

국가노출감시체계 구축을 위한 작업환경측정과 특수건강진단 자료의 노출 정보 입력 실태 평가

최상준^{1,7*} · 고통희² · 박주현³ · 박동욱⁴ · 김환철⁵ · 임대성⁶ · 성예지^{1,7} · 고경윤¹ · 임지선³ · 서회경⁸

¹가톨릭대학교 보건의료경영대학원, ²가톨릭관동대학교 국제성모병원, ³동국대학교 통계학과, ⁴한국방송통신대학교,
⁵인하대학교, ⁶한성보건안전기술원(주), ⁷가톨릭대학교 대학원, ⁸한국산업안전보건공단 산업안전보건연구원

Evaluation of the Input Status of Exposure-related Information of Working Environment Monitoring Database and Special Health Examination Database for the Construction of a National Exposure Surveillance System

Sangjun Choi^{1,7*} · Dong-Hee Koh² · Ju-Hyun Park³ · Donguk Park⁴ · Hwan-Cheol Kim⁵ ·
Dae Sung Lim⁶ · Yeji Sung^{1,7} · Kyoung Yoon Ko⁷ · Ji Seon Lim³ · Hoekyeong Seo⁸

¹Graduate School of Public Health and Healthcare Management, The Catholic University of Korea

²Department of Occupational and Environmental Medicine, International St.Mary's Hospital

³Department of Statistics, Dongguk University

⁴Department of Environmental Health, Korea National Open University

⁵Department of Occupational and Environmental Medicine, Inha University

⁶Hansung Health and Safety Technology Co., Ltd.

⁷Graduate School, The Catholic University of Korea

⁸Occupational Safety and Health Research Institute, Korea Occupational Safety and Health Agency

ABSTRACT

Objectives: The purpose of this study is to evaluate the input status of exposure-related information in the working environment monitoring database (WEMD) and special health examination database (SHED) for the construction of a national exposure surveillance system.


Methods: The industrial and process code input status of WEMD and SHED for 21 carcinogens from 2014 to 2016 was compared. Data from workers who performed both work environment monitoring and special health examinations in 2019 and 2020 were extracted and the actual status of input of industrial and process codes was analyzed. We also investigated the cause of input errors through a focus group interview with 12 data input specialists.


Results: As a result of analyzing WEMD and SHED for 21 carcinogens, the five-digit industrial code matching rate was low at 53.5% and the process code matching rate was 19% or less. Among the data that


*Corresponding author: Sangjun Choi, Tel:02-2258-7379, E-mail: junilane@gmail.com


Graduate School of Public Health and Healthcare Management, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul, 06591, Korea


Received: August 3, 2022, Revised: September 4, 2022, Accepted: September 23, 2022


 Sangjun Choi <http://orcid.org/0000-0001-8787-7216>


 Ju-Hyun Park <https://orcid.org/0000-0001-9675-6475>


 Hwan-Cheol Kim <https://orcid.org/0000-0002-3635-1297>


 Yeji Sung <https://orcid.org/0000-0003-2267-5490>

 Ji Seon Lim <https://orcid.org/0000-0002-6642-5031>

 Donghee Koh <http://orcid.org/0000-0002-2868-4411>

 Donguk Park <https://orcid.org/0000-0003-3847-7392>

 Daesung Lim <http://orcid.org/0000-0003-4190-0390>

 Kyoung Yoon Ko <https://orcid.org/0000-0001-6616-5164>

 Hoekyeong Seo <https://orcid.org/0000-0002-8069-3788>

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

simultaneously conducted work environment monitoring and special health examination in 2019 and 2020, the process code matching rate was very low at 18.1% and 5.2%, respectively. The main causes of exposure-related data input errors were the difference between the WEMD and SHED process code input systems from 2020, the number of standard process and job codes being too large, and the inefficiency of the standard code search system.

Conclusions: In order to use WEMD and SHED as a national surveillance system, it is necessary to simplify the number of standard code input codes and improve the search system efficiency.

Key words: National surveillance system, standardization, special health examination data, work environment monitoring data

I. 서 론

직업성 질환의 발생을 예방하고 필요한 경우 적절한 보상을 결정하기 위해서는 발생 질환의 직업적 노출 요인을 규명하는 것이 필요하며, 이를 위해서는 상시적인 노출요인과 건강영향을 모니터링 할 수 있는 국가 감시 체계(national surveillance system)가 필요하다. 특히 초기 노출 후 긴 잠복기를 갖는 암(cancer)과 같은 만성 질환자의 노출 요인을 규명하기 위해서는 질환자의 직무력에 따른 과거 노출 평가(retrospective exposure assessment)가 필요하며, 직무기간 동안 노출 수준을 파악할 수 있는 체계가 필요하다. 직업적 요인에 따른 질병과의 연관성 연구에서 Reed & Harcourt (1941)가 노출 평가를 위해 최초로 직무-노출매트릭스(Job-Exposure Matrix, JEM) 방법을 소개하였으며, Hoar et al.(1980)과 Partanen et al.(1985)이 직업적 유해인자의 영향에 대한 역학연구에서 노출 평가 방법으로 JEM을 사용한 이후 역학연구에서 과거 노출을 추정하는 방법으로 많이 사용되고 있다.

핀란드 산업보건연구소(Finnish Institute of Occupational Health, FIOH)는 국가 차원의 노출평가 자료를 활용하여 1945년 이후부터 직업적 노출요인들에 대해 FINJEM(Finnish Job-Exposure Matrix)을 개발하여 국가 노출 감시체계로 활용해 오고 있다(Kauppinen et al., 2014). 이렇게 개발된 FINJEM은 발암성 물질에 대해 유럽연합의 타 국가들과 협력 연구를 통해 새로운 발암성 물질에 대한 노출력 DB인 CAREX(CARcinogen EXposure)를 개발하는데도 활용되었다. CAREX는 유럽연합(EU)의 'Europe Against Cancer Program'의 지원에 의해 만들어진 발암물질에 대한 국제적 노출정보 시스템으로서 국가별, 발암물질별, 산업별로 분류된 노출자료와 노출된 근로자의 추정치를 제공한다. 이후 독일은 MEGA(Stamm, 2001; Gabriel, 2006), 프랑스는 COLCHIC

(Vincent & Jeandel, 2001)과 같은 정량적 노출평가 자료를 데이터베이스화하여 국가 노출 감시체계로 활용해 오고 있다.

우리나라는 산업안전보건법에 의해 근로자에 대한 보건관리 제도로써 작업환경측정(이하 측정)과 특수건강진단(이하 특검)을 실시하고 있다. 측정과 특검 결과는 2002년부터 'K2B'라는 전산 시스템(<https://k2b.kosha.or.kr/index.do>)을 통해 실시 기관들이 한국 산업안전보건공단으로 결과 자료를 입력하여 보고하고 있다. 측정 및 특검의 생물학적 노출지표 검사 자료는 전국을 대상으로 주기적인 정량적 노출수준을 평가하고 있고 K2B를 통해 하나의 데이터베이스로 축적되어 오고 있어 산업, 공정, 직종 등의 노출을 설명해 줄 수 있는 정보(이하 노출관련 K2B 정보)가 정확하다면 그 활용도가 매우 높다. 현재까지 측정 혹은 특검 자료를 이용하여 석면(Choi et al., 2017), 벤젠(Koh et al., 2015), 납(Koh et al., 2017; Koh et al., 2018)과 20개 인체 발암물질(human carcinogens)(Koh et al., 2021a; Koh et al., 2021b)에 대한 노출인구 비율(exposure prevalence)과 노출강도(exposure intensity) 추정 연구들이 수행되어 왔다. 그러나 선행 연구들은 모두 산업별 노출 특성을 평가하였고 구체적인 노출 특성을 파악할 수 있는 공정과 직종 정보를 추가로 활용하지 못한 제한점이 있다. 이는 K2B의 공정과 직종에 대한 표준화 코드 입력 결과에 대한 정확도가 낮고 구체적인 정보가 없는 '기타' 정보가 많기 때문이라고 추정되는데, 실제 입력 정보들의 입력 실태에 대해 구체적으로 평가된 연구는 없었다.

이에 본 연구는 노출관련 K2B 정보(산업, 공정, 직종)의 입력 실태를 비교 분석한 후 표준코드 입력 오류 원인을 평가하여 향후 국가 노출 감시체계로서 활용도를 높이기 위해 필요한 개선점을 파악하고자 실시되었다.

II. 대상 및 방법

1. 측정/특검 결과 입력 실태 평가를 위한 대상 자료

측정/특검 결과 입력 자료의 비교를 위해 K2B에 보고된 자료 중 다음과 같이 발암물질 자료와 측정, 특검 동시 수검 자료를 추출하여 분석하였다.

1) 발암물질 자료

발암물질은 노출 후 일정한 잠복기를 거쳐 발생하기 때문에 국가 노출 감시체계의 필요성이 크기 때문에 측정-특검 결과의 노출관련 정보 입력 실태 확인이 우선적으로 필요하다고 판단하였다. 특히 여러 발암물질들 중 국제암연구소(International Agency for Research on Cancer, IARC)에서 Group 1로 분류하고 있고, 국내에서 발암물질 감시체계 활용을 위한 Korean CAREX(CARcinogen EXposure) 데이터베이스 구축에 대한 연구(Koh et al., 2021a)가 실시된 21종(비소, 아스민, 석면, 벤젠, 베릴륨, 1,3-부타디엔, 카드뮴, 6가 크롬, 산화에틸렌, 포름알데히드, 염산, 불산, 미네랄오일미스트, 니켈, 질산, 유리규산, 황산, 트리클로로에틸렌, 염화비닐, 용접흄, 목분진)에 대해 2014년부터 2016년까지 실시되어 K2B로 보고된 측정과 특검 자료를 선정하였다.

2) 동시 수검자료

2019년과 2020년에 측정과 특검을 모두 수검 받은 작업자에 대한 자료를 선정하였다. 이는 비교적 측정-특검 결과에 입력된 노출관련 정보의 일치율이 높을 것으로 판단되기 때문에 이 자료를 분석함으로써 현재 측정-특검 정보를 최대한 연계할 수 있는 가능성이 어느 정도인지 확인하기 위함이다.

2. 노출관련 정보 입력 실태 분석

분석 대상 자료에 입력된 정보 중 노출 관련 K2B 정보인 산업, 공정, 직종에 대한 표준코드 입력 실태를 분석하였고, 측정과 특검 자료에 입력된 표준코드의 일치율을 분석함으로써 측정-특검 자료의 연계 가능성을 평가하였다. K2B 자료의 추출은 한국산업안전보건공단에 요청하여 실시되었으며, 측정/특검 대상 사업장명 및 피검자의 이름을 포함한 개인정보는 제외된 자료를 받았으며, 별도의 가공 없이 원자료에 측정과 특검 기관

이 입력한 결과를 상호 비교하여 분석하였다. 대상 자료는 사업장관리번호, 조사 년도, 물질명, 표준산업코드와 산업명, 표준공정코드와 공정명, 표준직업코드와 직업명 정보를 갖고 있다.

발암물질 자료의 경우 '연도+사업장관리번호+발암물질명'이 동일한 자료를 추출하여 동일한 사업장에서 동일한 년도에 동일한 발암물질에 대해 실시된 측정과 특검 자료 중 노출관련 코드의 일치율을 분석하였다. 산업분류 코드의 일치율은 연계변수(연도+사업장관리번호+발암물질명)를 기준으로 중복된 자료를 제외하고 분석하였고, 공정코드 일치율은 '연도+사업장관리번호+발암물질+공정코드'를 연계변수로 하고 중복을 제외하고 분석하였다.

측정-특검 동시 수검자료의 경우 2019년과 2020년에 측정과 특검 실시 자료 중 '사업장관리번호+물질코드+근로자명'으로 두 자료를 매칭하여 특정 유해물질에 대해 측정도 하고, 특검도 받은 근로자(건수)를 추출하였다. 이렇게 추출된 자료에 대해 측정과 특검 자료의 표준공정코드와 직종코드 정보를 비교 분석하였다.

3. 노출관련 정보 입력 오류 원인 조사

측정과 특검 기관에 종사하며 직접 측정결과와 특검 결과를 전산입력하는 담당자 12명(측정 10, 특검 2)을 대상으로 focus group interview(FGI)를 실시하여 노출관련 정보의 표준화 코드 입력 시스템 파악과 정보 입력시 애로사항 및 오류 발생 원인이 무엇인지 조사하였다.

III. 연구결과

1. 발암물질 자료에 대한 노출관련 정보 입력 실태

총 21개 발암물질에 대해 2014년부터 2016년까지 실시된 측정 및 특검 자료의 전체 건수와 사업장 수, 측정-특검 자료 사이의 표준산업코드와 표준공정코드 정보 일치 정도 등 주요 특성을 Table 1에 요약하였다.

전체 분석 대상 자료 중 측정 자료는 929,753건, 특검 자료는 1,662,224건이었고, 물질 중에는 용접 흄이 가장 많았고(측정 190,588건, 특검 428,485건), 베릴륨이 가장 적었다(측정 317건, 특검 1,341건). 연계변수를 이용하여 중복 자료를 제거하고 각 발암물질에 대해 최소 1건 이상 측정, 혹은 특검을 실시한 사업장 수

Table 1. Characteristics of WEM and SHE results for 21 carcinogens from 2014 to 2016

Substance	WEM database		SHE database		Data of places of business that performed both WEM and SHE		
	Number of records	Number of places of business	Number of records	Number of places of business	Number of places of business	Matching between WEM and SHE data	
						KSIC10 th 5 digit	SPC 5 digit
Arsenic	1,313	51	14,300	72	29	11	2
Arsine	2,697	287	11,620	278	118	46	15
Asbestos	652	189	11,101	811	64	23	17
Benzene	19,661	2,403	58,916	1,800	765	402	126
Beryllium	317	48	1,341	26	15	4	2
1,3-Butadiene	5,048	671	14,408	420	183	110	46
Cadmium	7,494	1,383	11,087	769	460	246	93
Crystalline silica	53,974	8,853	4,673	403	268	143	38
Ethylene oxide	11,443	4,084	10,201	1,956	1,634	1,025	409
Formaldehyde	51,631	7,222	81,201	3,773	2,854	1,673	621
Hexavalent chromium	40,513	10,019	28,055	3,209	1,943	1,182	393
Hydrochloric acid	63,502	12,269	149,135	7,103	4,630	2,661	840
Hydrofluoric acid	17,636	2,187	72,579	1,525	858	398	137
Mineral oil mist	136,027	35,179	174,419	14,913	11,426	5,607	1,618
Nickel	148,738	20,192	278,909	11,196	6,196	3,187	1,283
Nitric acid	45,667	9,683	110,700	5,454	3,735	2,089	662
Sulfuric acid	88,999	15,212	146,754	8,154	5,723	3,257	1,034
Trichloroethylene	18,295	5,517	27,361	2,634	1,854	1,163	290
Vinyl chloride	2,814	543	5,204	308	192	120	30
Welding fume	190,588	40,644	428,485	22,266	11,963	6,031	2,737
Wood dust	22,744	6,887	21,775	2,308	1,623	867	265
Total	929,753	183,523	1,662,224	89,378	56,533	30,245	10,658

WEM: Working environment monitoring, SHE: Special health examination, KSIC10th: Korea Standard Industrial Classification 10th revision, SPC: Standard process classification

는 측정 183,523개, 특검 89,378개였고 측정-특검 사업장 및 물질 정보가 동일하게 일치된 수는 56,533개였다. 즉, 측정을 한 경우 해당 물질에 대해 특검도 실시한 비율은 30.8%(56,533/183,523)이고, 특검을 받은 경우 측정도 실시한 비율은 63.3%(56,533/89,378)였다.

총 21개 발암물질에 대해 측정-특검의 일치된 자료 56,533개 중 제10차 개정된 한국표준산업분류 코드 세세분류(5 digit)가 일치하는 것은 30,245개로 53.5%였고, 표준공정 코드(5 digit)가 일치하는 것은 10,658개로 18.9%였다. 물질별 표준산업코드와 공정코드의 일치율을 비교하면 Figure 1과 같이 표준산업코드 일치율은 베릴륨이 26.7%로 가장 낮고 산화에틸렌의 경우

최대 62.7%까지 차이가 크게 분포한 반면 표준공정코드 일치율은 최대 26.6% 이하로 대부분 낮은 분포를 보였다. 표준산업코드 일치율의 경우 각 물질별 특검을 받은 사업장 중 측정도 실시한 비율이 높을수록 증가하는 양의 상관성($r=0.57$)을 보였다(Figure 2).

2. 측정/특검 동시 수검 자료에 대한 노출관련 정보 입력 실태

2019년과 2020년에 측정과 특검을 동시에 수검 받은 작업자의 자료 분석 결과 Table 2와 같이 2019년에는 239개 유해인자에 대해 859,782개, 2020년에는 234개 유해인자에 대해 809,786개 자료가 확인되었다. 분석 대상 자료의 산업분류 코드 자료는 받지 못해

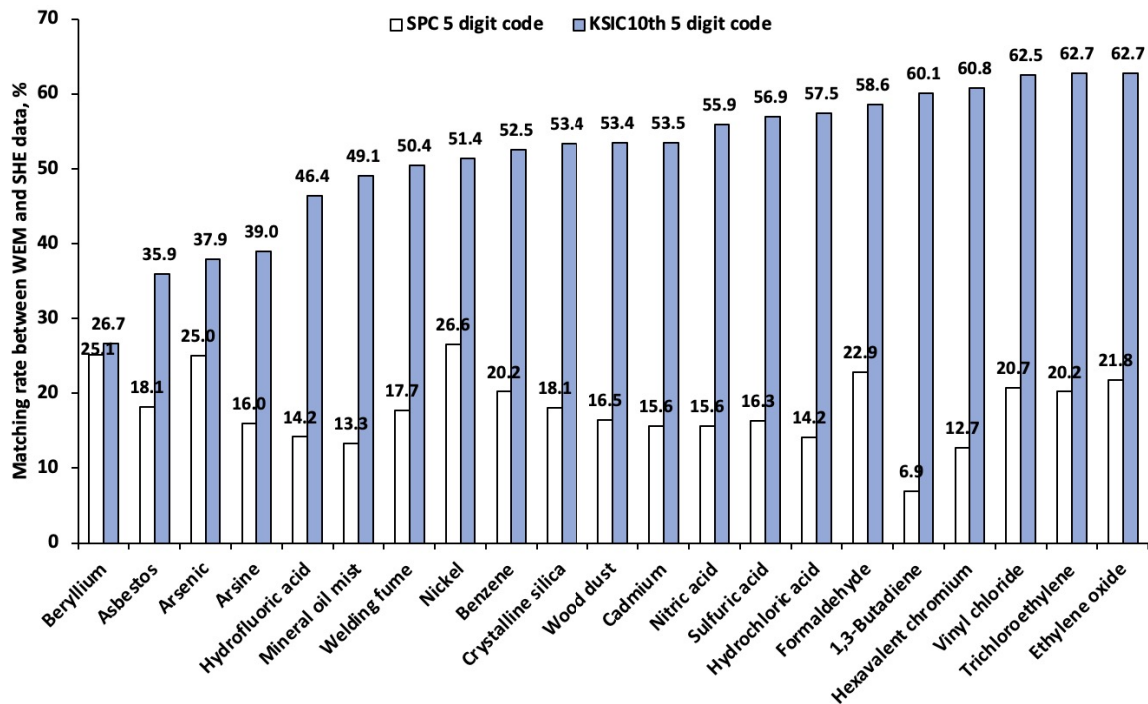


Figure 1. Comparison of matching rate of standard industry and process between WEM and SHE data for 21 carcinogens from 2014 to 2016 (WEM: Working environment monitoring, SHE: Special health examination, KSIC10th: Korea Standard Industrial Classification 10th revision, SPC: Standard process classification).

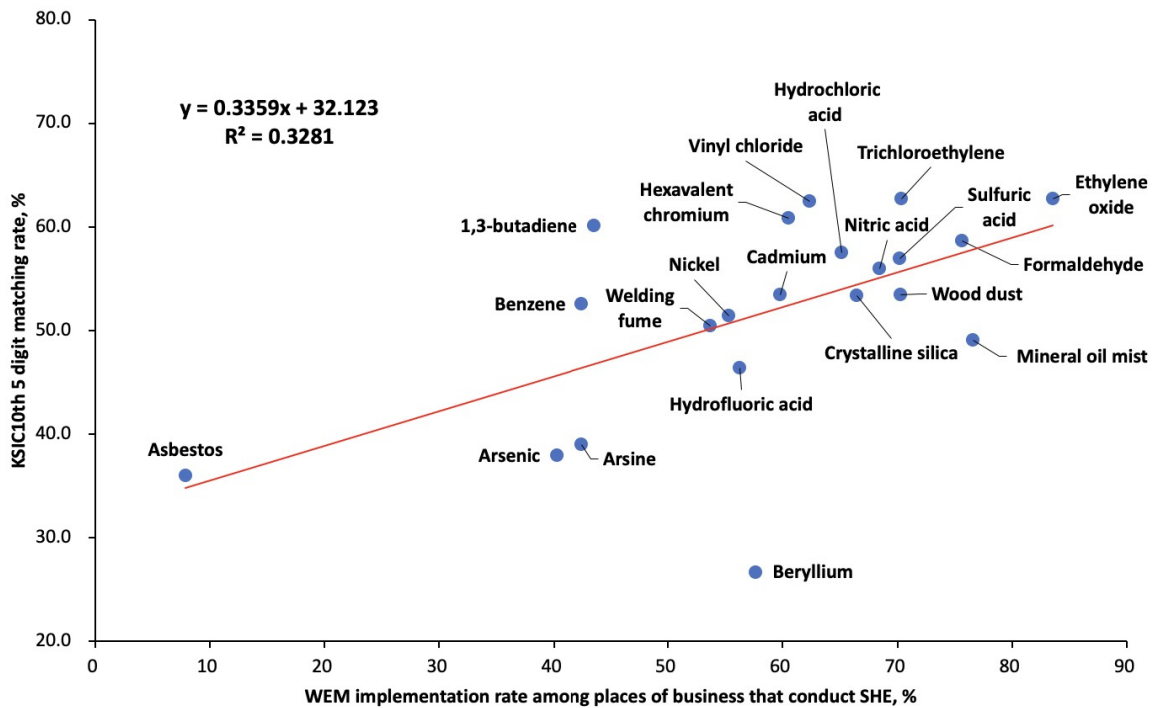


Figure 2. Correlation between WEM implementation rate among places of business that conduct SHE and matching rate of standard industrial codes (WEM: Working environment monitoring, SHE: Special health examination, KSIC10th: Korea Standard Industrial Classification 10th revision).

Table 2. Summary of data characteristics of workers who received both WEM and SHE from 2019 to 2020

Item	2019	2020
Total record	859,782	809,786
Number of chemical agents	239	234
Matching rate of SPC 5 digit	155,995 (18.1%)	41,710 (5.2%)
Matching rate of SPC 2 digit	232,301 (28.2%)	108,129 (13.4%)
Records entered as 'Other Process'	WEM: 197,136 (22.9%), SHE: 547,232 (63.6%)	WEM: 31,450 (3.9%), SHE: 527,746 (65.2%)

WEM: Working environment monitoring, SHE: Special health examination, SPC: Standard process classification

Table 3. Status of top-ten most frequent standard occupations entered in SHE data of workers who received both WEM and SHE in 2020

KSCO7 th 5 digit	KSCO7 th name	N	%
93009	Production related elementary workers n.e.c.	221,909	27.4
89909	Machine operators n.e.c.	48,885	6.0
79999	Technical occupations related workers n.e.c.	38,115	4.7
75105	Automobile paint mechanics	29,050	3.6
85432	Ship assemblers	21,196	2.6
74309	Welders n.e.c.	20,779	2.6
14139	Production managers n.e.c.	17,911	2.2
31241	Production management clerks	16,490	2.0
84219	Painting machine operators n.e.c.	16,317	2.0
84159	Metal processing machine operators n.e.c.	15,684	1.9

WEM: Working environment monitoring, SHE: Special health examination,
KSCO7th: Korea Standard Classification of Occupation 7th revision

표준산업코드의 일치율에 대해서는 분석을 할 수 없었고, 모든 자료에 누락 없이 입력되어 있는 표준공정코드에 대한 일치율 분석이 가능하였다.

측정자료의 공정코드와 특검의 공정코드가 일치하는 경우는 2019년은 155,995건으로 18.1%였으나, 2020년에는 41,710건으로 5.2%로 더 낮았다. 입력되어 있는 5자리 공정코드를 앞 2자리로 묶어서 비교하면 일치율이 2019년은 28.2%, 2020년은 13.4%로 높아졌다. 입력된 공정코드가 특정 공정을 확인할 수 없는 '기타'로 입력된 비율은 측정의 경우 2019년 22.9%에서 2020년에는 3.9%로 낮아졌지만, 특검의 경우는 2019년과 2020년 각각 63.6%, 65.2%로 측정의 경우보다 높은 수준으로 큰 변화가 없었다.

2020년 특검 자료 중 한국표준직업분류(Korea Standard Classification of Occupation, KSCO) 제 7차 개정안으로 입력되어 있는 표준직종코드 현황을 분석한 결과 Table 3과 같이 다빈도 직종은 27.4%가 '그

외 제조 관련 단순 종사원(KSCO=93009)', 6.0%는 '그 외 기계 조작원(KSCO=89909)', 4.7%는 '그 외 기능 관련 종사원(KSCO=79999)' 등 비특이적 직종이었다.

3. 노출관련 정보 입력 오류 원인

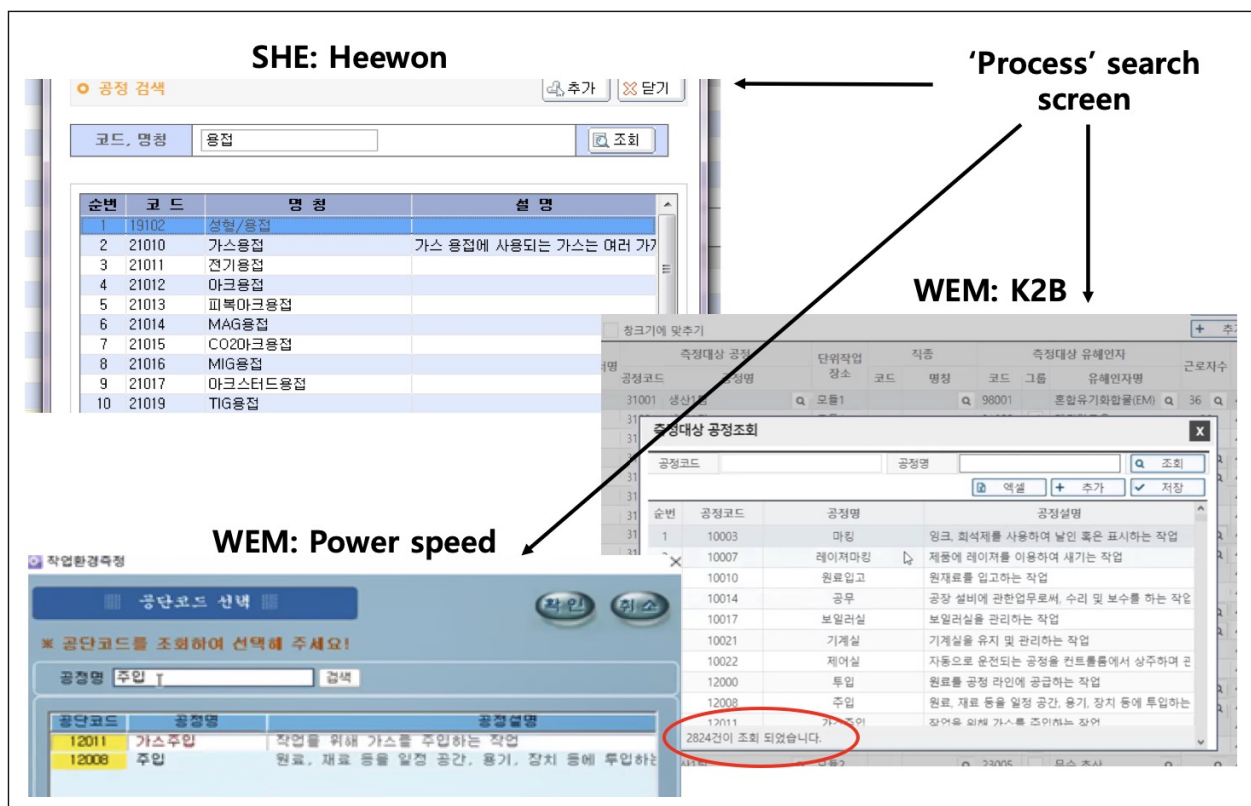
노출관련 정보 입력 오류 원인 조사를 위해 실시한 FGI 대상자 기본 특성과 결과는 Table 4에 요약하였다. 인터뷰에 참여한 총 12개 기관 중 K2B 시스템을 이용하는 곳은 4곳이었고, 8곳은 상업용 프로그램을 활용하고 있었고, 각 기관별 인터뷰에 참여한 전산 입력 경력은 5년에서 12년이였다.

노출관련 세 가지 정보(산업, 공정, 직종) 중 특히 공정코드 입력 오류가 발생하고, 측정-특검 자료 사이의 정보가 불일치되는 주요 원인으로는 첫째, 측정과 특검 결과의 표준공정코드 입력시스템이 2020년부터 서로 다르기 때문인 것으로 확인되었다. 2019년까지는 측정-특검 모두 약 1390여개의 5자리 표준공정코드를 활용

Table 4. Basic characteristics of focus group interview subjects and summary of interview results on the cause of standard code input error

Item	WEM	SHE
Interviewee		
N	10	2
Career, year	6~9	5~12
Results input program		
K2B	4	0
Commercial program	6	2
Major cause of input error	<ul style="list-style-type: none"> • Difference between the WEM and SHE process code input systems from 2020 • Standard code is too classified • Inefficiency of the standard code search system 	

WEM: Working environment monitoring, SHE: Special health examination

**Figure 3.** Example of a 'process' search screen according to the WEM and SHE results input program (WEM: Working environment monitoring, SHE: Special health examination).

하여 입력해 왔으나, 2020년부터 K2B 시스템을 이용한 측정결과를 입력하는 경우에는 '80000' 이후 코드 번호의 경우 입력자가 직접 공정명을 입력할 수 있는 개방형 입력이 가능한 코드를 추가하였으나, 특검 기관들은 기존 공정코드 시스템을 그대로 활용하고 있었다. 주요 오염원인 두 번째는 산업, 공정, 직종 모두 표준코드의 수가 지나치게 세분화되고 많기 때문에 적절한 코

드를 선택하여 입력하기가 어렵다는 평가였다. 세 번째는 프로그램에서 표준코드 검색 방법이 비효율적이라는 평가였는데, Figure 3과 같이 표준공정코드를 검색할 경우 검색어와 일치되는 단어가 '공정명' 컬럼의 내용과만 비교되어 선택될 수 있도록 하고 있었다. 이는 표준코드명과 정확히 일치하는 검색어를 활용하지 않는다면 적절한 코드를 찾아서 입력하기 어려운 구조이다.

IV. 고 찰

암과 같이 잠복기를 갖고 발생하는 비가역적 직업성 질환에 대한 과거 노출 평가를 위해 JEM을 활용해 오고 있으며, 이때 활용 가능한 자료 중 대표적인 것이 JEM 구축 관련 유해인자에 대한 선행연구 후 발표된 문헌자료(peer-reviewed literature)이다. 그러나 문헌자료마다 노출관련 정보가 제한적인 경우가 많아(예를 들어, 산업에 대한 정보는 있으나, 공정에 대한 정보가 없거나, 직종에 대한 정보가 없는 경우) 역학연구 대상 집단을 대상으로 설문조사를 하거나 전문가의 판단에 기초한 정성적 노출수준 평가결과를 활용하기도 한다. 이렇게 구축한 대표적인 JEM이 캐나다에서 만든 CANJEM(Siemiatycki & Lavoué, 2018)이다. 정성적 노출수준에 기초한 JEM의 활용도가 높긴 하지만 정량적인 노출수준과 관련 노출정보를 활용할 수 있는 자료가 있다면 보다 활용도가 높은 JEM 구축이 가능하다. 우리나라의 측정과 특검 결과에 대한 K2B 데이터베이스는 산업, 공정, 직종 등 노출특성을 설명해 줄 수 있는 정보들이 표준화된 코드로 입력되어 있기 때문에 각 정보의 입력 정확도가 높다면 JEM 구축 활용도가 높을 것이다. 이에 본 연구에서는 측정-특검 자료의 K2B 노출관련 정보 입력 실태를 분석하고 입력오류의 원인이 무엇인지 확인하였다.

발암물질 21종에 대한 3개년 측정-특검 자료를 분석한 결과 산업에 대한 표준코드 일치율은 53.5%로 낮게 나타났으며, 물질에 따라 26.7%부터 62.7%까지 차이가 컸다(Figure 1). 산업에 대한 표준코드 선택은 측정, 특검 기관에서 대부분 조사대상 사업장으로부터 사업자 등록증 정보를 받거나 알려주는 정보에 기초해서 하기 때문에 비교적 측정과 특검 자료 사이에 일치될 가능성이 크다고 예측했으나 높지 않았다.

흥미로운 점은 발암물질 별 특검을 실시한 사업장 중 측정도 실시한 사업장이 많을수록 측정과 특검 자료의 산업정보 일치율이 높게 나타났는데(Figure 2), 이는 측정이 실시된 사업장의 경우 특검 대상자를 선정하고 사업장의 업종 정보 등을 확인하는데 있어 측정보고서 자료 활용이 가능하기 때문에 측정이 실시되지 않은 사업장 보다 상대적으로 측정-특검 산업정보의 일치율이 높다고 판단되었다. 작업환경측정 제도의 경우 산업안전보건법에 의해 예비조사를 의무화 하고 있으나 특검의 경우는 사전 예비조사가 의무화되지 않았기 때문에

해당 사업장의 정보와 특검 대상자의 노출정보를 사업장 담당자가 보내주는 정보에 의존하는 경우가 많다. 따라서 측정을 통해 확인된 사업장의 산업과 공정 등 노출관련 정보의 정확도가 높다고 할 수 있으며, 이러한 측정 정보를 참고하여 특검이 실시되는 경우 측정-특검 자료의 노출정보 일치율이 높아질 수 있다.

표준공정 정보의 경우 21개 발암물질에 대한 평균 일치율이 18.9%로 산업코드보다 더욱 낮았고, 물질별 차이도 크지 않으며 산업 정보의 일치율이 가장 높았던 산화에틸렌의 경우도 공정 정보 일치율은 21.8%로 낮았다. 또한 특검 실시 사업장 중 측정실시율과도 무관하였다. 이는 측정결과를 K2B에 전송할 경우엔 표준공정 코드 정보를 입력하도록 하고 있으나, 사업장 제출용으로 출력한 보고서에는 측정기관이 각 사업장마다 사용하는 부서, 공정, 단위작업장소 정보들만 출력되고 표준공정코드와 공정명은 출력되지 않도록 프로그램화되어 있어 특검 기관이 측정정보를 참고해도 표준공정 정보의 확인은 불가능하기 때문이다. 따라서 향후 K2B에 입력된 표준공정 정보도 보고서 출력시 함께 출력될 수 있도록 한다면 특검에 활용할 수도 있고 보다 정확한 공정 정보 입력이 가능할 것으로 기대된다.

측정-특검 자료의 표준공정 코드 정보의 일치율이 낮은 특성은 측정-특검 동시 수검자료만 추출하여 분석한 결과에서도 동일하게 나타났다(Table 2). 특히 2019년(18.1%)보다 2020년(5.2%)이 더욱 낮게 나타났는데 이는 FGI결과 2020년부터 표준공정코드 입력시스템의 변화 때문으로 확인되었다. 2020년부터 K2B 시스템을 이용해 측정결과를 입력하는 경우에는 입력자가 직접 공정명을 입력할 수 있는 개방형 코드(80000~)를 추가하였으나, 특검 기관들은 기존 공정코드 시스템을 그대로 활용하고 있었다. 따라서 2020년의 측정과 특검 자료 중 공정코드 일치율은 더욱 낮게 나타났고, 측정의 경우 적절한 공정코드가 없다고 판단될 경우 입력자가 직접 코드명을 자유롭게 입력 가능하도록 되었기 때문에 '기타'로 입력된 비율은 낮아졌지만, 특검의 경우는 2019년과 2020년 각각 63.6%, 65.2%로 측정의 경우보다 높은 수준으로 큰 변화가 없었다. 이렇게 2020년부터 변경된 표준공정코드 입력 시스템은 산업 및 공정의 변화에 따른 새로운 공정명에 대한 탄력적 대응이 가능하다는 장점이 있으나, 너무 많은 공정명이 자율적으로 생성될 수 있다는 단점이 있다. 따라서 주기적으로 새롭게 생성된 코드를 재 분류하여 표준화하여 갱신

하는 과정이 필요하며, 이러한 갱신 시스템이 동반되지 못한다면 너무 많은 표준화 코드 생성으로 인해 자료 활용이 불가능해진다. 2020년 검색 결과 표준공정 코드 수가 이미 2,800개가 넘는 것을 확인할 수 있었으며, 향후 개선이 필요하다고 판단된다.

직종은 측정의 경우 의무 입력 사항이 아니며, 전산 시스템에서도 K2B의 경우엔 2020년부터 전산 입력 메뉴가 추가되었으나, 기관에서 상용 프로그램을 활용하는 경우엔 직종 입력 메뉴가 없었다. 특검은 2020년부터 직종코드 입력이 필수 사항인데, 한국표준직업분류 세세분류(5자리) 코드를 입력하고 있다. 그래서 2020년 측정-특검 동시 수집자료 중 특검 자료(n=809,786)에 대해서 표준직종코드 입력현황을 분석하였다(Table 3). 분석결과 전체 자료의 약 1/3 이상이 특화되지 않은 비특이적 단순 종사원, 기계조작원, 관련 종사원 등으로 입력되어 있었다. 사업장 담당자가 표준직업분류 코드를 입력해서 특검기관에 명단을 주지 않기 때문에, 특검기관 담당자가 임의로 직종을 입력하게 된다. 특검기관 담당자가 작업장을 잘 모르기 때문에 단순 종사원으로 입력하는 경우가 많은 것으로 추정된다. 반면 일부에서는 자동차 도장 정비원(KSCO=75105), 선박 조립

원(KSCO=85432), 용접원(KSCO=74309) 등의 특이적인 직종을 식별할 수 있었다. 하지만 이렇듯 직종명을 입력하게 함으로써 일부에서는 직종의 특이성이 높아졌으나, 비특이적인 기타 직종이 비율도 높아 도리어 민감도가 떨어질 수 있기 때문에, 직종 정보를 활용하는데 있어서는 제한이 있을 수 있을 것으로 판단된다. 이는 현재 사용중인 한국표준직업분류 제7차 개정안의 세세분류는 1,231개로 매우 많기 때문에 정확한 직종 코드를 찾아 입력하는데 어려움이 있기 때문이다. 예를 들어, 간호사의 경우 현재 세세분류 코드에서는 일반간호사(KSCO=24302), 전문간호사(KSCO=24301)로 구분되어 있는데 기관 입력자가 구체적인 정보를 갖고 있지 않다면 선택하기 어려우며 노출관점에서는 둘 다 간호사로 분류해도 무방할 것이다. 따라서 유사 직종을 그룹화 하여 보다 간소화된 직종코드로 표준화될 필요가 있다. 최상준 등(2019)은 2014년부터 2016년까지 대한산업보건협회에서 측정한 화학적 인자에 대해 측정된 180개 사업장 자료를 활용하여 총 70개 산업과 161개 공정, 13개 직무로 표준화를 시도한 바 있다(Choi et al., 2019). 이 연구에서는 직종을 13개 직무로 매우 큰 직무그룹으로 표준화하였는데, 이 경우에는

Scheme: Census 2012/NAICS 2012/SOC 2010							
Industry Title:		Occupation Title:	He repaired automobile.	Autocode			
Census Industry:	8770 - Automotive Repair and Mair	Census Occupation:	7200 - Automotive Service Technicians and Me				
NAICS:	8111 - Automotive Repair and Mair	SOC:	49-3023 - Automotive Service Technicians and				
Suggested Candidates (Ind 5/Occ 5)							
Title	Ind Code	Title	NAICS Score	Title	Occ Code	Title	SOC Score
Automotive Repair and Maintenance	8770	Automotive Repair and Maintenance	8111 0.982	Automotive Service Technicians and Mechanics	7200	Automotive Service Technicians and Mechanics	49-30230.992
Personal and Household Goods Repair and Maintenance	8880	Personal and Household Goods Repair and Maintenance	8114 0.014	Automotive Body and Related Repairers	7150	Automotive Body and Related Repairers	49-30210.003
NIOSH-Insufficient Information	9990	Insufficient information - NIOSH	009990 0.003	Insufficient Information	9900	Insufficient Information - NIOSH	00-99000.003
Motor Vehicles and Motor Vehicle Equipment Manufacturing	3570	Motor Vehicle Manufacturing	3361 <0.001	Electric Motor, Power Tool, and Related Repairers	7040	Electric Motor, Power Tool, and Related Repairers	49-20920.002
Motor Vehicles and Motor Vehicle Equipment Manufacturing	3570	Motor Vehicle Body and Trailer Manufacturing	33621 <0.001	Small Engine Mechanics	7240	Motorcycle Mechanics	49-30520.001

Figure 4. Standard industry and occupational auto-coding results for the statement 'He repaired automobile' using NIOCCS (NIOSH Industry and Occupation Computerized Coding System).

표준화된 공정 정보와 직무군이 함께 연계되어 해석되어야 활용도가 있기 때문에 직무군 정보만으로는 충분한 직종 특성을 파악하기는 어려운 한계가 있다. 따라서 향후 현재의 한국표준직업분류코드를 바탕으로 노출관점에서 유사한 직무군들을 재분류하여 적절한 수의 표준직종 코드 개발이 필요하다.

측정-특검 자료 입력 업무를 수행해 온 12개 기관 담당자들을 대상으로 FGI를 통해 확인한 표준정보 입력 오류의 주요 원인은 크게 세 가지였다(Table 4). 가장 큰 두 가지 원인은 위에서 고찰하였듯이 2020년부터 측정-특검 결과에 대한 표준공정 입력 시스템의 차이가 있었다는 점과 공정, 직종 등이 1,000여개 이상의 표준코드로 세분화되어 있어 정확한 선택이 어렵다는 점이었다. 따라서 향후 이 두 가지 요인을 해결하기 위해서는 측정-특검 사이에 동일한 표준코드 입력시스템을 갖도록 해야 하고, 표준코드를 유사노출관점에서 보다 간소하게 재분류 할 필요가 있다. 이외에 추가적으로 지적하고 있는 주요 입력오류 원인은 프로그램상의 표준코드 검색 방법이 비효율적이라는 점이었다. 공정코드 검색창에 검색어를 입력하면 해당 검색어와 일치하는 단어가 공정명에 있는 경우만 검색되고 있다. 이를 공정명 뿐만 아니라 해당 공정에 대한 설명 글, 해당 공정으로 분류 가능한 유사 공정 색인어들과도 검색되어 관련 공정코드를 찾을 수 있도록 시스템이 개선될 필요가 있다. 이와 관련하여 벤치마킹 할 수 있는 시스템이 있다.

미국의 국립산업안전보건연구원(National Institute for Occupational Safety and Health, NIOSH)에서는 연구자들이 각종 서베이, 사망진단서, 의료 기록 등의 자료 정보를 이용하여 산업과 직업에 대한 표준코드로 쉽게 전환해 주는 무료 웹용 어플리케이션인 'NIOSH Industry and Occupation Computerized Coding System(NIOCCS, <https://csams.cdc.gov/nioccs/Default.aspx>)'를 개발하였다. NIOCCS는 표준산업과 표준직업코드와 각 코드명을 설명해 줄 수 있는 유사 색인어 데이터베이스를 기반으로 문장 혹은 단어를 입력하면 자동으로 표준산업 및 직업코드를 찾아준다. 예를 들어 Figure 4와 같이 '그는 자동차를 수리했다(He repaired automobile)'라는 문장을 입력하자 색인어 데이터베이스와 검색결과 가장 적합도가 높은 표준산업과 표준직업 코드를 점수와 함께 5개씩 제시해 준다. 이 경우 '자동차 정비 및 수리업'과 '자동차 정비

및 수리원'이 가장 높은 점수인 0.982, 0.992를 나타내고 있어 적합한 표준코드라는 것을 검색자가 확인하는데 도움을 주고 있다. 측정-특검 자료의 표준공정 혹은 직업 코드를 검색하는 방법도 이와 유사하게 공정명 혹은 직종명의 단어와 검색어가 일치될 때만 찾아지게 하지 말고 각 표준코드명을 설명할 수 있는 다양한 색인어 데이터베이스를 활용한 최적의 코드명이 찾아질 수 있도록 개선될 필요가 있다.

V. 결 론

발암물질 21종에 대해 2014년부터 2016년까지 실시된 측정, 특검 자료를 매칭하여 분석한 결과 표준산업 코드 입력결과 사이의 일치율은 53.5%, 표준공정 코드 일치율은 18.9%로 낮았다. 2019년과 2020년에 측정과 특검을 같은 시기에 받은 근로자에 대한 자료를 분석한 결과 표준공정 일치율은 18.1%에서 5.2%로 더욱 낮아졌음이 확인되었다. 이렇게 산업, 공정의 표준코드 일치율이 낮은 주요 원인은 첫째, 측정-특검 자료에 대한 표준공정 입력 시스템이 2020년부터 서로 달라진 계통적 원인과 둘째, 산업, 공정, 직종 등 표준화코드가 1,000여개 이상 너무 세분화되어 있다는 점, 그리고 K2B 프로그램에서 표준화된 코드명을 검색하는 방법이 비효율적이라는 점 등으로 조사되었다. 따라서, 향후 측정-특검 데이터베이스를 국가노출감시체계로서 활용도를 높이기 위해서는 측정-특검에 대한 각 정보의 입력 시스템을 동일하게 하고, 노출관점에서 적절한 코드 수로 재표준화가 필요하며, 코드명 검색이 보다 빠르고 쉽게 될 수 있도록 개선이 필요하다.

감사의 글

본 연구는 2022년 한국산업안전보건공단 산업안전보건연구원 학술용역 지원 사업에 의해 수행되었으며, 표준정보 입력오류 원인 조사에 참여해 준 측정, 특검 기관 종사자분들께 감사드립니다.

References

- Choi S, Jeong JY, Im S, Lim D, Koh DH et al. The standardization of work environment measurement information for constructing exposure surveillance

- system. J Korean Soc Occup Environ Hyg 2019; 29(3):322-335
- Choi S, Kang D, Park D, Lee H, Choi B. Developing asbestos job exposure matrix using occupation and industry specific exposure data (1984-2008) in republic of korea. Saf Health Work 2017;8(1): 105-115(<https://doi.org/10.1016/j.shaw.2016.09.002>)
- Gabriel S. The BG measurement system for hazardous substances (BGMG) and the exposure database of hazardous substances (MEGA). Int J Occup Saf Ergon 2006;12(1):101-104(<https://doi.org/10.1080/10803548.2006.11076673>)
- Hoar SK, Morrison AS, Cole P, Silverman DT. An occupation and exposure linkage system for the study of occupational carcinogenesis. JOM 1980; 22(11):722-726
- Kauppinen T, Uuksulainen S, Saalo A, Mäkinen I, Pukkala E. Use of the Finnish Information System on Occupational Exposure (FINJEM) in epidemiologic, surveillance, and other applications. Ann Occup Hyg 2014;58(3):380-396
- Koh DH, Jeon HK, Lee SG, Ryu HW. The relationship between low-level benzene exposure and blood cell counts in korean workers. Occup Environ Med 2015;72(6):421-427(<http://dx.doi.org/10.1136/occup-med-2014-102227>)
- Koh DH, Park JH, Lee SG, Kim HC, Choi S et al. Development of korean CARcinogen EXposure: an initiative of the occupational carcinogen surveillance system in korea. Ann Work Expo Health 2021(a); 65(5):528-538(<https://doi.org/10.1093/annweh/wxaa135>)
- Koh DH, Park JH, Lee SG, Kim HC, Choi S et al. Combining lead exposure measurements and experts' judgment through a bayesian framework. Ann Work Expo Health 2017;61(9):1054-1075 (<https://doi.org/10.1093/annweh/wxx072>)
- Koh DH, Park JH, Lee SG, Kim HC, Choi S et al. Estimation of lead exposure prevalence in korean population through combining multiple experts' judgment based on objective data sources. Ann Work Expo Health 2018;62(2):210-220(DOI:10.1093/annweh/wxx106)
- Koh DH, Park JH, Lee SG, Kim HC, Jung H et al. Estimation of lead exposure intensity by industry using nationwide exposure databases in korea. Saf Health Work 2021(b);12(4):439-444(<https://doi.org/10.1016/j.shaw.2021.07.008>)
- Partanen T, Kauppinen T, Nurminen M, Nickels J, Hernberg S et al. Formaldehyde exposure and respiratory and related cancer: A case-reference study among finnish woodworkers. Scandinavian J Work Environ Health 1985; 11:409-415
- Reed JV, Harcourt AK. The essentials of occupational diseases. Baltimore: CC Thomas 1941
- Siemiatycki J, Lavoué J. Availability of a new job-exposure matrix (CANJEM) for epidemiologic and occupational medicine purposes. J Occup Environ Med 2018;60(7):e324-e328
- Stamm R. MEGA-database: one million data since 1972. Appl Occup and Environ Hyg 2001;16(2):159-163
- Vincent R, Jeandel B. COLCHIC-occupational exposure to chemical agents database: current content and development perspectives. Appl Occup Environ Hyg 2001;16(2):115-121(<https://doi.org/10.1080/104732201460190>)

<저자정보>

최상준(교수), 고동희(교수), 박주현(교수), 박동욱(교수), 김환철(교수), 임대성(대표이사), 성예지(대학원생), 고경윤(대학원생), 임지선(대학원생), 서희경(연구원)