

Original Article

Epidemiology, Occupation  
& Environmental Medicine



Received: Jul 18, 2017  
Accepted: Nov 13, 2017

Address for Correspondence:

Young-Ho Khang, MD, PhD  
Department of Health Policy and  
Management, Seoul National University  
College of Medicine, 103 Daehak-ro,  
Jongno-gu, Seoul 03080, Korea.  
E-mail: yhkhang@snu.ac.kr

© 2018 The Korean Academy of Medical  
Sciences.

This is an Open Access article distributed  
under the terms of the Creative Commons  
Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>)  
which permits unrestricted non-commercial  
use, distribution, and reproduction in any  
medium, provided the original work is properly  
cited.

ORCID iDs

Ikhan Kim   
<https://orcid.org/0000-0002-6428-4159>  
Jinwook Bahk   
<https://orcid.org/0000-0002-7715-9955>  
Yeon-Yong Kim   
<https://orcid.org/0000-0003-2179-8931>  
Jeehye Lee   
<https://orcid.org/0000-0002-9522-1919>  
Hee-Yeon Kang   
<https://orcid.org/0000-0001-8530-8087>  
Juyeon Lee   
<https://orcid.org/0000-0001-6235-6043>  
Sung-Cheol Yun   
<https://orcid.org/0000-0001-8503-109X>  
Jong Heon Park   
<https://orcid.org/0000-0002-4749-5878>  
Soon-Ae Shin   
<https://orcid.org/0000-0001-8858-0801>

# Comparison of District-level Smoking Prevalence and Their Income Gaps from Two National Databases: the National Health Screening Database and the Community Health Survey in Korea, 2009–2014

Ikhan Kim <sup>1</sup>, Jinwook Bahk <sup>2,3</sup>, Yeon-Yong Kim <sup>4</sup>, Jeehye Lee <sup>1</sup>,  
Hee-Yeon Kang <sup>2</sup>, Juyeon Lee <sup>5</sup>, Sung-Cheol Yun <sup>6</sup>, Jong Heon Park <sup>4</sup>,  
Soon-Ae Shin <sup>4,7</sup> and Young-Ho Khang <sup>1,2</sup>

<sup>1</sup>Department of Health Policy and Management, Seoul National University College of Medicine, Seoul, Korea

<sup>2</sup>Institute of Health Policy and Management, Seoul National University Medical Research Center, Seoul, Korea

<sup>3</sup>Department of Public Health, Keimyung University, Daegu, Korea

<sup>4</sup>Big Data Steering Department, National Health Insurance Service, Wonju, Korea

<sup>5</sup>The People's Health Institute, Seoul, Korea

<sup>6</sup>Department of Clinical Epidemiology and Biostatistics, University of Ulsan College of Medicine, Asan  
Medical Center, Seoul, Korea

<sup>7</sup>Gwanak Branch Office, National Health Insurance Service, Seoul, Korea

## ABSTRACT

**Background:** We compared age-standardized prevalence of cigarette smoking and their income gaps at the district-level in Korea using the National Health Screening Database (NHSD) and the Community Health Survey (CHS).

**Methods:** Between 2009 and 2014, 39,049,485 subjects participating in the NHSD and 989,292 participants in the CHS were analyzed. The age-standardized prevalence of smoking and their interquintile income differences were calculated for 245 districts of Korea. We examined between-period correlations for the age-standardized smoking prevalence at the district-level and investigated the district-level differences in smoking prevalence and income gaps between the two databases.

**Results:** The between-period correlation coefficients of smoking prevalence for both genders were 0.92–0.97 in NHSD and 0.58–0.69 in CHS, respectively. When using NHSD, we found significant income gaps in all districts for men and 244 districts for women. However, when CHS was analyzed, only 167 and 173 districts for men and women, respectively, showed significant income gaps. While correlation coefficients of district-level smoking prevalence from two databases were 0.87 for men and 0.85 for women, a relatively weak correlation between income gaps from the two databases was found.

**Conclusion:** Based on two databases, income gaps in smoking prevalence were evident for nearly all districts of Korea. Because of the large sample size for each district, NHSD may provide stable district-level smoking prevalence and its income gap and thus should be considered as a valuable data source for monitoring district-level smoking prevalence and its socioeconomic inequality.

**Keywords:** Health Surveys; Income; Korea; Sample Size; Smoking; Socioeconomic Factors

Young-Ho Khang   
<https://orcid.org/0000-0002-9585-8266>

### Funding

This study was funded and administratively supported by the National Health Insurance Service of Korea (2016-2-0010; NHIS-2017-1-213). Kim I and Khang YH were supported by Mid-career Researcher Program of the National Research Foundation (NRF) funded by the Ministry of Science & ICT (No. NRF-2014R1A2A1A11051392 and NRF-2017R1A2A2A05069746). Khang YH was also supported by a grant from Seoul National University Hospital.

### Disclosure

The authors have no potential conflicts of interest to disclose.

### Author Contributions

Conceptualization: Khang YH. Data curation: Bahk J, Kim YY. Formal analysis: Bahk J, Kim I. Investigation: Kim I, Bahk J, Kim YY, Lee J, Kang HY, Lee J, Yun SC, Park JH, Shin SA, Khang YH. Writing - original draft: Kim I, Khang YH. Writing - review & editing: Kim I, Bahk J, Kim YY, Lee J, Kang HY, Lee J, Yun SC, Park JH, Shin SA, Khang YH.

## INTRODUCTION

Cigarette smoking is known as a crucial risk factor that causes inequalities in health outcomes. A prior UK study showed that smoking was the most important risk factor that contributes to socioeconomic inequality in mortality from coronary heart disease.<sup>1</sup> A Korean study also have reported that the eradication of cigarette smoking would reduce absolute income gaps in all-cause mortality by 30%–40%.<sup>2</sup> In addition, behavioral factors including cigarette smoking, alcohol consumption and physical activity explained 16.4% of relative inequality and 84.0% of absolute inequality in all-cause mortality.<sup>3</sup> Smoking prevalence is higher in lower socioeconomic groups. Prior studies in Korea have shown that in both men and women, the inequalities in smoking prevalence exist according to education, occupation, and income.<sup>4–8</sup> The extent of this inequality was found to be even greater in younger age groups.<sup>9</sup>

Achieving equity in smoking prevalence is one of the most important goals of Korean national public health policy. Reducing inequality in smoking prevalence across different income levels was first established as a goal in the Health Plan 2010 established in 2005.<sup>10</sup> The subsequently released Third National Health Plan 2020 (2011–2020) and the recent Fourth National Health Plan 2020 (2016–2020) also included the reduction of gaps in smoking prevalence by geographical area and income as a goal.<sup>11,12</sup> Policies and intervention programs to reduce inequality in smoking prevalence need to be carried out not only by the central government, but also by local governments at the district-level. To that end, data on the income gap in smoking prevalence in local districts (si, gun, and gu) are needed.<sup>13</sup>

Since 2008, the Korea Centers for Disease Control and Prevention have implemented the Community Health Survey (CHS) to measure health status and health behaviors in si, gun, and gu. This survey has enabled various health indicators and health behaviors to be measured at the district-level. However, to measure health gaps associated with socioeconomic position in a district, the sample sizes must be sufficiently large. This necessitates combining several rounds of CHS data.<sup>13</sup> According to a prior study examining income gap in smoking prevalence at the district-level by combining CHS data for 7 years (2008–2014), there were some districts where the smoking prevalence for lowest 20% of income levels were lower than the smoking prevalence for highest 20% of income levels.<sup>13</sup> Considering that previous studies on the income gap in smoking prevalence in Korea have consistently reported pro-rich inequalities in smoking prevalence,<sup>2,4–6,9</sup> the relatively low smoking prevalence found in low-income level in some districts is likely due to small sample sizes in the CHS.

Besides the CHS, another Korean database that could be used to demonstrate income gaps in smoking prevalence at the si, gun, and gu level is the National Health Screening Database (NHSD) formed by the National Health Insurance Service (NHIS).<sup>14</sup> This database is based on general health screening data and transitional age health screening data obtained from the self-employed insured, insured employees, and their dependents. The NHSD includes self-reported health behaviors such as cigarette smoking and alcohol consumption assessed via a questionnaire, information from anthropometric examination such as height and body weight, and blood test and urine analysis. The sample of the 2014 NHSD was 17,379,951 individuals.<sup>14</sup> Health insurance premium levels, which can be used as a proxy for income, can be computed based on the National Health Information Database (NHID) eligibility database.<sup>14</sup> However, it has not yet been established whether it is possible to monitor

smoking prevalence and the magnitude of inequality in smoking in si, gun, and gu using the NHSD.

A good health-related indicator should have high validity and reliability. In monitoring of specific indicators over time, good reliability (or reproducibility) is of the highest importance. In this study, to monitor smoking prevalence and income-related inequality in smoking prevalence in si, gun, and gu, we aimed to identify whether the NHSD could be used, and therefore conducted the following analyses. First, we divided the study period between 2009 and 2014 into three periods and compared the magnitudes of between-period correlations of district-level smoking prevalence based on the NHSD and CHS. Second, we compared the smoking prevalence and the extents of income gap in smoking prevalence in si, gun, and gu, using the NHSD and CHS.

## METHODS

### Data

Data were obtained from the NHSD, which consists of general health screening data and transitional age health screening data, and the CHS from 2009–2014. The NHIS annually conducts general health screening program and transitional age health screening program for the self-employed insured, insured employees and their dependents. In the case of general health screening program, the heads of household and the household members aged 40 years or older are eligible for a biennial health examination. For insured employees, employees in non-office work are eligible for general health screening once a year while employees in office work are eligible for general health screening every two years, regardless of age. Dependents of insured employees aged 40 years or older are eligible for a biennial health examination. Among subjects for general health screening program, those aged 40 and 66 are entitled to transitional age health screening program.<sup>14</sup>

As a result, when using annual data from the NHSD, particular subpopulations such as employees in non-office work or subjects for transitional age health screening program are overrepresented. Dependents of insured employees aged less than 40 years and household members who are not household heads under 40 years of age are not eligible for health screening programs. All health insurance subscribers aged 40 years or older are eligible for health screening programs at least once every two years. Considering these eligibility criteria for health screening in Korea, we have reconstructed the two-year NHSD into one study period to secure the representativeness of data as follows. First, we have limited the study subjects to those aged 40 years or older because certain subpopulations are not eligible for health screening programs. Second, for the subjects who participated in the health screening programs twice during the two-year period, we used information on the smoking status, districts, and age from the most recent year while national health insurance premium level was calculated as the average of the two years. In this study, we analyzed 39,049,485 individuals (18,882,425 men and 20,167,060 women) aged 40 years or older in the NHSD for 2009–2014.

CHS is a nationwide, community-based survey conducted annually on approximately 900 residents aged 19 years or older, who have been sampled using stratified cluster sampling methods. Other details about CHS have been described elsewhere.<sup>13</sup> This study utilized 989,292 participants (443,570 men and 545,722 women) aged 40 years or older in the CHS from 2009 to 2014.

### Reclassification of si, gun, and gu

In this study, we reclassified the 252 si, gun, and gu according to the 2015 governmental administrative classification into 245 units of analysis in the same manner as described elsewhere,<sup>13</sup> taking into account changes in local administrative district over the study period.

### Cigarette smoking

Information on cigarette smoking was obtained in response to a question about current smoking status. In the NHSD, current smokers were defined as respondents who answered “yes, currently smoking” to the question “have you ever smoked 5 packs or more of cigarettes in the course of your lifetime?” In the CHS, respondents who smoked five or more packs (100 or more cigarettes) in their lifetime and answered “yes” in 2008 or “smokes daily,” “occasionally smokes” in 2009 and beyond for the question “do you smoke currently?” were defined as the current smokers.

### Income

In the NHSD, income data was calculated from national health insurance premiums level at the end of the year (December 31) by linking with the NHID eligibility database. In case year-end health insurance premium information was not available, we used health insurance premium levels in the most recent month. The number of population from NHID was very close to the number of population from Korean Statistical Information Service (KOSIS). In addition, because the concordance correlation coefficient of mortality rate between NHID and KOSIS was over 0.995, NHID is supposed to be sufficient to measure the health status at the district-level.<sup>15</sup> In spite of different methods of calculating national health insurance premium levels between the self-employed insured and insured employees, prior studies have shown that national health insurance premium level can be used as an appropriate proxy for income level when measuring health gap.<sup>16,17</sup> In this study, we used the equivalized health insurance premium levels of the NHID eligibility database when calculating the income quintiles. The equation is shown below.

$$\text{Equivalized health insurance premium} = \frac{\text{Health insurance premium}}{(\text{Numbers of household members})^{0.5}}$$

In the CHS, equivalized household income, self-reported annual or monthly household income adjusted by numbers of household members, was used to calculate the income quintiles. Details of income data for CHS have been described elsewhere.<sup>13</sup>

### Statistical analysis

In this study, we have treated two-year NHSD data into data for a single period, combining data for 2009–2010, 2011–2012, and 2013–2014. In a recent international collaborative study, two years of NHSD were combined into a single period of data as well.<sup>18</sup> After merging two years of data into a single period, we calculated the income quintiles at the nearest quintile points in equivalized health insurance premiums according to gender, 245 si, gun, and gu, and 5-year age groups.

We estimated the age-standardized smoking prevalence and its interquintile income gap for each study period with similar method as described elsewhere.<sup>13</sup> In addition, after combining the NHSD for six years (2009–2014), we calculated overall age-standardized smoking prevalence, age-standardized smoking prevalence by income quintiles and interquintile income gap in smoking prevalence according to gender and district.

The age-standardized smoking prevalence and its interquintile income gap based on three periods (2009–2010, 2011–2012, and 2013–2014) and six years of data for 2009–2014 were also calculated for CHS. The methods for calculating income quintiles in CHS have been described in detail elsewhere.<sup>13</sup> Age-standardized smoking prevalence and its interquintile income gap were estimated in the same manner as the NHSD. When calculating age-standardized prevalence of the two databases, 2010 mid-year resident population in Korea was used as the standard population.

We conducted correlation analyses of the age-standardized smoking prevalence for three periods based on the two databases. We also performed correlation analyses and linear regressions of age-standardized smoking prevalence and interquintile income gaps using two databases in 2009–2014 and estimated Pearson's correlation coefficients, concordance correlation coefficients, and regression coefficients. Analyses were performed in SAS 9.4 (SAS Institute, Cary, NC, USA).

### Ethics statement

The present study protocol was reviewed and approved by the Institutional Review Board of Seoul National University College of Medicine (IRB No. E-1605-006-758). Informed consent was waived by the board.

## RESULTS

**Table 1** shows the characteristics of the study subjects. Among the total 39,049,485 subjects of NHSD from 2009 to 2014, 18,882,425 were men (48.4%) and 20,167,060 were women (51.6%). The number of subjects by study period increased over time. The mean age for men was 54.9 years (standard deviation [SD], 10.3 years) and 55.9 years for women (SD, 10.6 years). The crude smoking prevalence for men was 36.3% and 2.9% for women. The crude

**Table 1.** Characteristics of study subjects according to study periods, data from the NHSD and the CHS in Korea, 2009–2014

Data	Sex	Variables	2009–2010	2011–2012	2013–2014	2009–2014
NHSD	Men	No. of study subjects	5,625,455	6,450,961	6,806,009	18,882,425
		Age, yr	54.6 (10.2)	54.8 (10.3)	55.2 (10.4)	54.9 (10.3)
		No. of current smokers	2,069,495	2,347,341	2,435,865	6,852,701
		Crude smoking rate, %	36.8	36.4	35.8	36.3
		Age-standardized smoking rate, %	36.1 (36.1–36.2)	35.8 (35.7–35.8)	35.5 (35.5–35.5)	35.3 (35.3–35.3)
	Women	No. of study subjects	5,991,596	6,896,988	7,278,476	20,167,060
		Age, yr	55.6 (10.5)	55.8 (10.6)	56.2 (10.7)	55.9 (10.6)
		No. of current smokers	171,046	206,678	215,030	592,754
		Crude smoking rate, %	2.9	3.0	3.0	2.9
		Age-standardized smoking rate, %	2.9 (2.9–2.9)	3.0 (3.0–3.0)	3.0 (3.0–3.0)	3.0 (3.0–3.1)
CHS	Men	No. of study subjects	144,876	148,096	150,598	443,570
		Age, yr	57.4 (11.8)	58.2 (12.0)	58.8 (12.1)	58.2 (12.0)
		No. of current smokers	60,808	58,647	57,793	177,248
		Crude smoking rate, %	42.0	39.6	38.4	40.0
		Age-standardized smoking rate, %	43.8 (43.5–44.1)	42.2 (41.9–42.6)	41.6 (41.3–42.0)	42.5 (42.3–42.7)
	Women	No. of study subjects	174,257	183,962	187,503	545,722
		Age, yr	59.3 (12.6)	59.8 (12.7)	60.4 (12.8)	59.9 (12.7)
		No. of current smokers	6,489	5,857	5,703	18,049
		Crude smoking rate, %	3.7	3.2	3.0	3.3
		Age-standardized smoking rate, %	3.6 (3.5–3.6)	3.1 (3.0–3.2)	3.0 (2.9–3.1)	3.2 (3.2–3.3)

Values are presented as mean (SD) or percent (95% CI). Age-standardized smoking prevalence was estimated with the 2010 mid-year resident population as the standard population.

NHSD = National Health Screening Database, CHS = Community Health Survey, SD = standard deviation, CI = confidence interval.

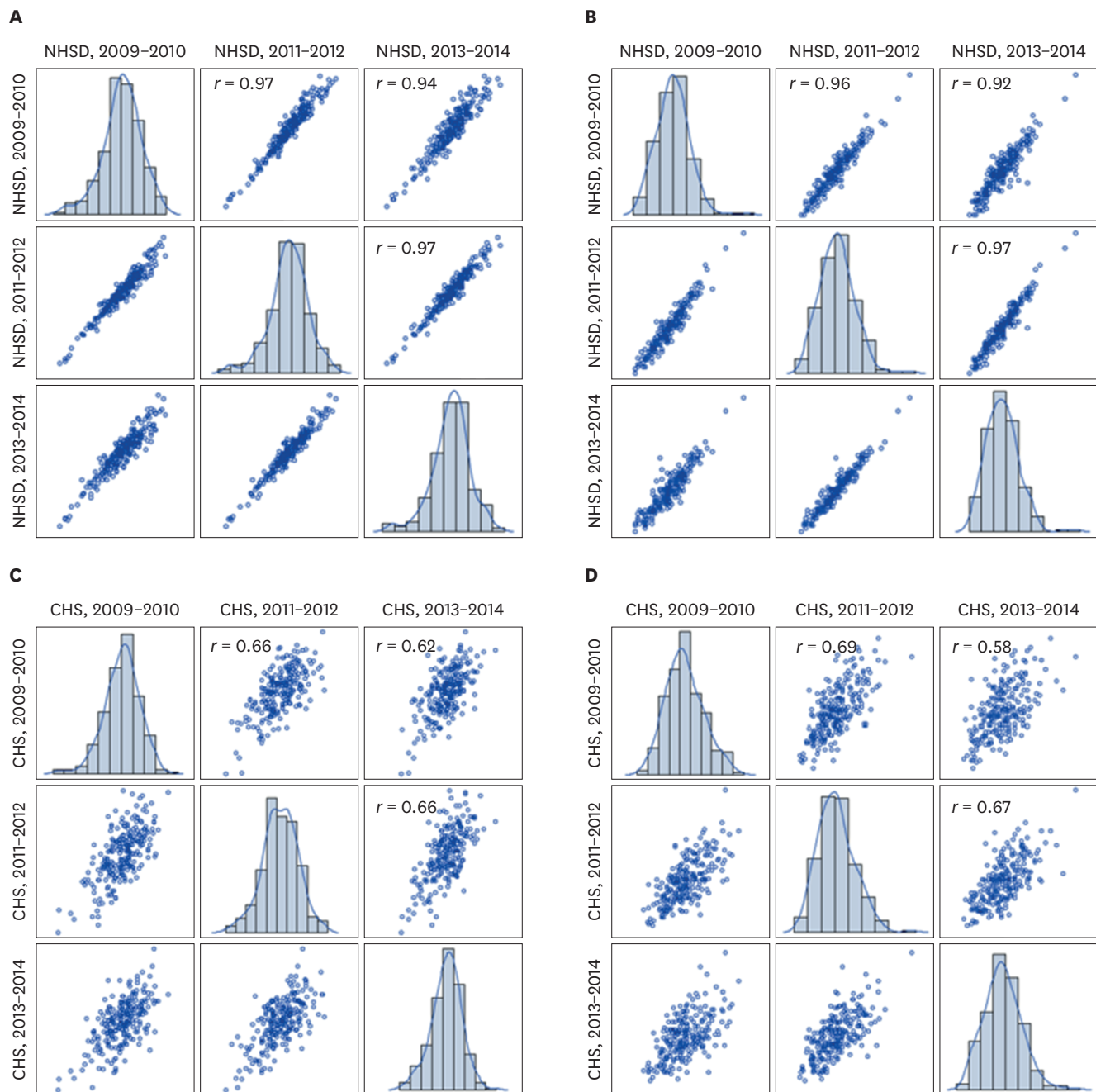


smoking prevalence in men decreased over time, while that of women did not vary. The age-standardized smoking prevalence for men was 35.3% (95% confidence interval [CI], 35.3%, 35.3%) and 3.0% for women (95% CI, 3.0%, 3.1%). In the 2009–2014 CHS, 443,570 were men (44.8%) and 545,722 were women (55.2%). The mean age of men was 58.2 years (SD, 12.0 years) and that of women was 59.9 years (SD, 12.7 years). The crude smoking prevalence in men was 40.0%, and 3.3% in women. The CHS data showed that the crude smoking prevalence in both genders gradually decreased over time. The age-standardized smoking prevalence of men was 42.5% (95% CI, 42.3%, 42.7%), and that of women was 3.2% (95% CI, 3.2%, 3.3%).

**Fig. 1** presents the between-period correlations for district-level age-standardized smoking prevalence based on both databases. Overall, the correlation was higher for the smoking prevalence based on the NHSD rather than the CHS. This was true for both genders. In the NHSD, the correlation coefficient ( $r$ ) ranged from 0.92 (correlation coefficient for women's smoking prevalence between 2009–2010 and 2013–2014) to 0.97 (correlation coefficient for men's smoking prevalence between 2011–2012 and 2013–2014). In contrast, in the CHS, the correlation coefficient ( $r$ ) was between 0.58 (correlation coefficient for women's smoking prevalence between 2009–2010 and 2013–2014) and 0.69 (correlation coefficient for women's smoking prevalence between 2009–2010 and 2011–2012).

**Table 2** shows the distribution of overall age-standardized smoking prevalence and age-standardized smoking prevalence by income quintiles in 245 si, gun, and gu in Korea based on the NHSD and the CHS. For men, the median age-standardized smoking prevalence from the NHSD was 36.9% (95% CI, 36.7%, 37.2%). The SD was 3.1 percentage points (%p). The minimum value was 26.0% in Gwangyang-si, Jeollanam-do (95% CI, 25.6%, 26.4%). The maximum value was 44.1% (95% CI, 43.7%, 44.6%) in Seo-gu, Daegu. In the CHS, the median smoking prevalence was 42.9% (95% CI, 39.8%, 46.1%) and the SD was 3.8%p. The minimum value was 28.1% (95% CI, 25.5%, 30.7%) in Gwacheon-si, Gyeonggi-do while the maximum value was 51.5% (95% CI, 48.2%, 54.8%) from Taebaek-si, Gangwon-do. In women, the median age-standardized smoking prevalence from the NHSD was 2.8% (95% CI, 2.7%, 2.9%). The SD was 1.0%p. The minimum value was 0.8% (95% CI, 0.6%, 0.9%) in Goheung-gun, Jeollanam-do. The maximum value was 7.4% (95% CI, 7.1%, 7.7%) in Dongducheon-si, Gyeonggi-do. In the CHS, the median age-standardized smoking prevalence was 3.2% (95% CI, 2.5%, 4.0%). The SD was 1.1%p. The minimum value was 0.7% (95% CI, 0.4%, 1.1%) from districts including Jangheung-gun, Jeollanam-do. The maximum value was 7.7% (95% CI, 6.4%, 8.9%) in Dongducheon-si, Gyeonggi-do. In both the NHSD and CHS, the median age-standardized smoking prevalence decreased with higher income levels and this was true for both genders.

**Table 2** also shows the distribution of interquintile income gaps in smoking prevalence in si, gun, and gu of Korea. In men, the median income gap in smoking prevalence based on the NHSD was 9.4%p (95% CI, 7.6%p, 11.2%p) and the SD was 2.1%p. The minimum value was 4.0%p (95% CI, 2.1%p, 5.9%p), in Gapyeong-gun, Gyeonggi-do while the maximum value was 19.0%p (95% CI, 18.1%p, 19.9%p) from Nam-gu, Pohang-si, Gyeongsangbuk-do. In the CHS, the median income gap in smoking prevalence for men was 9.4%p (95% CI, 2.4%p, 16.4%p). The SD was 4.6%p. The minimum value was -1.8%p (95% CI, -8.6%p, 5.0%p) in Goheung-gun, Jeollanam-do. The maximum value was 23.9%p (95% CI, 16.9%p, 30.9%p) in Nam-gu, Pohang-si, Gyeongsangbuk-do. In women, the median income gap in smoking prevalence based on the NHSD was 2.3%p (95% CI, 2.0%p, 2.6%p), and the SD was 0.9%p.



**Fig. 1.** Scatter plots and correlation coefficients ( $r$ ) for correlations of age-standardized smoking prevalence between biennial time periods of analysis, findings from the NHSD and the CHS in Korea, 2009–2014. **(A)** Results from the NHSD in men. **(B)** Results from the NHSD in women. **(C)** Results from the CHS in men. **(D)** Results from the CHS in women.

NHSD = National Health Screening Database, CHS = Community Health Survey.

The minimum value was 0.0%p (95% CI, -0.6%p, 0.6%p) in Gurye-gun, Jeollanam-do, while the maximum value was 5.6%p (95% CI, 4.5%p, 6.7%p) in Jung-gu, Busan. In the CHS, the median income gap in smoking prevalence for women was 3.1%p (95% CI, 0.1%p, 6.1%p). The SD was 1.9%p. The minimum value was -1.1%p (95% CI, -2.5%p, 0.3%p) in Gurye-gun, Jeollanam-do. The maximum value was 9.9%p (95% CI, 6.0%p, 13.8%p) in Dongducheon-si, Gyeonggi-do.

**Table 2.** Distribution of overall age-standardized smoking prevalence, smoking prevalence by income quintiles and interquintile income gaps in smoking prevalence, findings for 245 local districts using the NHSD and the CHS in Korea, 2009–2014

Variables		NHSD, 2009–2014					CHS, 2009–2014				
		Median	SD	Min	Max	IQR	Median	SD	Min	Max	IQR
Men	Overall	36.9%	3.1%p	26.0%	44.1%	3.6%p	42.9%	3.8%p	28.1%	51.5%	4.7%p
	Income Q1	41.3%	3.0%p	31.5%	49.0%	3.3%p	47.5%	4.3%p	33.3%	57.5%	5.5%p
	Income Q2	39.7%	3.2%p	28.0%	46.5%	3.2%p	44.6%	4.5%p	29.1%	56.5%	5.9%p
	Income Q3	37.2%	3.6%p	24.7%	45.9%	4.5%p	42.4%	4.8%p	26.6%	54.0%	6.4%p
	Income Q4	34.1%	3.4%p	20.4%	42.8%	4.1%p	40.6%	4.7%p	25.5%	52.9%	5.9%p
	Income Q5	31.6%	3.2%p	19.4%	39.3%	3.9%p	37.8%	4.7%p	25.3%	51.2%	6.3%p
	Q1–Q5	9.4%p	2.1%p	4.0%p	19.0%p	2.6%p	9.4%p	4.6%p	–1.8%p	23.9%p	5.9%p
Women	Overall	2.8%	1.0%p	0.8%	7.4%	1.3%p	3.2%	1.1%p	0.7%	7.7%	1.5%p
	Income Q1	4.2%	1.4%p	1.1%	10.1%	1.7%p	5.2%	2.1%p	0.8%	13.4%	2.8%p
	Income Q2	3.0%	1.1%p	0.7%	7.5%	1.6%p	3.3%	1.6%p	0.1%	9.6%	2.2%p
	Income Q3	2.7%	1.0%p	0.8%	7.1%	1.4%p	2.7%	1.3%p	0.1%	7.0%	1.7%p
	Income Q4	2.2%	0.9%p	0.6%	6.9%	1.1%p	2.2%	1.2%p	0.1%	7.4%	1.7%p
	Income Q5	1.9%	0.8%p	0.5%	5.3%	1.0%p	1.9%	1.1%p	0.2%	6.5%	1.3%p
	Q1–Q5	2.3%p	0.9%p	0.0%p	5.6%p	1.1%p	3.1%p	1.9%p	–1.1%p	9.9%p	2.3%p

NHSD = National Health Screening Database, CHS = Community Health Survey, SD = standard deviation, IQR = interquartile range.

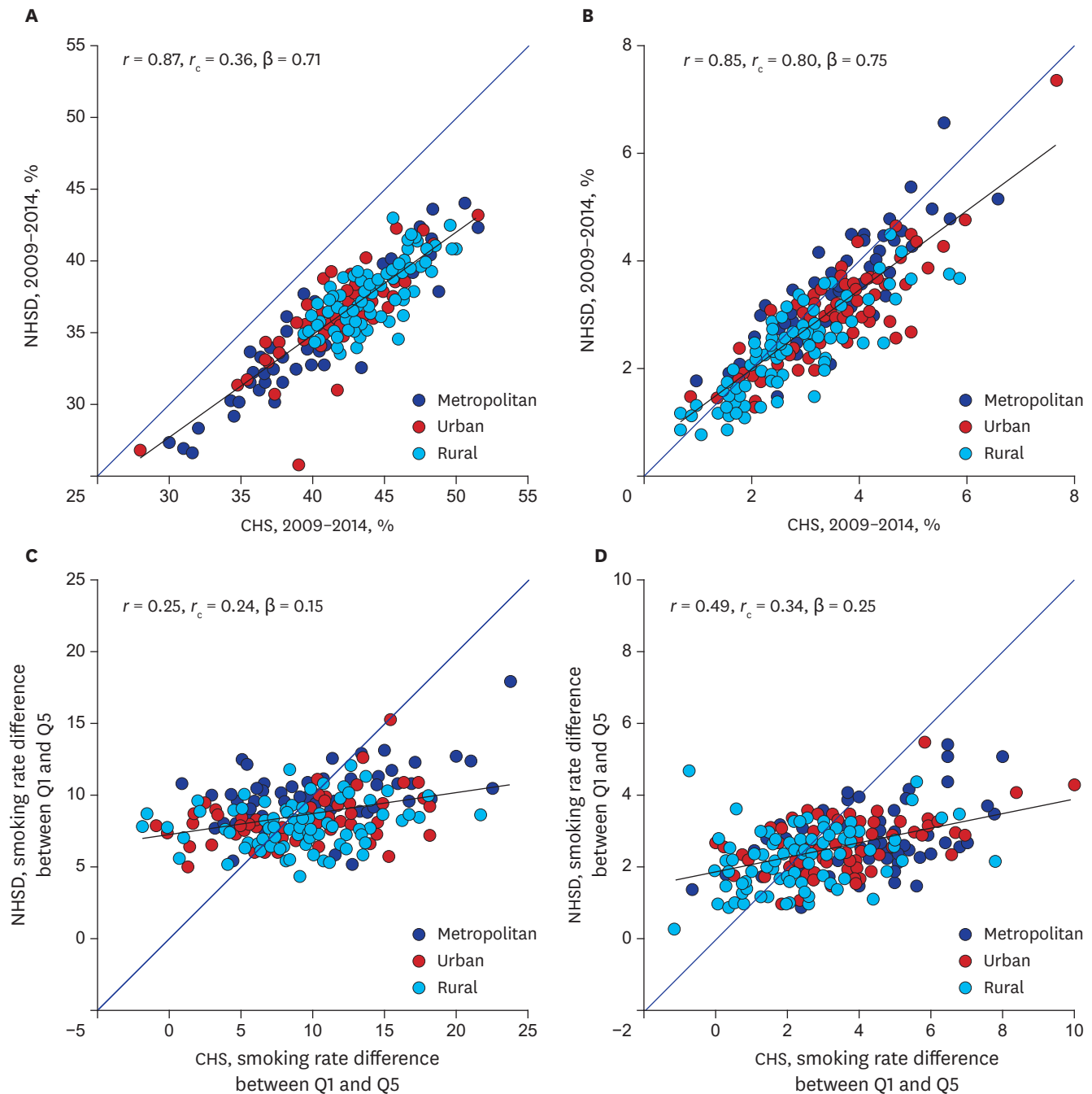
Supplementary Tables 1 and 2 show the age-standardized smoking prevalence and its interquintile income gaps based on the NHSD and CHS. According to the NHSD, the income gaps in smoking prevalence was over 0.0%p in all 245 si, gun, and gu for both men and women. Based on the CHS, the income gaps in smoking prevalence was over 0.0%p in 240 si, gun, and gu for men and 242 si, gun, and gu for women. Considering the 95% CI of the income gaps in smoking prevalence, the number of districts where the lower limit of the 95% CI was higher than 0.0%p was 245 for men and 244 for women based on the NHSD, and 167 for men and 173 for women based on the CHS. In both databases, there was no district where the upper limit of the 95% CI was lower than 0.0%p. This was true for both genders.

Fig. 2A and B show the age-standardized smoking prevalence in 245 si, gun, and gu based on the 2009–2014 NHSD and CHS. The correlation in smoking prevalence from the two databases was quite high: the correlation coefficient ( $r$ ) was 0.87 in men and 0.85 in women. In men, the smoking prevalence estimated from the CHS tended to be lower than those calculated from the NHSD. With higher smoking prevalence, the difference in smoking prevalence between the two databases tended to increase: the regression coefficient ( $\beta$ ) was 0.71 for men and 0.75 for women. For women, the agreement between the two databases was high: the concordance correlation coefficient ( $r_c$ ) was 0.80. Fig. 2C and D show the income gaps in smoking prevalence in 245 si, gun, and gu. The CHS showed greater variation in the income gaps in smoking prevalence rather than the NHSD. Based on the CHS, there were a few districts with an income gap in smoking prevalence lower than 0.0%p, but no such districts were found in the NHSD.

## DISCUSSION

In this study, we compared the age-standardized smoking prevalence and the magnitude of the income gaps in smoking prevalence in 245 si, gun, and gu in Korea, using the NHSD and CHS, which enabled us to examine the extent of health inequalities at the district-level. The study results showed a higher between-period correlation for district-level smoking prevalence in the NHSD compared to the CHS, in both genders. The correlation coefficient of the age-standardized smoking prevalence across periods ranged between 0.92 and 0.97





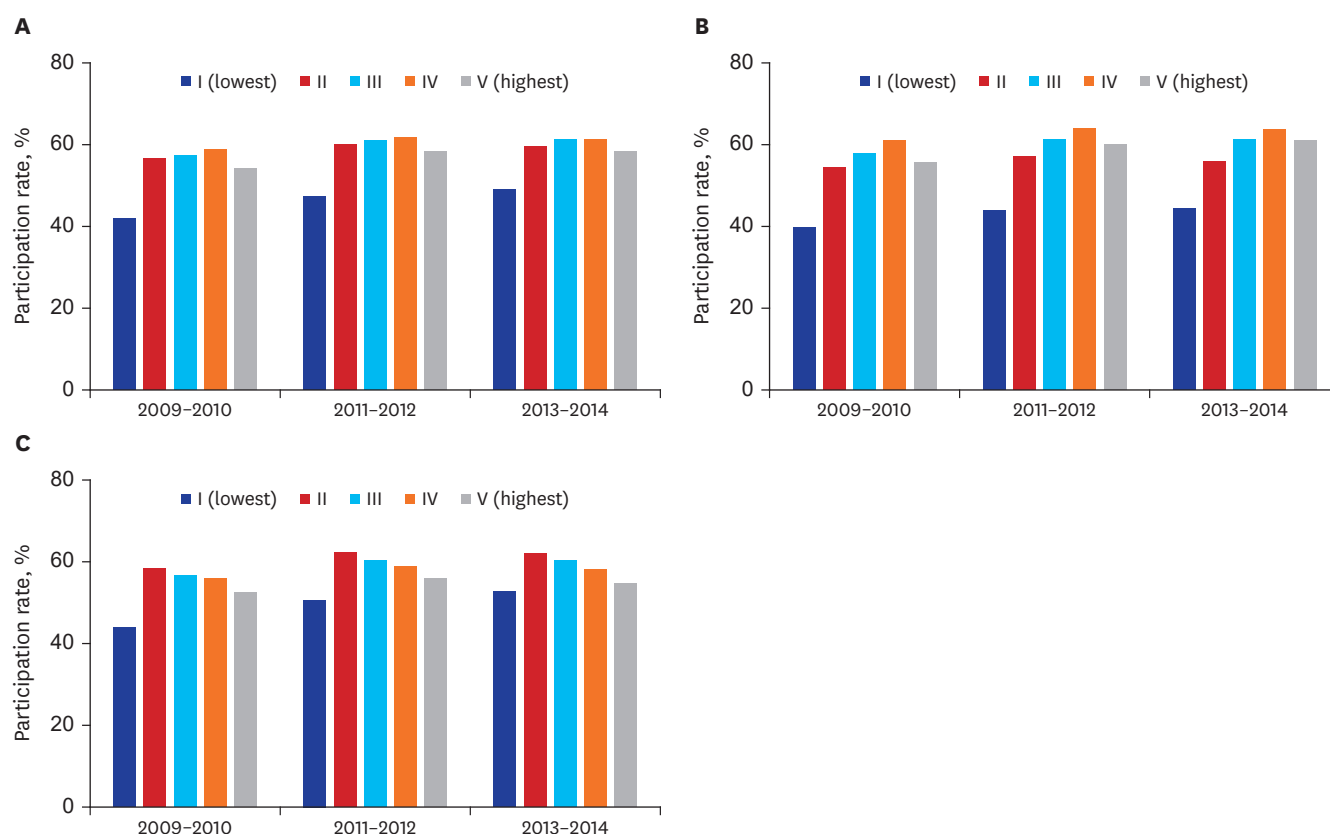
**Fig. 2.** Scatter plots and correlation coefficients ( $r, r_c$ ) as well as regression coefficients ( $\beta$ ) for correlations of district-level age-standardized smoking prevalence and their interquintile income differences from the NHSD and the CHS among the 245 local districts in Korea, 2009–2014. **(A)** Correlations of smoking prevalences in men. **(B)** Correlations of smoking prevalences in women. **(C)** Correlations of interquintile income differences in men. **(D)** Correlations of interquintile income differences in women. NHSD = National Health Screening Database, CHS = Community Health Survey.

in the NHSD and between 0.58 and 0.69 in the CHS. This finding indicates that the CHS exhibited a higher fluctuation in smoking prevalence over time. The key factor that explains this difference is different sample sizes in each district. In the CHS in 2013–2014, the median sample size for a district was 1,312, with a range of 1,001 to 4,204. In contrast, in the NHSD in the same period, the median sample size for a district was 46,535, with a range of 3,223

to 171,658. The larger sample size of the NHSD than the CHS is likely to have lowered the stochastic variation in measurements of age-standardized smoking prevalence, leading to the higher correlation coefficients across different periods.

The analysis also showed that the SD of the income gaps in smoking prevalence based on the NHSD from 2009 to 2014 was 2.1%p and 0.9%p for men and women, respectively. In the CHS, the SD was 4.6%p and 1.9%p for men and women, respectively. The variation in the income gaps in smoking prevalence based on the NHSD was lower than was observed based on the CHS. Moreover, based on the CHS, there were a few districts in which the smoking prevalence for the highest 20% of income levels was higher than the smoking prevalence for the lowest 20% of income levels. However, there was no such tendency in the NHSD. Based on the NHSD, the number of districts with a lower limit of the 95% CI for the income gap in smoking prevalence being higher than 0.0%p was 245 in men and 244 in women. For the CHS, the corresponding numbers of districts were 167 for men and 173 for women. These findings show clearer pro-rich inequality in smoking prevalence in using the NHSD than using the CHS. Such findings suggest that when calculating income gaps in smoking prevalence in si, gun, and gu, the CHS, due to its limited sample size, might yield high stochastic variations. Thus, it is necessary to use the NHSD (in addition to the CHS) as one source for monitoring income gaps in smoking prevalence in si, gun, and gu. Furthermore, the NHSD is useful in time-series monitoring of smoking inequality, as the income gap in smoking prevalence in si, gun, and gu can be estimated from the combined NHSD for just two years, and as the correlation coefficients for smoking prevalence across different periods were higher in the NHSD. It should be noted that income gaps in smoking prevalence are hardly be presented using two-year CHS data.

This study showed that the correlation of smoking prevalence in men and women between the two databases (**Fig. 2A and B**) was high: the correlation coefficients were 0.87 and 0.85 for men and women, respectively. The results of comparisons of smoking prevalence suggests that NHSD can be valuable for policy use. Nonetheless, although the prevalence of smoking in men tended to be higher in the CHS than in the NHSD, smoking prevalence in women did not show a clear difference between the two databases. Such findings may be partly attributed to participation rates in national health screening programs (both general health screening programs and transitional age health screening program) by income level. As **Fig. 3** shows, age-standardized participation rates for men in national health screening programs among the fourth fifth and highest fifth of the income quintiles were higher than their counterparts in the lowest fifth and second fifth. In contrast, in women, the participation rates in national health screening programs by those in the second fifth and third fifth of the income quintiles were relatively higher than the rates among those who belonged to other income quintile groups. In the lowest fifth group of income quintiles, women also showed higher participation rates in national health screening programs than men. Considering that the age-standardized smoking prevalence in low-income groups is high in both genders in Korea, the income gap in national health screening programs participation rates may partially explain the relatively low smoking prevalence in the NHSD. Moreover, the different survey methods and questionnaires for obtaining information on current smoking status may help explain the discrepancy in the smoking prevalence between the two databases. A study has reported that face-to-face interviews may yield lower prevalence of socially undesirable behaviors such as cigarette smoking than the prevalence obtained using a self-administered questionnaire.<sup>19</sup> The CHS was conducted using face-to-face interviews in the subjects' homes. Thus, it is possible that the CHS might have yielded lower smoking prevalence among respondents, particularly



**Fig. 3.** Age-standardized participation rate in the national health screening programs by interquintile income level in Korea, 2009–2014. (A) Participation rate in total. (B) Participation rate in men. (C) Participation rate in women.

women, than the NHSD. However, the results of this study showed that the prevalence of smoking in women was relatively similar across the two databases, whereas the prevalence in men was not so. Therefore, it is hard to explain the difference in the smoking prevalence between these databases based on the different survey mode alone. Considering that the survey questions about current smoking are not identical between the NHSD and the CHS, the differences in the smoking prevalence between the two databases might have resulted from differences in use of survey questions and response categories.

The results of this study show that income inequality in smoking clearly exists in almost all si, gun, and gu in Korea. Based on the NHSD, in all 245 districts, there was a 0.0%p or higher income gap in smoking prevalence. When limiting subjects to those in their 40s, the CHS also showed that there were 240 districts and 242 districts for men and women, respectively, where a 0.0%p or higher income gap in smoking prevalence existed. Such a tendency is similar to the results of a prior study that reported an income gap in smoking prevalence (among individuals aged  $\geq 19$  years) in most si, gun, and gu in Korea, based on seven years of CHS data in 2008–2014.<sup>13</sup>

This study has the following limitations. First, the income data of the two databases are not identical. In the NHSD, health insurance premium levels were used as a proxy for income. In comparison, the CHS used self-reported annual or monthly household income data. Moreover, while the monthly household income data of the 2014 CHS were categorical, in the present study, we used the median in each category. Second, the questions on current

smoking status were not identical in the two databases. Third, this study did not provide data on the annual smoking prevalence, as data from two years were combined for analysis to resolve the overrepresentation of subpopulations in the NHSD.

Despite such limitations, this study has meaningful implications. It has demonstrated that, in addition to the CHS, the NHSD can be used as a data source for monitoring smoking inequality at the district-level. It is also worth noting that the smoking prevalence by si, gun, and gu in Korea as calculated from the NHSD demonstrated higher reproducibility over time than the smoking prevalence estimated based on the CHS. The CHS is an important and valuable source of data that provides information on a wide range of health indicators, including smoking. However, its sample sizes are too small to produce reliable health inequality indicators at the district-level. Therefore, the CHS needs to be improved in the future in terms of sample size, and health inequality studies at the district-level using the NHSD should be more actively conducted.

## SUPPLEMENTARY MATERIALS

### Supplementary Table 1

Age-standardized smoking prevalence, smoking prevalence by income quintiles and interquintile income gaps in smoking prevalence in men, findings for 245 local districts using the NHSD and the CHS in Korea, 2009–2014

[Click here to view](#)

### Supplementary Table 2

Age-standardized smoking prevalence, smoking prevalence by income quintiles and interquintile income gaps in smoking prevalence in women, findings for 245 local districts using the NHSD and the CHS in Korea, 2009–2014

[Click here to view](#)

## REFERENCES

1. Kivimäki M, Shipley MJ, Ferrie JE, Singh-Manoux A, Batty GD, Chandola T, et al. Best-practice interventions to reduce socioeconomic inequalities of coronary heart disease mortality in UK: a prospective occupational cohort study. *Lancet* 2008;372(9650):1648-54.  
[PUBMED](#) | [CROSSREF](#)
2. Khang YH, Lynch JW, Jung-Choi K, Cho HJ. Explaining age-specific inequalities in mortality from all causes, cardiovascular disease and ischaemic heart disease among South Korean male public servants: relative and absolute perspectives. *Heart* 2008;94(1):75-82.  
[PUBMED](#) | [CROSSREF](#)
3. Khang YH, Lynch JW, Yang S, Harper S, Yun SC, Jung-Choi K, et al. The contribution of material, psychosocial, and behavioral factors in explaining educational and occupational mortality inequalities in a nationally representative sample of South Koreans: relative and absolute perspectives. *Soc Sci Med* 2009;68(5):858-66.  
[PUBMED](#) | [CROSSREF](#)
4. Khang YH, Yun SC, Cho HJ, Jung-Choi K. The impact of governmental antismoking policy on socioeconomic disparities in cigarette smoking in South Korea. *Nicotine Tob Res* 2009;11(3):262-9.  
[PUBMED](#) | [CROSSREF](#)

5. Park EJ, Kim H, Kawachi I, Kim IH, Cho SI. Area deprivation, individual socioeconomic position and smoking among women in South Korea. *Tob Control* 2010;19(5):383-90.  
[PUBMED](#) | [CROSSREF](#)
6. Cho HJ, Song YM, Smith GD, Ebrahim S. Trends in socio-economic differentials in cigarette smoking behaviour between 1990 and 1998: a large prospective study in Korean men. *Public Health* 2004;118(8):553-8.  
[PUBMED](#) | [CROSSREF](#)
7. Choi S, Kim Y, Park S, Lee J, Oh K. Trends in cigarette smoking among adolescents and adults in South Korea. *Epidemiol Health* 2014;36:e2014023.  
[PUBMED](#) | [CROSSREF](#)
8. Yun WJ, Rhee JA, Kim SA, Kweon SS, Lee YH, Ryu SY, et al. Household and area income levels are associated with smoking status in the Korean adult population. *BMC Public Health* 2015;15:39.  
[PUBMED](#) | [CROSSREF](#)
9. Khang YH, Cho HJ. Socioeconomic inequality in cigarette smoking: trends by gender, age, and socioeconomic position in South Korea, 1989–2003. *Prev Med* 2006;42(6):415-22.  
[PUBMED](#) | [CROSSREF](#)
10. Ministry of Health and Welfare (KR). *Health Plan 2010*. Seoul, Korea: Ministry of Health and Welfare; 2005.
11. Ministry of Health and Welfare (KR). *Health Plan 2020 (2011–2015)*. Seoul, Korea: Ministry of Health and Welfare; 2011.
12. Ministry of Health and Welfare (KR). *Health Plan 2020 (2016–2020)*. Cheongju, Korea: Ministry of Health and Welfare; 2015.
13. Kim I, Bahk J, Yoon TH, Yun SC, Khang YH. Income differences in smoking prevalences in 245 districts of South Korea: patterns by area deprivation and urbanity, 2008–2014. *J Prev Med Public Health* 2017;50(2):100-26.  
[PUBMED](#) | [CROSSREF](#)
14. Seong SC, Kim YY, Khang YH, Park JH, Kang HJ, Lee H, et al. Data resource profile: the National Health Information Database of the National Health Insurance Service in South Korea. *Int J Epidemiol* 2017;46(3):799-800.  
[PUBMED](#)
15. Khang YH, Yun SC, Bahk J, Kang HY, Kim I, Seo S, et al. *Developing Indicators for Evaluating and Monitoring the Health Plan 2020 Using the National Health Information Database of the National Health Insurance Service in Korea*. Wonju, Korea: National Health Insurance Service; 2016.
16. Khang YH, Bahk J, Yi N, Yun SC. Age- and cause-specific contributions to income difference in life expectancy at birth: findings from nationally representative data on one million South Koreans. *Eur J Public Health* 2016;26(2):242-8.  
[PUBMED](#) | [CROSSREF](#)
17. Khang YH, Yang S, Cho HJ, Jung-Choi K, Yun SC. Decomposition of socio-economic differences in life expectancy at birth by age and cause of death among 4 million South Korean public servants and their dependents. *Int J Epidemiol* 2010;39(6):1656-66.  
[PUBMED](#) | [CROSSREF](#)
18. Zhou B, Bentham J, Di Cesare M, Bixby H, Danaei G, Cowan MJ, et al. Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants. *Lancet* 2017;389(10064):37-55.  
[PUBMED](#) | [CROSSREF](#)
19. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)* 2005;27(3):281-91.  
[PUBMED](#) | [CROSSREF](#)