3	0.9	39	-6	0.87	0.09	0.1
34	산업용 기계 및 장비 수리업	9	0	0.96	12	-2	0.84	0.18	0.79
35	전기, 가스, 증기 및 공기 조절 공급업	33	6	1.21	47	7	1.18	0.08	0.39
36	수도업	2	0	1.31	3	0	1.17	0.43	1
37	하수, 폐수 및 분뇨 처리업	11	6	2.17	19	12	2.5	0.07	0.17
38	폐기물 수집, 운반, 처리 및 원료 재생업	25	1	1.06	34	-2	0.94	0.03	0.04
39	환경 정화 및 복원업	1	1	2.23	2	1	2.67	0.16	0.58
41	종합 건설업	387	-8	0.98	630	31	1.05	0.02	0.07
42	전문직별 공사업	8	-3	0.7	10	-7	0.61	0.2	0.37
45	자동차 및 부품 판매업	22	1	1.06	35	4	1.14	0.06	0.17
46	도매 및 상품 중개업	400	-13	0.97	633	3	1	0.03	0.13
47	소매업; 자동차 제외	201	-3	0.98	290	-17	0.95	0.08	0.2
49	육상 운송 및 파이프라인 운송업	234	68	1.41	400	156	1.64	0.02	0.07
50	수상 운송업	7	-1	0.85	9	-4	0.71	0.05	0.12
51	항공 운송업	6	-1	0.9	9	-2	0.83	0.25	0.76

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	116	18	1.18	169	17	1.11	0.08	0.22
55	숙박업	24	-5	0.83	37	-7	0.84	0.07	0.23
56	음식점 및 주점업	114	-2	0.98	163	-6	0.96	0.11	0.34
58	출판업	131	-59	0.69	208	-84	0.71	0.04	0.04
59	영상·오디오 기록물 제작 및 배급업	7	-6	0.53	13	-7	0.65	0.09	0.1
60	방송업	14	4	1.42	19	5	1.32	0.17	0.19
61	우편 및 통신업	21	-5	0.82	35	-4	0.9	0.06	0.08
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	26	0	1	38	-3	0.94	0.08	0.37
63	정보서비스업	19	-12	0.61	27	-21	0.56	0.08	0.09
64	금융업	100	-2	0.98	138	-17	0.89	0.11	0.31
65	보험 및 연금업	17	-6	0.74	27	-8	0.78	0.02	0.09
66	금융 및 보험관련 서비스업	32	2	1.06	48	0	1	0.11	0.15
68	부동산업	219	75	1.52	368	151	1.69	0.01	0.02
70	연구개발업	21	-12	0.64	34	-17	0.67	0.09	0.19
71	전문 서비스업	123	-1	0.99	177	-14	0.93	0.08	0.3
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	133	-10	0.93	198	-23	0.9	0.06	0.06

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	16	-2	0.9	25	-2	0.91	0.08	0.2
74	사업시설 관리 및 조경 서비스업	166	51	1.45	271	94	1.53	0.03	0.03
75	사업 지원 서비스업	323	6	1.02	465	-26	0.95	0.08	0.29
76	임대업; 부동산 제외	7	-1	0.84	10	-3	0.79	0.17	0.3
84	공공 행정, 국방 및 사회보장 행정	101	30	1.41	158	50	1.46	0.05	0.24
85	교육 서비스업	133	-13	0.91	190	-33	0.85	0.1	0.34
86	보건업	232	-16	0.94	342	-36	0.9	0.07	0.1
87	사회복지 서비스업	173	10	1.06	238	-8	0.97	0.17	0.59
90	창작, 예술 및 여가관련 서비스업	12	1	1.04	18	1	1.05	0.08	0.21
91	스포츠 및 오락관련 서비스업	26	-9	0.74	37	-17	0.68	0.15	0.5
94	협회 및 단체	33	-2	0.94	49	-5	0.91	0.07	0.3
95	개인 및 소비용품 수리업	60	6	1.11	97	13	1.16	0.02	0.02
96	기타 개인 서비스업	36	5	1.17	57	10	1.22	0.03	0.04

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-9〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	13	-4	0.75	16	-7	0.68	0.05	0.11
02	임업	1	0	0.69	1	-1	0.56	0.07	0.31
03	어업	4	-1	0.74	6	-1	0.81	0.09	0.18
05	석탄, 원유 및 천연가스 광업	18	7	1.62	27	12	1.84	0.04	0.06
06	금속 광업	1	0	1.74	2	1	2.15	0.53	0.81
07	비금속광물 광업; 연료용 제외	12	4	1.5	18	7	1.69	0.03	0.06
08	광업 지원 서비스업	0	0	0	1	0	19.22		
10	식료품 제조업	150	25	1.2	210	41	1.24	0.02	0.07
11	음료 제조업	36	-1	0.98	48	-1	0.97	0.03	0.04
12	담배 제조업	14	7	1.88	18	8	1.79	0.11	0.32
13	섬유제품 제조업; 의복 제외	100	-26	0.79	128	-43	0.75	0.02	0.03
14	의복, 의복 액세서리 및 모피제품 제조업	38	-10	0.79	53	-13	0.81	0.05	0.09
15	가죽, 가방 및 신발 제조업	43	10	1.3	57	12	1.26	0.06	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	31	8	1.32	51	19	1.59	0.08	0.14
17	펄프, 종이 및 종이제품 제조업	67	13	1.25	92	19	1.26	0.01	0.01
18	인쇄 및 기록매체 복제업	26	-12	0.69	31	-21	0.6	0.08	0.08
19	코크스, 연탄 및 석유정제품 제조업	21	4	1.23	28	5	1.22	0.08	0.2
20	화학 물질 및 화학제품 제조업; 의약품 제외	149	-9	0.94	197	-18	0.92	0.01	0.05
21	의료용 물질 및 의약품 제조업	35	-4	0.89	44	-9	0.83	0.02	0.03
22	고무 및 플라스틱제품 제조업	92	-13	0.88	128	-14	0.9	0.02	0.02
23	비금속 광물제품 제조업	107	-4	0.97	147	-3	0.98	0.04	0.03
24	1차 금속 제조업	161	22	1.16	218	30	1.16	0.01	0.02
25	금속 가공제품 제조업; 기계 및 가구 제외	178	28	1.19	243	39	1.19	0.02	0.07
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	265	-86	0.76	371	-106	0.78	0.02	0.04
27	의료, 정밀, 광학 기기 및 시계 제조업	46	0	1.01	59	-4	0.94	0.01	0.05
28	전기장비 제조업	133	-15	0.9	167	-34	0.83	0.01	0.03
29	기타 기계 및 장비 제조업	234	-4	0.98	337	12	1.04	0.01	0.02
30	자동차 및 트레일러 제조업	264	-57	0.82	373	-62	0.86	0.02	0.03
31	기타 운송장비 제조업	104	-3	0.97	140	-5	0.97	0.02	0.02

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	28	3	1.13	40	7	1.2	0.06	0.21
33	기타 제품 제조업	31	-3	0.92	46	0	1	0.09	0.09
34	산업용 기계 및 장비 수리업	6	-1	0.9	8	-1	0.9	0.04	0.06
35	전기, 가스, 증기 및 공기 조절 공급업	100	-6	0.94	145	2	1.01	0.06	0.16
36	수도업	5	1	1.25	8	2	1.41	0.13	0.18
37	하수, 폐수 및 분뇨 처리업	6	1	1.13	8	1	1.13	0.04	0.08
38	폐기물 수집, 운반, 처리 및 원료 재생업	44	19	1.77	56	22	1.65	0.02	0.02
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	418	36	1.1	557	36	1.07	0.01	0.01
42	전문직별 공사업	20	-5	0.81	24	-10	0.7	0.03	0.03
45	자동차 및 부품 판매업	29	-13	0.7	36	-21	0.64	0.07	0.1
46	도매 및 상품 중개업	321	9	1.03	430	2	1	0.01	0.01
47	소매업; 자동차 제외	90	-30	0.75	120	-45	0.73	0.02	0.07
49	육상 운송 및 파이프라인 운송업	419	27	1.07	544	16	1.03	0.01	0.02
50	수상 운송업	17	-2	0.9	24	-2	0.94	0.04	0.08
51	항공 운송업	17	-1	0.97	21	-2	0.9	0.03	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	114	9	1.08	155	12	1.09	0.01	0.01
55	숙박업	34	-8	0.8	48	-10	0.83	0.05	0.07
56	음식점 및 주점업	29	-5	0.85	37	-10	0.8	0.03	0.04
58	출판업	109	-26	0.81	139	-46	0.75	0.02	0.02
59	영상·오디오 기록물 제작 및 배급업	5	-3	0.64	7	-4	0.67	0.09	0.18
60	방송업	21	-3	0.88	27	-6	0.83	0.03	0.04
61	우편 및 통신업	161	-12	0.93	209	-24	0.9	0.04	0.05
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	4	-4	0.5	5	-6	0.43	0.1	0.21
63	정보서비스업	10	-5	0.68	13	-7	0.63	0.06	0.07
64	금융업	277	-9	0.97	387	0	1	0.02	0.05
65	보험 및 연금업	78	-2	0.97	100	-8	0.92	0.03	0.13
66	금융 및 보험관련 서비스업	38	-2	0.94	52	-3	0.94	0.02	0.03
68	부동산업	254	87	1.52	339	113	1.5	0.01	0.01
70	연구개발업	29	-2	0.92	42	-1	0.98	0.04	0.04
71	전문 서비스업	95	1	1.01	135	8	1.06	0.02	0.03
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	146	12	1.09	208	25	1.14	0.04	0.06

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	3	-4	0.44	5	-4	0.56	0.11	0.11
74	사업시설 관리 및 조경 서비스업	88	24	1.38	115	29	1.33	0.01	0.01
75	사업 지원 서비스업	157	30	1.24	225	53	1.31	0.02	0.03
76	임대업; 부동산 제외	13	7	2.01	20	11	2.22	0.06	0.07
84	공공 행정, 국방 및 사회보장 행정	64	7	1.12	86	8	1.1	0.07	0.13
85	교육 서비스업	37	-7	0.84	55	-5	0.91	0.03	0.14
86	보건업	74	3	1.04	107	10	1.11	0.03	0.04
87	사회복지 서비스업	12	-1	0.93	17	-1	0.94	0.04	0.11
90	창작, 예술 및 여가관련 서비스업	16	8	2.11	20	9	1.91	0.12	0.32
91	스포츠 및 오락관련 서비스업	26	-4	0.88	38	-2	0.94	0.03	0.06
94	협회 및 단체	35	0	1	49	1	1.03	0.04	0.09
95	개인 및 소비용품 수리업	43	-9	0.82	53	-19	0.74	0.03	0.03
96	기타 개인 서비스업	17	4	1.32	25	7	1.42	0.06	0.12

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-10〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	8	2	1.38	11	2	1.24	0.1	0.37
02	임업	2	0	0.98	3	-1	0.83	0.42	0.8
03	어업	1	-2	0.3	2	-3	0.34	0.14	0.27
05	석탄, 원유 및 천연가스 광업	2	1	2.27	4	3	3.39	0.23	0.36
06	금속 광업	0	0	0	0	0	0		
07	비금속광물 광업; 연료용 제외	6	2	1.49	9	4	1.61	0.07	0.07
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	53	-7	0.88	71	-20	0.78	0.12	0.51
11	음료 제조업	8	-1	0.93	11	-2	0.85	0.16	0.79
12	담배 제조업	1	0	1.71	1	0	1.18	0.02	0.05
13	섬유제품 제조업; 의복 제외	46	5	1.11	65	4	1.07	0.04	0.05
14	의복, 의복 액세서리 및 모피제품 제조업	15	-2	0.9	19	-6	0.75	0.04	0.04
15	가죽, 가방 및 신발 제조업	9	-1	0.94	12	-2	0.89	0.19	0.66

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	23	9	1.63	34	12	1.58	0.08	0.2
17	펄프, 종이 및 종이제품 제조업	19	-2	0.91	29	-3	0.92	0.06	0.09
18	인쇄 및 기록매체 복제업	26	-2	0.93	41	-1	0.97	0.05	0.16
19	코크스, 연탄 및 석유정제품 제조업	3	-1	0.72	4	-3	0.61	0.41	1
20	화학 물질 및 화학제품 제조업; 의약품 제외	75	-5	0.94	123	1	1.01	0.04	0.28
21	의료용 물질 및 의약품 제조업	29	4	1.18	45	7	1.19	0.05	0.23
22	고무 및 플라스틱제품 제조업	51	-23	0.69	67	-47	0.59	0.2	0.53
23	비금속 광물제품 제조업	60	17	1.4	87	23	1.35	0.02	0.09
24	1차 금속 제조업	47	6	1.16	67	5	1.08	0.1	0.13
25	금속 가공제품 제조업; 기계 및 가구 제외	114	-6	0.95	181	-4	0.98	0.02	0.09
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	168	-39	0.81	229	-91	0.71	0.11	0.37
27	의료, 정밀, 광학 기기 및 시계 제조업	29	-9	0.77	40	-18	0.68	0.09	0.1
28	전기장비 제조업	62	-21	0.74	97	-30	0.76	0.06	0.15
29	기타 기계 및 장비 제조업	165	-16	0.91	230	-50	0.82	0.08	0.35
30	자동차 및 트레일러 제조업	110	-21	0.84	186	-12	0.94	0.01	0.05
31	기타 운송장비 제조업	63	-2	0.97	103	1	1.01	0.02	0.16

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	22	8	1.59	34	14	1.65	0.08	0.12
33	기타 제품 제조업	18	-3	0.84	28	-5	0.85	0.07	0.08
34	산업용 기계 및 장비 수리업	8	-1	0.93	11	-3	0.79	0.17	0.77
35	전기, 가스, 증기 및 공기 조절 공급업	32	8	1.32	45	10	1.29	0.08	0.38
36	수도업	2	1	1.49	3	1	1.33	0.42	1
37	하수, 폐수 및 분뇨 처리업	11	7	2.45	19	12	2.79	0.07	0.18
38	폐기물 수집, 운반, 처리 및 원료 재생업	22	2	1.08	29	-2	0.94	0.04	0.1
39	환경 정화 및 복원업	0	0	0	0	-1	0		
41	종합 건설업	329	-14	0.96	530	11	1.02	0.03	0.07
42	전문직별 공사업	7	-3	0.67	9	-7	0.56	0.19	0.33
45	자동차 및 부품 판매업	18	1	1.05	28	2	1.08	0.05	0.12
46	도매 및 상품 중개업	282	1	1	447	17	1.04	0.02	0.15
47	소매업; 자동차 제외	118	-1	0.99	172	-8	0.96	0.08	0.11
49	육상 운송 및 파이프라인 운송업	217	60	1.38	376	145	1.63	0.02	0.06
50	수상 운송업	6	-1	0.83	8	-3	0.7	0.06	0.13
51	항공 운송업	3	0	0.89	5	-1	0.85	0.25	0.56

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	93	19	1.26	134	20	1.18	0.08	0.24
55	숙박업	15	-4	0.81	21	-7	0.74	0.08	0.25
56	음식점 및 주점업	50	-3	0.95	74	-4	0.95	0.09	0.39
58	출판업	103	-43	0.71	170	-55	0.76	0.04	0.04
59	영상·오디오 기록물 제작 및 배급업	7	-2	0.76	13	-1	0.92	0.09	0.1
60	방송업	11	4	1.47	15	5	1.41	0.14	0.17
61	우편 및 통신업	16	-5	0.76	26	-6	0.82	0.05	0.12
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	16	-5	0.77	24	-8	0.74	0.13	0.61
63	정보서비스업	9	-9	0.51	12	-15	0.44	0.2	0.2
64	금융업	56	-1	0.98	71	-16	0.82	0.16	0.42
65	보험 및 연금업	12	-1	0.93	18	-1	0.93	0.06	0.19
66	금융 및 보험관련 서비스업	20	0	1.02	31	0	0.99	0.07	0.09
68	부동산업	157	49	1.45	268	105	1.64	0.01	0.03
70	연구개발업	18	-6	0.74	30	-8	0.79	0.1	0.2
71	전문 서비스업	77	6	1.08	106	-4	0.96	0.12	0.4
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	108	-11	0.91	157	-26	0.86	0.07	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	8	-3	0.74	13	-4	0.79	0.07	0.1
74	사업시설 관리 및 조경 서비스업	99	26	1.35	163	50	1.45	0.03	0.06
75	사업 지원 서비스업	184	17	1.1	257	-2	0.99	0.09	0.36
76	임대업; 부동산 제외	5	-2	0.75	7	-3	0.68	0.19	0.33
84	공공 행정, 국방 및 사회보장 행정	75	25	1.5	121	49	1.69	0.05	0.23
85	교육 서비스업	37	-11	0.77	54	-19	0.74	0.11	0.4
86	보건업	79	18	1.3	110	17	1.18	0.13	0.26
87	사회복지 서비스업	19	-4	0.82	26	-10	0.73	0.23	0.63
90	창작, 예술 및 여가관련 서비스업	7	0	1.01	11	0	1.03	0.08	0.15
91	스포츠 및 오락관련 서비스업	20	-4	0.84	28	-9	0.75	0.22	0.55
94	협회 및 단체	12	-8	0.61	16	-15	0.52	0.11	0.42
95	개인 및 소비용품 수리업	52	2	1.04	82	6	1.07	0.02	0.02
96	기타 개인 서비스업	19	5	1.35	32	10	1.48	0.04	0.07

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-11〉 중분류별 백혈병 발병자수 추계 결과, 2000년도까지의 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	6	4	2.66	7	4	2.4	0.1	0.15
02	임업	0	0	0	0	0	0		
03	어업	1	0	1.77	1	0	1.45	0.05	0.13
05	석탄, 원유 및 천연가스 광업	1	0	1.99	2	1	3.08	0.18	0.45
06	금속 광업	0	0	0	0	0	0		
07	비금속광물 광업; 연료용 제외	1	0	1.54	1	0	1.56	0.42	1
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	63	5	1.08	83	6	1.08	0.02	0.03
11	음료 제조업	4	-1	0.74	6	-1	0.87	0.1	0.2
12	담배 제조업	3	2	3.3	4	3	3.76	0.12	0.2
13	섬유제품 제조업; 의복 제외	82	12	1.17	101	8	1.09	0.01	0.01
14	의복, 의복 액세서리 및 모피제품 제조업	65	5	1.08	89	9	1.12	0.05	0.09
15	가죽, 가방 및 신발 제조업	27	10	1.58	35	13	1.57	0.03	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	4	-1	0.88	6	0	0.93	0.24	0.42
17	펄프, 종이 및 종이제품 제조업	10	0	0.98	14	1	1.07	0.1	0.12
18	인쇄 및 기록매체 복제업	18	8	1.88	24	11	1.84	0.06	0.06
19	코크스, 연탄 및 석유정제품 제조업	3	2	2.88	4	3	3.15	0.14	0.34
20	화학 물질 및 화학제품 제조업; 의약품 제외	24	-4	0.85	33	-5	0.87	0.04	0.13
21	의료용 물질 및 의약품 제조업	10	-2	0.84	12	-4	0.74	0.03	0.03
22	고무 및 플라스틱제품 제조업	25	2	1.08	35	4	1.13	0.04	0.16
23	비금속 광물제품 제조업	17	1	1.05	23	1	1.06	0.03	0.16
24	1차 금속 제조업	12	3	1.32	18	6	1.52	0.05	0.08
25	금속 가공제품 제조업; 기계 및 가구 제외	13	-11	0.55	19	-12	0.61	0.09	0.16
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	134	-12	0.91	176	-18	0.91	0.01	0.03
27	의료, 정밀, 광학 기기 및 시계 제조업	16	2	1.17	22	4	1.2	0.04	0.09
28	전기장비 제조업	30	-8	0.79	38	-12	0.76	0.03	0.03
29	기타 기계 및 장비 제조업	22	-3	0.88	29	-4	0.88	0.01	0.05
30	자동차 및 트레일러 제조업	45	9	1.26	62	14	1.31	0.03	0.03
31	기타 운송장비 제조업	6	-1	0.87	7	-2	0.81	0.06	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	6	0	0.97	9	0	1.06	0.09	0.2
33	기타 제품 제조업	18	4	1.32	24	6	1.33	0.06	0.06
34	산업용 기계 및 장비 수리업	0	-1	0	0	-1	0		
35	전기, 가스, 증기 및 공기 조절 공급업	1	-5	0.16	1	-7	0.12	0.04	0.13
36	수도업	0	0	0	0	0	0		
37	하수, 폐수 및 분뇨 처리업	0	-1	0	0	-1	0		
38	폐기물 수집, 운반, 처리 및 원료 재생업	2	0	0.84	3	0	1.05	0.16	0.16
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	39	-3	0.94	52	-3	0.94	0.05	0.05
42	전문직별 공사업	1	-2	0.38	2	-2	0.47	0.12	0.3
45	자동차 및 부품 판매업	3	-1	0.72	4	-2	0.71	0.1	0.1
46	도매 및 상품 중개업	82	-22	0.79	102	-36	0.74	0.02	0.02
47	소매업; 자동차 제외	60	-5	0.92	80	-6	0.93	0.03	0.05
49	육상 운송 및 파이프라인 운송업	13	-2	0.85	16	-5	0.77	0.07	0.22
50	수상 운송업	2	0	0.92	2	0	0.85	0.75	0.83
51	항공 운송업	5	-2	0.72	5	-4	0.55	0.1	0.1

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	18	0	0.97	26	1	1.04	0.06	0.09
55	숙박업	24	3	1.16	32	4	1.16	0.04	0.09
56	음식점 및 주점업	37	-1	0.97	55	4	1.09	0.03	0.03
58	출판업	23	-11	0.68	28	-17	0.63	0.04	0.04
59	영상·오디오 기록물 제작 및 배급업	1	-1	0.48	1	-1	0.48	0.58	0.82
60	방송업	4	0	1.12	5	0	1.07	0.05	0.1
61	우편 및 통신업	34	5	1.17	39	1	1.04	0.01	0.01
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	0	-2	0	0	-2	0		
63	정보서비스업	3	-2	0.65	4	-2	0.71	0.12	0.2
64	금융업	101	-5	0.95	140	0	1	0.04	0.05
65	보험 및 연금업	27	-8	0.76	32	-15	0.68	0.07	0.13
66	금융 및 보험관련 서비스업	9	-5	0.62	12	-7	0.61	0.13	0.54
68	부동산업	38	8	1.28	50	10	1.26	0.02	0.06
70	연구개발업	8	3	1.77	12	6	2.07	0.09	0.09
71	전문 서비스업	30	-6	0.83	42	-7	0.86	0.04	0.05
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	17	0	1.01	21	-1	0.95	0.02	0.05

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	3	0	0.94	4	0	1.03	0.1	0.33
74	사업시설 관리 및 조경 서비스업	80	48	2.47	102	59	2.38	0.03	0.03
75	사업 지원 서비스업	71	17	1.32	91	19	1.26	0.02	0.06
76	임대업; 부동산 제외	0	-1	0	0	-1	0		
84	공공 행정, 국방 및 사회보장 행정	13	0	0.97	16	-2	0.91	0.03	0.03
85	교육 서비스업	50	-2	0.96	71	1	1.02	0.02	0.07
86	보건업	109	-20	0.84	149	-23	0.87	0.03	0.03
87	사회복지 서비스업	30	0	1	42	2	1.04	0.03	0.04
90	창작, 예술 및 여가관련 서비스업	3	0	0.96	4	0	1.08	0.09	0.1
91	스포츠 및 오락관련 서비스업	14	4	1.36	16	3	1.19	0.03	0.04
94	협회 및 단체	8	-7	0.53	12	-8	0.59	0.08	0.19
95	개인 및 소비용품 수리업	3	-1	0.75	5	-1	0.86	0.1	0.14
96	기타 개인 서비스업	13	3	1.34	17	4	1.29	0.03	0.06

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-12〉 중분류별 백혈병 발병자수 추계 결과, 2000년도 이후 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	0	-2	0	0	-3	0		
02	임업	0	0	0	0	-1	0		
03	어업	0	0	0	0	-1	0		
05	석탄, 원유 및 천연가스 광업	0	0	0	0	0	0		
06	금속 광업	0	0	0	0	0	0		
07	비금속광물 광업; 연료용 제외	0	0	0	0	-1	0		
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	46	4	1.09	70	3	1.05	0.09	0.38
11	음료 제조업	3	1	1.27	4	1	1.25	0.25	0.57
12	담배 제조업	0	0	0	0	0	0		
13	섬유제품 제조업; 의복 제외	28	5	1.23	42	8	1.23	0.05	0.13
14	의복, 의복 액세서리 및 모피제품 제조업	34	10	1.42	55	19	1.52	0.06	0.08
15	가죽, 가방 및 신발 제조업	9	3	1.63	15	7	1.82	0.13	0.13

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	3	0	1	6	1	1.26	0.22	0.32
17	펄프, 종이 및 종이제품 제조업	5	0	0.96	7	-1	0.88	0.16	0.66
18	인쇄 및 기록매체 복제업	8	-2	0.83	12	-3	0.79	0.08	0.14
19	코크스, 연탄 및 석유정제품 제조업	1	0	1.31	2	1	1.57	0.29	0.99
20	화학 물질 및 화학제품 제조업; 의약품 제외	13	-6	0.67	19	-11	0.63	0.07	0.09
21	의료용 물질 및 의약품 제조업	2	-6	0.25	3	-9	0.23	0.22	0.32
22	고무 및 플라스틱제품 제조업	36	12	1.51	59	22	1.6	0.09	0.2
23	비금속 광물제품 제조업	6	-2	0.78	9	-2	0.82	0.07	0.13
24	1차 금속 제조업	3	-2	0.64	4	-3	0.6	0.08	0.16
25	금속 가공제품 제조업; 기계 및 가구 제외	16	-8	0.66	25	-13	0.65	0.1	0.27
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	81	-10	0.89	133	-1	0.99	0.02	0.09
27	의료, 정밀, 광학 기기 및 시계 제조업	8	-4	0.67	11	-7	0.62	0.15	0.31
28	전기장비 제조업	36	10	1.38	64	24	1.61	0.07	0.08
29	기타 기계 및 장비 제조업	28	1	1.04	42	1	1.02	0.05	0.28
30	자동차 및 트레일러 제조업	31	7	1.27	51	14	1.36	0.05	0.07
31	기타 운송장비 제조업	7	0	1.04	10	-1	0.95	0.09	0.12

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	2	-1	0.68	3	-2	0.65	0.22	0.31
33	기타 제품 제조업	9	1	1.06	12	-1	0.91	0.16	0.18
34	산업용 기계 및 장비 수리업	1	0	1.42	1	0	1.37	0.56	0.96
35	전기, 가스, 증기 및 공기 조절 공급업	1	-2	0.34	1	-3	0.32	0.57	0.92
36	수도업	0	0	0	0	0	0		
37	하수, 폐수 및 분뇨 처리업	0	-1	0	0	-1	0		
38	폐기물 수집, 운반, 처리 및 원료 재생업	3	0	0.93	5	0	0.94	0.11	0.44
39	환경 정화 및 복원업	1	1	11.51	2	2	13.99	0.17	0.38
41	종합 건설업	58	6	1.11	100	20	1.25	0.05	0.08
42	전문직별 공사업	1	0	1.11	2	0	1.13	0.39	0.71
45	자동차 및 부품 판매업	4	1	1.15	7	2	1.42	0.2	0.36
46	도매 및 상품 중개업	118	-13	0.9	185	-14	0.93	0.05	0.09
47	소매업; 자동차 제외	83	-2	0.98	117	-9	0.93	0.11	0.36
49	육상 운송 및 파이프라인 운송업	17	8	1.93	25	12	1.87	0.05	0.16
50	수상 운송업	1	0	1	1	0	0.74	0.05	0.11
51	항공 운송업	3	0	0.9	4	-1	0.81	0.28	1

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	23	-2	0.94	35	-3	0.91	0.11	0.16
55	숙박업	9	-1	0.87	16	1	1.04	0.15	0.25
56	음식점 및 주점업	64	1	1.01	90	-2	0.97	0.15	0.31
58	출판업	28	-16	0.64	38	-29	0.56	0.08	0.1
59	영상·오디오 기록물 제작 및 배급업	0	-4	0	0	-6	0		
60	방송업	3	1	1.26	4	0	1.01	0.42	0.42
61	우편 및 통신업	5	0	1.07	9	2	1.28	0.16	0.18
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	10	5	1.91	14	6	1.73	0.06	0.07
63	정보서비스업	10	-3	0.75	15	-6	0.72	0.05	0.11
64	금융업	44	0	0.99	66	-1	0.99	0.06	0.17
65	보험 및 연금업	5	-5	0.5	9	-6	0.59	0.07	0.13
66	금융 및 보험관련 서비스업	12	1	1.13	17	0	1.02	0.19	0.41
68	부동산업	62	27	1.75	101	46	1.84	0.02	0.02
70	연구개발업	3	-6	0.34	4	-9	0.33	0.12	0.17
71	전문 서비스업	46	-7	0.87	71	-10	0.88	0.04	0.13
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	25	1	1.04	41	4	1.1	0.05	0.06

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	8	1	1.15	11	1	1.1	0.23	0.41
74	사업시설 관리 및 조경 서비스업	67	25	1.61	108	44	1.68	0.04	0.05
75	사업 지원 서비스업	139	-11	0.93	208	-24	0.9	0.07	0.19
76	임대업; 부동산 제외	2	0	1.2	3	1	1.24	0.18	0.28
84	공공 행정, 국방 및 사회보장 행정	26	4	1.21	37	0	1.01	0.11	0.3
85	교육 서비스업	96	-2	0.98	136	-14	0.9	0.11	0.32
86	보건업	153	-34	0.82	232	-53	0.81	0.05	0.05
87	사회복지 서비스업	154	14	1.1	213	2	1.01	0.17	0.59
90	창작, 예술 및 여가관련 서비스업	5	0	1.09	8	1	1.1	0.08	0.3
91	스포츠 및 오락관련 서비스업	6	-5	0.52	9	-8	0.53	0.15	0.4
94	협회 및 단체	21	6	1.38	33	10	1.42	0.06	0.25
95	개인 및 소비용품 수리업	8	4	1.83	14	8	2.16	0.1	0.17
96	기타 개인 서비스업	17	0	1.02	25	0	1	0.05	0.05

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율



〈표 부-13〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적함오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	214	27	1.14	294	35	1.13	0.03	0.16
B	광업(05~08)	407	260	2.77	550	349	2.73	0.02	0.07
C	제조업(10~34)	16999	-3251	0.84	23616	-4555	0.84	0.02	0.06
D	전기, 가스, 증기 및 공기 조절 공급업(35)	631	-191	0.77	928	-206	0.82	0.01	0.01
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	561	330	2.43	763	443	2.38	0.02	0.02
F	건설업(41~42)	3050	-50	0.98	4274	-59	0.99	0.02	0.06
G	도매 및 소매업(45~47)	2539	-1507	0.63	3544	-2132	0.62	0.03	0.1
H	운수 및 창고업(49~52)	6267	2249	1.56	8727	3176	1.57	0.02	0.05
I	숙박 및 음식점업(55~56)	764	4	1.01	1094	29	1.03	0.01	0.07
J	정보통신업(58~63)	1818	-1074	0.63	2612	-1414	0.65	0.01	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	2521	-976	0.72	3538	-1309	0.73	0.02	0.08
L	부동산업(68)	4087	2771	3.11	5537	3716	3.04	0.01	0.01
M	전문, 과학 및 기술 서비스업(70~73)	1537	-604	0.72	2161	-829	0.72	0.02	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	3656	1926	2.11	5084	2675	2.11	0.01	0.01
O	공공 행정, 국방 및 사회보장 행정(84)	889	430	1.94	1248	605	1.94	0.01	0.01
P	교육 서비스업(85)	537	33	1.06	748	37	1.05	0.02	0.11
Q	보건업 및 사회복지 서비스업(86~87)	881	-285	0.76	1236	-399	0.76	0.02	0.1
R	예술, 스포츠 및 여가관련 서비스업(90~91)	289	-28	0.91	381	-60	0.86	0.04	0.17

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-14〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적함오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	142	74	2.08	212	110	2.09	0.08	0.34
B	광업(05~08)	63	35	2.21	92	52	2.31	0.09	0.09
C	제조업(10~34)	6829	-1513	0.82	9986	-2361	0.81	0.08	0.27
D	전기, 가스, 증기 및 공기 조절 공급업(35)	107	-38	0.74	181	-24	0.88	0.02	0.05
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	328	183	2.26	466	251	2.17	0.1	0.26
F	건설업(41~42)	2270	238	1.12	3263	270	1.09	0.08	0.26
G	도매 및 소매업(45~47)	1917	-998	0.66	2720	-1549	0.64	0.11	0.33
H	운수 및 창고업(49~52)	2745	1354	1.97	3958	1931	1.95	0.09	0.28
I	숙박 및 음식점업(55~56)	606	-8	0.99	865	-2	1	0.09	0.29
J	정보통신업(58~63)	513	-896	0.36	753	-1339	0.36	0.08	0.27

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	299	-385	0.44	428	-579	0.43	0.1	0.26
L	부동산업(68)	2173	1489	3.18	3539	2539	3.54	0.02	0.07
M	전문, 과학 및 기술 서비스업(70~73)	860	-629	0.58	1306	-907	0.59	0.06	0.22
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	3503	1574	1.82	5123	2261	1.79	0.08	0.26
O	공공 행정, 국방 및 사회보장 행정(84)	396	63	1.19	697	217	1.45	0.04	0.09
P	교육 서비스업(85)	545	-27	0.95	789	-39	0.95	0.1	0.34
Q	보건업 및 사회복지 서비스업(86~87)	1114	-394	0.74	1520	-631	0.71	0.19	0.52
R	예술, 스포츠 및 여가관련 서비스업(90~91)	164	-50	0.77	254	-62	0.8	0.06	0.23

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-15〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	198	22	1.13	273	29	1.12	0.03	0.14
B	광업(05~08)	402	259	2.82	542	346	2.77	0.02	0.07
C	제조업(10~34)	14415	-3604	0.8	20029	-5012	0.8	0.02	0.06
D	전기, 가스, 증기 및 공기 조절 공급업(35)	613	-186	0.77	900	-202	0.82	0.01	0.01
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	550	330	2.5	749	444	2.45	0.02	0.02
F	건설업(41~42)	2957	12	1	4153	40	1.01	0.02	0.06
G	도매 및 소매업(45~47)	2155	-1283	0.63	2963	-1853	0.62	0.03	0.11
H	운수 및 창고업(49~52)	6157	2290	1.59	8573	3233	1.61	0.02	0.05
I	숙박 및 음식점업(55~56)	386	-168	0.7	543	-230	0.7	0.01	0.07
J	정보통신업(58~63)	1631	-999	0.62	2356	-1302	0.64	0.01	0.04

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	2175	-774	0.74	3082	-998	0.76	0.02	0.07
L	부동산업(68)	3913	2701	3.23	5290	3617	3.16	0.01	0.01
M	전문, 과학 및 기술 서비스업(70~73)	1428	-501	0.74	2009	-680	0.75	0.02	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	3010	1586	2.11	4144	2165	2.09	0.01	0.01
O	공공 행정, 국방 및 사회보장 행정(84)	844	432	2.05	1185	610	2.06	0.01	0.01
P	교육 서비스업(85)	385	64	1.2	534	84	1.19	0.02	0.07
Q	보건업 및 사회복지 서비스업(86~87)	506	-101	0.83	699	-146	0.83	0.02	0.11
R	예술, 스포츠 및 여가관련 서비스업(90~91)	260	-10	0.96	345	-30	0.92	0.04	0.16

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-16〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	126	68	2.16	190	102	2.17	0.07	0.33
B	광업(05~08)	61	34	2.27	89	51	2.37	0.08	0.09
C	제조업(10~34)	5342	-1686	0.76	7882	-2621	0.75	0.07	0.25
D	전기, 가스, 증기 및 공기 조절 공급업(35)	99	-36	0.73	167	-24	0.88	0.02	0.02
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	303	171	2.29	428	231	2.18	0.1	0.28
F	건설업(41~42)	2136	278	1.15	3055	307	1.11	0.08	0.28
G	도매 및 소매업(45~47)	1333	-867	0.61	1913	-1357	0.59	0.09	0.3
H	운수 및 창고업(49~52)	2627	1358	2.07	3782	1931	2.04	0.09	0.28
I	숙박 및 음식점업(55~56)	249	-125	0.67	359	-184	0.66	0.09	0.31
J	정보통신업(58~63)	399	-771	0.34	587	-1167	0.33	0.08	0.27

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	202	-271	0.43	287	-421	0.41	0.11	0.28
L	부동산업(68)	1939	1370	3.41	3215	2379	3.85	0.01	0.03
M	전문, 과학 및 기술 서비스업(70~73)	693	-493	0.58	1057	-730	0.59	0.05	0.19
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	2506	1207	1.93	3735	1776	1.91	0.07	0.25
O	공공 행정, 국방 및 사회보장 행정(84)	334	71	1.27	603	234	1.63	0.03	0.07
P	교육 서비스업(85)	236	-16	0.94	360	-15	0.96	0.06	0.25
Q	보건업 및 사회복지 서비스업(86~87)	268	-174	0.61	366	-294	0.55	0.16	0.48
R	예술, 스포츠 및 여가관련 서비스업(90~91)	124	-38	0.77	193	-50	0.79	0.07	0.24

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-17〉 대분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
A	농업, 임업 및 어업(01~03)	16	5	1.43	22	6	1.37	0.15	0.44
B	광업(05~08)	5	1	1.16	8	2	1.35	0.13	0.31
C	제조업(10~34)	2584	353	1.16	3586	456	1.15	0.02	0.08
D	전기, 가스, 증기 및 공기 조절 공급업(35)	18	-5	0.78	28	-4	0.86	0.08	0.11
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	11	1	1.06	14	-1	0.92	0.06	0.08
F	건설업(41~42)	93	-62	0.6	121	-99	0.55	0.06	0.06
G	도매 및 소매업(45~47)	384	-224	0.63	581	-278	0.68	0.01	0.01
H	운수 및 창고업(49~52)	110	-40	0.73	154	-57	0.73	0.04	0.18
I	숙박 및 음식점업(55~56)	378	172	1.83	550	259	1.89	0.02	0.07
J	정보통신업(58~63)	187	-75	0.71	256	-112	0.7	0.02	0.02

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	346	-202	0.63	456	-311	0.59	0.05	0.16
L	부동산업(68)	174	70	1.67	247	99	1.68	0.02	0.05
M	전문, 과학 및 기술 서비스업(70~73)	109	-104	0.51	152	-149	0.51	0.03	0.03
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	646	341	2.12	940	510	2.18	0	0.01
O	공공 행정, 국방 및 사회보장 행정(84)	45	-2	0.95	63	-5	0.93	0.04	0.1
P	교육 서비스업(85)	152	-31	0.83	215	-47	0.82	0.04	0.2
Q	보건업 및 사회복지 서비스업(86~87)	375	-184	0.67	536	-253	0.68	0.03	0.09
R	예술, 스포츠 및 여가관련 서비스업(90~91)	29	-18	0.62	36	-30	0.54	0.12	0.29

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-18〉 대분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
A	농업, 임업 및 어업(01~03)	16	6	1.61	22	8	1.56	0.28	0.49
B	광업(05~08)	2	0	1.26	3	1	1.32	0.25	0.32
C	제조업(10~34)	1487	174	1.13	2104	260	1.14	0.12	0.34
D	전기, 가스, 증기 및 공기 조절 공급업(35)	8	-2	0.77	14	-1	0.96	0.17	0.65
E	수도, 하수 및 폐기물 처리, 원료 재생업(36~39)	25	12	1.96	38	20	2.11	0.07	0.1
F	건설업(41~42)	134	-40	0.77	208	-37	0.85	0.06	0.1
G	도매 및 소매업(45~47)	584	-132	0.82	806	-191	0.81	0.15	0.41
H	운수 및 창고업(49~52)	118	-5	0.96	175	0	1	0.08	0.29
I	숙박 및 음식점업(55~56)	357	117	1.49	506	182	1.56	0.1	0.28
J	정보통신업(58~63)	114	-125	0.48	166	-173	0.49	0.08	0.25

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
K	금융 및 보험업(64~66)	97	-115	0.46	141	-158	0.47	0.09	0.22
L	부동산업(68)	234	119	2.03	324	160	1.97	0.13	0.44
M	전문, 과학 및 기술 서비스업(70~73)	167	-136	0.55	249	-177	0.58	0.09	0.35
N	사업시설 관리, 사업 지원 및 임대 서비스업(74~76)	997	367	1.58	1388	485	1.54	0.13	0.31
O	공공 행정, 국방 및 사회보장 행정(84)	62	-8	0.88	94	-17	0.85	0.11	0.17
P	교육 서비스업(85)	309	-12	0.96	430	-25	0.95	0.14	0.42
Q	보건업 및 사회복지 서비스업(86~87)	846	-220	0.79	1154	-338	0.77	0.2	0.53
R	예술, 스포츠 및 여가관련 서비스업(90~91)	40	-12	0.76	61	-12	0.84	0.06	0.18

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) / (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015~2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015~2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-19〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	143	9	1.07	201	16	1.09	0.03	0.15
02	임업	9	-3	0.76	12	-5	0.71	0.32	0.78
03	어업	62	21	1.5	81	24	1.42	0.04	0.11
05	석탄, 원유 및 천연가스 광업	250	168	3.04	332	219	2.96	0.03	0.08
06	금속 광업	24	20	5.47	31	25	5.19	0.03	0.03
07	비금속광물 광업; 연료용 제외	133	73	2.21	188	104	2.25	0.02	0.08
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	1202	95	1.09	1643	104	1.07	0.02	0.11
11	음료 제조업	250	-34	0.88	346	-48	0.88	0.01	0.05
12	담배 제조업	128	71	2.25	173	95	2.21	0.04	0.06
13	섬유제품 제조업; 의복 제외	1334	176	1.15	1907	294	1.18	0	0.02
14	의복, 의복 액세서리 및 모피제품 제조업	486	-75	0.87	653	-130	0.83	0.03	0.07
15	가죽, 가방 및 신발 제조업	380	79	1.26	513	96	1.23	0.02	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	302	116	1.63	416	157	1.61	0.03	0.07
17	펄프, 종이 및 종이제품 제조업	446	21	1.05	589	-1	1	0.03	0.06
18	인쇄 및 기록매체 복제업	255	-53	0.83	354	-77	0.82	0.02	0.12
19	코크스, 연탄 및 석유정제품 제조업	113	-15	0.88	159	-17	0.9	0.05	0.05
20	화학 물질 및 화학제품 제조업; 의약품 제외	1068	-180	0.86	1459	-274	0.84	0.03	0.05
21	의료용 물질 및 의약품 제조업	195	-131	0.6	277	-177	0.61	0.02	0.06
22	고무 및 플라스틱제품 제조업	739	-100	0.88	1039	-131	0.89	0.01	0.04
23	비금속 광물제품 제조업	1191	332	1.39	1680	491	1.41	0.01	0.05
24	1차 금속 제조업	1153	114	1.11	1657	221	1.15	0.01	0.03
25	금속 가공제품 제조업; 기계 및 가구 제외	1256	86	1.07	1732	100	1.06	0.02	0.07
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	1316	-1740	0.43	1867	-2397	0.44	0.03	0.07
27	의료, 정밀, 광학 기기 및 시계 제조업	238	-141	0.63	322	-209	0.61	0.05	0.05
28	전기장비 제조업	784	-422	0.65	1083	-596	0.64	0.02	0.06
29	기타 기계 및 장비 제조업	1290	-526	0.71	1839	-695	0.73	0.01	0.04
30	자동차 및 트레일러 제조업	1514	-938	0.62	2026	-1377	0.6	0.04	0.13
31	기타 운송장비 제조업	824	22	1.03	1151	41	1.04	0.02	0.09

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	192	-9	0.96	262	-17	0.94	0.03	0.03
33	기타 제품 제조업	319	28	1.1	432	25	1.06	0.03	0.08
34	산업용 기계 및 장비 수리업	24	-26	0.48	38	-33	0.54	0.02	0.04
35	전기, 가스, 증기 및 공기 조절 공급업	596	-195	0.75	878	-215	0.8	0.01	0.01
36	수도업	35	5	1.16	50	8	1.2	0.08	0.31
37	하수, 폐수 및 분뇨 처리업	86	46	2.13	116	59	2.05	0.03	0.03
38	폐기물 수집, 운반, 처리 및 원료 재생업	466	278	2.48	637	376	2.44	0.02	0.02
39	환경 정화 및 복원업	9	7	4.03	11	8	3.41	0.13	0.17
41	종합 건설업	2838	-73	0.97	3985	-84	0.98	0.02	0.07
42	전문직별 공사업	212	23	1.12	289	25	1.09	0.02	0.03
45	자동차 및 부품 판매업	122	-195	0.39	160	-280	0.36	0.08	0.27
46	도매 및 상품 중개업	1746	-884	0.66	2447	-1246	0.66	0.03	0.09
47	소매업; 자동차 제외	671	-428	0.61	937	-606	0.61	0.03	0.07
49	육상 운송 및 파이프라인 운송업	4942	2049	1.71	6893	2903	1.73	0.02	0.06
50	수상 운송업	205	60	1.41	282	80	1.4	0.02	0.05
51	항공 운송업	97	-54	0.64	133	-76	0.64	0.03	0.04

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	1023	194	1.23	1420	268	1.23	0.01	0.01
55	숙박업	359	-20	0.95	510	-17	0.97	0.02	0.08
56	음식점 및 주점업	405	24	1.06	583	46	1.09	0.02	0.06
58	출판업	480	-617	0.44	708	-832	0.46	0.01	0.04
59	영상·오디오 기록물 제작 및 배급업	49	-15	0.77	70	-20	0.78	0.02	0.05
60	방송업	168	-18	0.9	225	-33	0.87	0.03	0.1
61	우편 및 통신업	1021	-337	0.75	1462	-413	0.78	0.02	0.03
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	19	-45	0.3	30	-61	0.33	0.05	0.06
63	정보서비스업	81	-42	0.66	117	-56	0.68	0.03	0.05
64	금융업	1905	-544	0.78	2696	-696	0.79	0.01	0.07
65	보험 및 연금업	424	-280	0.6	581	-394	0.6	0.05	0.12
66	금융 및 보험관련 서비스업	192	-152	0.56	261	-218	0.55	0.03	0.13
68	부동산업	4087	2771	3.11	5537	3716	3.04	0.01	0.01
70	연구개발업	151	-93	0.62	217	-122	0.64	0.01	0.04
71	전문 서비스업	614	-192	0.76	880	-246	0.78	0.01	0.02
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	741	-290	0.72	1023	-417	0.71	0.03	0.05

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	31	-30	0.51	41	-44	0.48	0.03	0.06
74	사업시설 관리 및 조경 서비스업	1660	1085	2.89	2276	1477	2.85	0	0
75	사업 지원 서비스업	1940	835	1.76	2729	1188	1.77	0.01	0.01
76	임대업; 부동산 제외	56	6	1.12	79	9	1.14	0.02	0.02
84	공공 행정, 국방 및 사회보장 행정	889	430	1.94	1248	605	1.94	0.01	0.01
85	교육 서비스업	537	33	1.06	748	37	1.05	0.02	0.11
86	보건업	709	-259	0.73	1008	-347	0.74	0.02	0.08
87	사회복지 서비스업	172	-26	0.87	227	-52	0.81	0.04	0.16
90	창작, 예술 및 여가관련 서비스업	52	-14	0.79	63	-29	0.69	0.09	0.19
91	스포츠 및 오락관련 서비스업	237	-14	0.94	318	-32	0.91	0.04	0.17
94	협회 및 단체	367	61	1.2	515	88	1.21	0.01	0.01
95	개인 및 소비용품 수리업	233	-161	0.59	314	-238	0.57	0.05	0.15
96	기타 개인 서비스업	162	35	1.27	228	49	1.27	0.03	0.11

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-20〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남/여 합계

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	56	18	1.48	81	24	1.43	0.11	0.43
02	임업	28	16	2.33	37	18	2	0.21	0.65
03	어업	58	39	3.12	95	68	3.54	0.03	0.14
05	석탄, 원유 및 천연가스 광업	7	2	1.43	11	4	1.66	0.16	0.52
06	금속 광업	5	4	5.23	6	5	4.49	0.54	0.58
07	비금속광물 광업; 연료용 제외	51	29	2.28	75	43	2.37	0.08	0.11
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	491	35	1.08	704	34	1.05	0.1	0.36
11	음료 제조업	26	-27	0.49	37	-40	0.49	0.13	0.4
12	담배 제조업	1	-2	0.3	1	-3	0.31	0.41	0.95
13	섬유제품 제조업; 의복 제외	379	88	1.3	579	163	1.39	0.04	0.16
14	의복, 의복 액세서리 및 모피제품 제조업	192	27	1.16	286	50	1.21	0.06	0.19
15	가죽, 가방 및 신발 제조업	93	25	1.36	149	53	1.55	0.03	0.04

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	172	88	2.05	249	125	2.02	0.08	0.24
17	펄프, 종이 및 종이제품 제조업	137	10	1.08	207	20	1.11	0.07	0.26
18	인쇄 및 기록매체 복제업	127	-52	0.71	176	-86	0.67	0.14	0.37
19	코크스, 연탄 및 석유정제품 제조업	21	-3	0.86	30	-8	0.8	0.17	0.64
20	화학 물질 및 화학제품 제조업; 의약품 제외	359	-126	0.74	564	-150	0.79	0.04	0.13
21	의료용 물질 및 의약품 제조업	67	-89	0.43	95	-135	0.41	0.15	0.29
22	고무 및 플라스틱제품 제조업	443	-25	0.95	638	-54	0.92	0.08	0.28
23	비금속 광물제품 제조업	371	120	1.48	550	183	1.5	0.07	0.25
24	1차 금속 제조업	308	79	1.35	433	89	1.26	0.1	0.29
25	금속 가공제품 제조업; 기계 및 가구 제외	770	60	1.08	1103	38	1.04	0.08	0.29
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	580	-810	0.42	855	-1191	0.42	0.1	0.27
27	의료, 정밀, 광학 기기 및 시계 제조업	120	-118	0.51	171	-183	0.48	0.12	0.3
28	전기장비 제조업	327	-197	0.62	460	-314	0.59	0.11	0.33
29	기타 기계 및 장비 제조업	786	-253	0.76	1091	-475	0.7	0.12	0.37
30	자동차 및 트레일러 제조업	490	-277	0.64	749	-378	0.66	0.06	0.19
31	기타 운송장비 제조업	307	-57	0.84	476	-83	0.85	0.05	0.23

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	87	4	1.05	133	12	1.1	0.06	0.1
33	기타 제품 제조업	135	-5	0.96	190	-16	0.92	0.1	0.33
34	산업용 기계 및 장비 수리업	40	-8	0.84	59	-12	0.83	0.07	0.29
35	전기, 가스, 증기 및 공기 조절 공급업	102	-36	0.74	173	-21	0.89	0.02	0.03
36	수도업	5	-3	0.65	8	-4	0.7	0.2	0.65
37	하수, 폐수 및 분뇨 처리업	54	28	2.12	76	38	2.01	0.12	0.35
38	폐기물 수집, 운반, 처리 및 원료 재생업	272	155	2.32	387	214	2.23	0.09	0.24
39	환경 정화 및 복원업	2	0	0.92	3	-1	0.81	0.43	0.94
41	종합 건설업	2190	216	1.11	3140	231	1.08	0.08	0.28
42	전문직별 공사업	80	22	1.38	123	39	1.46	0.04	0.04
45	자동차 및 부품 판매업	43	-59	0.42	58	-90	0.39	0.19	0.57
46	도매 및 상품 중개업	1299	-610	0.68	1854	-959	0.66	0.1	0.32
47	소매업; 자동차 제외	575	-330	0.64	808	-499	0.62	0.12	0.35
49	육상 운송 및 파이프라인 운송업	2138	1284	2.5	3043	1818	2.49	0.09	0.3
50	수상 운송업	73	32	1.76	114	53	1.86	0.04	0.15
51	항공 운송업	5	-24	0.18	6	-37	0.14	0.34	0.36

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	529	61	1.13	795	97	1.14	0.07	0.23
55	숙박업	115	-16	0.88	175	-15	0.92	0.05	0.08
56	음식점 및 주점업	491	8	1.02	689	12	1.02	0.11	0.34
58	출판업	272	-639	0.3	390	-969	0.29	0.09	0.34
59	영상·오디오 기록물 제작 및 배급업	27	-34	0.44	41	-49	0.45	0.03	0.05
60	방송업	25	-22	0.53	37	-30	0.55	0.24	0.61
61	우편 및 통신업	46	-80	0.37	67	-119	0.36	0.04	0.07
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	44	-82	0.35	64	-126	0.34	0.1	0.33
63	정보서비스업	99	-38	0.72	154	-47	0.77	0.06	0.19
64	금융업	200	-246	0.45	282	-369	0.43	0.12	0.29
65	보험 및 연금업	45	-55	0.45	70	-77	0.47	0.05	0.05
66	금융 및 보험관련 서비스업	54	-84	0.39	76	-132	0.36	0.11	0.4
68	부동산업	2173	1489	3.18	3539	2539	3.54	0.02	0.07
70	연구개발업	49	-107	0.31	73	-160	0.31	0.08	0.41
71	전문 서비스업	315	-232	0.58	491	-317	0.61	0.05	0.14
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	459	-246	0.65	688	-368	0.65	0.07	0.25

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	37	-43	0.46	54	-63	0.46	0.09	0.3
74	사업시설 관리 및 조경 서비스업	1536	1016	2.96	2299	1526	2.97	0.07	0.21
75	사업 지원 서비스업	1943	575	1.42	2787	757	1.37	0.09	0.31
76	임대업; 부동산 제외	24	-17	0.59	38	-22	0.63	0.06	0.08
84	공공 행정, 국방 및 사회보장 행정	396	63	1.19	697	217	1.45	0.04	0.09
85	교육 서비스업	545	-27	0.95	789	-39	0.95	0.1	0.34
86	보건업	616	-313	0.66	845	-491	0.63	0.16	0.43
87	사회복지 서비스업	498	-81	0.86	675	-141	0.83	0.24	0.63
90	창작, 예술 및 여가관련 서비스업	30	-21	0.58	48	-27	0.64	0.05	0.13
91	스포츠 및 오락관련 서비스업	134	-29	0.82	206	-35	0.86	0.06	0.25
94	협회 및 단체	161	8	1.05	225	-1	0.99	0.12	0.37
95	개인 및 소비용품 수리업	159	-118	0.57	228	-185	0.55	0.1	0.28
96	기타 개인 서비스업	168	39	1.31	235	48	1.26	0.11	0.31

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-21〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 남성

0단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	130	4	1.03	185	11	1.06	0.03	0.12
02	임업	9	-2	0.85	12	-3	0.8	0.32	0.78
03	어업	59	20	1.5	76	21	1.39	0.04	0.12
05	석탄, 원유 및 천연가스 광업	248	168	3.08	328	218	2.99	0.03	0.08
06	금속 광업	23	19	5.54	30	24	5.26	0.01	0.02
07	비금속광물 광업; 연료용 제외	131	73	2.26	184	104	2.3	0.02	0.08
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	895	-9	0.99	1206	-46	0.96	0.03	0.12
11	음료 제조업	231	-34	0.87	322	-46	0.88	0.01	0.07
12	담배 제조업	121	67	2.25	163	90	2.21	0.04	0.07
13	섬유제품 제조업; 의복 제외	1035	123	1.13	1492	224	1.18	0.01	0.01
14	의복, 의복 액세서리 및 모피제품 제조업	228	-122	0.65	312	-175	0.64	0.02	0.02
15	가죽, 가방 및 신발 제조업	257	16	1.07	354	20	1.06	0.01	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	280	110	1.65	387	151	1.64	0.02	0.05
17	펄프, 종이 및 종이제품 제조업	408	19	1.05	541	1	1	0.03	0.05
18	인쇄 및 기록매체 복제업	218	-56	0.8	306	-78	0.8	0.02	0.12
19	코크스, 연탄 및 석유정제품 제조업	111	-13	0.89	157	-14	0.92	0.05	0.05
20	화학 물질 및 화학제품 제조업; 의약품 제외	979	-170	0.85	1331	-262	0.84	0.03	0.06
21	의료용 물질 및 의약품 제조업	164	-121	0.58	239	-156	0.61	0.01	0.03
22	고무 및 플라스틱제품 제조업	624	-134	0.82	878	-177	0.83	0.01	0.04
23	비금속 광물제품 제조업	1097	295	1.37	1542	433	1.39	0.01	0.05
24	1차 금속 제조업	1100	93	1.09	1588	196	1.14	0.01	0.02
25	금속 가공제품 제조업; 기계 및 가구 제외	1117	31	1.03	1543	29	1.02	0.02	0.06
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	1018	-1524	0.4	1428	-2113	0.4	0.03	0.08
27	의료, 정밀, 광학 기기 및 시계 제조업	164	-167	0.49	227	-236	0.49	0.06	0.08
28	전기장비 제조업	626	-447	0.58	846	-647	0.57	0.03	0.08
29	기타 기계 및 장비 제조업	1193	-535	0.69	1706	-705	0.71	0.01	0.04
30	자동차 및 트레일러 제조업	1344	-983	0.58	1796	-1430	0.56	0.04	0.15
31	기타 운송장비 제조업	794	17	1.02	1104	28	1.03	0.02	0.1

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	146	-33	0.82	199	-50	0.8	0.02	0.02
33	기타 제품 제조업	241	-2	0.99	322	-17	0.95	0.03	0.08
34	산업용 기계 및 장비 수리업	24	-24	0.5	38	-30	0.56	0.02	0.04
35	전기, 가스, 증기 및 공기 조절 공급업	580	-190	0.75	854	-208	0.8	0.01	0.01
36	수도업	33	4	1.14	46	6	1.16	0.07	0.26
37	하수, 폐수 및 분뇨 처리업	84	46	2.19	113	60	2.12	0.02	0.02
38	폐기물 수집, 운반, 처리 및 원료 재생업	457	277	2.54	625	376	2.51	0.02	0.02
39	환경 정화 및 복원업	9	7	4.18	11	8	3.53	0.13	0.17
41	종합 건설업	2749	-16	0.99	3870	8	1	0.02	0.07
42	전문직별 공사업	208	29	1.16	283	32	1.13	0.02	0.03
45	자동차 및 부품 판매업	117	-185	0.39	152	-268	0.36	0.08	0.28
46	도매 및 상품 중개업	1513	-752	0.67	2079	-1098	0.65	0.03	0.1
47	소매업; 자동차 제외	525	-345	0.6	733	-487	0.6	0.03	0.09
49	육상 운송 및 파이프라인 운송업	4894	2055	1.72	6823	2908	1.74	0.02	0.06
50	수상 운송업	202	64	1.47	278	88	1.46	0.02	0.05
51	항공 운송업	88	-39	0.69	120	-56	0.68	0.04	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	973	209	1.27	1353	293	1.28	0.01	0.02
55	숙박업	245	-61	0.8	342	-83	0.8	0.02	0.11
56	음식점 및 주점업	141	-106	0.57	201	-147	0.58	0.02	0.03
58	출판업	432	-547	0.44	644	-729	0.47	0.01	0.03
59	영상·오디오 기록물 제작 및 배급업	45	-12	0.8	63	-16	0.8	0.02	0.02
60	방송업	160	-14	0.92	215	-26	0.89	0.03	0.12
61	우편 및 통신업	906	-350	0.72	1303	-430	0.75	0.02	0.04
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	18	-40	0.31	29	-53	0.35	0.06	0.07
63	정보서비스업	70	-36	0.66	101	-48	0.68	0.04	0.06
64	금융업	1656	-420	0.8	2365	-505	0.82	0.01	0.06
65	보험 및 연금업	354	-226	0.61	490	-313	0.61	0.04	0.11
66	금융 및 보험관련 서비스업	165	-128	0.56	228	-180	0.56	0.02	0.08
68	부동산업	3913	2701	3.23	5290	3617	3.16	0.01	0.01
70	연구개발업	144	-84	0.63	209	-108	0.66	0.01	0.03
71	전문 서비스업	549	-130	0.81	786	-161	0.83	0.01	0.01
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	705	-267	0.73	974	-381	0.72	0.03	0.06

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	30	-20	0.6	40	-30	0.57	0.04	0.06
74	사업시설 관리 및 조경 서비스업	1255	794	2.72	1688	1048	2.64	0.01	0.01
75	사업 지원 서비스업	1699	783	1.85	2377	1102	1.87	0.01	0.01
76	임대업; 부동산 제외	56	9	1.2	79	14	1.21	0.02	0.02
84	공공 행정, 국방 및 사회보장 행정	844	432	2.05	1185	610	2.06	0.01	0.01
85	교육 서비스업	385	64	1.2	534	84	1.19	0.02	0.07
86	보건업	415	-99	0.81	578	-138	0.81	0.03	0.11
87	사회복지 서비스업	91	-2	0.98	122	-8	0.94	0.03	0.12
90	창작, 예술 및 여가관련 서비스업	47	-8	0.85	57	-20	0.74	0.09	0.2
91	스포츠 및 오락관련 서비스업	213	-2	0.99	289	-10	0.97	0.03	0.15
94	협회 및 단체	309	56	1.22	430	79	1.22	0.01	0.01
95	개인 및 소비용품 수리업	218	-162	0.57	293	-239	0.55	0.05	0.16
96	기타 개인 서비스업	127	34	1.36	177	46	1.35	0.04	0.12

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-22〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 남성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	41	10	1.34	60	14	1.31	0.08	0.4
02	임업	28	17	2.61	37	20	2.22	0.21	0.66
03	어업	57	40	3.31	93	68	3.75	0.03	0.15
05	석탄, 원유 및 천연가스 광업	7	2	1.51	11	5	1.74	0.16	0.53
06	금속 광업	4	3	4.45	5	4	3.81	0.48	0.51
07	비금속광물 광업; 연료용 제외	50	29	2.37	73	43	2.45	0.08	0.11
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	266	-52	0.84	394	-77	0.84	0.06	0.26
11	음료 제조업	23	-22	0.51	34	-32	0.51	0.1	0.41
12	담배 제조업	0	-3	0	0	-4	0		
13	섬유제품 제조업; 의복 제외	248	31	1.14	388	74	1.24	0.03	0.17
14	의복, 의복 액세서리 및 모피제품 제조업	82	-5	0.94	126	-2	0.99	0.05	0.15
15	가죽, 가방 및 신발 제조업	79	29	1.58	129	57	1.79	0.03	0.03

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	157	83	2.11	229	119	2.09	0.08	0.24
17	펄프, 종이 및 종이제품 제조업	114	4	1.03	172	10	1.06	0.07	0.23
18	인쇄 및 기록매체 복제업	98	-49	0.67	137	-81	0.63	0.13	0.38
19	코크스, 연탄 및 석유정제품 제조업	17	-5	0.77	24	-9	0.72	0.14	0.58
20	화학 물질 및 화학제품 제조업; 의약품 제외	310	-112	0.74	492	-131	0.79	0.04	0.1
21	의료용 물질 및 의약품 제조업	50	-80	0.39	72	-121	0.38	0.14	0.32
22	고무 및 플라스틱제품 제조업	329	-62	0.84	478	-104	0.82	0.07	0.28
23	비금속 광물제품 제조업	337	111	1.49	502	169	1.51	0.07	0.26
24	1차 금속 제조업	289	76	1.35	402	80	1.25	0.11	0.3
25	금속 가공제품 제조업; 기계 및 가구 제외	679	49	1.08	979	29	1.03	0.08	0.27
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	371	-721	0.34	571	-1073	0.35	0.06	0.17
27	의료, 정밀, 광학 기기 및 시계 제조업	84	-115	0.42	119	-180	0.4	0.11	0.31
28	전기장비 제조업	244	-195	0.56	343	-311	0.52	0.11	0.34
29	기타 기계 및 장비 제조업	688	-263	0.72	956	-484	0.66	0.12	0.34
30	자동차 및 트레일러 제조업	396	-291	0.58	607	-407	0.6	0.06	0.18
31	기타 운송장비 제조업	284	-59	0.83	441	-86	0.84	0.05	0.22

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	68	-5	0.93	105	-2	0.99	0.06	0.11
33	기타 제품 제조업	89	-24	0.79	124	-43	0.74	0.11	0.38
34	산업용 기계 및 장비 수리업	40	-5	0.88	59	-9	0.86	0.07	0.29
35	전기, 가스, 증기 및 공기 조절 공급업	94	-34	0.73	159	-21	0.88	0.02	0.05
36	수도업	5	-2	0.71	8	-3	0.75	0.2	0.65
37	하수, 폐수 및 분뇨 처리업	51	27	2.16	71	36	2.01	0.13	0.37
38	폐기물 수집, 운반, 처리 및 원료 재생업	250	143	2.34	355	196	2.24	0.09	0.26
39	환경 정화 및 복원업	2	0	1.05	3	0	0.91	0.43	0.94
41	종합 건설업	2058	254	1.14	2935	267	1.1	0.09	0.29
42	전문직별 공사업	78	23	1.42	120	40	1.5	0.04	0.04
45	자동차 및 부품 판매업	42	-49	0.46	57	-76	0.43	0.18	0.57
46	도매 및 상품 중개업	961	-521	0.65	1393	-818	0.63	0.08	0.29
47	소매업; 자동차 제외	330	-298	0.53	464	-464	0.5	0.1	0.31
49	육상 운송 및 파이프라인 운송업	2097	1272	2.54	2975	1791	2.51	0.09	0.3
50	수상 운송업	73	35	1.91	114	57	2.01	0.04	0.15
51	항공 운송업	5	-13	0.28	6	-21	0.22	0.34	0.36

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	452	64	1.17	688	105	1.18	0.06	0.2
55	숙박업	80	-17	0.82	122	-21	0.85	0.05	0.12
56	음식점 및 주점업	169	-108	0.61	237	-163	0.59	0.13	0.4
58	출판업	223	-546	0.29	321	-834	0.28	0.09	0.35
59	영상·오디오 기록물 제작 및 배급업	18	-31	0.37	28	-44	0.38	0.05	0.06
60	방송업	21	-18	0.53	30	-26	0.53	0.28	0.7
61	우편 및 통신업	37	-74	0.33	54	-111	0.33	0.03	0.07
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	34	-75	0.31	48	-118	0.29	0.12	0.41
63	정보서비스업	66	-27	0.71	106	-33	0.76	0.06	0.13
64	금융업	135	-167	0.45	190	-259	0.42	0.12	0.33
65	보험 및 연금업	30	-38	0.44	46	-54	0.46	0.05	0.05
66	금융 및 보험관련 서비스업	37	-66	0.36	51	-107	0.32	0.15	0.37
68	부동산업	1939	1370	3.41	3215	2379	3.85	0.01	0.03
70	연구개발업	36	-92	0.28	56	-137	0.29	0.05	0.3
71	전문 서비스업	226	-149	0.6	354	-212	0.63	0.05	0.11
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	408	-219	0.65	611	-332	0.65	0.06	0.23

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	23	-34	0.4	36	-50	0.42	0.04	0.26
74	사업시설 관리 및 조경 서비스업	1081	697	2.81	1656	1077	2.86	0.06	0.19
75	사업 지원 서비스업	1403	524	1.6	2045	716	1.54	0.08	0.29
76	임대업; 부동산 제외	22	-13	0.62	35	-17	0.67	0.05	0.06
84	공공 행정, 국방 및 사회보장 행정	334	71	1.27	603	234	1.63	0.03	0.07
85	교육 서비스업	236	-16	0.94	360	-15	0.96	0.06	0.25
86	보건업	183	-137	0.57	252	-228	0.52	0.14	0.48
87	사회복지 서비스업	85	-37	0.7	115	-66	0.64	0.21	0.48
90	창작, 예술 및 여가관련 서비스업	24	-12	0.66	40	-15	0.72	0.04	0.09
91	스포츠 및 오락관련 서비스업	100	-25	0.8	153	-35	0.81	0.08	0.28
94	협회 및 단체	114	10	1.1	161	5	1.03	0.09	0.28
95	개인 및 소비용품 수리업	149	-114	0.57	213	-180	0.54	0.11	0.31
96	기타 개인 서비스업	97	23	1.31	136	25	1.23	0.11	0.31

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-23〉 중분류별 폐암 발병자수 추계 결과, 2000년도까지의 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	13	5	1.64	16	5	1.44	0.21	0.58
02	임업	0	-1	0	0	-2	0		
03	어업	3	1	1.51	6	3	1.98	0.13	0.13
05	석탄, 원유 및 천연가스 광업	2	0	1.14	4	1	1.44	0.22	0.4
06	금속 광업	1	1	4.22	1	1	3.95	0.58	0.89
07	비금속광물 광업; 연료용 제외	2	0	0.88	3	0	1	0.17	0.17
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	307	103	1.51	436	151	1.53	0.01	0.1
11	음료 제조업	19	0	1.01	24	-2	0.92	0.08	0.19
12	담배 제조업	7	4	2.19	9	5	2.15	0.04	0.07
13	섬유제품 제조업; 의복 제외	299	53	1.22	415	70	1.2	0.03	0.12
14	의복, 의복 액세서리 및 모피제품 제조업	258	47	1.22	341	45	1.15	0.04	0.14
15	가죽, 가방 및 신발 제조업	123	63	2.05	159	75	1.9	0.05	0.09

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	22	6	1.37	29	7	1.29	0.19	0.4
17	펄프, 종이 및 종이제품 제조업	38	2	1.06	48	-2	0.97	0.06	0.17
18	인쇄 및 기록매체 복제업	37	3	1.1	48	1	1.02	0.03	0.13
19	코크스, 연탄 및 석유정제품 제조업	2	-2	0.55	2	-3	0.39	0.03	0.07
20	화학 물질 및 화학제품 제조업; 의약품 제외	89	-11	0.89	128	-12	0.92	0.02	0.02
21	의료용 물질 및 의약품 제조업	31	-11	0.74	37	-21	0.64	0.07	0.26
22	고무 및 플라스틱제품 제조업	115	34	1.41	161	46	1.41	0.01	0.03
23	비금속 광물제품 제조업	94	37	1.65	137	58	1.73	0.02	0.02
24	1차 금속 제조업	53	21	1.66	69	24	1.54	0.04	0.13
25	금속 가공제품 제조업; 기계 및 가구 제외	139	55	1.66	189	72	1.61	0.04	0.19
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	298	-216	0.58	438	-283	0.61	0.03	0.06
27	의료, 정밀, 광학 기기 및 시계 제조업	74	26	1.54	95	28	1.41	0.02	0.04
28	전기장비 제조업	158	25	1.19	237	50	1.27	0.01	0.02
29	기타 기계 및 장비 제조업	97	9	1.11	133	9	1.08	0.02	0.03
30	자동차 및 트레일러 제조업	170	44	1.35	229	53	1.3	0.02	0.04
31	기타 운송장비 제조업	30	6	1.23	47	13	1.39	0.06	0.1

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	46	24	2.13	63	33	2.09	0.07	0.09
33	기타 제품 제조업	78	30	1.63	109	42	1.62	0.03	0.1
34	산업용 기계 및 장비 수리업	0	-2	0	0	-3	0		
35	전기, 가스, 증기 및 공기 조절 공급업	16	-6	0.73	24	-6	0.8	0.07	0.2
36	수도업	2	1	1.76	3	2	2.1	0.27	0.94
37	하수, 폐수 및 분뇨 처리업	2	0	1.01	2	-1	0.74	0.3	0.37
38	폐기물 수집, 운반, 처리 및 원료 재생업	9	1	1.08	11	0	0.97	0.05	0.1
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	89	-57	0.61	115	-92	0.56	0.07	0.07
42	전문직별 공사업	4	-5	0.44	6	-7	0.44	0.09	0.18
45	자동차 및 부품 판매업	5	-10	0.34	9	-12	0.43	0.13	0.14
46	도매 및 상품 중개업	233	-132	0.64	368	-148	0.71	0.01	0.03
47	소매업; 자동차 제외	146	-83	0.64	204	-119	0.63	0.02	0.02
49	육상 운송 및 파이프라인 운송업	48	-6	0.9	70	-5	0.93	0.04	0.08
50	수상 운송업	3	-5	0.39	4	-7	0.33	0.14	0.15
51	항공 운송업	9	-15	0.37	13	-21	0.4	0.17	0.43

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	50	-15	0.77	67	-24	0.73	0.07	0.27
55	숙박업	114	42	1.57	168	67	1.66	0.02	0.03
56	음식점 및 주점업	264	130	1.97	382	193	2.01	0.02	0.09
58	출판업	48	-70	0.41	64	-103	0.38	0.1	0.1
59	영상·오디오 기록물 제작 및 배급업	4	-3	0.55	6	-4	0.61	0.16	0.65
60	방송업	8	-4	0.64	10	-7	0.6	0.04	0.12
61	우편 및 통신업	115	13	1.13	159	17	1.12	0.02	0.03
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	1	-5	0.15	1	-8	0.11	0.05	0
63	정보서비스업	11	-5	0.68	15	-8	0.67	0.05	0.06
64	금융업	249	-124	0.67	331	-192	0.63	0.04	0.13
65	보험 및 연금업	70	-54	0.56	92	-81	0.53	0.07	0.19
66	금융 및 보험관련 서비스업	27	-24	0.53	33	-38	0.47	0.13	0.41
68	부동산업	174	70	1.67	247	99	1.68	0.02	0.05
70	연구개발업	7	-9	0.44	8	-14	0.38	0.13	0.39
71	전문 서비스업	65	-61	0.51	94	-85	0.52	0.05	0.05
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	36	-23	0.61	49	-35	0.58	0.01	0.07

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	1	-10	0.09	1	-15	0.08	0.06	0.12
74	사업시설 관리 및 조경 서비스업	405	291	3.56	588	428	3.69	0.01	0.01
75	사업 지원 서비스업	241	53	1.28	353	86	1.32	0.02	0.02
76	임대업; 부동산 제외	0	-3	0	0	-5	0		
84	공공 행정, 국방 및 사회보장 행정	45	-2	0.95	63	-5	0.93	0.04	0.1
85	교육 서비스업	152	-31	0.83	215	-47	0.82	0.04	0.2
86	보건업	294	-160	0.65	431	-210	0.67	0.02	0.05
87	사회복지 서비스업	81	-24	0.77	106	-43	0.71	0.08	0.22
90	창작, 예술 및 여가관련 서비스업	5	-6	0.46	7	-9	0.44	0.15	0.16
91	스포츠 및 오락관련 서비스업	24	-12	0.66	29	-22	0.57	0.12	0.33
94	협회 및 단체	58	5	1.09	84	10	1.13	0.03	0.07
95	개인 및 소비용품 수리업	15	1	1.07	21	2	1.08	0.05	0.1
96	기타 개인 서비스업	35	1	1.03	51	3	1.06	0.01	0.06

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

〈표 부-24〉 중분류별 폐암 발병자수 추계 결과, 2000년도 이후 입사자, 여성

단위 : 명(발병자수 및 초과분), 비율(발병비, 적합오차)

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
01	농업	15	8	2.07	20	10	1.95	0.32	0.58
02	임업	0	-1	0	0	-2	0		
03	어업	1	0	0.72	2	0	0.96	0.14	0.31
05	석탄, 원유 및 천연가스 광업	0	0	0	0	0	0		
06	금속 광업	1	1	17.68	1	1	16.97	0.76	0.89
07	비금속광물 광업; 연료용 제외	1	0	0.81	1	0	0.88	0.1	0.35
08	광업 지원 서비스업	0	0	0	0	0	0		
10	식료품 제조업	225	87	1.63	310	111	1.56	0.16	0.49
11	음료 제조업	3	-5	0.39	4	-7	0.33	0.43	0.45
12	담배 제조업	1	1	4.6	1	1	4.95	0.4	0.95
13	섬유제품 제조업; 의복 제외	131	57	1.76	191	89	1.87	0.08	0.14
14	의복, 의복 액세서리 및 모피제품 제조업	110	32	1.41	160	52	1.48	0.08	0.22
15	가죽, 가방 및 신발 제조업	14	-4	0.78	21	-4	0.83	0.08	0.27

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
16	목재 및 나무제품 제조업; 가구 제외	15	5	1.53	20	6	1.45	0.1	0.29
17	펄프, 종이 및 종이제품 제조업	23	6	1.35	34	10	1.43	0.1	0.44
18	인쇄 및 기록매체 복제업	29	-2	0.92	39	-5	0.89	0.22	0.37
19	코크스, 연탄 및 석유정제품 제조업	4	2	1.61	5	2	1.54	0.37	0.93
20	화학 물질 및 화학제품 제조업; 의약품 제외	49	-14	0.77	72	-18	0.8	0.07	0.3
21	의료용 물질 및 의약품 제조업	17	-9	0.66	23	-14	0.62	0.17	0.23
22	고무 및 플라스틱제품 제조업	114	36	1.46	161	50	1.45	0.11	0.28
23	비금속 광물제품 제조업	34	9	1.36	49	14	1.4	0.11	0.18
24	1차 금속 제조업	19	4	1.24	31	9	1.42	0.05	0.15
25	금속 가공제품 제조업; 기계 및 가구 제외	91	12	1.15	123	9	1.08	0.16	0.48
26	전자 부품, 컴퓨터, 영상, 음향 및 통신장비 제조업	209	-88	0.7	284	-118	0.71	0.17	0.47
27	의료, 정밀, 광학 기기 및 시계 제조업	36	-3	0.93	52	-3	0.95	0.15	0.28
28	전기장비 제조업	83	-2	0.97	117	-3	0.98	0.12	0.28
29	기타 기계 및 장비 제조업	98	10	1.12	135	9	1.07	0.17	0.57
30	자동차 및 트레일러 제조업	94	14	1.18	142	29	1.26	0.08	0.24
31	기타 운송장비 제조업	23	1	1.05	35	3	1.1	0.15	0.34

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
32	가구 제조업	19	9	1.97	27	14	2.05	0.13	0.14
33	기타 제품 제조업	46	18	1.66	66	27	1.69	0.1	0.23
34	산업용 기계 및 장비 수리업	0	-2	0	0	-3	0		
35	전기, 가스, 증기 및 공기 조절 공급업	8	-2	0.82	14	0	1.03	0.17	0.65
36	수도업	0	-1	0	0	-1	0		
37	하수, 폐수 및 분뇨 처리업	3	1	1.55	6	3	2.03	0.13	0.12
38	폐기물 수집, 운반, 처리 및 원료 재생업	22	11	2.09	32	18	2.18	0.09	0.12
39	환경 정화 및 복원업	0	0	0	0	0	0		
41	종합 건설업	132	-39	0.77	205	-36	0.85	0.06	0.11
42	전문직별 공사업	2	-1	0.68	3	-1	0.72	0.1	0.17
45	자동차 및 부품 판매업	1	-10	0.09	1	-15	0.07	0.79	1
46	도매 및 상품 중개업	338	-89	0.79	461	-141	0.77	0.15	0.41
47	소매업; 자동차 제외	245	-32	0.88	344	-36	0.91	0.14	0.41
49	육상 운송 및 파이프라인 운송업	41	12	1.43	68	28	1.7	0.03	0.05
50	수상 운송업	0	-3	0	0	-5	0		
51	항공 운송업	0	-11	0	0	-16	0		

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [†]
52	창고 및 운송관련 서비스업	77	-3	0.96	108	-8	0.93	0.11	0.44
55	숙박업	35	1	1.04	53	7	1.14	0.07	0.09
56	음식점 및 주점업	322	116	1.56	452	175	1.63	0.11	0.32
58	출판업	49	-94	0.34	69	-134	0.34	0.11	0.33
59	영상·오디오 기록물 제작 및 배급업	9	-4	0.7	13	-5	0.74	0.12	0.14
60	방송업	4	-4	0.51	7	-4	0.66	0.16	0.23
61	우편 및 통신업	9	-6	0.59	14	-8	0.63	0.1	0.1
62	컴퓨터 프로그래밍, 시스템 통합 및 관리업	10	-7	0.59	16	-9	0.64	0.08	0.11
63	정보서비스업	33	-10	0.76	48	-13	0.78	0.09	0.32
64	금융업	65	-80	0.45	93	-110	0.46	0.11	0.2
65	보험 및 연금업	15	-18	0.46	24	-23	0.51	0.07	0.07
66	금융 및 보험관련 서비스업	17	-17	0.49	25	-25	0.5	0.1	0.48
68	부동산업	234	119	2.03	324	160	1.97	0.13	0.44
70	연구개발업	13	-16	0.45	17	-23	0.43	0.19	0.75
71	전문 서비스업	89	-84	0.51	138	-106	0.57	0.05	0.22
72	건축 기술, 엔지니어링 및 기타 과학기술 서비스업	51	-27	0.65	76	-36	0.68	0.12	0.48

코드	분류명	2018년			2022년 (추계)			예측오차	
		발병자수	초과분	발병비	발병자수	초과분	발병비	훈련 [†]	검증 [‡]
73	기타 전문, 과학 및 기술 서비스업	14	-9	0.61	18	-13	0.58	0.25	0.42
74	사업시설 관리 및 조경 서비스업	455	320	3.36	643	449	3.31	0.11	0.26
75	사업 지원 서비스업	540	51	1.1	742	41	1.06	0.15	0.37
76	임대업; 부동산 제외	2	-3	0.37	3	-5	0.34	0.23	0.28
84	공공 행정, 국방 및 사회보장 행정	62	-8	0.88	94	-17	0.85	0.11	0.17
85	교육 서비스업	309	-12	0.96	430	-25	0.95	0.14	0.42
86	보건업	433	-176	0.71	594	-263	0.69	0.17	0.41
87	사회복지 서비스업	413	-44	0.9	560	-75	0.88	0.24	0.66
90	창작, 예술 및 여가관련 서비스업	6	-9	0.4	8	-12	0.41	0.15	0.34
91	스포츠 및 오락관련 서비스업	34	-3	0.91	53	0	1.01	0.05	0.15
94	협회 및 단체	47	-3	0.95	64	-6	0.91	0.23	0.6
95	개인 및 소비용품 수리업	10	-4	0.7	15	-6	0.72	0.05	0.07
96	기타 개인 서비스업	71	17	1.31	99	23	1.31	0.11	0.3

* 초과분 = (해당업종 발병자수) - (표준군 발병자수); 발병비 = (해당업종 발병자수) - (표준군 발병자수); 표준군 = 전체업종 합계.

† 2015-2018년 발병자수를 전체(-2018년분)로부터 예측(적합)하였을 때, 예측값의 실제값대비 제곱근평균오차비율(root-mean-squared error ratio)

‡ 2015-2018년 발병자수를 과거(-2014년분)로부터 예측(추계)하였을 때, 예측값의 실제값대비 제곱근평균오차비율

연구진

연구기관 : 숙명여자대학교

연구책임자 : 최영근 (조교수, 숙명여자대학교)

연구원 : 김양우 (전임의, 한양대학교구리병원)

연구원 : 박성오 (조교수, 성신여자대학교)

연구원 : 이동환 (부교수, 이화여자대학교)

연구원 : 이우주 (부교수, 서울대학교)

연구보조원 : 김혜린 (석사과정, 이화여자대학교)

연구보조원 : 이윤아 (석사과정, 성신여자대학교)

연구보조원 : 이은경 (석사과정, 숙명여자대학교)

연구보조원 : 정승필 (석사과정, 서울대학교)

연구상대역 : 이경은 (선임연구위원, 중부권역학조사팀)

연구기간

2022. 05. 01. ~ 2022. 11. 30.

본 연구는 산업안전보건연구원의 2022년도 위탁연구 용역사업에 의한 것임

본 연구보고서의 내용은 연구책임자의 개인적 견해이며,
우리 연구원의 공식견해와 다를 수도 있음을 알려드립니다.

산업안전보건연구원장

연구제목

(2022-산업안전보건연구원-803)

발 행 일 : 2022년 11월 30일

발 행 인 : 산업안전보건연구원 원장 김은아

연구책임자 : 숙명여자대학교 조교수 최영근

발 행 처 : 안전보건공단 산업안전보건연구원

주 소 : (44429) 울산광역시 중구 종가로 400

전 화 : 032-510-0753

팩 스 : 032-510-0759

Homepage : <http://oshri.kosha.or.kr>

I S B N : 979-11-92782-48-5

공공안심글꼴 : 무료글꼴, 한국출판인회의, Kopub바탕체/돋움체