

Original Article
Preventive & Social Medicine



Spatio-temporal Analysis of District-level Life Expectancy from 2004 to 2017 in Korea

Hwa-Kyung Lim ,¹ Hee-Yeon Kang ,^{1,2} Ikhan Kim ,^{2,3*} and Young-Ho Khang ^{1,2}

¹Institute of Health Policy and Management, Seoul National University Medical Research Center, Seoul, Korea

²Department of Health Policy and Management, Seoul National University College of Medicine, Seoul, Korea

³Department of Health Policy and Management, Jeju National University College of Medicine and Graduate School of Medicine, Jeju, Korea



Received: Jul 16, 2020

Accepted: Oct 22, 2020

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

*Present address: Department of Medical Humanities and Social Medicine, Kosin University College of Medicine, Busan, Korea

© 2021 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

Hwa-Kyung Lim
<https://orcid.org/0000-0003-2330-718X>
Hee-Yeon Kang
<https://orcid.org/0000-0001-8530-8087>
Ikhan Kim
<https://orcid.org/0000-0002-6428-4159>
Young-Ho Khang
<https://orcid.org/0000-0002-9585-8266>

ABSTRACT

Background: Health indicators, such as mortality rates or life expectancy, need to be presented at the local level to improve the health of local residents and to reduce health inequality across geographic areas. The aim of this study was to estimate life expectancy at the district level in Korea through a spatio-temporal analysis.

Methods: Spatio-temporal models were applied to the National Health Information Database of the National Health Insurance Service to estimate the mortality rates for 19 age groups in 250 districts from 2004 to 2017 by gender in Korea. Annual district-level life tables by gender were constructed using the estimated mortality rates, and then annual district-level life expectancy by gender was estimated using the life table method and the Kannisto-Thatcher method. The annual district-level life expectancies based on the spatio-temporal models were compared to the life expectancies calculated under the assumption that the mortality rates in these 250 districts are independent from one another.

Results: In 2017, district-level life expectancy at birth ranged from 75.5 years (95% credible interval [CI], 74.0–77.0 years) to 84.2 years (95% CI, 83.4–85.0 years) for men and from 83.9 years (95% CI, 83.2–84.6 years) to 88.2 years (95% CI, 87.3–89.1 years) for women. Between 2004 and 2017, district-level life expectancy at birth increased by 4.57 years (95% CI, 4.49–4.65 years) for men and by 4.06 years (95% CI, 3.99–4.12 years) for women. To obtain stable annual life expectancy estimates at the district level, it is recommended to use the life expectancy based on spatio-temporal models instead of calculating life expectancy using observed mortality.

Conclusion: In this study, we estimated the annual district-level life expectancy from 2004 to 2017 in Korea by gender using a spatio-temporal model. Local governments could use annual district-level life expectancy estimates as a performance indicator of health policies to improve the health of local residents. The approach to district-level analysis with spatio-temporal modeling employed in this study could be used in future analyses to produce district-level health-related indicators in Korea.

Keywords: Geography; Life Expectancy; Mortality; Korea; Spatio-Temporal Analysis

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- 2019R111A1A01063964). This research was also supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant No. H118C0446).

Disclosure

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Khang YH. Data curation: Kang HY. Formal analysis: Lim HK, Kim IH. Writing - original draft: Lim HK, Khang YH. Writing - review & editing: Lim HK, Khang YH, Kang HY, Kim IH.

INTRODUCTION

Health indicators, such as the mortality rate or life expectancy (LE) in subnational areas, provide information that is necessary to improve community health and to address local health problems in vulnerable geographic areas. Local governments in Korea are establishing regional health care plans at the district level (i.e., 'si', 'gun', 'gu') to improve the quality of health care services for local residents and thereby to improve the health of local residents.¹ LE presented at the district level could be used as a performance indicator of health policies, and calculating the LE gaps across geographic areas at the district level each year could help achieve a long-term health equity policy.

The National Health Insurance Service (NHIS) in Korea is the mandatory health insurance scheme for the entire Korean population. The National Health Information Database (NHID) provided by the NHIS consists of sub-databases (DBs) including the eligibility DB, the healthcare utilization DB, the health screening DB, and the healthcare provider DB.² The NHID DB has been used to monitor health status and health equity at the national and local levels for the entire Korean population. Several Korean studies have examined subnational life expectancies using the NHID DB.³⁻⁸ Bahk et al.³ showed that the crude mortality rate and LE at birth across the national, provincial, and district levels calculated using the NHID DB were consistent with the corresponding results calculated using the Korean Statistical Information Service (KOSIS) data from Statistics Korea, indicating that the NHID DB was a good source for monitoring the mortality rate and LE with representation of the entire Korean population. Kim et al.⁴ presented town-level life expectancy and other mortality metrics using the NHID DB. Lim et al.⁵ used the NHID DB to measure differences in quality-adjusted life expectancy by income in Korea at the national and district levels. Bahk et al.⁶ presented LE for 253 electoral constituencies in Korea using the NHID DB. However, these studies on LE and mortality metrics at the local level in Korea have limitations because their analyses were based on the assumption that observations are independent from one another (i.e., that there is no spatial autocorrelation), despite similar mortality rates in adjacent regions. Similarly, mortality rates in adjacent years should not be assumed to be independent from one another.

Moreover, estimating LE at the local level in Korea may encounter difficulties because of the rarity of deaths and small populations, especially in rural areas, which lead to uncertainty in estimating mortality rates. A previous study reported that the minimum population required to calculate a stable LE was 5,000.⁹ In order to ensure the minimum population needed to calculate income-quintile specific LE in small areas, Bahk et al.³ used 6-year pooled data and Lim et al.⁵ used 7-year pooled data for the number of population and deaths in each region. As a result, it was impossible to estimate LE by year.

These issues could be addressed effectively within the framework of spatio-temporal modeling, which allows for dependence of mortality rates on both space and time and improves the robustness of mortality estimates in subnational populations. Studies on the estimation of LE and mortality metrics at the local level using spatio-temporal models have been recently conducted,¹⁰⁻¹⁴ but, to the best of our knowledge, no such studies have been conducted in Korea. This study aimed to estimate the annual district-level LE at birth in Korea based on the estimated mortality rates from spatio-temporal models using the NHID DB.

METHODS

Data sources

We obtained aggregated data from the NHID provided by the NHIS on the numbers of deaths and population by gender (both sexes, men, women) over the course of 14 years (2004–2017), divided into 19 age groups (0, 1–4, 5–9, ..., 80–84, and 85+ years), and 250 districts based on the administrative districts in Korea in 2017 (the most recent year available at the time of this study). As administrative district classification criteria have changed several times during the study period (2004–2017), we reclassified the districts by year according to this 2017 administrative district classification to maintain the consistency of administrative district classifications in the NHID DB. In particular, the Sejong Special Provincial Municipality, which was established in 2012, was reclassified using the previous administrative district ('dong' unit). These reclassification processes have been applied in the previous studies.^{3,5,7,8}

Because there is considerable variation in education levels among districts in Korea, a district-level covariate for the proportion of the population aged 25 years or older who have completed a 4-year college or higher (= the number of people over 25 years old who have graduated from a 4-year college or higher / the number of people over 25 years old) was used in the spatio-temporal model. The covariate was created using data from the Population and Housing Census of Korea, with the proportion of college graduates from 2004 to 2009 based on the 2005 census data, the proportion of college graduates from 2010 to 2014 based on the 2010 census data, and the proportion of college graduates from 2015 to 2017 based on the 2015 census data.

Statistical analysis

Spatial autocorrelation of mortality rates

Fig. 1 shows the distribution of age-standardized mortality rates at the district level by gender in 2017. The age-standardized mortality rates show little evidence of complete spatial randomness. Age-standardized mortality rates in neighboring districts were similar and showed clustered patterns. The extent to which similar (or dissimilar) districts are clustered can be quantified by a measure, such as Moran's *I* spatial statistic,¹⁵ a measure of global spatial autocorrelation that measures how the mortality rate in one district is similar to others surrounding it: Moran's *I* spatial statistic is close to +1 when similar values are spatially clustered, 0 when mortality rates are completely randomly distributed (i.e., no spatial autocorrelation), and -1 when dissimilar values are next to each other. Moran's *I* spatial statistic for age-standardized mortality rates at the district level by gender in 2017 was 0.584 for men and 0.399 for women, respectively, supporting the need to calculate LE based on mortality rates at the district level estimated using models that allow spatial dependence. The distribution of age-standardized mortality rates at the district level for both sexes in 2017 is given in **Supplementary Fig. 1**. Moran's *I* spatial statistic for both sexes was 0.611.

Spatio-temporal model

The allocation of numbers of deaths into units grouped by gender, 19 age groups, and 250 districts can result in few deaths in some units, leading to highly variable estimates of mortality rates. To overcome this issue, we used a Bayesian hierarchical model to estimate mortality rates by sharing information across age groups and regions and over time. In this approach, mortality rates for each age group, region, and year are informed by data in that age group-region-year unit, as well as by those in the adjacent age groups, adjacent districts, and adjacent years. The extent to which the estimated mortality rates are affected by adjacent

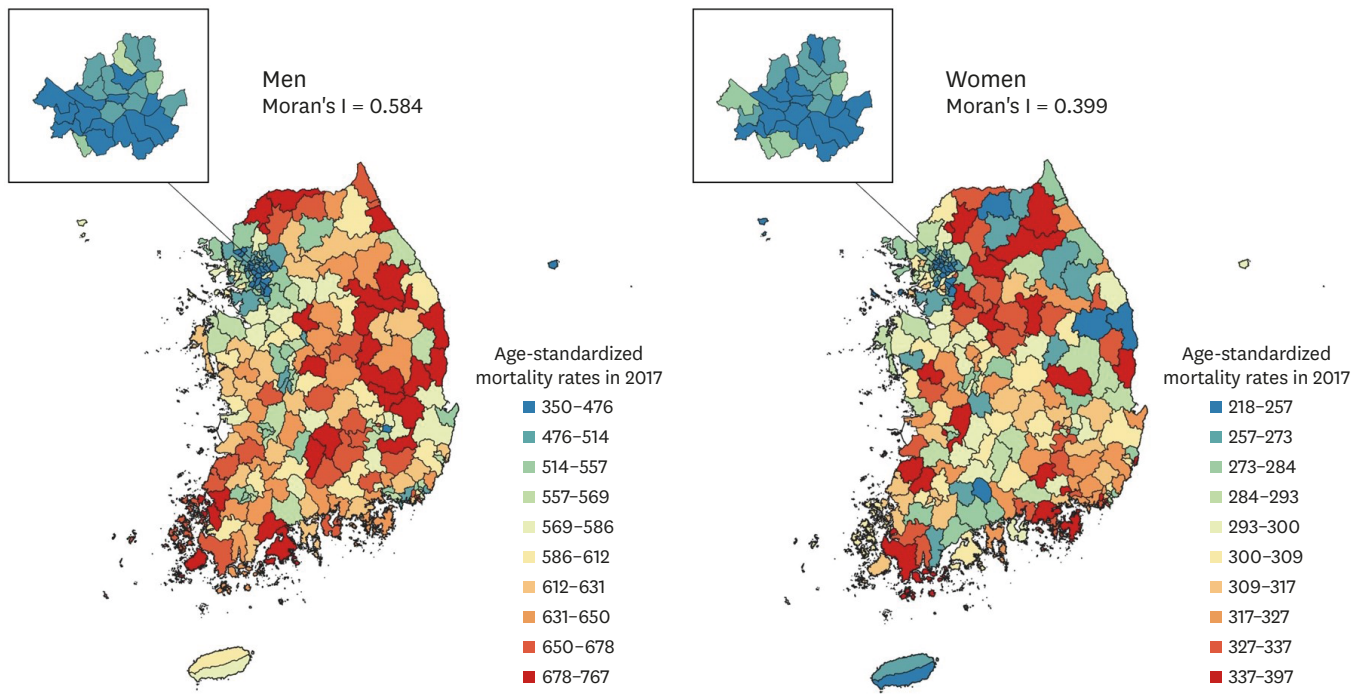


Fig. 1. Age-standardized mortality rates per 100,000 population in 250 districts for men (left) and women (right) in Korea, 2017, shown in 10 color groups with ranges for the categories determined by the deciles of age-standardized mortality rates per 100,000 population.

units depends on the number of deaths, with larger groups and those with higher mortality rates being mostly informed by their own data and smaller groups with lower mortality rates being affected by a combination of their own data and data in other units.

The following model¹³ was used to estimate mortality rates by age group, region, and year, and then LE was calculated based on the estimated mortality rates. All analyses were done separately by gender (men, women, and both sexes combined):

$$Deaths_{ita} \sim \text{Poisson}(m_{ita} \cdot Population_{ita}),$$

$$\log(m_{ita}) = \beta_0 + \beta_1 X_{it} + \gamma_{1,i} + (\gamma_{2,i} \cdot t + \gamma_{3,it}) + (\gamma_{4,i} \cdot \alpha + \gamma_{5,ia}) + \gamma_{6,ta},$$

where m_{ita} is the mortality rate in region i ($= 1, \dots, 250$ district), year t ($= 0, \dots, 13$, renumbered sequentially from 2004–2017), and age group α ($= 0, \dots, 18$, renumbered sequentially from 0, 1–4, 5–9, ..., 80–84, and 85+); β_0 is a fixed intercept; X_{it} is a vector of covariates for region i and year t , and β_1 is the associated vector of regression coefficients; $\gamma_{1,i}$ is a region-level random intercept; $\gamma_{2,i}$ and $\gamma_{4,i}$ are region-level random slopes for year and age group, respectively; $\gamma_{3,it}$, $\gamma_{5,ia}$, and $\gamma_{6,ta}$ are random intercepts on the region-year level, the region-age group level, and the year-age group level, respectively. $\gamma_{1,it}$, $\gamma_{2,it}$ and $\gamma_{4,i}$ were each assumed to follow a conditional autoregressive (CAR) distribution^{16,17} and these random effects allow for spatial variation beyond those already described by covariates at the overall level ($\gamma_{1,i}$), linear deviation from the overall time trend ($\gamma_{2,i}$), and linear deviation from the overall age group trend ($\gamma_{4,i}$). The spatial structure of the CAR prior was imposed through an adjacency matrix \mathbf{W} of size $n \times n$, with n being the number of regions. The diagonal entries of \mathbf{W} are 0, and the off-diagonal entries are specified such that $w_{ij} = 1$ if regions i and j are neighbors and $w_{ij} = 0$ otherwise. In this study, spatial weights based on Queen contiguity, which defines neighbors

as spatial regions sharing a common edge or a common vertex, were used. $\gamma_{3,it}$ and $\gamma_{5,ia}$ were assumed to follow $N(0, \sigma_{\gamma_3}^2)$ and $N(0, \sigma_{\gamma_5}^2)$, respectively. These random effects allow for non-linear region-level deviations in the year and age group trends that are not captured by other trend components (i.e., linear deviations are accounted for by $\gamma_{2,i}$ and $\gamma_{4,i}$). $\gamma_{6,ia}$ was assumed to follow a normal distribution, with the precision matrix specified as the Kronecker product of the structure matrices using a first-order random walk for each of year and age group (i.e., a type IV interaction in the Knorr-Held classification¹⁷), which allows for simultaneous smoothing over year and age groups. logGamma (1, 0.001) hyper-priors were specified for the logarithm of the precision of each random effect and normal (0, 1000) hyper-priors were used for the common intercept and slope.

Calculation of life expectancy at birth

LE at birth was calculated based on the life table method, using the mortality rates at the district level estimated from the spatio-temporal models. Separate life tables were established for each combination of gender (both sexes, men, and women), 250 districts, and 14 years. We employed the Kannisto-Thatcher method¹⁸ to expand the terminal age group 85+ to estimate the probability of dying for each of the 5-year age groups 85–89, 90–94, ..., 120–124, 125+. The 95% credible intervals (CIs) for LE at birth were calculated based on the 2.5th and 97.5th percentiles of the posterior distribution of the estimated mortality rates.

All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) and R version 3.4.1 (www.r-project.org). The models for estimating mortality rates were implemented in the R-INLA software. Subsequent procedures for estimating LE at birth were performed using SAS. The maps were created using QGIS version 3.12.0.

Ethics statement

The present study protocol was reviewed and approved by the National Health Insurance Service of Korea (NHIS-2020-1-257) and the Seoul National University Hospital Institutional Review Board (IRB No. E-2005-024-1121). The requirement for informed consent was exempted by the board.

RESULTS

Table 1 summarizes the results for death at the national and district levels from 2004 to 2017. The first section shows the number of population, the number of deaths, and age-standardized mortality rates at the national level: for example, in 2017, 145,935 of 25,062,848 men died, and the age-standardized mortality rate per 100,000 population was 548 in men, which was higher than the corresponding rate of 290 in women. The age-standardized mortality rate at the national level decreased by 333 deaths per 100,000 population in men and by 185 deaths per 100,000 population in women between 2004 and 2017. Differences in the age-standardized mortality gap between men and women tended to increase over time. The second section shows the distribution of age-standardized mortality rates in districts: for example, in 2017, the mean (standard deviation) of age-standardized mortality rates in men was 583 (80) deaths per 100,000 population, and the lowest mortality rate in men was 350 deaths per 100,000 population, compared with 537, 585, and 642 deaths per 100,000 population in districts in the 25th, 50th, and 75th percentiles, respectively, and 767 deaths per 100,000 population in the district with the highest rate. The difference between the minimum and maximum age-standardized mortality rate in districts decreased by 306 deaths

Table 1. National-level deaths, number of population, and age-standardized mortality rates, and district-level distribution of age-standardized mortality rates by gender in Korea, 2004–2017

Gender	Year	National-level deaths, population, and age-standardized mortality rate				District-level age-standardized mortality rate per 100 000 population									
		No. of population	No. of deaths	Mortality rate per 100,000	Change in mortality rates ^a	Mean	SD	Min	Q1	Q2	Q3	Max	Max – Min	Gap in Max – Min ^b	Moran's I
Men	2004	23,852,969	129,193	880		935	126	544	846	950	1,023	1,268	723		0.587
	2005	23,950,517	128,669	844	-36	898	120	552	824	910	983	1,223	671	-53	0.570
	2006	24,042,773	128,300	810	-70	860	117	506	784	867	937	1,138	632	-91	0.614
	2007	24,143,782	129,834	786	-94	833	114	478	753	839	917	1,122	645	-78	0.593
	2008	24,260,696	130,452	752	-128	802	114	433	728	811	892	1,100	667	-57	0.623
	2009	24,410,504	131,465	718	-162	768	111	439	694	776	851	1,030	592	-132	0.593
	2010	24,483,859	136,014	711	-170	759	108	452	689	762	833	1,030	578	-146	0.576
	2011	24,575,874	137,101	687	-194	732	102	425	668	741	804	997	572	-151	0.612
	2012	24,697,966	141,395	679	-201	725	103	409	662	730	796	995	586	-137	0.580
	2013	24,773,343	139,687	636	-244	677	92	394	622	685	734	927	533	-191	0.575
	2014	24,849,161	141,392	613	-267	654	91	389	600	657	711	952	562	-161	0.552
	2015	24,928,709	143,675	593	-287	635	91	371	579	633	696	921	550	-173	0.619
	2016	25,008,619	144,000	567	-313	606	85	362	553	611	664	793	431	-292	0.618
	2017	25,062,848	145,935	548	-333	583	80	350	537	585	642	767	417	-306	0.584
Women	2004	23,764,721	105,234	475		491	54	341	452	491	527	702	361		0.366
	2005	23,872,078	106,140	459	-16	472	50	325	441	475	507	637	311	-50	0.369
	2006	23,975,730	105,619	437	-38	451	47	317	420	449	481	608	291	-70	0.339
	2007	24,083,415	106,982	423	-52	435	45	324	407	434	461	557	233	-128	0.436
	2008	24,207,199	105,963	396	-79	410	44	287	380	413	438	542	256	-105	0.482
	2009	24,353,874	106,292	377	-98	389	43	244	365	389	415	496	252	-109	0.411
	2010	24,442,869	110,249	371	-104	384	41	236	356	386	411	491	255	-106	0.391
	2011	24,548,415	111,401	356	-119	368	41	231	341	369	395	514	283	-78	0.443
	2012	24,678,785	117,161	354	-121	366	40	217	344	370	390	479	262	-99	0.372
	2013	24,772,837	115,945	333	-142	344	41	226	322	345	369	500	275	-86	0.337
	2014	24,872,377	117,636	320	-155	330	34	230	311	333	351	409	179	-182	0.329
	2015	24,965,447	122,121	313	-162	325	35	235	305	326	346	420	185	-177	0.415
	2016	25,058,551	122,943	299	-176	308	32	217	288	310	329	412	195	-166	0.435
	2017	25,135,239	126,095	290	-185	299	32	218	279	300	322	397	179	-182	0.399

SD = standard deviation, Q1 = 25th percentile, Q2 = 50th percentile, Q3 = 75th percentile.

^aChange in mortality rates: the gap in age-standardized mortality rate per 100,000 population between 2004 and each year; ^bGap in Max – Min: the gap in the difference between the maximum and minimum age-standardized mortality rate per 100,000 population between 2004 and each year.

per 100,000 population in men and by 182 deaths per 100,000 population in women between 2004 and 2017. The last column in the second section shows Moran's I spatial statistic. From 2004 to 2017, Moran's I spatial statistic ranged from 0.552 to 0.623 for men and from 0.329 to 0.482 for women, indicating that there were geographic patterns of spatial clustering in age-standardized mortality rates across districts. The results for both sexes combined are provided in **Supplementary Table 1**.

In 2017, the LE at birth was 78.9 years (95% CI, 77.9–80.0 years) for men and 85.6 years (95% CI, 84.8–86.3 years) for women. For men, a high LE at birth was observed in the west (Seoul, Gyeonggi, Sejong, and Daejeon), while a low LE at birth was mainly observed in the northeast (Gangwon and Gyeongbuk) and south (Jeonnam and Gyeongnam). For women, high LE at birth was observed particularly in Sejong, Seoul, and Jeju, while low LE at birth was observed in some regions in Chungbuk, Ulsan, and Gyeongnam. Among 250 districts, LE at birth for men ranged from a minimum of 75.5 years (95% CI, 74.0–77.0 years) in Jindo-gun, Jeonnam to a maximum of 84.2 years (95% CI: 83.4–85.0) in Suji-gu, Yongin-si, Gyeonggi. LE at birth for women ranged from a minimum of 83.9 years (95% CI, 83.2–84.6 years) in Sasang-gu, Busan to a maximum of 88.2 years (95% CI, 87.3 to 89.1 years) in Gwacheon-si, Gyeonggi (**Fig. 2**). The

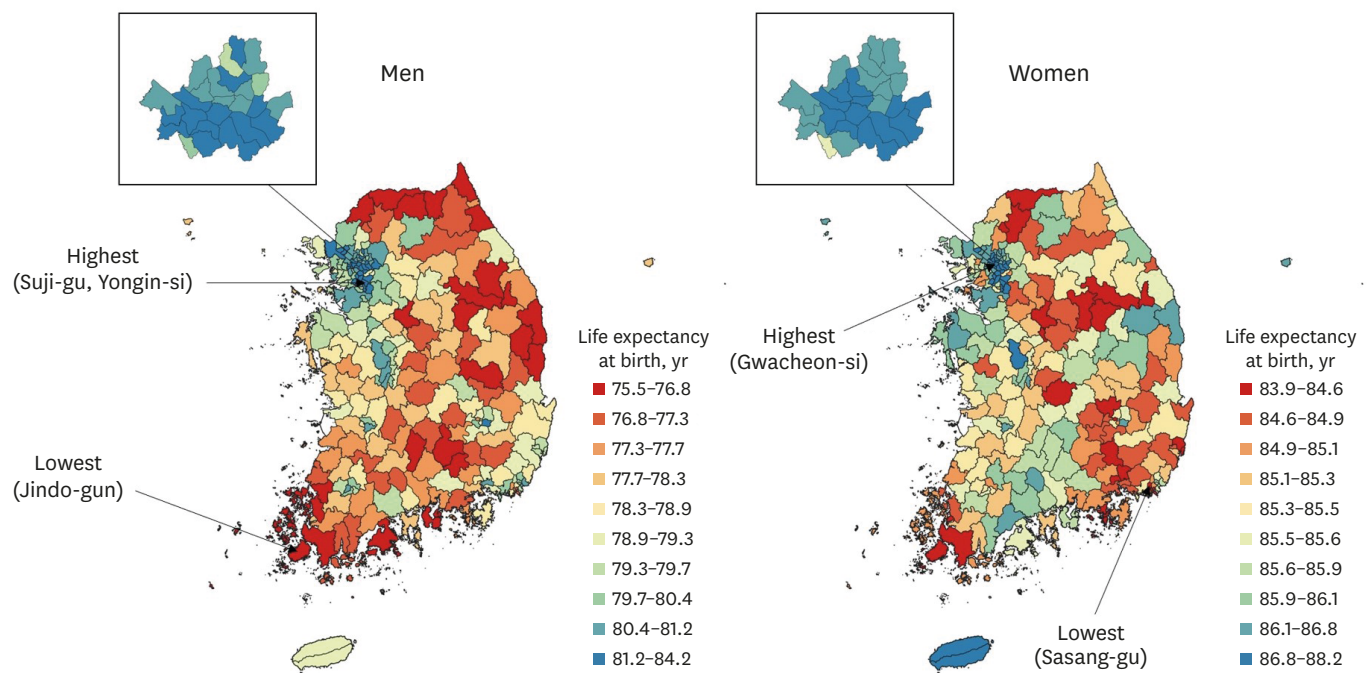


Fig. 2. Life expectancy at birth in 250 districts for men (left) and women (right) in Korea, 2017, shown in 10 color groups with ranges for the categories determined by the deciles of life expectancy at birth.

results of district-level LE at birth from 2004 to 2017 are presented in **Table 2**. The results for both sexes combined are given in **Supplementary Fig. 2** and **Supplementary Table 2**.

Between 2004 and 2017, district LE at birth increased by 4.57 years (95% CI, 4.49–4.65 years) for men (from 74.4 years [95% CI, 73.3–75.5 years] to 78.9 years [95% CI, 77.9–80.0 years]) and by 4.06 years (95% CI, 3.99–4.12 years) for women (from 81.5 years [95% CI, 80.7–82.4 years] to 85.6 years [95% CI, 84.8–86.3 years]) on average. LE at birth increased in all districts between 2004 and 2017, but the magnitude of those increases varied widely, from 2.42 years (95% CI, 2.35–2.44 years) in Yanggu-gun, Gangwon to 8.02 years (95% CI, 7.92–8.11 years) in Gangseo-gu, Busan for men, and from 2.87 years (95% CI, 2.81–2.90 years) in Gangjin-gun, Jeonnam to 5.70 years (95% CI, 5.60–5.78 years) in Sejong for women. For men, between 2004 and 2017, high increases in LE at birth were mainly observed in the southeast (Ulsan, Busan, Daegu, and nearby areas) and northwest (some areas of Gyeonggi and Sejong), while low increases in LE at birth were mainly observed in the north (Gangwon), east (Gyeongbuk), and southwest (near the border between Jeonnam and Jeonbuk). For women, high increases in LE at birth were mainly observed in the southeast (Busan, Gyeongnam, and Gyeongbuk) and northwest (Gyeonggi and Sejong), while low increases in LE at birth were mainly observed in band-shaped areas that cross the country from north to south (**Fig. 3**). The results for both sexes combined are given in **Supplementary Fig. 3**.

The gap in LE at birth between men and women decreased by 0.51 years (from 7.15 years in 2004 to 6.64 years in 2017). The gap between the lowest and highest LE at birth among districts decreased by 0.55 years (from 9.24 years to 8.69 years) for men and by 0.89 years (from 5.23 years to 4.34 years) for women between 2004 and 2017. The corresponding interquartile ranges increased by 0.12 years (from 2.28 years to 2.40 years) for men, but decreased by 0.09 years (from 1.08 years to 0.98 years) for women between 2004 to 2017.

Table 2. Distribution of district-level life expectancies at birth based on the spatio-temporal model by gender, 2004–2017

Gender	Year	District-level life expectancy at birth								
		Mean	SD	Min	Q1	Q2	Q3	Max	Max – Min	Gap in Max – Min ^a
Men	2004	74.4	1.7	71.1	73.0	74.0	75.3	80.3	9.2	
	2005	74.8	1.7	71.4	73.5	74.4	75.8	80.4	9.0	–0.3
	2006	75.1	1.7	72.0	73.8	74.8	76.1	80.8	8.7	–0.5
	2007	75.4	1.7	72.2	74.1	75.0	76.4	81.2	9.1	–0.2
	2008	75.8	1.8	72.5	74.4	75.5	76.9	81.5	9.0	–0.3
	2009	76.1	1.8	72.8	74.7	75.9	77.2	82.1	9.3	–0.0
	2010	76.4	1.7	73.4	75.1	76.2	77.4	82.2	8.8	–0.5
	2011	76.7	1.7	73.7	75.3	76.5	77.8	82.4	8.7	–0.6
	2012	76.8	1.7	73.8	75.5	76.7	77.9	82.6	8.8	–0.4
	2013	77.4	1.8	74.2	76.0	77.2	78.5	83.1	8.9	–0.4
	2014	77.6	1.8	74.4	76.2	77.5	78.8	83.2	8.8	–0.4
	2015	78.3	1.7	75.2	76.9	78.2	79.4	83.6	8.4	–0.8
	2016	78.7	1.7	75.4	77.2	78.5	79.7	84.0	8.6	–0.6
	2017	78.9	1.8	75.5	77.5	78.9	79.9	84.2	8.7	–0.6
Women	2004	81.5	0.9	79.3	80.9	81.4	82.0	84.5	5.2	
	2005	81.8	0.9	79.7	81.2	81.7	82.2	84.7	5.1	–0.2–0.5
	2006	82.0	0.8	80.1	81.5	81.9	82.5	84.9	4.8	–0.4
	2007	82.4	0.8	80.5	81.8	82.3	82.9	85.3	4.8	–0.3
	2008	82.9	0.9	80.9	82.3	82.8	83.3	85.9	5.0	–0.1
	2009	83.3	0.9	81.4	82.8	83.2	83.8	86.5	5.1	–0.7
	2010	83.5	0.9	81.8	83.0	83.4	83.9	86.4	4.6	–0.8
	2011	83.8	0.8	82.3	83.3	83.7	84.2	86.7	4.4	–0.8
	2012	83.9	0.8	82.2	83.3	83.7	84.3	86.6	4.4	–1.0
	2013	84.4	0.8	82.8	83.8	84.2	84.7	87.1	4.3	–0.9
	2014	84.6	0.8	83.0	84.0	84.5	85.0	87.4	4.4	–0.6
	2015	85.0	0.9	83.0	84.5	84.9	85.4	87.7	4.7	–0.7
	2016	85.4	0.9	83.6	84.8	85.3	85.8	88.1	4.5	–0.9
	2017	85.6	0.8	83.9	85.0	85.5	86.0	88.2	4.3	

SD = standard deviation, Q1 = 25th percentile, Q2 = 50th percentile, Q3 = 75th percentile.

^aGap in Max – Min: the gap in the difference between the maximum and minimum life expectancy at birth between 2004 and each year.

In 2017, very high LEs at birth, denoted by dots as outliers in the box plots, were observed in Suji-gu, Yongin-si, Gyeonggi; Bundang-gu, Seongnam-si, Gyeonggi; Seocho-gu, Seoul; Gangnam-gu, Seoul; Giheung-gu, Yongin-si, Gyeonggi; and Gwacheon-si, Gyeonggi for men, and observed in Gwacheon-si, Gyeonggi; Gangnam-gu, Seoul; Suji-gu, Yongin-si, Gyeonggi; Dongjak-gu, Seoul; Yongsan-gu, Seoul; Bundang-gu, Seongnam-si, Gyeonggi; Seocho-gu, Seoul; and Jung-gu, Seoul for women (Fig. 4). The results for both sexes combined are given in Supplementary Fig. 4.

Table 3 shows a comparison of LE at birth at the district level, calculated from the observational data of NHID and from the spatio-temporal model. The absolute median differences between the observed and model-based LE at birth at the district level from 2004 to 2017 were less than 0.4 years for men and 0.3 years for women. The correlation coefficients between the observed and model-based LE at birth at the district level from 2004 to 2017 were 0.879 to 0.943 for men and 0.624 to 0.781 for women. The reason for the low correlation coefficients of LE at birth was that in some districts, LE at birth calculated from the observed mortality rates was unstable due to the small population numbers in each age group in the life table. For example, the observed and model-based LEs at birth at the district level in 2017 were quite different in some districts with small populations, such as Ongjin-gun, Imsil-gun, and Yanggu-gun (Fig. 5). To explore this issue, based on the finding of a previous study⁹ that the minimum population required for a stable LE was 5000, we calculated the LE at birth for 250 districts using 7-year pooled data for population and deaths (2004–2010 and 2011–2017),

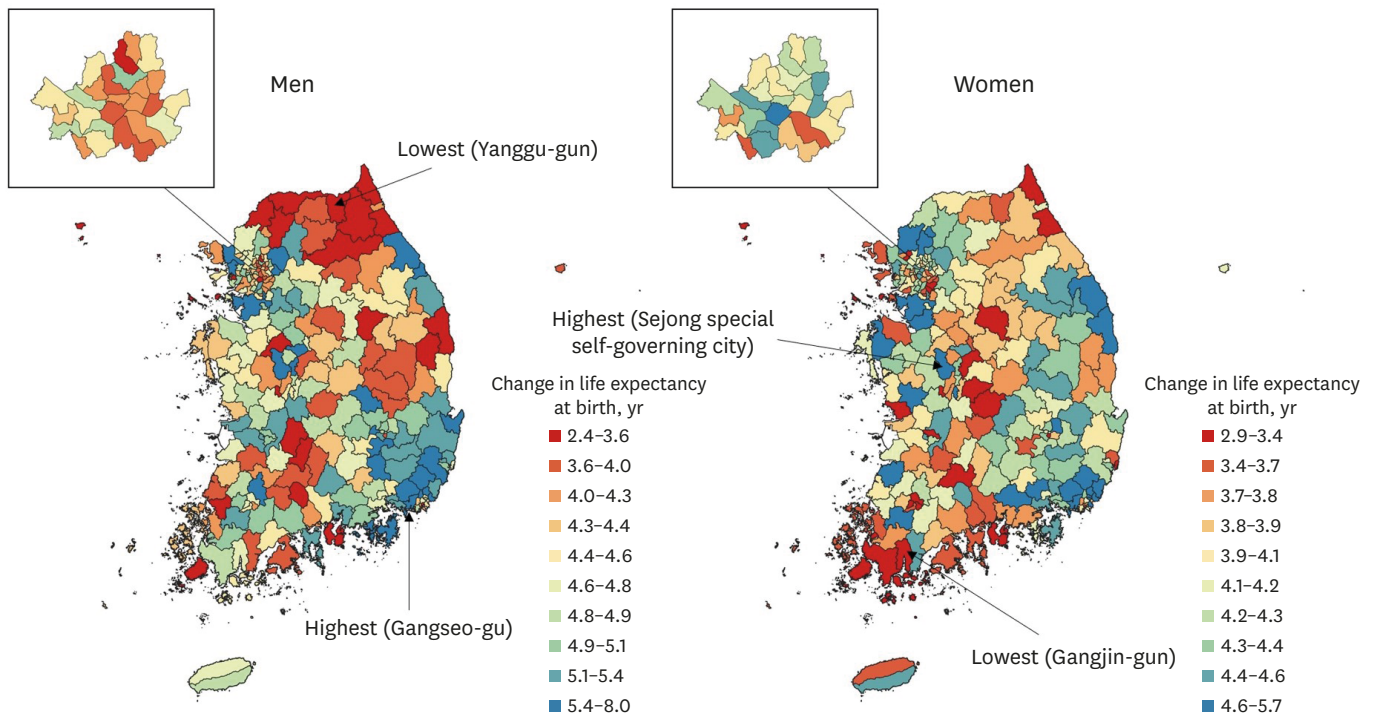


Fig. 3. Change in life expectancy at birth in 250 districts between 2004 and 2017 for men (left) and women (right) in Korea, shown in 10 color groups with ranges for the categories determined by the deciles of changes in life expectancy at birth.

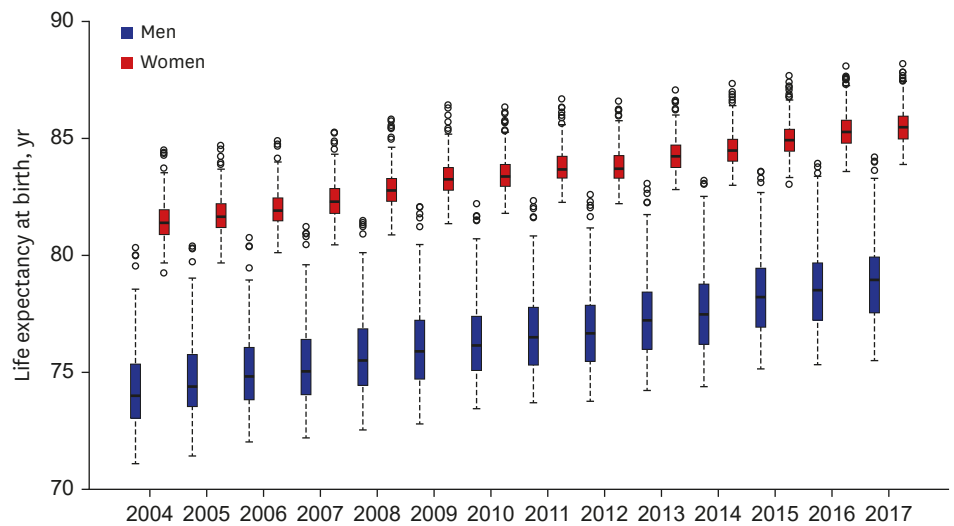


Fig. 4. Distribution of life expectancy at birth in 250 districts from 2004 to 2017 by gender in Korea; the bottom border, middle line, and top border of the boxes indicate the 25th, 50th, and 75th percentiles, respectively, across all districts; the dashed lines (whiskers) represent the full range across districts except outliers; and circles (outliers) represent the observations with the highest or lowest life expectancy at birth.

and the newly obtained correlation coefficients are in the last column in **Table 3**, showing that the observed and model-based LEs at birth were highly correlated. Thus, using LE obtained from spatio-temporal models, rather than LE calculated from the observed mortality rates, enables stable estimates of annual district-level LE to be obtained. The results for both sexes combined are given in **Supplementary Table 3**.

Spatio-temporal Analysis of District-level Life Expectancy

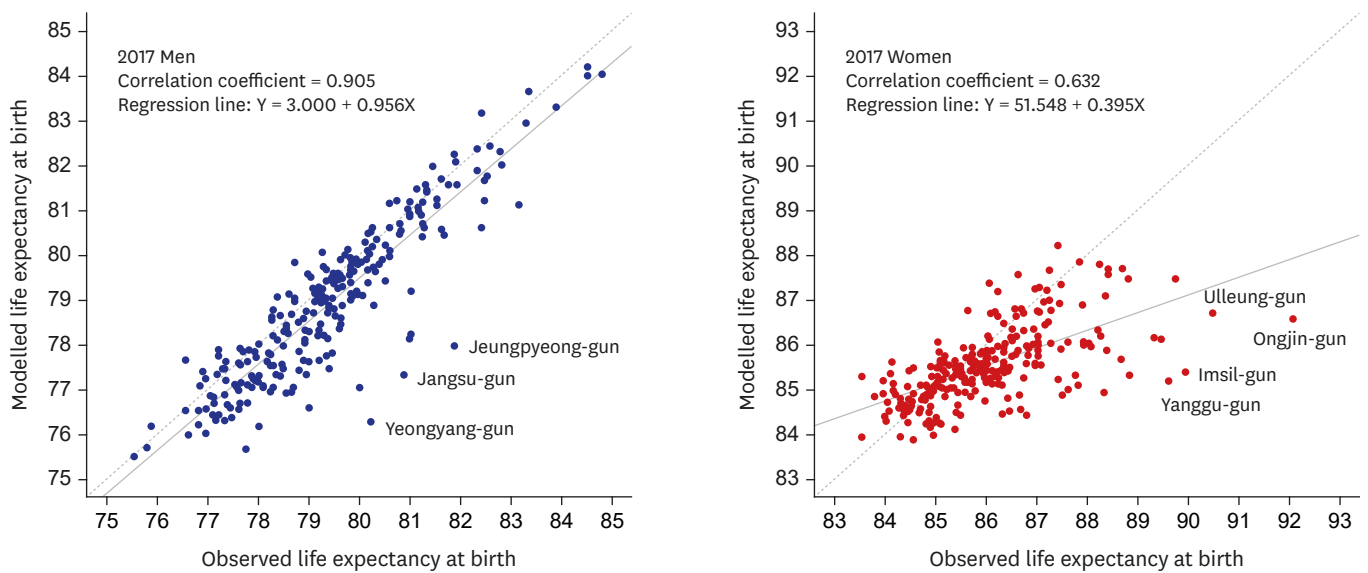


Fig. 5. Comparison of the life expectancy at birth calculated from observed mortality rates (x-axis) and the life expectancy at birth based on the spatio-temporal model (y-axis) in 2017; the solid lines represent the fitted regression lines and the dashed lines represent the straight lines with slope 1.

Table 3. Annual comparison of median results between the observed and model-based life expectancies at birth at the district level by gender

Gender	Year	Median life expectancy at birth				
		Observed ^a	Modeled ^b	Diff ^c	Corr ^d	Corr2 ^e
Men	2004	73.6	74.0	-0.4	0.930	0.992
	2005	74.1	74.4	-0.3	0.943	
	2006	74.9	74.8	-0.1	0.925	
	2007	75.2	75.0	-0.2	0.933	
	2008	75.7	75.5	0.0	0.911	
	2009	76.1	75.9	-0.0	0.917	
	2010	76.3	76.2	-0.1	0.899	
	2011	76.7	76.5	0.1	0.917	
	2012	77.0	76.7	0.1	0.923	
	2013	77.6	77.2	0.3	0.892	
	2014	78.1	77.5	0.4	0.891	
	2015	78.6	78.2	0.2	0.913	
	2016	78.9	78.5	0.3	0.879	
	2017	79.3	78.9	0.3	0.905	
Women	2004	81.2	81.4	-0.3	0.759	0.952
	2005	81.4	81.7	-0.3	0.781	
	2006	81.8	81.9	-0.1	0.682	
	2007	82.2	82.3	-0.1	0.685	
	2008	82.8	82.8	-0.0	0.742	
	2009	83.2	83.2	-0.1	0.688	
	2010	83.4	83.4	-0.0	0.624	
	2011	83.9	83.7	0.0	0.770	
	2012	84.0	83.7	0.1	0.687	
	2013	84.6	84.2	0.2	0.653	
	2014	84.8	84.5	0.2	0.624	
	2015	85.1	84.9	0.1	0.662	
	2016	85.7	85.3	0.3	0.635	
	2017	85.9	85.5	0.3	0.632	

All correlation coefficients in Corr and Corr2 are statistically significant (P value < 0.001).

^aObserved: median life expectancy at birth at the district level calculated from the observational data of the National Health Information Database; ^bModeled: median life expectancy at birth at the district level from the spatio-temporal model; ^cDiff: median difference between the observed and model-based life expectancies at birth at the district level; ^dCorr: correlation between the observed and model-based life expectancies at birth; ^eCorr2: correlation between the observed and model-based life expectancies at birth based on 7-year pooled data.

DISCUSSION

In this study, we estimated annual district-level LE using the NHID DB from 2004 to 2017 through spatio-temporal analysis. Compared to purely spatial analysis or time-series analysis, spatio-temporal analysis has the advantage of being able to simultaneously study the effects of space and time. Spatio-temporal models allow for various kinds of interactions and relationships (spatial dependencies, temporal dependencies, and space-time dependencies) that may be present in the data, thus avoiding inappropriate model fitting (e.g., by ignoring dependencies) that may lead to incorrect conclusions.

Estimating LE based on spatio-temporal modeling, as done in this study, is necessary for the following reasons. First, model-based LEs in districts with small populations are more stable than LEs calculated under the assumption of independent observations. Second, annual LEs can be calculated without pooling multi-year data to secure the minimum population number required for a stable LE in each geographic area. Third, since the mortality rates are similar in adjacent regions, it is statistically inappropriate to conduct an analysis assuming that the mortality rates in each area are independent of each other. This study is meaningful in that it is the first study, to the best of our knowledge, to estimate LE at the district level in Korea by addressing the above issues.

Table 2 and **Fig. 4** present decreases in the gaps in LE across districts in Korea. These findings are in line with prior Korean studies, which also reported similar temporal decreases in mortality metrics across districts in Korea.^{19,20} In addition, mortality inequalities by geographic deprivation index have decreased in Korea.^{21,22} These temporal patterns in geographic mortality inequalities are in contrast with the increasing inequalities in mortality metrics (LE and age-adjusted mortality rates) by individual socioeconomic position indicators, such as income and education, in Korea.²³⁻²⁵

Study results showed that the Moran's I was greater in men than in women. The gender difference in the Moran's I means that geographical clustering in mortality is more pronounced in men than in women in Korea. In Korea as well as in many other countries, the magnitude of socioeconomic inequalities in mortality and life expectancy is greater in men than in women.^{5,7,23-25} If the geographic areas in Korea reflect socio-economic position and such geographic areas tend to cluster, it is expected that the magnitude of clustering in mortality would be greater in men than in women.

This study has certain limitations that should be acknowledged. First, when constructing the adjacency weight matrix **W** in the model, it was not considered whether there were obstacles (e.g., rivers, mountains) between regions. However, most of those regions are actually connected by bridges or tunnels. Second, there are many islands in the west and south of Korea, so it was not possible to create adjacency weights based on the Queen criterion (i.e., neighbors if they share a common edge or a common vertex) in those regions. Therefore, if the island regions were nearby on the map, they were considered adjacent regions. In fact, most islands are connected by highway or ship routes, so people can come and go. In the case of Ulleungdo island on the east coast, neighboring regions were defined based on four ship routes.

Based on the results of this study, the annual district-level LE could be used as a performance indicator of the local government's health policies to improve the health of local residents, and information on the LE gap between districts will ultimately contribute to stimulating

health-equity policies and projects at the district level. It is also expected that annually monitoring and evaluating LE gaps across regions at the district level will help achieve a long-term health equity policy.

The approach to district-level analysis with a spatio-temporal modeling employed in this study could be used in future analyses to produce district-level health-related indicators (e.g., healthy LE) in Korea, such as forecasts of future mortality and LE, or cause-specific mortality rates.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

National-level deaths, number of population, age-standardized mortality rates, and district-level distribution of age-standardized mortality rates for both sexes combined, 2004-2017

[Click here to view](#)

Supplementary Table 2

Distribution of district-level life expectancies at birth based on the spatio-temporal model for both sexes combined, 2004–2017

[Click here to view](#)

Supplementary Table 3

Annual comparison of median results between the observed and model-based life expectancies at birth at the district level for both sexes combined

[Click here to view](#)

Supplementary Table 4

The observed and spatio-temporal model-based life expectancies at birth at the district level by gender in 2017

[Click here to view](#)

Supplementary Fig. 1

Age-standardized mortality rates per 100,000 population in 250 districts for both sexes in Korea, 2017, shown in 10 color groups with ranges for the categories determined by the deciles of age-standardized mortality rates per 100,000 population.

[Click here to view](#)

Supplementary Fig. 2

Life expectancy at birth in 250 districts for both sexes combined in Korea, 2017, shown in 10 color groups with ranges for the categories determined by the deciles of life expectancy at birth.

[Click here to view](#)

Supplementary Fig. 3

Change in life expectancy at birth in 250 districts between 2004 and 2017 for both sexes combined in Korea, shown in 10 color groups with ranges for the categories determined by the deciles of changes in life expectancy at birth.

[Click here to view](#)

Supplementary Fig. 4

Distribution of life expectancy at birth in 250 districts from 2004 to 2017 for both sexes combined in Korea; the bottom border, middle line, and top border of the boxes indicate the 25th, 50th, and 75th percentiles, respectively, across all districts; the dashed lines (whiskers) represent the full range across districts except outliers; and circles (outliers) represent the observations with the highest life expectancy at birth.

[Click here to view](#)

REFERENCES

1. Korea Health Promotion Institute. Guideline for establishment of the 7th local healthcare plan. http://www.khealth.or.kr/kps/publish/view?menuId=MENU00890&page_no=B2017003&pageNum=1&siteId=&srch_text=&srch_cate=28&srch_type=ALL&str_clft_cd_list=&str_clft_cd_type_list=&board_idx=10127. Updated 2018. Accessed February 15, 2020.
2. Seong SC, Kim YY, Khang YH, Heon Park J, 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](#) | [CROSSREF](#)
3. Bahk J, Kim YY, Kang HY, Lee J, Kim I, Lee J, et al. Using the National Health Information Database of the National Health Insurance Service in Korea for monitoring mortality and life expectancy at national and local levels. *J Korean Med Sci* 2017;32(11):1764-70.
[PUBMED](#) | [CROSSREF](#)
4. Kim I, Lim HK, Kang HY, Khang YH. Comparison of three small-area mortality metrics according to urbanity in Korea: the standardized mortality ratio, comparative mortality figure, and life expectancy. *Popul Health Metr* 2020;18(1):3.
[PUBMED](#) | [CROSSREF](#)
5. Lim D, Bahk J, Ock M, Kim I, Kang HY, Kim YY, et al. Income-related inequality in quality-adjusted life expectancy in Korea at the national and district levels. *Health Qual Life Outcomes* 2020;18(1):45.
[PUBMED](#) | [CROSSREF](#)
6. Bahk J, Kang HY, Khang YH. Life expectancy and inequalities therein by income from 2016 to 2018 across the 253 electoral constituencies of the National Assembly of the Republic of Korea. *J Prev Med Public Health* 2020;53(2):143-8.
[PUBMED](#) | [CROSSREF](#)
7. Khang YH, Lim D, Bahk J, Kim I, Kang HY, Chang Y, et al. A publicly well-accepted measure versus an academically desirable measure of health inequality: cross-sectional comparison of the difference between income quintiles with the slope index of inequality. *BMJ Open* 2019;9(6):e028687.
[PUBMED](#) | [CROSSREF](#)
8. Kim I, Bahk J, Yun SC, Khang YH. Income gaps in self-rated poor health and its association with life expectancy in 245 districts of Korea. *Epidemiol Health* 2017;39:e2017011.
[PUBMED](#) | [CROSSREF](#)
9. Eayres D, Williams ES. Evaluation of methodologies for small area life expectancy estimation. *J Epidemiol Community Health* 2004;58(3):243-9.
[PUBMED](#) | [CROSSREF](#)
10. Bennett JE, Li G, Foreman K, Best N, Kontis V, Pearson C, et al. The future of life expectancy and life expectancy inequalities in England and Wales: Bayesian spatiotemporal forecasting. *Lancet* 2015;386(9989):163-70.
[PUBMED](#) | [CROSSREF](#)

11. Bennett JE, Pearson-Stuttard J, Kontis V, Capewell S, Wolfe I, Ezzati M. Contributions of diseases and injuries to widening life expectancy inequalities in England from 2001 to 2016: a population-based analysis of vital registration data. *Lancet Public Health* 2018;3(12):e586-97.
[PUBMED](#) | [CROSSREF](#)
12. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Kutz MJ, Huynh C, et al. US county-level trends in mortality rates for major causes of death, 1980–2014. *JAMA* 2016;316(22):2385-401.
[PUBMED](#) | [CROSSREF](#)
13. Dwyer-Lindgren L, Stubbs RW, Bertozzi-Villa A, Morozoff C, Callender C, Finegold SB, et al. Variation in life expectancy and mortality by cause among neighbourhoods in King County, WA, USA, 1990–2014: a census tract-level analysis for the Global Burden of Disease Study 2015. *Lancet Public Health* 2017;2(9):e400-10.
[PUBMED](#) | [CROSSREF](#)
14. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Mackenbach JP, van Lenthe FJ, et al. Inequalities in life expectancy among US Counties, 1980 to 2014: temporal trends and key drivers. *JAMA Intern Med* 2017;177(7):1003-11.
[PUBMED](#) | [CROSSREF](#)
15. Moran PA. Notes on continuous stochastic phenomena. *Biometrika* 1950;37(1-2):17-23.
[PUBMED](#) | [CROSSREF](#)
16. Leroux BG, Lei X, Breslow N. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. New York, NY: Springer-Verlag; 2000, 179-91.
17. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. *Stat Med* 2000;19(17-18):2555-67.
[PUBMED](#) | [CROSSREF](#)
18. Thatcher AR, Kannisto V, Vaupel JW. *The Force of Mortality at Ages 80 to 120*. Odense, Denmark: Odense University Press; 1998.
19. Khang YH, Hwang IA, Yun SC, Lee MS, Lee SI, Jo MW, et al. Census population vs. registration population: which population denominator should be used to calculate geographical mortality. *J Prev Med Public Health* 2005;38(2):147-53.
[PUBMED](#)
20. Khang YH, Yoon SC, Bahk JW, Kang HY, Kim I, Seo SB, 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. National Health Insurance Service & Seoul National University R&DB Foundation, 2016. http://lib.nhis.or.kr/search/detail/CATXBZ000000031459?briefLink=/searchA/xbz?bk_2=jtj2016jtj. Updated 2016. Accessed March 20, 2020.
21. Choi MH, Yoon TH. Trends in amenable mortality to health care across the districts in Korea from 1993 to 2013. *J Crit Soc Welfare* 2015;(49): 404-32.
22. Khang YH. Monthly Community Health Report. Geographic health inequalities: time trends in mortality indicators and future research agenda. <https://chs.cdc.go.kr/chs/hd/hdMain.do>. Updated October 2019. Accessed March 20, 2020.
23. Khang YH, Bahk J, Lim D, Kang HY, Lim HK, Kim YY, et al. Trends in inequality in life expectancy at birth between 2004 and 2017 and projections for 2030 in Korea: multiyear cross-sectional differences by income from national health insurance data. *BMJ Open* 2019;9(7):e030683.
[PUBMED](#) | [CROSSREF](#)
24. Bahk J, Lynch JW, Khang YH. Forty years of economic growth and plummeting mortality: the mortality experience of the poorly educated in South Korea. *J Epidemiol Community Health* 2017;71(3):282-8.
[PUBMED](#) | [CROSSREF](#)
25. Son M, Cho Y, Oh J, Kawachi I, Yi J, Kwon S. Social inequalities in life expectancy and mortality during the transition period of economic crisis (1993–2010) in Korea. *Int J Equity Health* 2012;11(1):71.
[PUBMED](#) | [CROSSREF](#)