입원 폐암환자의 동반질환에 대한 연관성 규칙 분석

김성수

청주대학교 보건행정·헬스케어학부 교수

Association Rules of Comorbidities in Patients Hospitalized for Lung Cancer

Sung-Soo Kim

Professor, Department of Health Administration & Healthcare, Cheongju University, Cheongju, Korea

Objectives: The purpose of this study is to find out how the accompanying diseases of lung cancer patients are related. Through this, we aim to provide implications for the management of lung cancer patients and the national cancer policy. **Methods:** The subjects of this study were the Korean National Hospital Discharge in-depth Injury Survey, which was surveyed from 2006 to 2016, and among them, patients with lung cancer were selected. The major variable, comorbidities, were grouped based on the middle classification of the KCD-7 revision. The general characteristics were analyzed by frequency, and χ^2 -test was performed to test the difference in the distribution for the presence of comorbidity. In order to analyze the association rules of comorbidities, the support, confidence, and improvement of the association rules were presented and visualized using the arules package and apriori algorithm of the R program. **Results:** As a result of the analysis, the proportion of patients with comorbid diseases out of 21,882 discharged patients was 82.86%, and an average of 2.48 ± 2.19 comorbidities was diagnosed. As a result of applying the association rules criteria of support (> 0.01), confidence (> 0.1), and lift (> 1), 33 comorbidity patterns were found. The high association rules were “hypertensive diseases (I10-I15)” and “malignant neoplasms of ill-defined, secondary and unspecified sites (C76-C80)”, “diabetes mellitus (E10-E14)” and “hypertensive diseases (I10-I15)”, “malignant neoplasms of ill-defined, secondary and unspecified sites (C76-C80)” and “diabetes mellitus (E10-E14)”, “influenza and pneumonia (J09-J18)” and “hypertensive diseases (I10-I15)”, and so on. **Conclusions:** This study provides the rules related to comorbidities of lung cancer patients in Korean. It is used for the diagnosis and treatment of lung cancer patients, and it provides implications for effectively improving cancer policy. Future research on the effect on medical use or treatment results is needed, focusing on major comorbidities.

Key words: Lung cancer, Comorbidity, Association rules, Hospitalization, Inpatient

서론

폐암은 흔하게 진단되는 질환이며 사망률이 높은 질환으로 알려져 있다[1,2]. 최근 세계보건기구(World Health Organization, WHO)의 발표자료에 따르면, 2020년 새롭게 진단된 암환자 1,929만 명 중 220만 명이 폐암이며, 이는 전체 암 중에서 2위인 11.4%를 차지하고 있는 것으로 1위인 유방암 11.7%와 근소한 차이의 발생률을 나타내고 있다[3]. 폐암

으로 인한 사망은 전체 암 사망환자 약 996만 명 중 18.0%로 180만 명에 가까워 사망원인 1위를 차지하는 것으로 보고되었다. 우리나라 중앙암등록본부에서 발표한 최근 2017년 통계에 따르면 위암, 대장직장암에 이어서 폐암은 26,985명이 발생하여 3번째로 흔하게 발생하는 암으로 세계보건기구의 발생률 2위와 큰 차이가 없다[4]. 사망률에 있어서는 2위인 간암 10,721명보다 월등히 높은 17,980명으로 2015년에 이어서 1위를 차지하고 있다. 이는 2015년 11,311명보다 58%나 늘어난 수

Corresponding author: Sung-Soo Kim

298 Daeseong-ro, Cheongwon-gu, Cheongju 28503, Korea
Tel: +82-43-229-7998, E-mail: mra7033@naver.com

Received: February 06, 2021 Revised: February 20, 2021 Accepted: February 22, 2021

No potential conflict of interest relevant to this article was reported.

How to cite this article:

Kim SS. Association Rules of comorbidities in patients hospitalized for lung cancer. J Health Info Stat 2021;46(1):100-109. Doi: <https://doi.org/10.21032/jhis.2021.46.1.100>

© It is identical to the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permit unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2021 Journal of Health Informatics and Statistics

치료 흡연율은 감소하고 있으나 인구고령화와 라돈과 같은 환경오염의 지속적인 영향으로 인하여 앞으로 20년 간 증가추세가 예상된다[5]. 저선량 흉부 CT (low-dose computed tomography, LDCT)를 이용한 폐암의 조기진단 권고와 새로운 치료법 등 임상연구가 지속되고 있으나 단기간에 발생률과 사망률을 낮추기는 어려워 보인다.

암정책에 있어서 진단과 치료에 대한 연구도 중요하지만 세계적인 고령화 추세로 인해 국민의료비 증가와 의료이용의 변화에 관심이 커지고 있다. 선행연구를 살펴본 결과 폐암환자의 동반질환과 의료이용 및 의료비의 연관성을 분석한 결과에 따르면, 폐암환자의 동반질환은 잠재적인 치료선택을 제한하고 추가적인 의료비 부담을 증가시키는 것으로 나타났다[6]. 폐암환자의 생존율을 연구한 결과에서도 몇몇 질환을 제외하고는 동반질환이 생존율을 낮추는데 기여하는 것으로 분석되었다[7]. 폐암환자의 동반질환에 대한 다양한 연구들이 시도되고 있지만 연구대상의 규모와 수집방법에 따라 동반질환 유병률이 313%[6]에서 73.3%[7] 등 다른 결과들이 보고되고 있다[8]. 따라서 동반질환을 효과적으로 확인하기 위해서는 자가보고에 의한 코호트 연구보다는 지역사회나 국가차원의 등록데이터를 활용하는 것이 효율적이다. 개별 동반질환을 확인하기 위해서는 Charlson 동반질환 지수(Charlson Comorbidity Index, CCI)와 같은 보정지수를 사용하기 보다는 국제적인 질병분류 기준과 연구대상에 따라 분류수준을 고려하여야 한다. 최근 우리나라의 전체 암환자의 동반질환을 규명한 연구에서는 한국 표준질병사인분류의 대분류를 기준으로 하였기 때문에 상세한 동반질환을 분석하는 데는 한계가 있었다[9]. 분석방법에서도 기술통계적인 방법을 탈피하고 동반질환 간 패턴을 분석하여 숨겨진 의미를 확인할 필요가 있다. 최근 보건 의료 분야의 빅데이터에 관심이 커지면서 인공지능 기반의 머신러닝(machine learning), 딥러닝(deep learning), 데이터마이닝(data mining)을 활용한 연구들이 활발히 진행되고 있다[10-12]. 본 연구에서는 폐암환자의 동반질환 패턴을 분석하는데 연관성 규칙이 적합할 것이라 판단하였다. 연관성 규칙은 Agrawal et al. [13]에 의해 제안된 데이터마이닝 기법 중 하나로 모수통계에서 필요한 정규성, 선형성 등 가정을 필요로 하지 않는다. 목표변수가 없기 때문에 비지도학습 기법으로 분류되며 여러 변수들 사이에서 숨겨진 패턴을 탐색하는 방법으로 마케팅 분야에서는 장바구니 분석으로도 알려져 있다[14]. 연관성 규칙에서 Apriori 알고리즘이 보편적으로 널리 사용되고 있는데, 최소 지지도와 신뢰도를 만족하는 규칙을 선별하고 향상도 1을 초과하는 최종 연관규칙을 분석하는 두 단계를 거치게 된다. 본 연구에서는 동반질환을 1과 0으로 이분화하고 매트릭스 형태로 변환하여 연관성 규칙 분석 적용이 가능하며 동반질환의 숨겨진 패턴을 찾기에 유용하다[15]. 본 연구는 폐암환자의 동반질환 유병률을 평가하고 다빈도 동반질환을 확인하여 이들 동반질환 간에 어떠한 연관관

계가 있는지를 규명하는 것이다. 타당한 연구성과를 위해 전국 규모의 다년간 누적된 자료를 기반으로 연관성 규칙 분석을 적용하여 의미 있는 동반질환 패턴을 근거로 정책적 시사점을 제공하고자 한다.

연구 방법

연구대상 및 자료수집

본 연구의 대상은 의료기관에 입원하여 퇴원한 환자 중 폐암이 주진단(principal diagnosis)인 환자들이다. 연구목적에 따라 연구대상의 동반질환 연관규칙을 효과적으로 규명하기 위하여 질병관리청에서 매년 조사하고 있는 퇴원손상심층조사 자료를 활용하였다.

퇴원손상심층조사 자료는 통계청 승인번호 117060에 해당하는 국가승인통계로 2005년 첫 시행된 자료에는 부진단 코드가 7개까지 수집되었으나 이후 조사부터는 최대 20개까지 조사되었다[16]. 조사항목의 일관성을 유지하기 위하여 본 연구에서는 첫 조사자료인 2005년도를 제외하고 2006년부터 최근까지 공개된 2016년 자료인 총 11년 간의 자료를 활용하고자 하였다. 질병관리청에서 제시하는 자료제공 절차에 따라 이용계획서와 서약서를 제출하고 자료를 획득하였으며, 조사 자료 중 개인정보나 의료기관을 확인할 수 있는 정보는 포함되지 않아 연구활용에는 문제가 없었다.

퇴원손상심층조사 자료는 100명 이상 의료기관의 퇴원환자를 대상으로 표본병원과 병원내 표본환자를 추출 틀로 하는 층화이단집락 추출법이 적용되어 매년 20만 건 이상의 자료가 수집된다. 본 연구를 위해 데이터베이스 관리 소프트웨어인 MySQL을 이용하여 연도별 자료를 통합하였다. 동반질환의 연관성 규칙 분석이 용이한 매트릭스 형태로 추출하기 위해 정규화 과정을 거쳐 3개의 테이블로 이루어진 통합 데이터베이스를 구축하였다. 부진단을 보조적으로 설명하는 건강 상태 및 보건서비스 접촉에 영향을 주는 요인(Z00-Z99)과 독립된(일차성) 여러 부위의 악성 신생물(C97)은 제외하였다. 이러한 과정을 거쳐서 최종 연구대상에 포함된 폐암 퇴원환자 총 21,882명과 이들의 동반질환에 해당하는 부진단은 54,337건이 추출되었다.

변수정의

본 연구대상을 선정하는데 중요한 변수인 주진단은 질병을 진단하기 위해 검사 후 확인된 최종 진단을 말한다. 의료기관에 방문하거나 치료를 요하는 가장 중요한 질환을 말하며, 때로는 이와 관련없이 새로운 질병이 발견되고 의료자원 소모가 많은 주요 질병을 주진단으로 선정하기도 한다[9,17]. 주진단은 퇴원 시까지 확인된 한 개의 질병이 반드시 등록되는 중요한 병태이며, 주된병태(main condition)라 일컫는다. 부진단은 입원기간동안 발견된 주진단 이외의 질병을 말하며 환자

에 따라서는 부진단이 존재하지 않기도 한다[18,19].

의료기관의 퇴원환자는 한국표준질병사인분류 기준에 따라 주진단과 부진단을 코드화하여 퇴원요약자료에 입력하게 된다. 현재 우리나라 의료기관에서 적용되는 한국표준질병사인분류는 7차 개정판이며 이는 대분류 22개, 중분류 267개, 소분류 2,082개, 세분류 12,505개로 구성되어 있다. 본 연구에서 연구대상자인 폐암환자의 한국표준질병사인분류 코드는 C33-C34.9이다. 폐암환자의 동반질환을 분석하기 위해 퇴원손상심층조사 자료에 입력된 최대 20개의 부진단 코드를 활용하였는데, 한국표준질병사인분류 7차 개정판의 중분류를 기준으로 이들 질환을 그룹화하였다.

인구학적 특성에 해당하는 나이는 생애주기별 연령그룹을 기준으로 19-44세, 45-64세, 65-74세, 75세 이상으로 분류하였다. 재원일수는 퇴원일에서 입원일을 뺀 단위법을 적용하였으며, 이는 의료기관에서 환자실적통계 시 활용하는 방법이다. 수술여부는 주수술코드가 있는 경우와 없는 경우를 이원화하여 분류하였다. 기타 변수는 퇴원손상심층조사 자료의 분류기준을 따랐으며 일부 변수에서 빈도가 적은 것은 기타로 묶어서 분류하였다.

분석방법

앞서 설명한 통합 데이터베이스에서 structured query language (SQL)을 이용하여 연구대상과 관련 변수를 추출하였으며, 자료분석을 위하여 오픈소스인 R version 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria)를 활용하였다. 일반적 특성은 빈도분석을 시행하였고 동반질환과의 독립성 검정을 위해 카이제곱 검정을 시행하였다. 통계적 검정은 5% 유의수준을 기준으로 하였다. 본 연구의 목적을 달성하기 위하여 데이터마이닝 기법인 연관성 규칙을 분석에 활용하였는데, R프로그램의 arules package와 apriori algorithm을 이용하여 연관규칙에 대한 지지도, 신뢰도, 향상도를 분석하고 시각화하였다. Apriori algorithm은 연관성 규칙을 탐색하기 위해 가장 먼저 개발되었으며, 사건(event)들의 발생 빈도를 중심으로 한 대표적인 알고리즘이다. 보건의료분야에도 임상적 의사결정, 질병요인의 탐색, 약물이용 등에 활용되고 있어 본 연구에서도 이와 같은 방법을 적용하였다[20-22]. 타당한 규칙을 찾기 위한 지표로 지지도(support), 신뢰도(confidence), 향상도(lift)를 제시하고 있다[10]. 지지도는 사건 A와 B가 동시에 일어날 확률을 말한다. 본 연구에서는 전체 환자 중 동반질환 A와 B가 함께 진단된 비율로 계산될 수 있다(Equation 1). 예를 들어 지지도(고혈압→당뇨병)가 10%라는 의미는 전체 환자 중 고혈압과 당뇨병을 함께 진단받은 환자가 10%라는 의미이다.

$$\text{Support}(\text{disease A} \rightarrow \text{disease B}) = \frac{P(\text{disease A, disease B})}{\text{Total number of patient}} \dots\dots\dots \text{Equation 1}$$

신뢰도는 연관성 강도를 측정하는 것으로 A사건 중에서 B사건이 일어나는 확률을 의미한다. 본 연구에서는 동반질환 A를 가진 환자 중에서 동반질환 B를 가진 환자의 비율로 계산되며(Equation 2), 신뢰도(고혈압→당뇨병)가 50%라는 의미는 고혈압을 진단받은 환자 중에서 당뇨병도 진단받은 환자의 비율이 50%라는 의미이다.

$$\text{Confidence}(\text{disease A} \rightarrow \text{disease B}) = \frac{P(\text{disease A, disease B})}{P(\text{disease A})} \dots\dots\dots \text{Equation 2}$$

마지막으로 향상도는 연관규칙의 효용성을 판단할 수 있는 지표로, 연관규칙 ‘사건 A→사건 B’에서 사건 A와 B가 독립적으로 일어나는 경우에 비해 함께 일어나는 확률을 의미한다. 즉, 동반질환 A와 B가 독립적으로 일어나는 것과 비교해 동시에 발생하는 비율로 계산된다(Equation 3). 향상도(고혈압→당뇨병)가 1이라는 의미는 고혈압과 당뇨병이 독립적으로 진단받은 것에 비해 고혈압과 당뇨병을 함께 진단받은 비율이 같으므로 연관성이 없다는 의미이나 향상도가 2라면 두 질환이 독립적으로 진단받는 것에 비해 함께 진단받는 것이 2배로 긍정적이라는 것을 나타낸다.

$$\text{Lift}(\text{disease A} \rightarrow \text{disease B}) = \frac{P(\text{disease A, disease B})}{P(\text{disease A}) \cdot P(\text{disease B})} \dots\dots\dots \text{Equation 3}$$

동반질환의 분류기준인 중분류는 최대 267개의 질환군으로 분류되는 본 연구의 자료 특성을 감안하여 지지도가 >1%, 신뢰도는 >10% 기준을 적용하였다. 정제된 연관규칙을 찾기 위해서 향상도가 1을 초과하는 항목들만 최종 채택하였다. 향상도가 1인 경우는 우연한 패턴을 의미하며, 1미만인 경우 관련이 적은 패턴이므로 연구결과에서 제외하였다. 또한 향상도를 조절하여 주요한 연관규칙을 추가 탐색하였다.

연구 결과

연구대상자의 특성

연구대상자의 일반적 특성 및 동반질환 유무에 따른 분포를 분석한 결과 Table 1과 같다. 동반질환을 보유한 환자는 82.86%이며, 남자 83.13%, 여자 82.16%로 비슷한 수준으로 통계적으로 유의한 수준은 아니었다. 평균연령은 동반질환이 있는 폐암 환자가 다소 높게 나타났으며, 연령군별로는 65-74세 83.87%, 75세 이상 82.95%, 45-64세 82.30%, 19-44세 77.23% 순으로 나타나 차이를 보였다. 진료비 지불방법별로 동반질환을 가진 환자의 비율은 기타가 87.68%인 반면 의료급여 2종은 82.41%로 차이가 있었다. 응급실을 통해 입원한 환자의 88.67%는 동반질환이 있었으나 외래와 기타 입원경로 환자는 각각 80.85%, 44.83%로

Table 1. Baseline characteristics of lung cancer inpatients by comorbidity

Variables	Comorbidity		Total	χ^2 or t	p
	With	Without			
Sex				2.874	0.090
Male	13,143 (83.13)	2,668 (16.87)	15,811 (72.26)		
Female	4,988 (82.16)	1,083 (17.84)	6,071 (27.74)		
Age (y)	65.95 $\pm$ 10.50	65.09 $\pm$ 11.12	65.81 $\pm$ 10.61	-4.400	< 0.001
Age group (y)				25.401	< 0.001
19-44	607 (77.23)	179 (22.77)	786 (3.59)		
45-64	6,688 (82.30)	1,438 (17.70)	8,126 (37.14)		
65-74	7,076 (83.87)	1,361 (16.13)	8,437 (38.56)		
≥ 75	3,760 (82.95)	773 (17.05)	4,533 (20.72)		
Insurance type				3.629	0.304
National health	16,305 (82.77)	3,393 (17.23)	19,698 (90.02)		
Medicaid I	1,559 (83.24)	314 (16.76)	1,873 (8.56)		
Medicaid II	89 (82.41)	19 (17.59)	108 (0.49)		
Others	178 (87.68)	25 (12.32)	203 (0.93)		
Admission route				212.438	< 0.001
Emergency	5,110 (88.67)	653 (11.33)	5,763 (26.34)		
Outpatient	13,008 (80.85)	3,082 (19.15)	16,090 (73.53)		
Others	13 (44.83)	16 (55.17)	29 (0.13)		
Treatment outcome				128.477	< 0.001
Improved	13,911 (82.83)	2,883 (17.17)	16,794 (76.75)		
Not improved	1,814 (76.54)	556 (23.46)	2,370 (10.83)		
Death	2,262 (88.39)	297 (11.61)	2,559 (11.69)		
Others	144 (90.57)	15 (9.43)	159 (0.73)		
Numbers of comorbidity	3.00 $\pm$ 2.07	-	2.48 $\pm$ 2.19	N/A	N/A
Operation				25.236	< 0.001
Yes	2,356 (79.62)	603 (20.38)	2,959 (13.52)		
No	15,775 (83.36)	3,148 (16.64)	18,923 (86.48)		
ALOS (day)	13.23 $\pm$ 17.93	8.30 $\pm$ 9.82	12.39 $\pm$ 16.92	-23.667	< 0.001
Disposition				66.952	< 0.001
Return-home	14,734 (81.99)	3,237 (18.01)	17,971 (82.13)		
Other hospital	1,058 (84.24)	198 (15.76)	1,256 (5.74)		
Death	2,262 (88.39)	297 (11.61)	2,559 (11.69)		
Others	77 (80.21)	19 (19.79)	96 (0.44)		
Bed size				334.154	< 0.001
100-299	756 (66.26)	385 (33.74)	1,141 (5.21)		
300-499	1,297 (78.09)	364 (21.91)	1,661 (7.59)		
500-999	9,879 (82.64)	2,075 (17.36)	11,954 (54.63)		
$\geq 1,000$	6,199 (86.99)	927 (13.01)	7,126 (32.57)		
Total	18,131 (82.86)	3,751 (17.14)	21,882 (100.0)		

Unit: n (%) or mean $\pm$ standard deviation.

ALOS, average length of stay; N/A, not applicable.

차이가 있었다. 치료결과에서는 사망환자와 기타가 각각 88.39%, 90.57%로 평균보다 높고 완쾌와 치료 안함은 이보다 낮은 82.83%, 76.54%로 분석되었다. 동반질환의 개수는 전체 환자 평균 2.48개를 보였고, 동반질환을 진단받은 환자군에서는 평균 3개의 동반질환이 있었다. 입원 기간 중 수술을 받은 환자보다 그렇지 않은 환자의 동반질환 비율이

높게 나타나 통계적으로 유의한 차이를 보였다. 평균재원일수는 동반질환이 있는 환자가 그렇지 않은 환자에 비해 평균 5일 가량 높게 나타났다. 환자가 입원한 의료기관의 병상별로 분석한 결과 병상이 클수록 동반질환 비율이 높게 나타났으며, 통계적으로 유의하였다.

Table 2. Distribution of 20 comorbidities of study population base on multiple response

	Comorbidities	n	%
C76-C80	Malignant neoplasms of ill-defined, secondary and unspecified sites	16,131	73.72
I10-I15	Hypertensive diseases	5,213	23.82
E10-E14	Diabetes mellitus	3,175	14.51
J09-J18	Influenza and pneumonia	2,341	10.70
J40-J47	Chronic lower respiratory diseases	1,925	8.80
K20-K31	Diseases of esophagus, stomach and duodenum	1,522	6.96
J90-J94	Other diseases of pleura	981	4.48
I30-I52	Other forms of heart disease	949	4.34
N40-N51	Diseases of male genital organs	730	3.34
D60-D64	Aplastic and other anemias	688	3.14
B90-B94	Sequelae of infectious and parasitic diseases	630	2.88
I20-I25	Ischemic heart diseases	602	2.75
E70-E90	Metabolic disorders	601	2.75
I60-I69	Cerebrovascular diseases	572	2.61
D70-D77	Other diseases of blood and blood-forming organs	502	2.29
I80-I89	Diseases of veins, lymphatic vessels and lymph nodes, NEC	464	2.12
J95-J99	Other diseases of the respiratory system	456	2.08
J60-J70	Lung diseases due to external agents	432	1.97
C15-C26	Malignant neoplasms of digestive organs	422	1.93
N17-N19	Renal failure	395	1.81

동반질환 순위

연구대상자의 다빈도 동반질환을 분석한 결과 Table 2와 같다. 한국 표준질병사인분류에서 중분류를 기준으로 폐암환자의 동반질환을 0과 1로 코딩한 후 다중응답분석을 이용하였다. 연구대상자 21,882명 중 1% 이상 비율을 가진 동반질환을 정렬한 결과 20개의 질환이 분석되었다. 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 동반된 비율이 73.2%로 가장 높게 나타났으며, 다음으로 고혈압성 질환(I10-I15) 23.82%, 당뇨병(E10-E15) 14.51%, 인플루엔자 및 폐렴(J09-J18) 10.70% 순이었다. 이외에도 호흡계통, 위장계통, 심혈관질환을 포함하여 대사장애, 신부전, 소화기관의 악성 신생물 등이 함께 폐암과 함께 진단되는 것으로 나타났다.

동반질환 연관성 규칙 분석결과

폐암환자의 동반질환 특성을 분석하기 위해 데이터마이닝 기법 중의 하나인 연관성 규칙을 이용하였다. 초기 모델에서는 44개의 규칙이 분석되었으나 향상도(lift)가 1 이하인 규칙을 제외하고 지지도가 높은 순으로 정렬한 결과 Table 3과 같이 33개의 규칙이 정리되었다. 33개의 연관성 규칙에서 향상도는 최소 1.002 (규칙 5, 6)에서 최대 2.240 (규칙

Table 3. Association rules of comorbidities in patients hospitalized for lung cancer

No	Rules	Support	Confidence	Lift	Count
1	I10-I15→C76-C80	0.105	0.441	1.032	2,294
2	C76-C80→I10-I15	0.105	0.245	1.032	2,294
3	E10-E14→I10-I15	0.076	0.533	2.240	1,669
4	I10-I15→E10-E14	0.076	0.321	2.240	1,669
5	E10-E14→C76-C80	0.061	0.428	1.002	1,341
6	C76-C80→E10-E14	0.061	0.143	1.002	1,341
7	C76-C80, E10-E14→I10-I15	0.032	0.524	2.205	703
8	C76-C80, I10-I15→E10-E14	0.032	0.306	2.141	703
9	J09-J18→I10-I15	0.025	0.241	1.014	558
10	I10-I15→J09-J18	0.025	0.107	1.014	558
11	J40-J47→I10-I15	0.025	0.297	1.250	542
12	I10-I15→J40-J47	0.025	0.104	1.250	542
13	K20-K31→C76-C80	0.024	0.460	1.078	529
14	J09-J18→E10-E14	0.018	0.173	1.210	401
15	E10-E14→J09-J18	0.018	0.128	1.210	401
16	I30-I52→C76-C80	0.017	0.451	1.055	378
17	D60-D64→C76-C80	0.017	0.550	1.289	377
18	K20-K31→I10-I15	0.016	0.298	1.252	342
19	J40-J47→J09-J18	0.016	0.187	1.768	341
20	J09-J18→J40-J47	0.016	0.147	1.768	341
21	J40-J47→E10-E14	0.014	0.170	1.188	310
22	I30-I52→I10-I15	0.014	0.368	1.549	309
23	E70-E90→C76-C80	0.013	0.505	1.183	286
24	N40-N51→I10-I15	0.013	0.389	1.637	281
25	I20-I25→I10-I15	0.013	0.487	2.047	275
26	I60-I69→I10-I15	0.012	0.496	2.088	271
27	C76-C80, J09-J18→I10-I15	0.011	0.255	1.074	248
28	I10-I15, J09-J18→C76-C80	0.011	0.444	1.041	248
29	C76-C80, I10-I15→J09-J18	0.011	0.108	1.023	248
30	I60-I69→C76-C80	0.011	0.449	1.051	245
31	I80-I89→C76-C80	0.011	0.536	1.254	241
32	E70-E90→I10-I15	0.011	0.419	1.761	237
33	D70-D77→C76-C80	0.011	0.468	1.096	234

C76-C80, Malignant neoplasms of ill-defined, secondary and unspecified sites; D60-D64, Aplastic and other anemias; D70-D77, Other diseases of blood and blood-forming organs; E10-E14, Diabetes mellitus; E70-E90, Metabolic disorders; I10-I15, Hypertensive diseases; I20-I25, Ischemic heart diseases; I30-I52, Other forms of heart disease; I60-I69, Cerebrovascular diseases; I80-I89, Diseases of veins, lymphatic vessels and lymph nodes, NEC; J09-J18, Influenza and pneumonia; J40-J47, Chronic lower respiratory diseases; K20-K31, Diseases of esophagus, stomach and duodenum; N40-N51, Diseases of male genital organs.

3, 4)까지 분포하였다(Figure 1, Table 3). 1 이하의 향상도는 우연에 의한 규칙이거나 관계가 낮은 것을 의미한다.

고혈압성 질환(I10-I15)과 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 함께 동반된 규칙1과 규칙2의 지지도가 10.5%로 가장 높게 나타났다. 다음으로 당뇨병(E10-E14)과 고혈압성 질환

Figure 1. Scatter plot for association rules analysis.

(I10-I15)이 함께 동반한 경우(규칙 3, 4)로 지지도 7.6%이며, 당뇨병(E10-E14)과 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)의 지지도(규칙 5, 6)는 6.1%로 분석되었다. 규칙 1에서 6의 경우 서로 쌍방으로 지지도가 나타나는 특성을 가진다.

연관성 규칙에서 신뢰도 분석결과를 살펴보면, 혈액 및 조혈기관의 기타 질환(D70-D77)이 동반된 환자중에서 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)을 함께 가지고 있는 비율이 55.0%로 가장 높고(규칙 17), 달리 분류되지 않은 정맥, 림프관 및 림프절의 질환(I80-I89)이 있는 환자중에서 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 동반된 경우(규칙 31)가 53.6%로 나타났다. 다음은 당뇨병(E10-E14)이 있는 경우 고혈압성 질환(I10-I15)이 동반된 경우(규칙 3)로 신뢰도 53.3%이며 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)과 당뇨병(E10-E14)이 함께 동반된 환자 중에서 고혈압성 질환(I10-I15)이 동반된(규칙 7) 비율은 52.4%로 분석되었다. 대사장애(E70-E90)가 동반된 환자에서 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 동반된(규칙 23) 비율은 50.5%를 보였다.

연관성 규칙에 나타난 주요 동반질환은 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80), 무형성 및 기타 빈혈(D60-D64), 혈액 및 조혈기관의 기타 질환(D70-D77), 당뇨병(E10-E14), 대사장애(E70-E90), 고혈압성 질환(I10-I15), 허혈심장질환(I20-I25), 기타 형태의 심장병(I30-I52), 뇌혈관질환(I60-I69), 달리 분류되지 않은 정맥, 림프관 및 림프절의 질환(I80-I89), 인플루엔자 및 폐렴(J09-J18), 만성 하부호흡기질환(J40-J47), 식도, 위 및 십이지장의 질환(K20-K31), 남성생식기관의 질환(N40-N51)으로 요약된다.

연관성 규칙을 네트워크 그래프로 시각화한 결과 Figure 2와 같다. 원의 크기가 클수록 지지도가 높은 것을 의미하며, 원의 색상이 진할

Figure 2. Network graph for association rules analysis. (A) 33 rules for lift >1, (B) 11 rules for lift >1.

수록 향상도가 높은 것을 의미한다. 33개 규칙을 시각화한 Figure 2A에서 규칙 3과 4의 지지도가 2.240으로 가장 높고 다음으로 규칙 7이 2.205, 규칙 8이 2.141, 규칙 26이 2.088, 규칙 25가 2.047순으로 주요하게 높게 나타났다. 또한 다수의 연결 경로를 가지는 주요 동반질환은 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80), 당뇨병(E10-E14), 고혈압성 질환(I10-I15), 인플루엔자 및 폐렴(J09-J18)인 것을 확인할 수 있다. 향상도 1.5 이상의 11개 연관규칙을 요약한 네트워크 그래프는 Figure 2B와 같은데, 여기에서도 주요 동반질환이 이차성 및 상세불명 부위의 악성 신생물(C76-C80), 당뇨병(E10-E14), 고혈압성 질환(I10-I15)이 시각적으로 나타나는 것을 볼 수 있다.

고찰 및 결론

폐암은 암질환 중에서도 흔하게 발병하고 사망률이 매우 높은 질환으로 알려져 있으며, 우리나라뿐만 아니라 세계적으로도 주목하고 있는 실정이다[23,24]. 지금까지 예방과 치료중심의 연구가 지속되고 있지만 환경오염과 고령화에 의한 인구구조의 변화로 인하여 폐암 발생률과 사망률은 앞으로도 증가추세가 이어질 전망이다[25]. 이러한 상황은 의료비 증가를 가속화할 수 있기 때문에 폐암환자의 관리를 위한 다양한 해결방안을 모색할 필요가 있다. 동반질환은 재원기간이나 의료서비스 이용을 늘리거나[6] 생존율에도 영향을 미치고[7] 있기 때문에 질병 관리를 효과적으로 하기 위해서는 이를 규명할 필요가 있다. 이를 위해 본 연구에서는 퇴원손상심층조사 자료를 이용하여 폐암환자의 동반질환에 대한 연관규칙을 실증하고자 하였다.

우리나라 입원암환자의 동반질환 연구를 살펴보면, 75%의 환자가 동반질환을 가지고 있는 것으로 나타났으나[9] 본 연구에서는 이보다 8%p 가까이 높은 82.86%로 분석되어 폐암환자의 동반질환 관리가 더욱 필요함을 시사한다. 미국의 Nebraska Hospital Discharge Data (NHDD) 자료를 이용한 폐암환자의 동반질환 유병률은 73.3%이었으며[7], 중국의 보험청구 데이터를 이용한 연구에서는 이보다 훨씬 낮은 31.3%이었다[6]. 전자의무기록을 기반으로 65세 폐암환자 대상의 연구에서는 77.5%로 나타나[7] 이용자료와 대상에 따라 차이가 있음을 알 수 있다. 본 연구는 일부 지역이나 연령제한을 두거나 청구자료가 아닌 전국 추정이 가능한 의무기록 기반의 자료를 활용하였기에 선행연구보다 높은 결과를 보이는 것으로 판단된다. 인구학적 특성별로 동반질환의 분포를 살펴보면, 성별에 따른 차이는 없었으나 연령이 증가할수록, 응급실을 통해 입원한 경우, 사망환자, 입원 중 수술이 없었던 환자, 의료기관의 병상수가 클수록 폐암환자의 동반질환 비율이 유의하게 높은 것으로 나타났다. 전체 암질환을 대상으로 한 연구[9]에서도 본 연구와 같은 결과를 보이고 있어 선행연구를 지지하는 것으로 나타났다. 따라서 이들 특성을 고려하여 동반질환 발생을 예측하고 관리할 수 있는 협진체계를 구축하는 것이 필요하다.

폐암환자의 동반질환에 대하여 다중응답분석을 실시한 결과 불명확한 이차성 및 상세불명 부위의 악성 신생물(C76-C80) 73.72%, 고혈압성 질환(I10-I15) 23.82%, 당뇨병(E10-E14) 14.51%, 인플루엔자 및 폐렴(J09-J18) 10.70%, 만성 하부호흡기질환(J40-J47) 8.80% 등의 순으로 나타났다. Kim [9]의 연구에서도 전이성 암을 포함한 신생물의 비율이 가장 높고 순환기계질환, 내분비질환이 많이 동반되는 것으로 나타나 본 연구결과와 일치하였다. 선행연구에서 심혈관계질환은 폐암환자의 23-25%에서 동반되는 것으로 알려져 있는데[25,26], 흡연, 연령과 같은 폐암의 위험요인을 공통적으로 가지고 있기 때문에 설명된다[27,28].

본 연구에서는 기타 형태의 심장병(I30-I52)과 허혈심장질환(I20-I25)을 포함하여 고혈압성 질환(I10-I15)의 동반비율이 23.82%로 비슷한 수준을 보여 선행연구를 지지하는 것으로 나타났다. Leduc et al. [23]의 연구에서 암환자의 8-18%가 당뇨병이 동반되었는데, 이러한 이유는 나이, 식습관, 흡연과 같은 위험요인이 겹치기 때문이다. 본 연구에서도 당뇨병의 동반비율이 14.51%로 암환자와 비슷한 수준이었다. 폐암은 연령이 높을수록 발병 가능성이 높고 면역력 저하로 인하여 폐렴에 쉽게 노출될 수 있기 때문에 함께 동반되는 경우가 흔하며 폐구균성폐렴은 폐암발병을 높인다[29]. 본 연구에서도 폐암환자에서 10% 이상의 폐렴 유병률을 보이고 있어 이를 뒷받침하고 있는 것으로 나타났다. 만성폐쇄성폐질환이나 폐기종, 기관지확장증과 같은 만성 하부호흡기질환(J40-J47)은 폐암과 밀접하게 관련이 있고 유병률이 25-50%에 이른다[23]. 90% 이상 대부분의 원인은 앞선 공통 위험요인과 마찬가지로 흡연이지만 이외에도 면역기능장애, 폐 미생물 군집과 Deoxyribonucleic Acid (DNA) 메틸화 및 마이크로(Ribonucleic Acid) 발현을 포함하는 후성 유전학적 요인이 모두 관여하는 특징을 가지고 있다[30]. 기타 남성생식기관의 질환(N40-N51), 무형성 및 기타 빈혈(D60-D64), 대사장애(E70-E90) 등이 있었으나 유병률은 주요 동반질환 보다 낮은 수준이다.

본 연구는 기술통계분석을 이용한 동반질환 규명에 그치지 않고 연관관계 규칙 분석을 이용하여 이들 질환 간의 패턴을 확인하고자 하였다. 일반적으로 연관관계 규칙에서 최소 지지도와 신뢰도는 연구자료의 특성에 따라 다르게 설정한다. 응급실을 통해서 입원한 천식환자의 연관성 규칙을 분석한 선행연구에서는 지지도 $\geq 0.5\%$, 신뢰도 $\geq 80\%$ 기준을 적용하여 2개의 연관규칙을 도출하였다[31]. 65세 이상 폐암환자의 동반질환을 분석한 연구에서는 지지도 $>1\%$, 신뢰도 $>50\%$ 를 기준으로 하였다[10]. 본 연구에서도 이러한 선행연구를 참고하여 지지도 $>1\%$ 와 신뢰도 $>20\%$ 를 설정하여 결과의 타당성을 확보하고자 노력하였다. 분석결과 총 33개의 규칙이 도출되었고 신뢰도가 높은 순으로 주요 패턴을 요약하면 다음과 같다. 폐암환자 중 혈액 및 조혈기관의 기타 질환(D70-D77)이 동반될 경우, 불명확한 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 함께 나타날 확률이 55.0%, 달리 분류되지 않은 정맥, 림프관 및 림프절의 질환(I80-I89)이 있는 환자 중에서 불명확한 이차성 및 상세불명 부위의 악성 신생물(C76-C80) 동반 확률은 53.6%로 신뢰도가 높았다. 특히 빈혈은 폐암과 만성 폐 질환에서 가장 흔한 유형이기 때문에[32,33] 폐암 환자의 예후 지표로 유용하다는 근거가 된다. 폐암환자 중 당뇨병(E10-E14)과 고혈압성 질환(I10-I15)이 함께 나타나는 확률은 7.6%를 차지하지만 당뇨병이 있는 환자 중에서 고혈압의 동반비율은 53.3%에 달하는 것으로 확인되어 고혈압 진단을 선제적으로 예측해볼 수 있을 것이다. 폐암환자에서 고혈압을 포

함하여 뇌경색, 심부전, 심장 부정맥 등 심혈관 합병증 유병률은 12.9-43%에 이르며[7, 34] 뇌혈관질환과 동반된 비소세포폐암(Non-Small Cell Lung Cancer) 환자는 사망률이 20% 증가된다는 보고[35]가 있으므로 이들 동반질환에 유의해야 할 필요가 있다. 고혈압성 질환(I10-I15)은 이차성 및 상세불명 부위의 악성 신생물(C76-C80)과 당뇨병(E10-E14)이 있는 환자 중에서도 동반될 확률이 52.4%에 달하기 때문에 폐암환자의 고혈압과 당뇨병 관리는 매우 중요한 부분임을 입증하고 있다. 대사장애(E70-E90)가 있는 환자 중에서도 이차성 및 상세불명 부위의 악성 신생물(C76-C80)이 동반되는 비율이 50.5%에 달하는 것으로 분석되었다. 대사장애는 당뇨병과 더불어 폐암과 함께 동반할 경우 그렇지 않은 폐암환자에 비하여 사망률이 유의하게 높기[35] 때문에 더욱 철저한 환자관리가 필요하다. 다수의 연관규칙을 네트워크 그래프를 통해서 연관경로가 집중되는 주요 동반질환을 시각화하였다(Figure 2). 33개의 전체 연관규칙에서는 불명확한, 이차성 및 상세불명 부위의 악성 신생물(C76-C80), 당뇨병(E10-E14), 고혈압성 질환(I10-I15), 인플루엔자 및 폐렴(J09-J18)에 집중되고 향상도 >1.5의 11개 연관규칙을 나타낸 네트워크 그래프에서는 고혈압성 질환(I10-I15)에 경로가 집중되는 것을 쉽게 확인할 수 있다.

이상으로 임상적 근거를 토대로 본 연구결과의 연관규칙을 살펴본 바 고혈압, 당뇨, 뇌혈관 질환 등 주요 동반질환의 철저한 관리가 필요하다. 이를 위해 첫째, 의료기관은 이를 활용한 환자관리와 협진체계를 구축하고 임상 의사결정시스템에 이를 반영하여 지속적인 모니터링 적용을 제안한다. 둘째, 의료기관은 그 동안 정보화 수준이 향상되고 환자의 대용량 데이터를 확보하고 있기 때문에 폐암과 같은 주요 질환의 동반질환 관리전략을 마련하는데 이와 같은 연관성 규칙을 적극 활용할 수 있기를 기대한다. 셋째, 국가 암관리정책은 의료기관의 협진체계와 진료전달체계를 지원하는 방안을 고도화하고 향후 평생건강 기록(electronic health record, EHR)을 통해 질환별 정밀의료를 발전시킬 수 있을 것이다.

본 연구는 국가에서 체계적으로 수집된 퇴원손상심층조사 자료를 이용하여 폐암환자의 동반질환과 이들의 패턴을 실증하여 임상적 타당성을 고찰하고 시사점을 제공하였다는데 의의가 있음에도 불구하고 몇 가지 제한점을 가진다. 첫째, 퇴원손상심층조사의 특성상 100병상 미만 의료기관의 환자는 연구대상에서 제외되었기 때문에 다소 차이가 있을 수 있다. 둘째, 전국규모 추정자료임에도 불구하고 연관성 규칙 분석에서 가중치를 적용하지 못하였다. 향후 연구에서는 이러한 문제를 해결할 수 있는 알고리즘을 개발하여 분석의 타당성을 높이는 노력이 필요하다. 셋째, 폐암환자의 질병 에피소드를 추적관찰하여 포괄적인 동반질환의 연관성을 규명할 수 있는 코호트 연구가 진행될 필요가 있다. 마지막으로 발생확률이 다소 낮은 지지도나 신뢰도를 보이

는 동반질환이라고 하더라도 의료의 특성상 이를 간과해서는 안될 것이다.

ORCID

Sung-Soo Kim <https://orcid.org/0000-0002-7205-0758>

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6):394-424. DOI: 10.3322/caac.21492
2. Goo JM. Lung Cancer Screening with Low-Dose CT: current status in other countries. *J Korean Soc Radiol* 2019;80(5):849-859 (Korean). DOI: 10.3348/jksr.2019.80.5.849
3. World Health Organization. Globocan 2020. Available at <https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf> [accessed on January 17, 2021].
4. Hong S, Won YJ, Park YR, Jung KW, Kong HJ, Lee ES, et al. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2017. *Cancer Res Treat* 2020;52(2):335-350 (Korean). DOI: 10.4143/crt.2020.206
5. Jang SH. Korean national lung cancer screening. *Korean J Med* 2020; 95(2):95-103 (Korean). DOI: 10.3904/kjm.2020.95.2.95
6. Ding R, Zhu D, He P, Ma Y, Chen Z, Shi X. Comorbidity in lung cancer patients and its association with medical service cost and treatment choice in China. *BMC Cancer* 2020;20(1):250. DOI: 10.1186/s12885-020-06759-8
7. Islam KMM, Jiang X, Anggondowati T, Lin G, Ganti AK. Comorbidity and survival in lung cancer patients. *Cancer Epidemiol Biomarkers Prev* 2015;24(7):1079-1085. DOI: 10.1158/1055-9965.epi-15-0036
8. Sandfeld-Paulsen B, Meldgaard P, Aggerholm-Pedersen N. Comorbidity in lung cancer: a prospective cohort study of self-reported versus register-based comorbidity. *J Thorac Oncol* 2018;13(1):54-62. DOI: 10.1016/j.jtho.2017.10.002
9. Kim SS. Analysis of variation of length of stay according to comorbidities in cancer inpatient. *J Health Info Stat* 2020;45(1):80-90 (Korean). DOI: 10.21032/jhis.2020.45.1.80
10. Feng J, Mu XM, Ma LL, Wang W. Comorbidity patterns of older lung

- cancer patients in Northeast China: an association rules analysis based on electronic medical records. *Int J Environ Res Public Health* 2020; 17(23):9119. DOI: 10.3390/ijerph17239119
11. Lakshmi KS, Vadivu G. A novel approach for disease comorbidity prediction using weighted association rule mining. *J Ambient Intell Human Comput* 2019. DOI: 10.1007/s12652-019-01217-1
12. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017;2(4): 230-243. DOI: 10.1136/svn-2017-000101
13. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. In: *Proceedings of the 1993 ACM SIGMOD international conference on management of data, 1993*, p. 207-216.
14. Valle MA, Ruz GA, Morrás R. Market basket analysis: complementing association rules with minimum spanning trees. *Expert Sys Appl* 2018; 97:146-162. DOI: 10.1016/j.eswa.2017.12.028
15. Hahsler M. *arulesViz: Interactive visualization of association rules with R*. *R J* 2017;9(2):163-175. DOI: 10.32614/RJ-2017-047
16. Korea Centers for Disease Control and Prevention. *Korean National Hospital Discharge In-depth Injury Survey guidebook*. Cheongju: Korea Centers for Disease Control and Prevention; 2015 (Korean).
17. Sukanya C. Validity of principal diagnoses in discharge summaries and ICD-10 coding assessments based on national health data of Thailand. *Healthc Inform Res* 2017;23(4):293-303. DOI: 10.4258/hir.2017.23.4.293
18. Ngene NC, Moodley J. Assigning appropriate and comprehensive diagnosis for scientific report. *Med Hypotheses* 2014;83(6):681-684. DOI: 10.1016/j.mehy.2014.09.017
19. Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P. Impact of comorbidity on lung cancer survival. *Int J Cancer* 2003;103(6):792-802. DOI: 10.1002/ijc.10882
20. Nahar J, Imam T, Tickle KS, Chen YPP. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Syst Appl* 2013;40(4):1086-1093. DOI: 10.1016/j.eswa.2012.08.028
21. Cheng CW, Chanani N, Venugopalan J, Maher K, Wang MD. *icuARM- An ICU clinical decision support system using association rule mining*. *IEEE J Transl Eng Health Med* 2013;1(1):122-131. DOI: 10.1109/JTEHM.2013.2290113
22. Pattanapruteep O, McEvoy M, Attia J, Thakkestian A. Evaluation of rational nonsteroidal anti-inflammatory drugs and gastro-protective agents use; association rule data mining using outpatient prescription patterns. *BMC Med Inform Decis Mak* 2017;17(1):96. DOI: 10.1186/s12911-017-0496-3
23. Leduc C, Antoni D, Charloux A, Falcoz PE, Quoix E. Comorbidities in the management of patients with lung cancer. *Eur Respir J* 2017; 49(3):1601721. DOI: 10.1183/13993003.01721-2016
24. Xu C, Yang H, Sun L, Cao X, Hou Y, Cai Q, et al. Detecting lung cancer trends by leveraging real-world and internet-based data: infodemiology study. *J Med Internet Res* 2020;22(3):e16184-e16184. DOI: 10.2196/16184
25. Sigel K, Wisnivesky JP. Comorbidity profiles of patients with lung cancer: a new approach to risk stratification?. *Ann Am Thorac Soc* 2017; 14(10):1512-1513. DOI: 10.1513/AnnalsATS.201706-442ED
26. Janssen-Heijnen ML, Schipper RM, Razenberg PP, Crommelin MA, Coebergh JW. Prevalence of co-morbidity in lung cancer patients and its relationship with treatment: a population-based study. *Lung Cancer* 1998;21(2):105-113. DOI: 10.1016/s0169-5002(98)00039-7
27. Hatlen P, Langhammer A, Carlsen SM, Salvesen Ø, Amundsen T. Self-reported cardiovascular disease and the risk of lung cancer, the HUNT Study. *J Thorac Oncol* 2014;9(7):940-946. DOI: 10.1097/JTO.000000-0000000180
28. Lee SW, Kim HC, Lee HS, Suh I. Thirty-year trends in mortality from cardiovascular diseases in Korea. *Korean Circ J* 2015;45(3):202-209 (Korean). DOI: 10.4070/kcj.2015.45.3.202
29. Lin TY, Huang WY, Lin JC, Lin CL, Sung FC, Kao CH, et al. Increased lung cancer risk among patients with pneumococcal pneumonia: a nationwide population-based cohort study. *Lung* 2014;192(1):159-165. DOI: 10.1007/s00408-013-9523-z
30. Parris BA, O'Farrell HE, Fong KM, Yang IA. Chronic obstructive pulmonary disease (COPD) and lung cancer: common pathways for pathogenesis. *J Thorac Dis* 2019; 11(Suppl 17):S2155-S2172. DOI: 10.21037/jtd.2019.10.54
31. Han S, Suh HS. Hospitalization through emergency department visit in patients with asthma: association rule mining approach. *J Health Technol Assess* 2019;7(2):134-138 (Korean). DOI: 10.34161/johta.2019.7.2.005
32. Balducci L. Anemia, fatigue and aging. In: *Management of Hematological Cancer in Older People*. edn.: Springer; 2015:35-48.
33. Similowski T, Agustí A, MacNee W, Schönhofer B. The potential im-

- pact of anaemia of chronic disease in COPD. *Eur Respir J* 2006;27(2): 390-396. DOI: 10.1183/09031936.06.00143704
34. Al-Kindi SG, Oliveira GH. Prevalence of preexisting cardiovascular disease in patients with different types of cancer: the unmet need for onco-cardiology. *Mayo Clin Proc* 2016;91(1):81-83. DOI: 10.1016/j.mayocp.2015.09.009
35. Iachina M, Jakobsen E, Møller H, Luchtenborg M, Mellemgaard A, Krasnik M, et al. The effect of different comorbidities on survival of non-small cells lung cancer patients. *Lung* 2015;193(2):291-297. DOI: 10.1007/s00408-014-9675-5