scientific reports

OPEN

Sex differences in the risk of incident systemic sclerosis: a nationwide population-based study with subgroup analyses

Oh Chan Kwon[1], Kyungdo Han[2✉] & Min-Chan Park[1,3✉]

Although higher prevalence of systemic sclerosis (SSc) in women than in men is well-known, it is unclear to what degree women are at higher risk of developing incident SSc. This study aims to assess the risk of incident SSc comparing women vs. men, and to identify subsets of individuals in whom the risk difference according to sex is more prominent. An analysis of a Korean nationwide cohort of 9,894,996 individuals was conducted. All individuals were followed up from 2009 to 2019. Cox proportional hazard models were used to assess the risk of incident SSc comparing women vs. men. Subgroup analysis was conducted by stratifying individuals according to multiple covariates. The absolute risk of SSc in men (i.e., background risk) was 0.004% (95% CI 0.003–0.004), and the absolute risk of SSc in women was 0.025% (95% CI 0.023–0.027). The absolute risk difference between women and men was 0.021% (95% CI 0.019–0.024). Women had a significantly higher risk of incident SSc than men (adjusted hazard ratio [aHR], 5.275; 95% confidence interval, 4.346–6.403). The effect size was more pronounced in middle-aged individuals (aHR 5.020 [< 40 years] vs. 5.868 [40–64 years] vs. 2.734 [≥ 65 years]; p-interaction < 0.001); those without abdominal obesity (aHR 5.863 vs. 3.658; p-interaction = 0.005); those who did not exercise regularly (aHR 5.701 vs. 3.932; p-interaction = 0.030); and those without hypertension (aHR 5.996 vs. 4.053; p-interaction = 0.010) or dyslipidemia (aHR 5.857 vs. 3.330; p-interaction = 0.001). Women had a 5-fold higher risk of incident SSc than men. The higher risk was more prominent in middle-aged individuals, those without abdominal obesity, those who do not exercise regularly, and do not have hypertension or dyslipidemia.

Abbreviations

SSc	Systemic sclerosis
NHIS	National Health Insurance Service
ICD-10	International Classification of Diseases-Tenth Revision
RID	Rare intractable disease
BMI	Body mass index
WC	Waist circumference
BP	Blood pressure
HDL-C	High-density lipoprotein cholesterol
LDL-C	Low-density lipoprotein cholesterol
GFR	Glomerular filtration rate
IRB	Institutional Review Board
CKD	Chronic kidney disease
HR	Hazard ratio
CI	Confidence interval
aHR	Adjusted hazard ratio
RA	Rheumatoid arthritis

[1]Division of Rheumatology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Korea. [2]Department of Statistics and Actuarial Science, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 06978, Korea. [3]Gangnam Severance Hospital, Yonsei University College of Medicine, 211 Eonjuro, Gangnam-gu, Seoul 06273, Korea. ✉email: hkd917@naver.com; mcpark@yuhs.ac

Systemic sclerosis (SSc) is an autoimmune disease characterized by microvascular damage, dysregulated innate and adaptive immunity, and multi-organ fibrosis[1]. Consistent with other autoimmune diseases, SSc has a marked women predominance[2]. A number of epidemiological studies reported a women predominance in SSc, with a women to men prevalence ratio ranging from 4:1 to 10:1[3–8]. Based on these data, female sex is suggested as a risk factor for development of SSc[9,10]. Although higher prevalence of SSc in women than in men is well-known, it is still unclear to what degree the risk of developing incident SSc is higher in women than in men.

In addition, although women predominance in SSc has been consistently observed throughout studies, the women to men prevalence ratio (4:1 ~ 10:1) varies widely[3–8]. This suggests that there could be a subset of the population in which the influence of sex on the development of SSc is more prominent. However, there are currently no data assessing whether the effect of sex on development of SSc differs in a particular subset of individuals.

On this basis, using a Korean nationwide cohort, we aimed to evaluate (i) to what degree the risk of developing incident SSc is higher in women than in men, and (ii) whether the effect of sex on development of SSc differs among subsets of individuals.

Methods

Study cohort

The study cohort was obtained from the Korean National Health Insurance Service (NHIS) claims database. The NHIS provides medical coverage to approximately 97% of the Korean population and, therefore, is a representative Korean nationwide database. The database includes data on demographics, diagnoses encoded by the International Classification of Diseases-Tenth Revision (ICD-10) and rare intractable disease (RID) registration information[11,12], prescription, and data from health examinations. Individuals aged ≥ 40 years or employees of any age in the NHIS database undergo a standardized national health examination every 2 years. The health examination data include anthropometric data such as body mass index (BMI) and waist circumference (WC), blood pressure (BP), lifestyle factors (smoking status, alcohol consumption, and physical activity) based on standardized self-reporting questionnaires, and laboratory data (fasting glucose, total cholesterol, high-density lipoprotein cholesterol [HDL-C], low-density lipoprotein cholesterol [LDL-C], triglyceride, estimated glomerular filtration rate [GFR], etc.).

Individuals who underwent a national health examination in 2009 (referred to as baseline) were initially selected from the NHIS database ($N = 10,601,274$). Subsequently, individuals aged < 20 years ($n = 15,431$), those who were diagnosed with SSc before baseline ($n = 890$), those with missing data ($n = 663,828$), and those who developed SSc or died within 1 year from baseline (one-year lag period) ($n = 26,129$) were excluded (Fig. 1). The remaining 9,894,996 individuals were followed up from baseline to December 2019 for the occurrence of incident SSc.

This study was approved by the Institutional Review Board (IRB) of Gangnam Severance Hospital (IRB No: 3-2022-0338). Owing to the retrospective nature of this study, the requirement for informed consent was waived and approved by the IRB of Gangnam Severance Hospital. Data are reported according to the STROBE checklist (Supplementary material).

Definitions of covariates and outcomes

Comorbidities such as type 2 diabetes, hypertension, dyslipidemia, and chronic kidney disease (CKD) were defined as previously described[13]: type 2 diabetes was defined as (i) ICD-10 codes E11–14 with prescriptions for anti-diabetic agents or (ii) fasting blood glucose level ≥ 126 mg/dL; hypertension as (i) ICD-10 codes I10-I13 and I15 with prescriptions for anti-hypertensive agents or (ii) systolic BP ≥ 140 mmHg or diastolic BP ≥ 90 mmHg; dyslipidemia as (i) ICD-10 code E78 with prescriptions for lipid-lowering agents or (ii) total cholesterol level ≥ 240 mg/dL; and CKD as estimated GFR of < 60 mL/min/1.73 m^2 by the Modification of Diet in Renal Disease equation. The outcome variable was incident SSc during the follow-up period. SSc was defined as RID code V138, as previously described[8].

Statistical analysis

Continuous and categorical variables are expressed as mean ± standard deviation and numbers (%), respectively. For comparison of continuous and categorical variables between groups, independent Student's t test and Chi-square test were used, respectively. The incidence rate of SSc was calculated (number of events per 100,000 person-years). Cox proportional hazard models were used to assess the risk of incident SSc comparing women vs. men. Model 1 was a univariable model (no covariates adjusted). Model 2 was adjusted for age (a known risk factor of SSc[10]). Model 3 was adjusted for age, BMI, income, smoking status, alcohol consumption, and physical activity (socioeconomic and behavioral factors). Model 4 was further adjusted for type 2 diabetes, hypertension, dyslipidemia, and CKD (comorbidities). To investigate whether there is particular subset of individuals in which the influence of sex is more prominent, we stratified the individuals according to age (< 40 vs. 40–64 vs. ≥ 65 years), income (Q2–Q4 vs. Q1 [lowest 25%]), obesity (BMI < 25 kg/m^2 vs. ≥ 25 kg/m^2), abdominal obesity (WC < 90 cm vs. ≥ 90 cm for men; and WC < 85 cm vs. ≥ 85 cm for women), current smoking (no vs. yes), alcohol consumption (< 30 g/day vs. ≥ 30 g/day), physical activity (vigorous exercise > 3 days a week for at least 20 min/day: no vs. yes), and presence of type 2 diabetes (no vs. yes), hypertension (no vs. yes), dyslipidemia (no vs. yes), and CKD (no vs. yes). A Cox proportional hazard model, adjusted for age, BMI, income, smoking status, alcohol consumption, physical activity, type 2 diabetes, hypertension, dyslipidemia, and CKD, was performed for subgroup analysis. All p values were two-sided, and a p value < 0.05 was considered statistically significant. Statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).
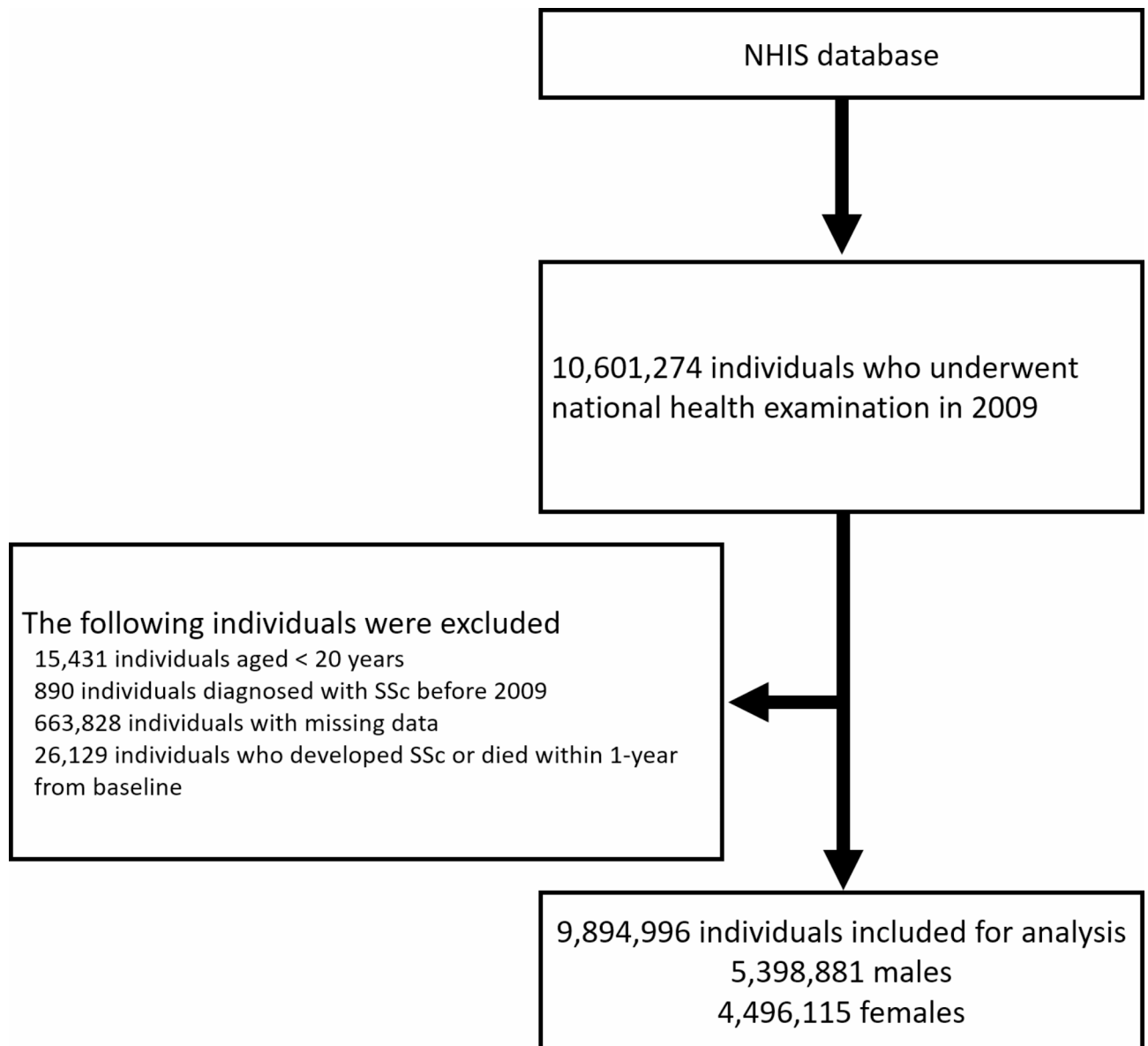
Fig. 1. Study population. NHIS, National Health Insurance Service; SSc, systemic sclerosis.

Results

Baseline characteristics

Of the 9,894,996 individuals, 5,398,881 and 4,496,115 were men and women, respectively. The comparisons of baseline characteristics between the men and women are summarized in Table 1. Compared with men, women were older; more commonly had low income (lowest 25%); more commonly were non-smokers and non-alcoholics; less commonly performed regular physical activity; less commonly had type 2 diabetes and hypertension; more commonly had dyslipidemia and CKD; had a lower BMI, WC, systolic and diastolic BP, and fasting glucose levels; had higher total cholesterol, HDL-C, and LDL-C levels; and had lower estimated GFR and triglyceride levels (all $p < .001$).

Risk of incident SSc comparing women and men

During follow-up (mean: 9.2 ± 1.1 years), incident SSc occurred in 1,355 individuals. The incidence rates of SSc for the total study population, men, and women were 1.49, 0.48, and 2.69 per 100,000 person-years, respectively. The absolute risk of SSc in men (i.e., background risk) was 0.004% (95% CI 0.003–0.004), and the absolute risk of SSc in women was 0.025% (95% CI 0.023–0.027). The absolute risk difference between women and men was 0.021% (95% CI 0.019–0.024). The results of Cox proportional hazard models estimating the risk of incident SSc in women compared with men are shown in Table 2. In model 1 (univariable model), women had a significantly higher risk of developing SSc (unadjusted hazard ratio [HR] 5.584; 95% confidence interval [CI] 4.854–6.424) than men. This association was consistently observed throughout the multivariable models (model 2: adjusted

	Men (N = 5,398,881)	Women (N = 4,496,115)
Age, years	45.96 ± 13.55	48.83 ± 14.47
Age groups, years, n (%)		
< 40	1,966,938 (36.43)	1,092,043 (24.29)
40–64	2,827,713 (52.38)	2,703,672 (60.13)
≥ 65	604,230 (11.19)	700,400 (15.58)
Income, lowest Q1, n (%)	641,724 (11.89)	888,967 (19.77)
Smoking, n (%)		
Non	1,641,846 (30.41)	4,258,713 (94.72)
Ex	1,336,973 (24.76)	82,212 (1.83)
Current	2,420,062 (44.83)	155,190 (3.45)
Drinking, n (%)		
Non	1,754,381 (32.50)	3,365,814 (74.86)
Mild	2,908,789 (53.88)	1,080,726 (24.04)
Heavy	735,711 (13.63)	49,575 (1.10)
Regular exercise, n (%)	1,075,489 (19.92)	701,966 (15.61)
Type 2 diabetes, n (%)	535,670 (9.92)	335,104 (7.45)
Hypertension, n (%)	1,506,730 (27.91)	1,184,564 (26.35)
Dyslipidemia, n (%)	900,139 (16.67)	903,474 (20.09)
CKD, n (%)	331,649 (6.14)	356,370 (7.93)
Height, cm	169.95 ± 6.44	156.46 ± 6.21
Weight, kg	69.79 ± 10.53	56.74 ± 8.42
BMI, kg/m^2	24.12 ± 3.06	23.2 ± 3.34
BMI Level, kg/m^2, n (%)		
< 18.5	121,967 (2.26)	245,293 (5.46)
< 23	1,819,059 (33.69)	2,044,165 (45.47)
< 25	1,454,236 (26.94)	981,911 (21.84)
< 30	1,808,144 (33.49)	1,069,158 (23.78)
≥ 30	195,475 (3.62)	155,588 (3.46)
Waist Circumference, cm	83.56 ± 7.83	76.26 ± 8.92
Systolic BP, mmHg	124.68 ± 14.14	119.83 ± 15.71
Diastolic BP, mmHg	78.03 ± 9.72	74.27 ± 10.10
Fasting glucose, mg/dL	99.15 ± 25.76	95.14 ± 21.16
Total Cholesterol, mg/dL	194.24 ± 36.14	195.94 ± 37.59
HDL-C, mg/dL	53.12 ± 25.3	59.65 ± 30.53
LDL-C, mg/dL	111.97 ± 38.8	115.45 ± 38.27
Triglyceride[a], mg/dL	128.61 (128.55–128.67)	95.94 (95.89–95.99)
Estimated GFR, mL/min/1.73 m^2	87.9 ± 51.45	87.16 ± 36.49

Table 1. Comparison of baseline characteristics according to sex. [a]Geometric mean (95% confidence interval). BMI, body mass index; BP, blood pressure; CKD, chronic kidney disease; GFR, glomerular filtration rate; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

Sex	N	Event	Duration (pyrs)	IR (/100,000 pyrs)	IR difference (/100,000 pyrs)	HR (95% CI)			
						Model 1	Model 2	Model 3	Model 4
Men	5,398,881	237	49,204,778.39	0.48	0 (Ref.)	1 (Ref.)	1 (Ref.)	1 (Ref.)	1 (Ref.)
Women	4,496,115	1,118	41,505,673.98	2.69	2.211 (2.213–2.211)	5.584 (4.854–6.424)	5.227 (4.541–6.018)	5.227 (4.307–6.343)	5.275 (4.346–6.403)

Table 2. Association between female sex and risk of incident systemic sclerosis. Model 1: Univariable model. Model 2: Adjusted for age. Model 3: Adjusted for age, BMI, income, smoking status, alcohol consumption, and physical activity. Model 4: Adjusted for age, BMI, income, smoking status, alcohol consumption, physical activity, type 2 diabetes, hypertension, dyslipidemia, and CKD. BMI, body mass index; CI, confidence interval; CKD, chronic kidney disease; HR, hazard ratio; IR, incidence ratio; pyrs, person-years.

		Sex	N	Event	Duration (pyrs)	IR (/100,000 pyrs)	HRᵃ (95% CI)	p-interaction
Age groups, years	< 40	Men	1,966,938	34	18,274,987.89	0.19	1 (Ref.)	<0.001
		Women	1,092,043	95	10,152,552.86	0.94	5.020 (3.307–7.620)	
	40–64	Men	2,827,713	161	25,976,999.50	0.62	1 (Ref.)	
		Women	2,703,672	888	25,185,972.53	3.53	5.868 (4.705–7.318)	
	≥65	Men	604,230	42	4,952,791.01	0.85	1 (Ref.)	
		Women	700,400	135	6,167,148.59	2.19	2.734 (1.894–3.947)	
Income	Q2-4	Men	4,757,157	197	43,432,203.82	0.45	1 (Ref.)	0.103
		Women	3,607,148	894	33,309,934.95	2.68	5.550 (4.522–6.810)	
	Q1	Men	641,724	40	5,772,574.57	0.69	1 (Ref.)	
		Women	888,967	224	8,195,739.03	2.73	4.077 (2.844–5.846)	
Obesity	No	Men	3,395,262	152	30,838,177.06	0.49	1 (Ref.)	0.364
		Women	3,271,369	824	30,204,718.03	2.73	5.525 (4.443–6.872)	
	Yes	Men	2,003,619	85	18,366,601.34	0.46	1 (Ref.)	
		Women	1,224,746	294	11,300,955.95	2.60	4.811 (3.642–6.355)	
Abdominal obesity	No	Men	4,243,012	174	38,710,109.95	0.45	1 (Ref.)	0.005
		Women	3,701,024	931	34,229,121.47	2.72	5.863 (4.754–7.232)	
	Yes	Men	1,155,869	63	10,494,668.45	0.60	1 (Ref.)	
		Women	795,091	187	7,276,552.51	2.57	3.658 (2.663–5.025)	
Current Smoking	No	Men	2,978,819	137	27,122,260.73	0.51	1 (Ref.)	0.146
		Women	4,340,925	1,090	40,086,655.58	2.72	5.708 (4.559–7.146)	
	Yes	Men	2,420,062	100	22,082,517.66	0.45	1 (Ref.)	
		Women	155,190	28	1,419,018.40	1.97	4.017 (2.634–6.127)	
Heavy drinking	No	Men	4,663,170	198	42,499,329.45	0.47	1 (Ref.)	0.396
		Women	4,446,540	1,109	41,045,968.72	2.70	5.369 (4.401–6.551)	
	Yes	Men	735,711	39	6,705,448.95	0.58	1 (Ref.)	
		Women	49,575	9	459,705.26	1.96	3.892 (1.873–8.089)	
Regular exercise	No	Men	4,323,392	176	39,384,790.08	0.45	1 (Ref.)	0.030
		Women	3,794,149	955	34,987,622.03	2.73	5.701 (4.627–7.024)	
	Yes	Men	1,075,489	61	9,819,988.32	0.62	1 (Ref.)	
		Women	701,966	163	6,518,051.95	2.50	3.932 (2.841–5.441)	
Type 2 diabetes	No	Men	4,863,211	205	44,512,760.55	0.46	1 (Ref.)	0.147
		Women	4,161,011	1,031	38,506,785.82	2.68	5.466 (4.471–6.681)	
	Yes	Men	535,670	32	4,692,017.84	0.68	1 (Ref.)	
		Women	335,104	87	2,998,888.16	2.90	3.966 (2.586–6.084)	
Hypertension	No	Men	3,892,151	149	35,800,979.13	0.42	1 (Ref.)	0.010
		Women	3,311,551	802	30,752,778.34	2.61	5.996 (4.810–7.475)	
	Yes	Men	1,506,730	88	13,403,799.27	0.66	1 (Ref.)	
		Women	1,184,564	316	10,752,895.64	2.94	4.053 (3.086–5.322)	
Dyslipidemia	No	Men	4,498,742	183	41,056,448.85	0.45	1 (Ref.)	0.001
		Women	3,592,641	906	33,210,089.52	2.73	5.857 (4.761–7.204)	
	Yes	Men	900,139	54	8,148,329.55	0.66	1 (Ref.)	
		Women	903,474	212	8,295,584.47	2.56	3.330 (2.394–4.632)	
CKD	No	Men	5,067,232	221	46,271,153.98	0.48	1 (Ref.)	0.786
		Women	4,139,745	1,023	38,315,942.47	2.67	5.303 (4.351–6.463)	
	Yes	Men	331,649	16	2,933,624.41	0.55	1 (Ref.)	
		Women	356,370	95	3,189,731.51	2.98	4.914 (2.847–8.482)	

Table 3. Subgroup analysis. ᵃAdjusted for age, BMI, income, smoking status, alcohol consumption, physical activity, type 2 diabetes, hypertension, dyslipidemia, and CKD. BMI, body mass index; CI, confidence interval; CKD, chronic kidney disease; HR, hazard ratio; IR, incidence ratio; pyrs, person-years.

HR [aHR] 5.227; 95% CI 4.541–6.018; model 3: aHR 5.227; 95% CI 4.307–6.343; and model 4: aHR 5.275; 95% CI 4.346–6.403).

Subgroup analysis
In the subgroup analysis, subsets of individuals, in which the effect of sex on the risk of incident SSc was more prominent, were identified (Table 3). The association between women and higher risk of incident SSc was more

pronounced in middle-aged individuals (aHR 5.020 [in individuals aged < 40 years] vs. 5.868 [in individuals aged 40–64 years] vs. 2.734 [in individuals aged ≥ 65 years]; p-interaction < 0.001), in individuals who did not have abdominal obesity than in those with abdominal obesity (aHR 5.863 vs. 3.658; p-interaction = 0.005), in those who do not exercise regularly than in those who exercise regularly (aHR 5.701 vs. 3.932; p-interaction = 0.030), in those without hypertension than in those with hypertension (aHR 5.996 vs. 4.053; p-interaction = 0.010), and in those without dyslipidemia than in those with dyslipidemia (aHR 5.857 vs. 3.330; p-interaction = 0.001).

Discussion

In this nationwide population-based cohort study, we showed that women are at a 5-fold higher risk of developing SSc than men, after adjusting for multiple covariates. The higher risk of incident SSc in women than in men was more prominent in several subgroups of individuals, including those aged 40–64 years, those without abdominal obesity, those who do not exercise regularly, those without hypertension, and those without dyslipidemia. These data are clinically meaningful as they provide comprehensive information that can lead to a better stratification of the risk of incident SSc in the perspective of sex.

Rheumatoid arthritis (RA) is another disease that has female predominance. In Korea, the incidence rate for RA is 16.5–42.0 per 100,000 person-years, with a women-to-men ratio of 3.4–3.5:1[14]. In our study, the incidence rate for SSc was 1.49 per 100,000 person-years, with a women-to-men ratio of 5.6:1. Compared with RA, the incidence rate of SSc was lower, but the women predominance was more pronounced.

Data from EUSTAR database have compared disease characteristics of SSc between men with SSc and women with SSc[15]. The study reported that men with SSc had more severe disease characteristics (diffuse subtype, higher frequency of digital ulcers, and pulmonary hypertension) than women with SSc. However, the risk of incident SSc according to sex in general population has not been assessed. In our study, we compared the risk of incident SSc between men and women in the general population, and showed that women have a 5-fold higher risk of incident SSc than men in the general population.

In the multivariable analysis adjusted for multiple covariates including age and comorbidities, women had an approximately 5-fold higher risk of incident SSc than men (model 4: aHR 5.275; 95% CI 4.346–6.403). This association could be attributable to estrogen. Mechanistically, by binding to estrogen receptor α, estradiol exerts a pro-fibrotic effect by inducing expression of extracellular matrix protein (such as fibronectin) and enhancing cell growth in the dermal fibroblast, and is thus involved in the pathogenesis of SSc[16,17]. Moreover, studies have reported that serum estradiol levels are higher in patients with SSc than in matched controls[16,18]. As women have higher levels of estrogen than men, this may have led to a higher risk of incident SSc. In the age-based subgroups, the impact of sex on the incidence of SSc was most pronounced in individuals aged 40–64 years (aHR 5.868; 95% CI 4.705–7.318). This is twice the effect size compared with that in individuals aged ≥ 65 years (aHR 2.734; 95% CI 1.894–3.947). Notably, middle age (45–64 years) is considered as a risk factor of incident SSc[10]. The effect size (aHR 5.868) of sex on the incidence of SSc in the individuals aged 40–64 years is considerable, given that the middle-aged individuals are already at a higher risk of incident SSc than individuals of other ages. Hence, women aged 40–64 years should be stratified as having a very high risk of incident SSc.

Sex had a more robust impact on the incidence of SSc in individuals without abdominal obesity (aHR 5.863) than in those with abdominal obesity (aHR 3.658). However, there were no significant differences in the effect of sex on the incidence of SSc between those without and with obesity (aHR 5.525 vs. 4.811; p-interaction = 0.364). Studies have shown that higher estradiol levels are inversely associated with abdominal obesity in women[19,20]. In other words, women without abdominal obesity are likely to have higher estradiol levels than those with abdominal obesity. Therefore, the larger effect of sex on the risk of incident SSc in those without abdominal obesity than in those with abdominal obesity could be attributable to the higher estradiol levels in the former group. On the other hand, there is no significant inverse association between estradiol levels and BMI in women[20]. This could be a possible explanation for the significant difference in the effect of sex observed only between individuals without and with abdominal obesity, but not between those without and with obesity. An inverse association between exercise and estrogen levels in women has also been repeatedly reported in studies[21–23]. This could explain the more prominent effect of sex on the risk of incident SSc in individuals who did not exercise regularly (aHR 5.701) compared with those who did (aHR 3.932). That is, women who do not exercise regularly are likely to have higher estrogen levels than those who exercise regularly. Therefore, the effect size of sex on incidence of SSc may have been higher in those who do not exercise regularly.

Other subgroups in which the effect of sex was more pronounced were those without hypertension (aHR 5.996) and without dyslipidemia (aHR 5.857) as compared with their respective counterparts (aHRs 4.053 and 3.330, respectively). As exercise influences blood pressure and lipid levels, the possibility of the influence of exercise on the different risk of incident SSc between women with/without hypertension and women with/without dyslipidemia should be considered. Women who do not exercise regularly would more likely have hypertension and dyslipidemia than women who exercise regularly. Therefore, given that the higher risk of incident SSc was more prominent in women who did not exercise regularly (aHR 5.701) than women who did (aHR 3.932), one could expect that the higher risk of incident SSc would be more prominent in women with hypertension and women with dyslipidemia, compared with their respective counterparts. However, the higher risk of incident SSc was more prominent in women without hypertension and women without dyslipidemia, compared with their respective counterparts. This suggests that the influence of exercise is not likely the cause of the different risk of incident SSc between women with/without hypertension and women with/without dyslipidemia. The larger effect size in individuals without hypertension and dyslipidemia could rather be attributable to the difference in estrogen levels with their respective counterparts. In women, estradiol has a BP-lowering effect[24,25]. Therefore, women without hypertension are more likely to have higher estradiol levels. In addition, estrogen is an important regulator of lipid metabolism[26]. Estrogen increases HDL-C and decreases LDL-C levels[26–28]. Hence, those without dyslipidemia are more likely to have higher estrogen levels.

Taken together, the more prominent effect of sex in those without hypertension and dyslipidemia could be due to the higher estrogen levels in these groups than their respective counterparts.

Some limitations should be noted in this study. First, the possibility of bias owing to unmeasured confounders exists. For instance, data on exposure to silica, which is a known environmental risk factor for the SSc development[10] were lacking. Nonetheless, as this was a large-scale population-based study consisting of approximately 10 million individuals, the influence by the unmeasured confounders would be negligible. Second, although, based on several translational and observational studies, we presumed that the results of our study were attributable to estrogen, the exact mechanism underlying the associations observed in our study remains elusive. Third, the RID code does not distinguish diffuse SSc and limited SSc, and we were unable to stratify SSc based on subtypes. These data, if available, would have been more informative. Fourth, as this study was based exclusively on the Korean population, results may not be generalizable to other ethnic populations. Further studies are needed to reveal the mechanism underlying the higher risk of incident SSc in women, and to confirm our findings in other ethnic groups.

In conclusion, we showed that women have a 5-fold higher risk of incident SSc than men using a nationwide cohort. This association was more prominent in middle-aged individuals, those without abdominal obesity, those who do not exercise regularly, and those without hypertension, and those without dyslipidemia. Considering that women have a higher risk of developing SSc, these subgroups of women should be stratified as having a very high risk of incident SSc.

Data availability

The data that support the findings of this study are available from the Korean National Health Insurance Services but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Korean National Health Insurance Services.

References

1. Denton, C. P. & Khanna, D. Systemic sclerosis. *Lancet* **390**, 1685–1699. https://doi.org/10.1016/s0140-6736(17)30933-9 (2017).
2. Whitacre, C. C. Sex differences in autoimmune disease. *Nat. Immunol.* **2**, 777–780. https://doi.org/10.1038/ni0901-777 (2001).
3. Lo Monaco, A., Bruschi, M., La Corte, R., Volpinari, S. & Trotta, F. Epidemiology of systemic sclerosis in a district of northern Italy. *Clin. Exp. Rheumatol.* **29**, S10–14 (2011).
4. Meier, F. M. et al. Update on the profile of the EUSTAR cohort: an analysis of the EULAR Scleroderma Trials and Research group database. *Ann. Rheum. Dis.* **71**, 1355–1360. https://doi.org/10.1136/annrheumdis-2011-200742 (2012).
5. Ferri, C. et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. *Med. (Baltim).* **81**, 139–153. https://doi.org/10.1097/00005792-200203000-00004 (2002).
6. Scussel-Lonzetti, L. et al. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. *Med. (Baltim).* **81**, 154–167. https://doi.org/10.1097/00005792-200203000-00005 (2002).
7. Roberts-Thomson, P. J. et al. Scleroderma in South Australia: further epidemiological observations supporting a stochastic explanation. *Intern. Med. J.* **36**, 489–497. https://doi.org/10.1111/j.1445-5994.2006.01125.x (2006).
8. Kang, G. W. et al. Incidence, prevalence, mortality and causes of death in systemic sclerosis in Korea: a nationwide population-based study. *Br. J. Dermatol.* **178**, e37–e39. https://doi.org/10.1111/bjd.15838 (2018).
9. Barnes, J. & Mayes, M. D. Epidemiology of systemic sclerosis: incidence, prevalence, survival, risk factors, malignancy, and environmental triggers. *Curr. Opin. Rheumatol.* **24**, 165–170. https://doi.org/10.1097/BOR.0b013e32834ff2e8 (2012).
10. Abbot, S., Bossingham, D., Proudman, S., de Costa, C. & Ho-Huynh, A. Risk factors for the development of systemic sclerosis: a systematic review of the literature. *Rheumatol. Adv. Pract.* **2**, rky041. https://doi.org/10.1093/rap/rky041 (2018).
11. Kim, J. A., Yoon, S., Kim, L. Y. & Kim, D. S. Towards actualizing the Value Potential of Korea Health Insurance Review and Assessment (HIRA) data as a resource for Health Research: strengths, limitations, applications, and strategies for optimal use of HIRA data. *J. Korean Med. Sci.* **32**, 718–728. https://doi.org/10.3346/jkms.2017.32.5.718 (2017).
12. Kwon, O. C. et al. Effects of immune-mediated inflammatory diseases on cardiovascular diseases in patients with type 2 diabetes: a nationwide population-based study. *Sci. Rep.* **12**, 11548. https://doi.org/10.1038/s41598-022-15436-8 (2022).
13. Hong, S. et al. Association between Obesity and Cardiovascular Disease in Elderly patients with diabetes: a retrospective cohort study. *J. Clin. Endocrinol. Metab.* **107**, e515–e527. https://doi.org/10.1210/clinem/dgab714 (2022).
14. Kim, H. & Sung, Y. K. Epidemiology of rheumatoid arthritis in Korea. *J. Rheum. Dis.* **28**, 60–67. https://doi.org/10.4078/jrd.2021.28.2.60 (2021).
15. Elhai, M. et al. A gender gap in primary and secondary heart dysfunctions in systemic sclerosis: a EUSTAR prospective study. *Ann. Rheum. Dis.* **75**, 163–169. https://doi.org/10.1136/annrheumdis-2014-206386 (2016).
16. Aida-Yasuoka, K. et al. Estradiol promotes the development of a fibrotic phenotype and is increased in the serum of patients with systemic sclerosis. *Arthritis Res. Ther.* **15**, R10. https://doi.org/10.1186/ar4140 (2013).
17. Soldano, S. et al. Endothelin and sex hormones modulate the fibronectin synthesis by cultured human skin scleroderma fibroblasts. *Ann. Rheum. Dis.* **68**, 599–602. https://doi.org/10.1136/ard.2008.097378 (2009).
18. Baker Frost, D. et al. Estradiol levels are elevated in older men with diffuse cutaneous SSc and are associated with decreased survival. *Arthritis Res. Ther.* **21**, 85. https://doi.org/10.1186/s13075-019-1870-6 (2019).
19. Jasieńska, G., Ziomkiewicz, A., Ellison, P. T., Lipson, S. F. & Thune, I. Large breasts and narrow waists indicate high reproductive potential in women. *Proc. Biol. Sci.* **271**, 1213–1217. https://doi.org/10.1098/rspb.2004.2712 (2004).
20. Mondragón-Ceballos, R., García Granados, M. D., Cerda-Molina, A. L., Chavira-Ramírez, R. & Hernández-López, L. E. Waist-to-hip ratio, but not body Mass Index, is Associated with testosterone and estradiol concentrations in Young Women. *Int. J. Endocrinol.* **2015** (654046). https://doi.org/10.1155/2015/654046 (2015).
21. Chan, M. F. et al. Usual physical activity and endogenous sex hormones in postmenopausal women: the European prospective investigation into cancer-norfolk population study. *Cancer Epidemiol. Biomarkers Prev.* **16**, 900–905. https://doi.org/10.1158/1055-9965.Epi-06-0745 (2007).
22. Bertone-Johnson, E. R., Tworoger, S. S. & Hankinson, S. E. Recreational physical activity and steroid hormone levels in postmenopausal women. *Am. J. Epidemiol.* **170**, 1095–1104. https://doi.org/10.1093/aje/kwp254 (2009).

23. Ennour-Idrissi, K., Maunsell, E. & Diorio, C. Effect of physical activity on sex hormones in women: a systematic review and meta-analysis of randomized controlled trials. *Breast Cancer Res.* **17**, 139. https://doi.org/10.1186/s13058-015-0647-3 (2015).
24. Chapman, A. B. et al. Systemic and renal hemodynamic changes in the luteal phase of the menstrual cycle mimic early pregnancy. *Am. J. Physiol.* **273**, F777–782. https://doi.org/10.1152/ajprenal.1997.273.5.F777 (1997).
25. Dubey, R. K., Oparil, S., Imthurn, B. & Jackson, E. K. Sex hormones and hypertension. *Cardiovasc. Res.* **53**, 688–708. https://doi.org/10.1016/s0008-6363(01)00527-2 (2002).
26. Barton, M. Cholesterol and atherosclerosis: modulation by oestrogen. *Curr. Opin. Lipidol.* **24**, 214–220. https://doi.org/10.1097/MOL.0b013e3283613a94 (2013).
27. Knopp, R. H. et al. Sex differences in lipoprotein metabolism and dietary response: basis in hormonal differences and implications for cardiovascular disease. *Curr. Cardiol. Rep.* **8**, 452–459. https://doi.org/10.1007/s11886-006-0104-0 (2006).
28. Meyer, M. R., Haas, E. & Barton, M. Need for research on estrogen receptor function: importance for postmenopausal hormone therapy and atherosclerosis. *Gend. Med.* **5 Suppl A**, 19–33. https://doi.org/10.1016/j.genm.2008.03.004 (2008).

Acknowledgements

Author contributions

OCK contributed to conceptualization, data curation, investigation, methodology, project administration, resources, software, validation, visualization, writing - original draft, and writing - review & editing. KH contributed to conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, and writing - review & editing. M-CP contributed to conceptualization, data curation, investigation, methodology, project administration, resources, software, supervision, validation, visualization, and writing - review & editing. All authors read and approved the final manuscript.

Funding

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board (IRB) of Gangnam Severance Hospital (No: 3-2022-0338). Owing to the retrospective nature of this study, the requirement for informed consent was waived by the IRB of Gangnam Severance Hospital. This study conformed the ethical guidelines laid out by the 1964 Helsinki declaration.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-84856-5.

Correspondence and requests for materials should be addressed to K.H. or M.-C.P.

Reprints and permissions information is available at www.nature.com/reprints.

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