Independent effect of physical activity and resting heart rate on the incidence of atrial fibrillation in the general population

Yeon Woo Choi1, Minsu Park2, Young-Hyo Lim3, Jisun Myung3, Byung Sik Kim3, Yonggu Lee4, Jeong-Hun Shin5, Hwan-Cheol Park4, Jinho Shin3, Chun Ki Kim5 & Jin-Kyu Park1

While physical activity (PA) may influence resting heart rate (RHR), and a low RHR may be a risk factor for atrial fibrillation (AF), controversy exists regarding the association between PA and development of AF. Using data from a Korean, prospective population cohort, we investigated the independent effect of PA and RHR on the incidence of AF in the general population. A total of 8,811 participants aged 40–69 years were analyzed. Total PA assessed based on questionnaires was divided into quartiles, with the lowest to the highest being Q1, Q2, Q3, and Q4. During a median follow-up of 139 months, AF developed in 167 participants (1.9%). Q3 of total PA was associated with a significantly lower risk of AF than Q1 even after adjusting for RHR as a covariate, but Q4 was not. The risk of AF was higher in participants with RHR < 60 bpm than in those with RHR 70–85 bpm, and the significance persisted after adjusting for PA as a covariate. This study showed that a moderate amount of total PA was associated with a lower risk of incident AF independent of RHR and that low RHR was an independent risk factor for AF in the general Korean population.

Atrial fibrillation (AF) is the most common arrhythmia and is an independent risk factor for cardiovascular events including heart failure, and ischemic stroke. Risk factors for the development of AF include hypertension, diabetes, obesity, smoking, ischemic heart disease, and heart failure.

Association between physical activity (PA) and AF has also been studied, and moderate PA is reported to be associated with a reduced risk of AF; while higher PA is not. On the other hand, a meta-analysis showed that the trend toward lower risk of AF in subjects practicing PA is non-significant and that the extreme heterogeneity of published studies does not allow conclusive results.

There are other issues related to PA. While previous studies divided PA into the light, moderate, and vigorous activities, none of them quantified the effect of total daily life PA. In addition, it is still not completely clear whether the PA and resting heart rate (RHR) have an independent effect on the development of AF. Endurance athletes are more likely to develop AF than non-athletes, with sinus bradycardia and increased vagal tone as possible pathophysiological mechanisms. Therefore, PA and RHR may presumably also influence each other in the general population, and a relation between RHR and incidence of AF has indeed been suggested. However, there are limited data about the interplay between PA and RHR in the development of AF. Further, there have been no large-scale studies analyzing the effect of PA and RHR on the incidence of AF in an Asian population.

Our study aimed to assess (1) whether the total PA and RHR have an independent effect on incident AF and (2) whether the amount of total daily activity has any influence on the risk of AF in the general population based on a prospective population study carried out in South Korea.

1Division of Cardiology, Department of Internal Medicine, Hanyang University Medical Center, Seoul, Republic of Korea. 2Statistics and Data Center, Samsung Biomedical Research Institute, Samsung Medical Center, Seoul, Republic of Korea. 3Department of Preventive Medicine, School of Medicine, Kyunghee University, Seoul, Republic of Korea. 4Division of Cardiology, Department of Internal Medicine, Hanyang University Guri Hospital, Guri City, Gyounggi-do, Republic of Korea. 5Department of Nuclear Medicine, Hanyang University Medical Center, Seoul, Republic of Korea. Yeon Woo Choi, Minsu Park and Young-Hyo Lim contributed equally. Correspondence and requests for materials should be addressed to J.-K.P. (email: cardiohy@gmail.com)
Methods

Study population. The Ansung-Ansan cohort is a prospective study population that consists of South Koreans aged 40–69 years residing in two cities, Ansan and Ansung. This study was started in 2001 and embedded within the Korean Genome and Epidemiology Study (KoGES), a population-based cohort study, to assess the incidence and risk factors for various chronic disorders including cardiovascular diseases.

The baseline survey, conducted from 2001 to 2002, included 10,030 adults. Following the baseline visit, these adults were examined every two years with the longest follow-up duration being 12 years. At each of the baseline and biennial follow-up visits, data on lifestyle characteristics, cardiovascular profiles, medical history, and disease incidence were collected. All medical tests and the interview-based surveys were conducted by health professionals who were trained to follow a standardized protocol, and the same instruments were used to collect the data. More detailed information regarding the enrollment of cohort members and study procedures is available in a previous report9.

Analytical sample. Among the 10,030 subjects, we analyzed 8,811 (4,235 males and 4,576 females), excluding participants whose heart rate measurement was not available (n = 15), those with AF at baseline (n = 40), those with no follow-up information (n = 965), and those with missing information on covariates (n = 199). All of the study protocol were approved by the Ethics Committee of KoGES at the Korean National Institute of Health and adhered to the principles of the Declaration of Helsinki. Informed consent was obtained from all subjects. This study was approved by the Institutional Review Board of Hanyang University Medical Center (IRB file No. HYUH 2017-12-033).

Physical activity measurement. PA was assessed using a questionnaire to quantify activities in occupational or leisure time domains and was categorized into 5 levels: steady state, sitting state, and light, moderate, and high activity. A corresponding metabolic equivalent (MET) was assigned for each level as follows: (1) steady state (MET 0) - lying down; (2) sitting state (MET 1.5) - typing, driving, office work, playing a musical instrument, sewing, studying, ironing, writing, and cooking; (3) light activity (MET 3) - walking, cleaning, laundry, child care, bathing, and exercise as entertainment (bicycle, table tennis); (4) moderate activity (MET 5) - fast walking, woodworking, lawn mowing, snow removal, and regular exercise (badminton, swimming, tennis); and (5) high activity (MET 7) - athletics, climbing, running, woodcutting, agriculture, forestry, and mining. The questionnaire included questions about whether participants performed each of the above 5 levels of PA ≥ 1 h/day or < 1 h/day. The total PA in MET-h/day was based on the PA intensity and the time commitment for each PA.

Measurement of resting heart rate and other covariates. RHR was recorded by measuring the pulse for 60 s. Covariates including baseline age, sex, urban or rural residence, education level, past medical history, hypertension, smoking status, and alcohol consumption were evaluated based on self-reports. Blood pressure (BP) was measured by well-trained nurses after a 5-min rest. Hypertension was defined as systolic BP > 140 mmHg, diastolic BP > 90 mmHg, or the use of antihypertensive medication. We defined a participant as diabetic if any of the following criteria were met: current use of an oral hypoglycemic agent or insulin or a serum fasting glucose ≥ 126 mg/dl. A participant was categorized as having hyperlipidemia if any of the following criteria were met: current use of lipid-lowering medication, serum total cholesterol ≥ 240 mg/dl, or serum triglyceride ≥ 200 mg/dl.

Follow-up and detection of AF. Biennial follow-up examinations and electrocardiograms (ECGs) were performed until 2014. A resting supine 12-lead ECG (MAC 5000®; GE Marquette Inc., Milwaukee, WI, USA) was obtained during quiet respiration. The ECG was recorded at 25 mm/s with 0.1 mV/mm standardization and was interpreted by a cardiologist to confirm the diagnosis and classified according to the Minnesota code. AF was defined as a composite of AF or atrial flutter in the standard 12-lead ECG.

AF at baseline was identified by ECG performed during the baseline visit and/or the self-reported history of physician-determined diagnosis made prior to the baseline visit. Similarly, AF newly developed after the baseline visit was also identified by biennial performed ECGs and/or the self-reported history of physician-determined diagnosis that was made between each biennial visit. If AF was identified by both ECGs and the history in the same subject, the earlier identification date was considered the time of AF development.

Statistical analyses. Continuous variables were presented as a median and interquartile range; Kruskal-Wallis test and Dunn’s multiple comparison test (post-hoc analysis), comparing all groups to controls were used to compare continuous variables among groups. Categorical variables were presented as numbers and percentages. The chi-squared test was used for categorical variables. A Cox proportional hazards regression model was used to assess the association between PA, RHR, and incidence of AF. The collected total PA (MET-h/day) was divided into quartiles, with the lowest and the highest being Q1 and Q4, respectively. Q1 was used as the reference. RHR was categorized into 5 groups (< 50, 50–59, 60–70, 70–85, and > 85 bpm). The RHR range of 70–85 bpm was used as a reference. The covariates were age, sex, residence, education, body mass index (BMI), comorbidity, smoking status, and alcohol consumption in models 1 and 2. In model 3, we added RHR or PA to evaluate the independent association with the incidence of AF. All statistical analyses were performed using R-3.4.0 software. A p < 0.05 was considered significant.

Results

Baseline demographic and health-related characteristics. The baseline characteristics according to the PA category are listed in Tables 1 and 2. A total of 8,811 participants (48.1% male) were followed up for a median period of 139 (97, 141) months. The median age was 50 (44, 60) years. There were several significant differences between participants in the lowest quartile (Q1) and the highest quartile (Q4): The median age of participants was higher, there were more male participants, the % of urban residents (=urban residents/(urban + rural...
RHR remained identical throughout all follow-up visits (Supplementary Fig. 2). This section and also because initial rankings of 5 RHR groups determined based on median values of baseline Q3). More detailed components of PA in each of the four quartiles of total PA are shown in Supplementary Fig. 1.

Table 1. Demographic characteristics according to quartiles of physical activity. Data shown with percentages in parentheses represent numbers of participants, and data shown with two numbers in parentheses represent median values with interquartile ranges in parentheses. †p < 0.01 vs. Q1. Kruskal-Wallis test (Dunn’s post hoc).

| Occupation                        | Total | Q1   | Q2   | Q3   | Q4   | p-value |
|-----------------------------------|-------|------|------|------|------|---------|
| Housewife                         | 2,453 | 728  | 711  | 742  | 722  | 0.001   |
| Office work                       | 391   | 123  | 151  | 98   | 19   | 0.001   |
| Agriculture                       | 2,486 | 343  | 175  | 410  | 1,558| 0.001   |
| Self-employment                   | 1,250 | 349  | 403  | 381  | 117  | 0.001   |
| Sales                             | 110   | 29   | 42   | 28   | 11   | 0.001   |
| Factory production                | 470   | 114  | 125  | 138  | 93   | 0.001   |
| Professional work                 | 357   | 106  | 101  | 109  | 41   | 0.001   |
| Others                            | 1,258 | 415  | 336  | 343  | 164  | 0.001   |

There was a gradual decrease in the incidence rate and risk of AF from Q1 to Q3 (Table 3); Q3 had the lowest incidence rate of AF (1.42/1,000 person-years) and a significantly lower risk of AF, after adjusting for covariates, than Q1 and Q4. This trend remained consistent after adjusting for RHR (hazard ratio [HR] 0.58, 95% confidence interval [CI] 0.37–0.90, p = 0.016). On the other hand, Q4 had a high incidence rate of AF (2.12/1,000 person-years); the risk of AF was not even lower than in residents (Table 1). There was a significant difference in the distribution of occupations. The large majority (62.7%) of participants engaged in agriculture belonged in Q4. Conversely, the smallest proportion of participants with all other occupations belonged in Q4 (Table 1).

BMI was significantly lower in Q4 than in Q1. However, the differences in age and BMI between Q3 and Q1 were not significant. The frequencies of hyperlipidemia and thyroid disease were lower among participants in Q4 than those in the other three quartiles. There were more current or former smokers in Q4 than in other quartiles but no significant differences in alcohol consumption among four quartiles (Table 2).

**Physical activity and resting heart rate.** Only the total PA measured at baseline was analyzed for the following reasons: 1) The total number of follow-up visits varied among participants; therefore, using the baseline data that were available in every participant would likely provide the most generalizable results. Further, including all data from varying numbers of follow-up visits could complicate the analysis, presentation, and interpretation of data. 2) After the baseline visit, there was some modification in the questionnaire for assessing PA for reasons unrelated to the present study; therefore, METs/day could not be calculated throughout all visits in a consistent manner.

For each component of PA according to total PA quartiles (Table 2), light activity ≥ 1 h/day was performed by more than half of participants in all quartiles except Q1. The percentage of participants performing moderate activity ≥ 1 h/day was highest (45.4%) in Q3. While the percentage of participants in Q4 performing moderate PA ≥ 1 h/day (33.5%) was not as high, 90.4% of participants in Q4 performed high PA ≥ 1 h/day (vs. 24.3% in Q3). More detailed components of PA in each of the four quartiles of total PA are shown in Supplementary Fig. 1.

Only baseline RHR data was analyzed for the same reason as “Reason 1” mentioned in the first paragraph of this section and also because initial rankings of 5 RHR groups determined based on median values of baseline RHR remained identical throughout all follow-up visits (Supplementary Fig. 2).

Comparing the baseline RHR among quartiles (Table 2), Q3 included the highest percentage of subjects with RHR < 60 bpm (27%), and the RHR was significantly lower in Q3 than in Q1 (62.0 [58.7, 67] bpm, p < 0.001). On the other hand, Q4 had the lowest percentage with RHR < 60 bpm (18%).

**Physical activity, resting heart rate and incident AF.** There was a gradual decrease in the incidence rate and risk of AF from Q1 to Q3 (Table 3); Q3 had the lowest incidence rate of AF (1.42/1,000 person-years) and a significantly lower risk of AF, after adjusting for covariates, than Q1 and Q4. This trend remained consistent after adjusting for RHR (hazard ratio [HR] 0.58, 95% confidence interval [CI] 0.37–0.90, p = 0.016). On the other hand, Q4 had a high incidence rate of AF (2.12/1,000 person-years); the risk of AF was not even lower than in
Q1 (HR 0.88, 95% CI 0.57–1.36). These associations between PA and incident AF were similar in both male and female subjects, but were more significant in females (Fig. 1).

However, when participants in Q3 and Q4 were divided into subgroups, i.e., high PA ≥ 1 hr/day versus high PA < 1 hr/day (Fig. 2), subgroup of Q4 with high PA ≥ 1 hr/day had a significantly higher risk of AF compared with the subgroup of Q3 with high PA < 1 hr/day. Conversely, the risk of AF in the subgroup of Q4 with high PA < 1 hr was not higher than in the subgroup of Q3 with high PA < 1 hr.

Table 4 shows that subjects with RHR < 60 had a higher risk of AF than those with RHR 70–85 bpm. Subjects with RHR 50–59 bpm had a significant risk of AF (HR 1.97, 95% CI 1.17–3.32, p = 0.011). These relationships remained consistent after adjusting for PA.

Discussion

Analysis of a prospective cohort based on the general population showed that moderate total PA was associated with a reduced risk of incident AF compared with that for lowest total PA and that the protective effect of PA was attenuated in subjects with highest total PA. Also, low RHR (< 60 bpm) was a risk factor for incident AF. These findings remained consistent after adjusting for PA.

Endurance athletes have a lower RHR and are also predisposed to AF. Suggested mechanisms include atrial remodeling, atrial ectopy, and increased vagal tone. The risk of AF increases substantially in the aging athlete with
Table 3. Hazard ratio for incident atrial fibrillation according to quartiles of total physical activity. Model 1: adjusted for age, sex, residence, education. Model 2: Model 1 + BMI, comorbidity, alcohol and smoking. Model 3: Model 2 + RHR. AF, atrial fibrillation; Incidence rate, per 1,000 person-years; HR, hazard ratio; CI, confidence interval; BMI, body mass index; RHR, resting heart rate.

| Physical activity (METs/day) | Q1      | Q2      | Q3      | Q4      |
|-----------------------------|---------|---------|---------|---------|
| No. with AF                 | 51 (2.3%) | 38 (1.9%) | 31 (1.4%) | 47 (2.1%) |
| Incidence rate              | 2.39    | 1.91    | 1.42    | 2.12    |
| Crude                       | HR (95% CI) |        |        |         |         |
|                            | 1       | 0.79 (0.52–1.20) | 0.59 (0.38–0.92) | 0.87 (0.59–1.29) |
|                            | p-value | 0.263  | 0.020  | 0.489  |
| Model 1                     | HR (95% CI) |        |        |         |         |
|                            | 1       | 0.78 (0.51–1.21) | 0.59 (0.38–0.92) | 0.90 (0.58–1.39) |
|                            | p-value | 0.273  | 0.021  | 0.622  |
| Model 2                     | HR (95% CI) |        |        |         |         |
|                            | 1       | 0.78 (0.51–1.19) | 0.58 (0.37–0.90) | 0.88 (0.57–1.36) |
|                            | p-value | 0.246  | 0.016  | 0.564  |

Figure 1. Risk of atrial fibrillation according to the level of physical activity.

Figure 2. Incidence of atrial fibrillation in relation to presence or absence of high activity (MET 7) ≥ 1 hr/day in Q3 and Q4. *p < 0.05 vs. Q3 with high PA < 1 hr/day. Kruskal-Wallis test (Dunn’s post hoc).
may predispose to AF. Increased heart rate was also found to be associated with risk of AF in the Losartan trial, atrial dilation, and fibrosis may play a role. One possible mechanism is subclinical sinus node dysfunction, which our study were not trained athletes. Therefore, it is conceivable that mechanisms other than autonomic change, attenuated the benefit of moderate PA. Several pathways have been suggested to explain the mechanism by which accurately reflected the level of PA. While we tried to identify AF using both ECGs and self-reported history of physician-determined diagnosis.

Table 4. Hazard ratio according to resting heart rate and incident atrial fibrillation. Model 1: adjusted for age, sex, residence, education. Model 2: Model 1 + BMI, comorbidity, alcohol, and smoking. Model 3: Model 2 + physical activities. AF, atrial fibrillation; Incidence rate, per 1000 person-year; HR, hazard ratio; CI, confidence interval; BMI, body mass index.

| No. with AF | Resting Heart Rate (bpm) | 50–59 | 60–70 | 70–85 | >85 |
|-------------|--------------------------|-------|-------|-------|-----|
| Crude       | HR (95% CI)              | 2.69 (1.01–7.12) | 2.00 (1.21–3.32) | 1.24 (0.77–2.00) | 1.93 (0.45–8.24) |
| p-value     | 0.047                    | 0.007 | 0.370 | 0.374 |
| Model 1     | HR (95% CI)              | 2.54 (0.95–6.79) | 1.86 (1.11–3.12) | 1.24 (0.77–2.00) | 1.75 (0.41–7.46) |
| p-value     | 0.064                    | 0.018 | 0.375 | 0.452 |
| Model 2     | HR (95% CI)              | 2.54 (0.94–6.88) | 1.93 (1.14–3.26) | 1.27 (0.79–2.06) | 1.82 (0.43–7.83) |
| p-value     | 0.066                    | 0.015 | 0.327 | 0.419 |
| Model 3     | HR (95% CI)              | 2.56 (0.95–6.88) | 1.97 (1.17–3.32) | 1.27 (0.79–2.06) | 1.83 (0.43–7.87) |
| p-value     | 0.062                    | 0.011 | 0.323 | 0.415 |

Our research has some limitations. In baseline data collection, we did not distinguish between occupational and leisure time activities when calculating total daily PA. In addition, self-reporting questionnaires may not have accurately reflected the level of PA. While we tried to identify AF using both ECGs and self-reported history of physician-determined diagnosis, we could not obtain participants' International Classification of Disease (ICD)
because ICD was considered protected personal information. Therefore, some paroxysmal AF that waxed and waned between each visit could still have been missed. Information about antihypertensive drugs such as beta blockers was not available for this cohort. However, there were no significant differences in the rate of antihypertensive use according to quartiles.

In conclusion, this prospective cohort study of a Korean general population suggested independent associations between total daily PA and incident AF as well as between RHR and incident AF. A moderate amount of total PA was associated with a lower risk of AF. Higher total PA was not associated with a lower RHR. RHR was also associated with incident AF, showing a U-shaped relationship, independent of the amount of total PA. Further prospective research is warranted to clarify whether the intervention of PA and RHR would be beneficial to prevent AF.

Data Availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References
1. Oduotayo, A. et al. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: a systematic review and meta-analysis. BMJ (Clinical research ed.) 354, i4482 (2016).
2. Huxley, R. R. et al. Physical activity, obesity, weight change and risk of atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. Circulation: Arrhythmia and Electrophysiology 11, 601–604 (2018).
3. Morseth, B. et al. Physical activity, resting heart rate, and atrial fibrillation: the Tromso Study. European heart journal 37, 2307–2313 (2016).
4. Brunetti, N. D. et al. Incidence of atrial fibrillation is associated with age and gender in subjects practicing physical exercise: a meta-analysis and meta-regression analysis. International journal of cardiology 221, 1056–1060 (2016).
5. Flannery, M. D., Kalman, J. M., Sanders, P. & La Gerche, A. State of the art review: atrial fibrillation in athletes. Heart, Lung and Circulation 26, 983–989 (2017).
6. Böhm, M. et al. Low resting heart rates are associated with new-onset atrial fibrillation in patients with vascular disease: results of the ONTARGET/TRANSCEND studies. Journal of internal medicine 278, 303–312 (2015).
7. Wang, W. et al. Relation of Resting Heart Rate to Incident Atrial Fibrillation (From ARIC [Atherosclerosis Risk in Communities] Study). The American journal of cardiology 121, 1169–1176 (2018).
8. Thelle, D. S. et al. Resting heart rate and physical activity as risk factors for lone atrial fibrillation: a prospective study of 309 540 men and women. Heart 99, 1755–1760 (2013).
9. Kim, Y. & Han, B. G. Cohort Profile: The Korean Genome and Epidemiology Study (KoGES) Consortium. Int J Epidemiol 46, e20, https://doi.org/10.1093/ije/dyx316 (2017).
10. Molina, L. et al. Long-term endurance sport practice increases the incidence of lone atrial fibrillation in men: a follow-up study: EP Europe 10, 618–623, https://doi.org/10.1093/eurheartj/eun071 (2009).
11. Matthias, W. Atrial fibrillation in endurance athletes. European journal of preventive cardiology 21, 1040–1048, https://doi.org/10.1177/204748731456414 (2013).
12. Ofman, P. et al. Regular physical activity and risk of atrial fibrillation: a systematic review and meta-analysis. Circ Arrhythm Electrophysiol 6, 252–256, https://doi.org/10.1161/circep.113.000147 (2013).
13. Gjesdal, K. & Grundvold, I. Atrial fibrillation and exercise in women: some answers given, some questions remain. Heart 101, 1605–1606, https://doi.org/10.1136/heartjnl-2015-307836 (2015).
14. Mora, S., Cook, N., Buring, J. E., Ridker, P. M. & Lee, I.-M. Physical activity and fibrinogen concentrations in 23,201 men and women in the EPIC-Norfolk population-based study. Atherosclerosis 198, 419–425 (2008).
15. O’NEAL, W. T., Almahmoud, M. F. & Soliman, E. Z. Resting heart rate and incident atrial fibrillation in the elderly. Pacing and Clinical Electrophysiology 38, 591–597 (2015).
16. Grundvold, I. et al. Low heart rates predict incident atrial fibrillation in healthy middle-aged men. Circ Arrhythm Electrophysiol 6, 726–731, https://doi.org/10.1161/circep.113.000267 (2013).
17. Chang, H.-Y. et al. Sinus node dysfunction in atrial fibrillation patients: the evidence of regional atrial substrate remodelling. EP Europace 15, 205–211, https://doi.org/10.1093/epj/eus219 (2013).
18. Okin, P. M. et al. Incidence of Atrial Fibrillation in Relation to Changing Heart Rate Over Time in Hypertensive PatientsCLINICAL PERSPECTIVE: The LIFE Study. Circulation: Arrhythmia and Electrophysiology 1, 337–343 (2008).
19. Aladin, A. L. et al. Relation of resting heart rate to incident atrial fibrillation (from the Henry Ford Hospital Exercise Testing Project). American journal of cardiology 119, 262–267 (2017).
20. Rahman, F., Kwan, G. F. & Benjamin, E. J. Global epidemiology of atrial fibrillation. Nature Reviews Cardiology 11, 639 (2014).
21. Schnabel, R. B. et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. Lancet (London, England) 386, 154–162, https://doi.org/10.1016/s0140-6736(14)61774-8 (2015).
22. Chung, S. S. et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 129, 837–847, https://doi.org/10.1161/circulationaha.111.645036 (2014).
23. Lee, S. B., Choi, E. K., Han, K. D., Cha, M. J. & Oh, S. Trends in the incidence and prevalence of atrial fibrillation and estimated thromboembolic risk using the CHA2DS2-VASc score in the entire Korean population. International journal of cardiology 236, 226–231, https://doi.org/10.1016/j.ijcard.2017.02.039 (2017).

Acknowledgements
The authors thank all participants, research staff of the Institute of Human Genomic Study at Ansan Hospital of the Korea University, and the Biostatistical Consulting and Research Lab, Hanyang University, for assistance with statistical analysis. This work was supported by a grant from the National Research Foundation of Korea (NRF), funded by the Korean government (MSIT; Ministry of Science and ICT) (No. NRF-2019R1F1A1046443).

Author Contributions
J.K.P and Y.W.C contributed to the idea and design of this study; Y.G.L., J.M. and M.P prepared and checked clinical coding, undertook the data analysis; Y.H.L., Y.W.C. and J.K.P wrote the draft; B.S.K. and J.H.S contributed to data collection and analysis; and C.K.K., H.C.P. and J.S. revised the manuscript; all authors have seen and approved the final version of the report.
Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-47748-7.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019