

Heat-related illness and chronic obstructive pulmonary disease: causal inference using marginal structural models

Received: 8 December 2025

Accepted: 12 May 2026

Published online: 23 May 2026

Cite this article as: Jin T., Kim S., Lee S. *et al.* Heat-related illness and chronic obstructive pulmonary disease: causal inference using marginal structural models. *BMC Public Health* (2026). <https://doi.org/10.1186/s12889-026-27788-1>

Taiyue Jin, Seyoung Kim, Seulbi Lee, Soontae Kim, Jin-Kyoung Oh, Bomi Park & Byungmi Kim

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Heat-related illness and chronic obstructive pulmonary disease: Causal inference using marginal structural models

Taiyue Jin ¹, Seyoung Kim ¹, Seulbi Lee ², Soontae Kim ³, Jin-Kyoung Oh ^{1,4},
Bomi Park ^{5,*}, and Byungmi Kim ^{1,4,6,*}

¹ Division of Cancer Prevention, National Cancer Control Institute, National Cancer Center, Goyang, Republic of Korea

² Department of Big Data Strategy, National Health Insurance Service, Wonju, Republic of Korea

³ Department of Environmental and Safety Engineering, Ajou University, Suwon, Republic of Korea

⁴ Department of Public Health & AI, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, Republic of Korea

⁵ Department of Preventive Medicine, College of Medicine, Chung-Ang University, Seoul, Republic of Korea

⁶ Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

* Correspondence:

Bomi Park

Department of Preventive Medicine, College of Medicine, Chung-Ang University, 84 Heukseouk-ro, Dongjak-gu, Seoul, 06974, Republic of Korea.

E-mail: bpark@cau.ac.kr; Tel.: +82-2-820-5873.

Byungmi Kim

Division of Cancer Prevention, National Cancer Control Institute, National

Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang, Gyeonggi 10408, Republic of Korea.

E-mail: kbm5369@ncc.re.kr; Tel.: +82-31-920-2962.

ABSTRACT

Background Heat-related illness (HRI) may serve as a more reliable individual-level indicator of extreme heat exposure compared to ambient temperature, which has spatial autocorrelation problems. This study aimed to provide causal evidence for the association between HRI and the risk of chronic obstructive pulmonary disease (COPD).

Methods A retrospective cohort study was conducted involving 1,392,876 Korean individuals who regularly participated in health screenings provided by the National Health Insurance Services. Using data from 2009–2021, both HRI and COPD were identified using primary diagnosis codes. Marginal structural models were employed to evaluate the causal relationship between HRI history and COPD risk. Standardized inverse probability of treatment and censoring weights were calculated based on time-dependent exposure and confounders. The risk of COPD was assessed according to the occurrence and frequency of HRI.

Results Over an average follow-up period of 11.5 years, 5,173 participants (0.4%) experienced HRI at least once. One year after study entry, 26,919 (1.9%) were newly diagnosed with COPD. A history of HRI was associated with an elevated COPD risk ranging from 2.09 to 2.13, regardless of whether it occurred during the warm season (April to September) or throughout the

entire year. Heat stroke, the most severe form of HRI, showed a stronger association than other HRI types. Older individuals, women, and rural residents were particularly vulnerable.

Conclusions These findings underscore the priority of thermal regulation during heat exposure for the prevention of COPD.

Keywords: heat-related illness; heat exposure; chronic obstructive pulmonary disease; marginal structural model; causal inference.

BACKGROUND

Chronic obstructive pulmonary disease (COPD) is a lifelong chronic respiratory condition [1]. Although not always fatal, it can affect quality of life and may lead to premature death in severe cases. The 2021 Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) reported that COPD was the sixth leading cause of global disability-adjusted life-years (DALYs) in 2021, up from ninth place in 2010 [2]. Furthermore, by 2050, the number of people living with COPD is projected to rise by 23% globally compared to that in 2020 [3]. In this projection, patients with COPD in East Asia and the Pacific are expected to represent the second largest contributing region among the seven global regions. A breakdown analysis by countries and territories revealed that the age-standardized prevalence of COPD in South Korea increased by 10.9 per 100,000 people from 1990 to 2019, making it the only country in East Asia to show an increasing trend [4].

Heat-related illness (HRI) refers to a spectrum of abnormal

thermoregulatory responses that result directly from exposure to extreme heat [5]. Mild forms of HRI, such as heat rash, heat cramps, and heat edema, can be alleviated with prompt cooling, rest, and hydration [5, 6]. However, heat stroke, which is defined by a rectal core temperature of 40 °C (104 °F) or above along with central nervous system symptoms, can be life-threatening. In a United States time-series study, the hospitalization rate of HRI peaked on days with extremely high ambient temperatures (31–38 °C) during the warm season (from May to September), suggesting that HRI may serve as a useful proxy for assessing individual heat exposure [7]. A recent prospective cohort study from China identified that individuals with a history of HRI had a 2.5-fold higher incident risk of cardiovascular disease (CVD) than did those without such a history [8]. However, to date, no epidemiological study has investigated whether HRI leads to the occurrence of COPD.

South Korea is geographically located in a mid-latitude temperate climate zone and is bordered by the sea on three sides. According to a report by the Korean Ministry of Environment, climate warming has emerged as a critical issue in South Korea over the past few decades, with the average ambient temperature rising by approximately 0.18 °C per decade [9]. During the summer from June to August, the average maximum temperature increased by about 1 °C in the periods 1999–2008 and 2009–2018 [10]. For instance, in Seoul, the capital of South Korea where over half of the population resides, the average summer maximum temperature

rose from 28.5 °C to 29.5 °C over this period. The incidence of HRI in South Korea has also increased sharply. In accordance with the Korea Heat-related Illness Surveillance System, the reported annual number of HRI cases rose from 443 in 2011 to 2,818 in 2023, representing a 536% increase over the 13-year period [11].

As climate change is occurring more rapidly than expected, heat exposure may become a key factor in the prevention and management of COPD. Clarifying its impact on the incidence of COPD is therefore of significant importance for global public health. In this large retrospective cohort study, we aimed to examine the causal relationship between HRI and COPD, taking into account all recorded HRI events over a long-term period.

MATERIALS AND METHODS

Study population

The participants in this study were recruited from the National Health Information Database (NHID), which is constructed by the National Health Insurance Services (NHIS) in South Korea [12]. The NHIS is a mandatory health insurance system that covers nearly the entire Korean population. Individuals insured under the NHIS are eligible for a biennial health screening program. The NHID includes a wide range of information, such as demographic characteristics, socioeconomic status, medical records, and health screening results.

This retrospective cohort study initially included 2,477,131 adults who

underwent health screening through the NHIS in 2009-2010 (**Supplementary Figure 1**). Among them, individuals who did not participate in regular screenings (defined as fewer than once every 4 years) were excluded (n=723,859). Additional exclusions were made for the following reasons: 35,170 participants had a history of HRI or COPD prior to enrollment; 213,038 had a history of lung cancer or asthma prior to enrollment; 2,972 were diagnosed with COPD within one year after enrollment; and 109,216 had missing baseline covariate data. After applying these exclusion criteria, a total of 1,392,876 participants were included in the final analysis.

Exposure and outcome assessments

Medical records including dates of visits, diagnostic codes, and type of hospital utilization were obtained from the NHID. Within a single claim, one primary diagnosis code and, if applicable, multiple secondary diagnosis codes were recorded for each visit.

To minimize the inclusion of HRI events related to underlying chronic conditions, inpatient or outpatient visits with a primary diagnosis code of T67 were defined as HRI events [13]. For each month, only the first diagnosis of HRI was counted, and any subsequent records occurring within one week of a previous visit were not considered new events.

In this study, two time scales were defined. The warm season time scale represents HRI events occurring from April to September, whereas the year-round represents events occurring throughout the entire year. For

each time scale, three exposure indicators were constructed. First, HRI history was defined as whether a participant experienced any HRI event during the given time scale (ever or never). Second, the frequency of HRI events was categorized as none, once, or two or more times. Third, HRI severity was classified into severe (heat stroke [T67.0]), moderate (heat exhaustion [T67.3–67.5]), and mild (heat syncope [T67.1], heat cramps [T67.2], heat fatigue [T67.6], heat edema [T67.7], and other effects of heat and light [T67.8–67.9]) [13]. All HRI events were ascertained from baseline to one year prior to the end of follow-up, and exposure variables were treated as time-varying variables.

Participants were followed from enrollment until the date of COPD diagnosis, death, or December 31, 2021, whichever came first. COPD was defined as the first inpatient or outpatient visit with a primary diagnosis code of J44 based on the ICD-10 [13]. To reduce the potential for reverse causation, follow-up for COPD began one year after enrollment.

Covariates

All covariates, except for time-fixed variables such as age at baseline and sex, were treated as time-dependent confounders. Residential areas were categorized into 16 regions according to the *si-do* administrative divisions (metropolitan cities and provinces) in South Korea. Household income level (in quintiles) was indirectly estimated using the insurance premium charged by the NHIS, which is calculated based on employment status and monthly income. Regarding lifetime tobacco use, participants

were classified as never, former, or current users, and further stratified into five groups by 10 pack-year intervals. Alcohol consumption (drinks per day) was calculated by multiplying the frequency (times per day) by the amount consumed per occasion (drinks per time). Vigorous and moderate physical activity (days per week) was assessed through self-reported frequency of activities that were hard or comfortable to breathe during.

Additional covariates including comorbidities and environmental factors were further considered in sensitivity analysis. Histories of asthma, hypertension, and type 2 diabetes were identified based on ICD-10 codes J45-J46, I10-I15, and E11-E14, respectively [13]. Spatially, annual ambient temperature ($^{\circ}\text{C}$) and $\text{PM}_{2.5}$ concentrations ($\mu\text{g}/\text{m}^3$) were also included as confounding factors. Meteorological and air pollution data were calculated using the Weather Research and Forecast - Sparse Matrix Operator Kernel Emission - Community Multiscale Air Quality modeling system at a 9-km resolution [14, 15]. These data were linked to each participant's residential address at the *si-gun-gu* administrative distinct level (counties and districts).

Statistical analysis

A causal framework incorporating both time-fixed and time-dependent covariates was developed to reflect the underlying relationship between HRI and COPD (**Supplementary Figure 2**) [16]. As shown in the directed acyclic graph, time-dependent covariates at time t can affect the exposure at the same time, which in turn may influence both the exposure and

covariate status at the subsequent time point ($t+1$)—a structure known as treatment-confounder feedback [17, 18]. The feedback loop may introduce bias when applying conventional models in longitudinal studies. To address this issue, a marginal structural model (MSM) based on inverse probability of treatment and censoring weighting (IPTW) was applied in this study [19, 20]. MSM is designed to account for treatment-confounder feedback, as weights are constructed using both the past and current status of time-dependent variables. By assigning these weights, a pseudo-population under a counterfactual scenario is created, enabling causal inference by emulating a randomized controlled trial [19, 21].

In this study, inverse probability weights for treatment (i.e., history of HRI) and censoring (i.e., dropout due to death) at each follow-up year were calculated using logistic regression models. The numerator of the weight was estimated using both time-fixed variables (age at baseline and sex) and the baseline values of time-dependent confounders (residential area, household income level, body mass index, smoking status, pack-years, alcohol consumption, and vigorous/moderate physical activity), whereas the denominator extended this by additionally incorporating both past and current status of these time-dependent confounders. The unstabilized weight was calculated as the reciprocal of the product of the denominators for treatment and censoring [22, 23]. Stabilized weight was derived by replacing the numerator of the unstabilized weight with the product of the corresponding numerators. As shown in **Figure 1a** and **Supplementary**

Figure 3, the stabilized weights showed lower variance than unstabilized weights. To handle extreme outliers in the stabilized weights, values below the 0.01 and above the 99.99 percentiles were replaced with the respective cut-off values (**Figure 1b**).

The association between HRI and the risk of COPD was evaluated using marginal structural Cox regression models with stabilized weights. Time-varying exposures were included in the model, reflecting HRI history, frequency, and severity. For comparison, conventional time-dependent Cox regression models without weighting were also performed. All analyses were conducted separately for warm-season and year-round HRI.

Missing information on covariates was handled using the last observation carried forward method [24]. For participants who did not undergo health screening at a given time point, values were imputed using the most recent available measurement within the previous one to three years. The primary model was adjusted for age at baseline, sex, residential area, household income level, body mass index, smoking status, pack-years, alcohol consumption, and vigorous/moderate physical activity. Variance inflation factors (VIFs) were calculated to assess multicollinearity among the exposure and covariates. Most VIFs ranged from 1.0 to 1.8, with the highest value observed for smoking status (3.1), and all values were below the threshold of 5.0 (data not shown) [25]. Subgroup analyses were carried out to examine whether the association between a history of HRI and COPD risk was modified by age (<60 versus ≥ 60 years), sex (men versus women),

and urbanity (urban versus rural). For the urbanity analysis, only 1,111,492 (79.8% of the total population) participants who did not relocate during the study period were included. Metropolitan cities in South Korea, including Seoul, Busan, Daegu, Incheon, Gwangju, Daejeon, and Ulsan, were classified as urban areas. In the sensitivity analysis, four models were applied, extending the primary model by additionally controlling for ambient temperature ($^{\circ}\text{C}$), $\text{PM}_{2.5}$ concentrations ($\mu\text{g}/\text{m}^3$), history of asthma, and hypertension or type 2 diabetes. To assess the robustness to potential unmeasured confounding factors, an E-value was calculated for the primary estimate [26]. Furthermore, a sensitivity analysis was conducted using stabilized weights without replacing extreme outliers. All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) and R software version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

During an average follow-up period of 11.5 years (standard deviation [SD]: 1.1), 5,173 participants (0.4%) had at least one episode of HRI throughout the year, of whom 3,991 cases occurred during the warm season. In total, 861 participants experienced HRI two or more times year-round, and 473 did so during the warm season. Over the same period, 26,919 participants (1.9%) developed COPD (**Table 1**). A total of 11,588 participants (0.8%) dropped out owing to death. At study entry, the mean age of the study population was 46.6 years (SD: 11.8). Participants with

COPD were older than those without COPD, with mean ages of 57.4 and 46.4 years, respectively. Moreover, COPD patients tended to be male and reside in rural areas. Detailed characteristics of the study population are provided in **Supplementary Table 1**. Briefly, the proportions of current tobacco users, especially those with heavy use, and consumers of alcohol were higher among COPD patients than among non-patients. Participants who developed COPD also had higher rates of comorbid hypertension and type 2 diabetes.

A 2.13-fold increased risk of COPD (95% confidence interval [CI]: 1.43–3.17) was observed among participants who experienced HRI at least once during the warm season, compared to those without such a history, as estimated using marginal structural Cox models (**Figure 2**). The risk was slightly attenuated for year-round HRI (hazard ratio [HR]: 2.09; 95% CI: 1.46–2.99). When stratifying HRI cases by frequency, the significant association persisted only among participants with a single HRI episode (HRs [95% CIs]: 1.64 [1.32–2.02] during the warm season and 1.71 [1.41–2.07] year-round), but not among those with two or more episodes (HRs [95% CIs]: 1.07 [0.52–2.18] during the warm season and 1.17 [0.69–2.00] year-round). Severe HRI showed the highest risk of COPD, with HRs of 2.93 for warm season HRI (95% CI: 1.49–5.78) and 3.31 for year-round HRI (95% CI: 1.80–6.08) (**Figure 3**). A lower risk was found for moderate HRI (HRs [95% CIs]: 2.16 [1.29–3.60] during the warm season and 2.04 [1.29–3.21] year-round), whereas the association was no longer evident in the mild HRI

group (HRs [95% CIs]: 0.58 [0.08–4.14] during the warm season and 0.47 [0.07–3.37] year-round). Results from the time-dependent Cox models were consistent with those from the marginal structural Cox models.

Among participants aged 60 years or older, a history of HRI was associated with an excess risk of COPD, with risks of 2.33-fold (95% CI: 1.43–3.81) during the warm season and 2.22-fold (95% CI: 1.41–3.48) year-round (**Figure 4**). By contrast, younger participants with year-round HRI also showed an increased risk (HR: 1.91; 95% CI: 1.05–3.46), although the association was attenuated and became null when HRI occurred during the warm season (HR: 1.80; 95% CI: 0.90–3.61). Both sexes demonstrated significant associations between HRI history and COPD risk, with a stronger effect in women (HRs [95% CIs]: 2.53 [1.35–4.72] during the warm season and 2.49 [1.44–4.30] year-round) compared to men (HRs [95% CIs]: 1.94 [1.15–3.26] during the warm season and 1.89 [1.18–3.04] year-round). Furthermore, when excluding participants who relocated during the study period, those living in rural areas exhibited a greater susceptibility to COPD development following HRI. The estimated risk was slightly higher for warm-season HRI (HR: 2.53; 95% CI: 1.64–3.91) than for year-round HRI (HR: 2.46; 95% CI: 1.65–3.67). However, no significant association was found among urban residents. Sensitivity analyses using various model specifications and including extreme weight outliers yielded results consistent with the primary models (**Supplementary Tables 2 and 3**). Furthermore, the E-value for the primary estimate was 3.60, with a lower

limit of 2.28.

DISCUSSION

In this large population-based cohort study of over one million adults, an excess risk of COPD following a history of HRI was observed, as estimated using MSM. The estimated risk was slightly higher when HRI occurred during warm weather compared to year-round occurrences. Participants who experienced heat stroke may have had a higher risk of developing COPD compared with those with mild or moderate types of HRI. The association between HRI history and COPD risk was stronger among individuals aged 60 years or older and those residing in rural areas. Additionally, women showed greater susceptibility than men. These aforementioned findings suggest a potential association between HRI and COPD risk and provide new insights into COPD prevention and the identification of vulnerable populations. However, the lack of clear dose-response relationship should be interpreted with caution.

Given its adverse effects on respiratory health, extreme heat exposure has increasingly been recognized as a potential risk factor for COPD [27]. In a Chinese study, daily COPD outpatient visits increased by 1.1-fold following heatwaves (defined as days with ambient temperature exceeding 30.3 °C) within the prior 0–7 days, compared to days with optimal temperatures (19.6 °C) [28]. Additionally, a United Kingdom study using high-resolution (1 km) meteorological data found that, on days with temperatures above 23.2 °C, the risk of COPD hospitalization was significantly elevated, with a

47% increase per 1 °C rise in ambient temperature [29]. Notably, a recent meta-analysis of 20 studies reported that heat exposure increases the morbidity and mortality risk of COPD by 19% [30]. However, most observational studies on COPD risk have used case-crossover or time-series designs, which focus on the acute impacts of extreme heat exposure and are limited in their ability to provide causal evidence [29, 31, 32]. Only a few cohort and cross-sectional studies have documented the association between long-term heat exposure and reduced lung function [33, 34]. For instance, a cohort study of young Chinese adults found that the lifetime average summer temperature from birth was inversely associated with lung function [33]. Further longitudinal studies are needed to clarify the potential role of heat exposure in the development of incident COPD.

Although definitions may vary slightly, most observational studies have characterized heat exposure at the regional level as ambient temperatures exceeding a certain threshold or percentile. However, between-person variance may be reduced, limiting the power to detect individual-level associations that incorporate personal confounding factors, and potentially underestimating the effect size [35, 36]. On the other hand, no epidemiological study has yet elucidated the long-term effects of heat exposure on COPD, a typical non-communicable disease that is not immediately fatal but can lead to prolonged impairment in quality of life [37]. The use of HRI allows for easier tracking of heat exposure in the population, even for events that occurred several years earlier.

Under extreme heat conditions, mild forms of HRI may initially present with elevated body temperature, electrolyte loss, and muscle soreness [5, 38, 39]. To maintain thermal homeostasis, cardiac output rises through heart rate acceleration to enhance blood flow to the skin surface, which may lead to an increased respiratory rate and, in severe cases, syncope [40]. Without appropriate intervention, these symptoms may worsen as the body's thermoregulatory capacity becomes impaired, leading to a rise in core temperature and severe dehydration. Once the core temperature exceeds 38–40 °C, irreversible organ dysfunction and cellular damage may occur, often accompanied by inflammatory reactions [27, 38]. In the lungs, heat-induced airway inflammation may contribute to respiratory stress [41]. However, these mechanisms likely reflect acute physiological responses to heat and may not be sufficient to explain the development of COPD, which develops over decades of exposure to multiple risk factors. In this context, HRI may function as a trigger or indicator of underlying susceptibility, as individuals with subclinical COPD may be more vulnerable to heat stress and therefore more likely to experience HRI prior to diagnosis.

In this study, the significant association between experiencing HRI and the risk of COPD persisted only in older people. Owing to altered thermal perception, extreme heat exposure may have a greater adverse impact on older adults than younger individuals [42]. In a Japanese study, older adults (mean age: 71 years) exhibited a higher threshold for detecting warmth than younger adults (mean age: 23 years), potentially indicating an

attenuated ability to sense heat and take preventive actions against HRI under the same hot conditions [42, 43]. A Canadian study also observed that adults aged 55–73 years had lower heat dissipation and greater heat storage [44]. In addition, age-related decline in lung function and the presence of common chronic diseases make older adults more vulnerable to heat exposure [45].

The stronger association in women than in men can be explained by sex differences in thermoregulatory responses to heat exposure. Due to their smaller body mass and skin surface area, women often require greater physiological effort to maintain thermal balance and are therefore less able to adapt to heat [46, 47]. Additionally, an increase in core temperature during the luteal phase of the menstrual cycle may further accelerate the onset of HRI symptoms [48]. A recent meta-analysis of 79 studies also suggested that women were more susceptible to CVD mortality following extreme heat events [49].

In light of our findings, urban-rural differences in heat-related damage may exist and tend to be more pronounced in rural regions. A spatial ecological study from the United States reported that the higher prevalence of HRI in rural areas compared to urban areas can largely be attributed to agricultural activities and a high proportion of mobile homes [50]. Furthermore, rural residents were more likely to spend extended time outdoors, work in outdoor-based industries, and live without air conditioning [51].

This study provides additional evidence on the association between heat-induced HRI and the risk of COPD over a long-term follow-up period. The large, representative sample and well-structured study design, along with the application of MSM for causal inference, strengthen the validity of our findings. To minimize misclassification, both HRI and COPD were ascertained based on primary diagnosis codes, excluding secondary diagnoses that may have resulted from severe health conditions. In addition, by including recurrent HRI records during the follow-up period, an analysis was made to examine how the risk varied with the diagnosis frequency. Furthermore, a series of sensitivity analyses were conducted to ensure the robustness of the estimates.

However, notwithstanding these strengths, several limitations should be noted. First, medical record-based HRI ascertainment may underestimate the true extent of heat exposure in the population, as individuals without clinical recognition were not included. As a result, the small number of HRI cases limited the statistical power to detect associations and made it difficult to evaluate whether the effect on COPD development differed between inpatients and outpatients. In addition, it was not possible to distinguish between classic and exertional heat stroke, which are caused by environmental heat exposure and vigorous physical activity, respectively [52]. Furthermore, due to limited information in the NHIS dataset, we were unable to perform occupational classification, although HRI may also be influenced by occupational exposure. Further

investigations addressing these limitations are warranted to replicate and expand upon the findings of this study. Second, more stringent COPD definition requiring at least one inpatient visit or two or more outpatient visits could not be applied due to the small number of COPD cases available for analysis. Additionally, the absence of information on medication use and spirometry measurements limited the ability to further refine case ascertainment. Nevertheless, physician-diagnosed COPD in the NHIS is likely to have acceptable sensitivity and specificity. Third, as deaths were treated as censoring events in this study, competing risks that may have precluded the development of COPD were not accounted for, potentially leading to an overestimation of the risk [53]. Nevertheless, follow-up for COPD events seems to have remained stable, given the extremely low death rate (0.8%) in the study population. Finally, the use of logistic regression models to estimate propensity scores in the IPTW may have introduced model misspecification [54]. To improve the accuracy, future research should adopt multiple statistical learning algorithms, such as the Super Learner, which identifies the optimal combination of models [55].

CONCLUSIONS

In conclusion, this study suggests an association between HRI episodes and the risk of COPD, regardless of whether HRI occurred during the warm season or throughout the entire year. Among the types of HRI, heat stroke exhibited the strongest risk. Older adults, rural residents, and women were found to be particularly susceptible. These findings highlight the need for

heat stress management in COPD prevention, especially among vulnerable populations.

Abbreviations

HRI: Heat-related illness

COPD: Chronic obstructive pulmonary disease

GBD: Global Burden of Diseases

CVD: Cardiovascular disease

NHID: National Health Information Database

NHIS: National Health Insurance Service

MSM: Marginal structural model

IPTW: Inverse probability of treatment and censoring weighting

VIFs: Variance inflation factors

HR: Hazard ratio

CI: Confidence interval

Supplementary Information

The online version contains supplementary material available at

Acknowledgements

Not applicable.

Authors' Contributions

Taiyue Jin: Conceptualization, Formal analysis, Methodology, Writing – original draft. **Seyoung Kim:** Writing – review & editing. **Seulbi Lee:** Writing – review & editing. **Soontae Kim:** Resources, Writing – review & editing. **Jin-Kyoung Oh:** Writing – review & editing. **Bomi Park:** Funding

acquisition, Writing – review & editing. **Byungmi Kim:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Funding

This work was supported by the National Cancer Center grant (24H1051-2) and the National R&D Program for Cancer Control through the National Cancer Center funded by the Ministry of Health & Welfare, Republic of Korea (HA23C0083).

Declarations

Ethics approval and consent to participate

This study was performed in accordance with the principles of the Declaration of Helsinki. Informed consent was submitted by all participants when they enrolled. The protocol was approved by the Institutional Review Board of the National Cancer Center Korea (NCC2022-0168).

Data Availability

The data that support the findings of this study are available from National Health Information Database of Korea but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

REFERENCES

1. Viegi G, Pistelli F, Sherrill DL, Maio S, Baldacci S, Carrozzi L: **Definition, epidemiology and natural history of COPD.** *Eur Respir J* 2007, **30**(5):993-1013.
2. Ferrari AJ, Santomauro DF, Aali A, Abate YH, Abbafati C, Abbastabar H, Abd ElHafeez S, Abdelmasseh M, Abd-Elsalam S, Abdollahi A *et al*: **Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021.** *The Lancet* 2024, **403**(10440):2133-2161.
3. Boers E, Barrett M, Su JG, Benjafield AV, Sinha S, Kaye L, Zar HJ, Vuong V, Tellez D, Gondalia R *et al*: **Global Burden of Chronic Obstructive Pulmonary Disease Through 2050.** *JAMA Network Open* 2023, **6**(12):e2346598-e2346598.
4. Safiri S, Carson-Chahhoud K, Noori M, Nejadghaderi SA, Sullman MJM, Ahmadian Heris J, Ansarin K, Mansournia MA, Collins GS, Kolahi A-A *et al*: **Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019.** *BMJ* 2022, **378**:e069679.
5. Gauer R, Meyers BK: **Heat-Related Illnesses.** *Am Fam Physician* 2019, **99**(8):482-489.
6. Becker JA, Stewart LK: **Heat-related illness.** *Am Fam Physician* 2011, **83**(11):1325-1330.
7. Sugg MM, Konrad CE, Fuhrmann CM: **Relationships between maximum temperature and heat-related illness across North Carolina, USA.** *International Journal of Biometeorology* 2016, **60**(5):663-675.
8. Huang Q, Ke L, Liu L, Liu Y, Li Y, Feng G, Yin B, Xiang W, Li J, Lyu K *et al*: **Heatwave warnings mitigate long-term cardiovascular diseases risk from heat-related illness: a real-world prospective cohort study.** *The Lancet Regional Health - Western Pacific* 2025, **55**.
9. **Climate Change**
[<https://eng.me.go.kr/eng/web/index.do?menuId=463#:~:text=Average%20annual%20temperature%20in%20Korea,workers%20from%20extreme%20weather%20conditions.>]
10. Park J, Chae Y: **Analysis of time-dependent effects of ambient temperatures on health by vulnerable groups in Korea in 1999-2018.** *Scientific Reports* 2023, **13**(1):922.
11. Lee J, An Daeshik, Ahn Younjhin: **Results of Operating the 2023 Heat-related Illness Surveillance System.** *Public Health Weekly Report* 2024, **17**(34):1421-1431.
12. Sang Cheol S, Yeon-Yong K, Sue KP, Young Ho K, Hyeon Chang K, Jong Heon P, Hee-Jin K, Cheol-Ho D, Jong-Sun S, Eun-Joo L *et al*:

- Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea.** *BMJ Open* 2017, 7(9):e016640.
13. World Health Organization: **International statistical classification of diseases and related health problems, 10th revision, Fifth edition, 2016.** In. Geneva; 2015.
 14. Kim B-U, Bae C, Kim HC, Kim E, Kim S: **Spatially and chemically resolved source apportionment analysis: Case study of high particulate matter event.** *Atmospheric Environment* 2017, 162:55-70.
 15. Jin T, Lee S, Seo J, Ye S, Kim S, Oh J-K, Kim S, Kim B: **Long-term ambient ozone exposure and lung cancer mortality: A nested case-control study in Korea.** *Environmental Pollution* 2025, 375:126299.
 16. Shrier I, Platt RW: **Reducing bias through directed acyclic graphs.** *BMC Medical Research Methodology* 2008, 8(1):70.
 17. VanderWeele TJ: **Principles of confounder selection.** *European Journal of Epidemiology* 2019, 34(3):211-219.
 18. Chesnaye NC, Stel VS, Tripepi G, Dekker FW, Fu EL, Zoccali C, Jager KJ: **An introduction to inverse probability of treatment weighting in observational research.** *Clinical Kidney Journal* 2022, 15(1):14-20.
 19. Robins JM, Hernán MÁ, Brumback B: **Marginal Structural Models and Causal Inference in Epidemiology.** *Epidemiology* 2000, 11(5).
 20. Robins JM: **Marginal Structural Models versus Structural nested Models as Tools for Causal inference.** In: *Statistical Models in Epidemiology, the Environment, and Clinical Trials: 2000// 2000;* New York, NY: Springer New York; 2000: 95-133.
 21. Héroux J, Moodie EEM, Strumpf E, Coyle N, Tousignant P, Diop M: **Marginal structural models for skewed outcomes: identifying causal relationships in health care utilization.** *Statistics in Medicine* 2014, 33(7):1205-1221.
 22. Hernán MÁ, Brumback B, Robins JM: **Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-Positive Men.** *Epidemiology* 2000, 11(5).
 23. Fewell Z, Hernán MA, Wolfe F, Tilling K, Choi H, Sterne JAC: **Controlling for Time-dependent Confounding using Marginal Structural Models.** *The Stata Journal* 2004, 4(4):402-420.
 24. Shao J, Zhong B: **Last observation carry-forward and last observation analysis.** *Statistics in Medicine* 2003, 22(15):2429-2441.
 25. O'brien RM: **A Caution Regarding Rules of Thumb for Variance Inflation Factors.** *Quality & Quantity* 2007, 41(5):673-690.
 26. VanderWeele TJ, Ding P: **Sensitivity Analysis in Observational Research: Introducing the E-Value.** *Annals of Internal Medicine*

- 2017, **167**(4):268-274.
27. Xu J, Xu G, Liu Z, Hou R, Ji JS, Huang C: **COPD risk due to extreme temperature exposure: combining epidemiological evidence with pathophysiological mechanisms.** *eBioMedicine* 2025, **116**.
 28. Zhao Y, Huang Z, Wang S, Hu J, Xiao J, Li X, Liu T, Zeng W, Guo L, Du Q *et al*: **Morbidity burden of respiratory diseases attributable to ambient temperature: a case study in a subtropical city in China.** *Environmental Health* 2019, **18**(1):89.
 29. Konstantinoudis G, Minelli C, Vicedo-Cabrera AM, Ballester J, Gasparrini A, Blangiardo M: **Ambient heat exposure and COPD hospitalisations in England: a nationwide case-crossover study during 2007-2018.** *Thorax* 2022, **77**(11):1098.
 30. Tran HM, Tsai F-J, Lee K-Y, Wang Y-H, Yang F-M, Ho S-C, Bui HTM, Hoang LNN, Bui LTM, Ho K-F *et al*: **Extreme temperature increases the risk of COPD morbimortality: A systematic review and meta-analysis.** *Science of The Total Environment* 2025, **958**:178087.
 31. Scheerens C, Nurhussien L, Aglan A, Synn AJ, Coull BA, Koutrakis P, Rice MB: **The impact of personal and outdoor temperature exposure during cold and warm seasons on lung function and respiratory symptoms in COPD.** *ERJ Open Research* 2022, **8**(1):00574-02021.
 32. Rau A, Tarr Gillian AM, Baldomero Arianne K, Wendt Chris H, Alexander Bruce H, Berman Jesse D: **Heat and Cold Wave-Related Mortality Risk among United States Veterans with Chronic Obstructive Pulmonary Disease: A Case-Crossover Study.** *Environmental Health Perspectives* 2024, **132**(2):027004.
 33. Miao J, Feng S, Wang M, Jiang N, Yu P, Wu Y, Ye T, Wen B, Lu P, Li S *et al*: **Life-time summer heat exposure and lung function in young adults: A retrospective cohort study in Shandong China.** *Environment International* 2022, **160**:107058.
 34. Collaco JM, Appel LJ, McGready J, Cutting GR: **The relationship of lung function with ambient temperature.** *PLOS ONE* 2018, **13**(1):e0191409.
 35. Haining R: **Bivariate Correlation with Spatial Data.** *Geographical Analysis* 1991, **23**(3):210-227.
 36. Deater-Deckard K, Godwin J, Lansford JE, Bacchini D, Bombi AS, Bornstein MH, Chang L, Di Giunta L, Dodge KA, Malone PS *et al*: **Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data.** *Journal of Adolescence* 2018, **62**:207-217.
 37. World Health Organization: **Global status report on noncommunicable diseases 2014.** In. World Health Organization, Geneva, Switzerland; 2014.
 38. Sorensen C, Hess J: **Treatment and Prevention of Heat-Related Illness.** *New England Journal of Medicine* 2022, **387**(15):1404-

- 1413.
39. Krau SD: **Heat-related illness: a hot topic in critical care.** *Crit Care Nurs Clin North Am* 2013, **25**(2):251-262.
 40. Crandall CG, González-Alonso J: **Cardiovascular function in the heat-stressed human.** *Acta Physiologica* 2010, **199**(4):407-423.
 41. Deng L, Ma P, Wu Y, Ma Y, Yang X, Li Y, Deng Q: **High and low temperatures aggravate airway inflammation of asthma: Evidence in a mouse model.** *Environmental Pollution* 2020, **256**:113433.
 42. Takeda R, Imai D, Suzuki A, Ota A, Naghavi N, Yamashina Y, Hirasawa Y, Yokoyama H, Miyagawa T, Okazaki K: **Lower thermal sensation in normothermic and mildly hyperthermic older adults.** *European Journal of Applied Physiology* 2016, **116**(5):975-984.
 43. Wu Y, Zhang Z, Liu H, Li B, Chen B, Kosonen R, Jokisalo J: **Age differences in thermal comfort and physiological responses in thermal environments with temperature ramp.** *Building and Environment* 2023, **228**:109887.
 44. Kenny GP, P. PM, S. MG, Pierre B, Sheila D, J. FB, Janine M, J. SR, E. SAJ, and Flouris AD: **Hyperthermia and cardiovascular strain during an extreme heat exposure in young versus older adults.** *Temperature* 2017, **4**(1):79-88.
 45. Balmain BN, Sabapathy S, Louis M, Morris NR: **Aging and Thermoregulatory Control: The Clinical Implications of Exercising under Heat Stress in Older Individuals.** *BioMed Research International* 2018, **2018**(1):8306154.
 46. Burse RL: **Sex Differences in Human Thermoregulatory Response to Heat and Cold Stress.** *Human Factors* 1979, **21**(6):687-699.
 47. Wickham KA, Wallace PJ, Cheung SS: **Sex differences in the physiological adaptations to heat acclimation: a state-of-the-art review.** *European Journal of Applied Physiology* 2021, **121**(2):353-367.
 48. Marsh SA, Jenkins DG: **Physiological Responses to the Menstrual Cycle.** *Sports Medicine* 2002, **32**(10):601-614.
 49. Zhou Y, Larochele L, Khan FA, Pilote L: **Sex differences in the impact of extreme heat on cardiovascular disease outcomes: a systematic review and meta-analysis.** *Environmental Health* 2025, **24**(1):20.
 50. Kovach MM, Konrad CE, Fuhrmann CM: **Area-level risk factors for heat-related illness in rural and urban locations across North Carolina, USA.** *Applied Geography* 2015, **60**:175-183.
 51. Matz CJ, Stieb DM, Brion O: **Urban-rural differences in daily time-activity patterns, occupational activity and housing characteristics.** *Environmental Health* 2015, **14**(1):88.
 52. Bouchama A, Abuyassin B, Lehe C, Laitano O, Jay O, O'Connor FG,

- Leon LR: **Classic and exertional heatstroke**. *Nature Reviews Disease Primers* 2022, **8**(1):8.
53. Austin PC, Lee DS, Fine JP: **Introduction to the Analysis of Survival Data in the Presence of Competing Risks**. *Circulation* 2016, **133**(6):601-609.
54. Austin PC, Stuart EA: **The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes**. *Statistical Methods in Medical Research* 2015, **26**(4):1654-1670.
55. Pirracchio R, Petersen ML, van der Laan M: **Improving Propensity Score Estimators' Robustness to Model Misspecification Using Super Learner**. *American Journal of Epidemiology* 2015, **181**(2):108-119.

Figures

Figure 1. Distribution of stabilized weights (natural log-transformed) in each follow-up year. (A) without extreme outlier handling; (B) after replacing extreme outliers with the values at the 0.01 and 99.99 percentiles.

Figure 2. Association between history of heat-related illness (HRI) and risk of chronic obstructive pulmonary disease (COPD) estimated using time-dependent Cox regression and marginal structural Cox models. “Ever” indicates any history of HRI, further divided into “Once” and “Twice or more” based on diagnosis frequency. Models were adjusted for age at baseline, sex, residential area, household income, body mass index, smoking status, pack-years, alcohol consumption, and vigorous/moderate physical activity.

Figure 3. Association between severity of heat-related illness (HRI) and risk of chronic obstructive pulmonary disease (COPD) estimated using time-dependent Cox regression and marginal structural Cox models. Severe, heat stroke; Moderate, heat exhaustion; Mild, heat syncope, cramps, fatigue, edema, and other effects of heat and light. Models were adjusted for age at

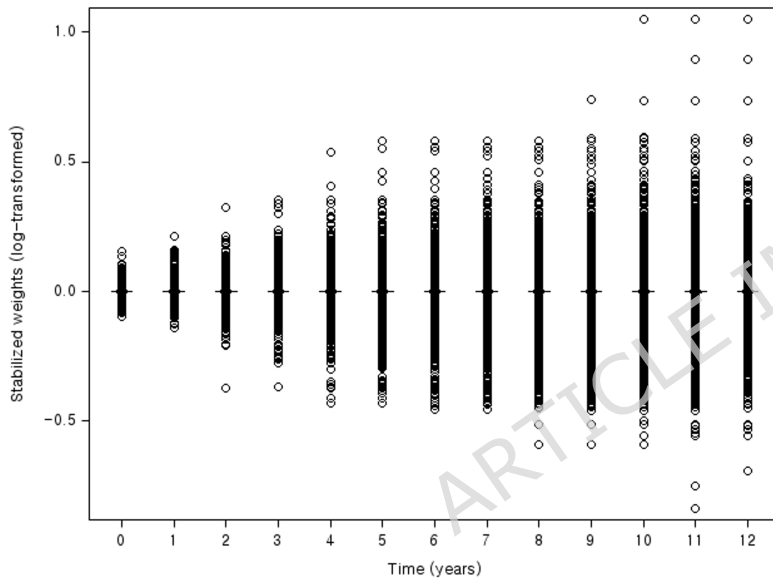
baseline, sex, residential area, household income, body mass index, smoking status, pack-years, alcohol consumption, and vigorous/moderate physical activity.

Figure 4. Subgroup analyses by age group, sex, and residential area for the association between history of heat-related illness (HRI) and risk of chronic obstructive pulmonary disease (COPD), estimated using the marginal structural Cox model. Age group was classified based on age at baseline. For residential area analysis, participants who did not relocate were included (n=1,111,492). Models were adjusted for age at baseline, sex, residential area, household income, body mass index, smoking status, pack-years, alcohol consumption, and vigorous/moderate physical activity.

Table 1. Baseline characteristics of the study population.

	Overall (n = 1392876)	COPD	
		Cases (n = 26919)	Non-cases (n = 1365957)
Follow-up duration, years	11.5 ± 1.1	5.4 ± 3.4	11.6 ± 0.5
Age at baseline			
<60 years	1165482 (83.7)	13802 (51.3)	1151680 (84.3)
≥60 years	227394 (16.3)	13117 (48.7)	214277 (15.7)
Sex			
Men	754009 (54.1)	18072 (67.1)	735937 (53.9)
Women	638867 (45.9)	8847 (32.9)	630020 (46.1)
Residential area at baseline			
Urban	666031 (47.8)	11211 (41.7)	654820 (47.9)
Rural	726845 (52.2)	15708 (58.4)	711137 (52.1)

Data was presented as mean ± standard deviation and frequency (percentage) for continuous and categorical variables, respectively.

A.**B.**