scientific reports

OPEN

Long-term ambient hydrocarbon exposure and incidence of urinary bladder cancer

Han-Wei Zhang [1,2,3,4,12], Zhi-Ren Tsai [5,6,7,12], Victor C. Kok [8,9]✉, Hsiao-Ching Peng[4], Yau-Hung Chen [10], Jeffrey J. P. Tsai[7,8] & Chung Y. Hsu[11]

Particulate matter and volatile organic compounds, including total hydrocarbons (THCs), are major ambient air pollutants. Primary nonmethane hydrocarbons (NMHCs) originate from vehicle emissions. The association between air pollution and urinary bladder cancer (UBC) is debatable. We investigated whether long-term exposure to ambient hydrocarbons increases UBC risk among people aged ≥ 20 years in Taiwan. Linkage dataset research with longitudinal design was conducted among 589,135 initially cancer-free individuals during 2000–2013; 12 airborne pollutants were identified. Several Cox models considering potential confounders were employed. The study outcomes were invasive or in situ UBC incidence over time. The targeted pollutant concentration was divided into three tertiles: T1/T2/T3. The mean age of individuals at risk was 42.5 (SD 15.7), and 50.5% of the individuals were men. The mean daily average over 10 years of airborne THC concentration was 2.25 ppm (SD 0.13), and NMHC was 0.29 ppm (SD 0.09). Both pollutants show long-term monotonic downward trend over time using the Mann–Kendall test. There was a dose-dependent increase in UBC at follow-up. UBC incidence per 100,000 enrollees according to T1/T2/T3 exposure to THC was 60.9, 221.2, and 651.8, respectively; it was 170.0/349.5/426.7 per 100,000 enrollees, corresponding to T1/T2/T3 exposure to NMHC, respectively. Without controlling for confounding air pollutants, the adjusted hazard ratio (adj.HR) was 1.83 (95% CI 1.75–1.91) per 0.13-ppm increase in THC; after controlling for $PM_{2.5}$, adj.HR was even higher at 2.09 (95% CI 1.99–2.19). The adj.HR was 1.37 (95% CI 1.32–1.43) per 0.09-ppm increase in ambient NMHC concentration. After controlling for SO_2 and CH_4, the adj.HR was 1.10 (95% CI 1.06–1.15). Sensitivity analyses showed that UBC development risk was not sex-specific or influenced by diabetes status. Long-term exposure to THC and NMHC may be a risk factor for UBC development. Acknowledging pollutant sources can inform risk management strategies.

Abbreviations

CI	Confidence interval
EPA	Environmental Protection Administration
HC	Hydrocarbons
HR	Hazard ratio
IDC-9-CM	International Classification of Disease, Ninth Revision, Clinical Modification
NHIRD	National Health Insurance Research Database
NMHC	Nonmethane hydrocarbon
OR	Odds ratio

[1]Program for Aging, China Medical University, Taichung, Taiwan. [2]Institute of Population Health Sciences, National Health Research Institutes, Miaoli, Taiwan. [3]Institute of Electrical Control Engineering, Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan. [4]Biomedica Corporation, New Taipei, Taiwan. [5]Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan. [6]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan. [7]Center for Precision Medicine Research, Asia University, Taichung, Taiwan. [8]Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan. [9]Division of Medical Oncology, Kuang Tien General Hospital Cancer Center, 117 Shatien Rd Shalu Dist., Taichung 43303, Taiwan. [10]Department of Chemistry, Tamkang University, New Taipei City 25137, Taiwan. [11]Graduate Institute of Biomedical Science, China Medical University, Taichung, Taiwan. [12]These authors contributed equally: Han-Wei Zhang and Zhi-Ren Tsai. ✉email: victorkok@asia.edu.tw; vckok@outlook.com

1

PAHs	Polycyclic aromatic hydrocarbons
PM	Particulate matter
T	Tertile
THC	Total hydrocarbon
UBC	Urinary bladder cancer
VOCs	Volatile organic compounds

Anthropogenic environmental pollutants are believed to account for a sizable portion of the worldwide incidence of cancer[1]. Over the past decades, hundreds of confirmed and suspected environmental carcinogens have been identified. Industrialization has led to air pollution from dust storms, smoke, fumes, and toxic gas emissions from thermal power plants, coal mines, petroleum, and chemicals. As a result, ambient air pollution has become one of the most significant environmental risks to health. In addition, exposure to outdoor air pollution poses an urgent public health challenge worldwide because it is ubiquitous, affects everyone, and has numerous adverse human health effects, including cancer[2,3]. According to the International Agency for Research on Cancer, air pollution is a Group 1 carcinogen. In addition to causing lung cancer, it is also associated with an increased risk of other types of cancer[4]. Therefore, studies on the toxicological effects of these anthropogenic ambient air pollutants and their impact on human organs are urgently needed.

Examining the pollutant profiles can inform policy-making on forming environmental and public health counteractive measures; exhaust emissions from transportation include total hydrocarbons (THC) and nitrogen oxides (NO_X). Exhaust gases emitted by industrial plants contain a variety of pollutants, such as volatile organic compounds (VOCs), nitrogen dioxide (NO_2), and carbon monoxide (CO)[1]. Outdoors, petrochemical solvent spillage, evaporated fuels, and biogenic emissions produce air pollutants in the form of gases, VOCs, and PM. Although regarded as low-carbon emissions, electricity generation from natural gas power plants produces airborne nonmethane hydrocarbons (NMHC) and nitrogen oxides (NOx) in the long run. Notably, after a series of photochemical reactions between VOCs (including hydrocarbons) and NO_X, the concentration of ground-level ozone (O_3) increases, leading to poor air quality[5]. Therefore, effective control of hydrocarbons (HC) can indirectly reduce the ground-level O_3 concentration to improve regional air quality.

VOCs are defined in several ways. Total hydrocarbons (THCs) are also VOCs when used in a broader sense. Oxygenated hydrocarbons, such as alcohols and aldehydes, are not considered THCs. HCs are the most toxic organic gases in vehicle emissions. Methane, a HC, is neither photoreactive nor toxic; conversely, "nonmethane hydrocarbons" are known to be reactive in ambient air. Although slight differences may occur among different geographical regions[5], a source apportionment study of non-biogenic, anthropogenic NMHC as an air pollutant in Delhi, India, disclosed that the primary source of NMHC was traffic vehicle emissions (petrol and diesel), with 38% from petroleum, 32% from liquefied petroleum gas, 16% from solid fuel combustion, and 14% from diesel[6].

The global cancer burden estimation report, GLOBOCAN 2020, was published in 2021, showing urinary bladder cancer (UBC) as the 12th most common cancer worldwide, accounting for 3% of all cancer burden[7]. In 2020, there were 573,278 newly diagnosed UBCs, with 212,536 new deaths the same year[7]. Furthermore, although the overall number of UBC cases is relatively minor compared with lung cancer, colorectal cancer, and liver cancer, the recurrence rate of bladder cancer is high at up to 70%. Fortunately, over the years, through government-led air pollution control, workplace health promotion, industrial occupational hazard exposure prevention, bans on aristolochic acid-containing medicines starting from November 4, 2003[8], and a substantial reduction in the cigarette smoking rate in Taiwan, the age-specific incidence rates of UBC in Taiwan are expected to decrease by > 25% from 2016 to 2025[9–11]. Despite all these efforts, it is still important to examine the magnitude of UBC development risk due to specific ambient air pollutants.

Studies on the association between air pollution and UBC have revealed contradictory results. A few studies with positive results suggest that air pollution, particularly PM, is associated with a higher incidence of UBC[12–19]. In contrast, other studies did not detect this causal relationship[20–23]. People living near chemical factories[24] harbor a higher risk of developing UBC. However, the culprit pollutants are thought to be airborne polycyclic aromatic hydrocarbons (PAHs) and motor vehicle engine exhaust. In addition, several studies have suggested that air pollution (ambient $PM_{2.5}$, traffic air pollution, and petrochemical air pollutant emissions) is associated with increased bladder cancer mortality in Taiwan[16,18,20]. However, to our knowledge, no studies have investigated the ambient air hydrocarbon pollutant, THC, and NMHC.

We hypothesized that exposure to the ambient air pollutants THC and NMHC would increase the risk of developing UBC. Data from the National Health Insurance Research Database (NHIRD) and government environmental databases were used to examine whether long-term exposure to HC in ambient air increased UBC risk among people aged ≥ 20 years in Taiwan. This is one of the first studies to investigate the risk of UBC associated with exposure to ambient air THC and NMHC pollution.

Results

Study population characteristics.
This study tracked 589,135 initially cancer-free individuals aged 20 years and above (Fig. 1). The mean age was 42.5 ± 15.7. Males accounted for 50.5% of the total population. The most prevalent medical comorbidities were dyslipidemia (32.5%), hypertension (24.0%), diabetes mellitus (23.5%), chronic liver disease (21.6%), and gout (18.7%). We also examined risk factors for UBC in the entire cohort, which included smoking-related diagnoses with a prevalence of 12.3%, alcohol use disorder (4.1%), morbid obesity (2.1%), spinal cord injury (1.2%), chronic cystitis (0.8%), chronic kidney disease (4.8%), and pesticide exposure (0.3%). Furthermore, 54.3%, 31.4%, 6.4%, and 0.9% of the participants lived in high, medium-high, medium, or low levels of urbanization, respectively. A total of 5416 study subjects originated from the six black-foot disease endemic regions, accounting for 0.9% of the entire study cohort. The demographic data and
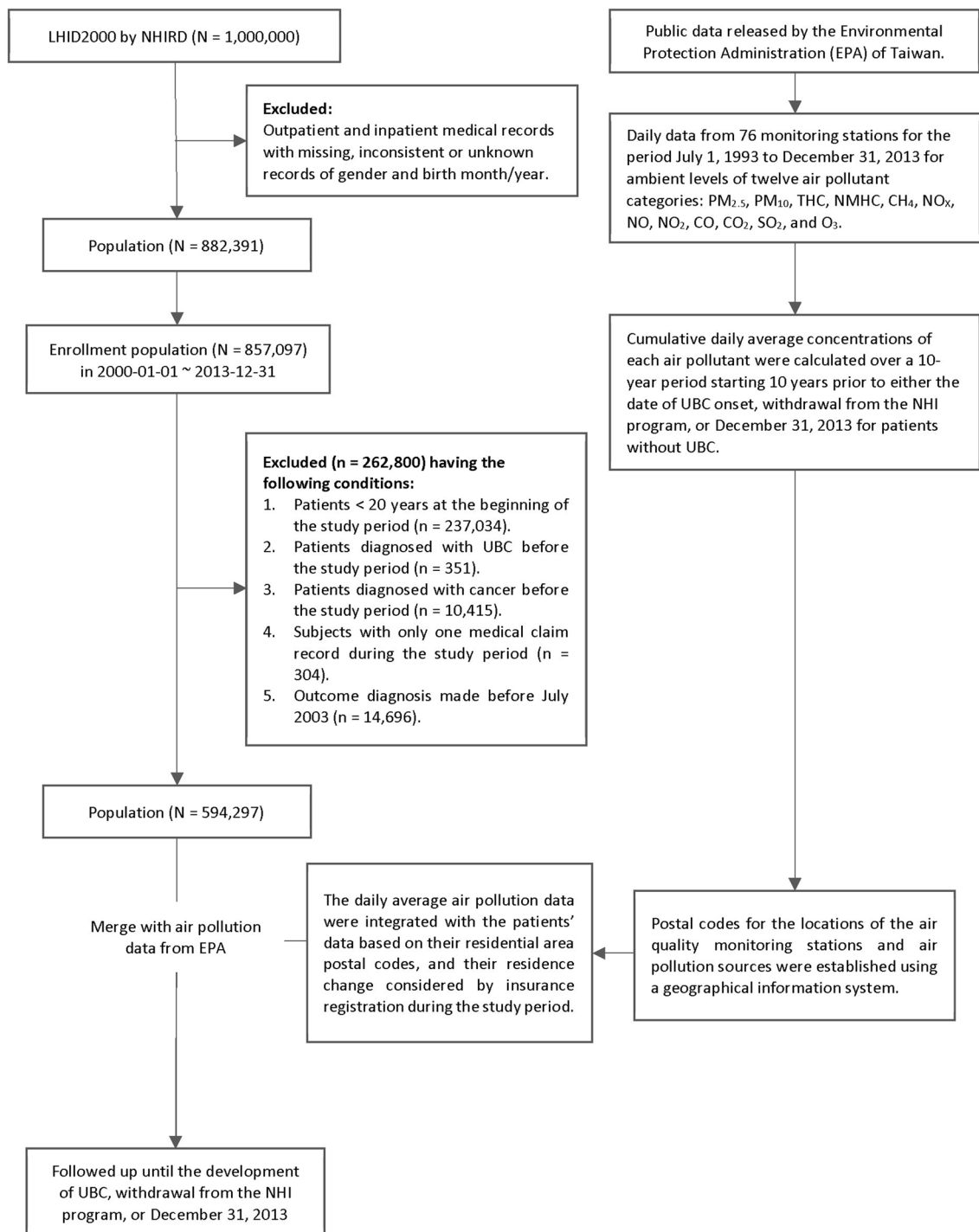
Figure 1. The study flow diagram.

comorbid states among tertiles of THC and NMHC are presented in Tables 1 and 2, respectively, with T1 and T3 being the lowest and highest levels of the daily average of the respective pollutant. Individuals exposed to any pollutant in T1 had more comorbidities and risk factors than those under T3 exposure (Tables 1 and 2).

Air pollution exposure. In the current study over the 10-year exposure period, the mean daily average of THC concentration was 2.25 ppm (SD = 0.13); NMHC was 0.29 ppm (SD = 0.09). Table S1 shows the Pearson's correlation analysis for the 12 air pollutants over a 10-year exposure period. The absolute value of the correlation coefficients < |0.3| denote a low correlation strength, which qualifies as the controlling pollutants in the multiple-pollutant models of the targeted pollutants, THC and NMHC. THC with SO_2 (r = 0.089), PM_{10} (r = − 0.238), and $PM_{2.5}$ (r = − 0.287), while NMHC with SO_2 (r = 0.159) and CH_4 (r = 0.243) (Supplementary Table S1). The summary statistics of the air pollutants over a 10-year exposure period are shown in Supplementary Table S2.

Characteristics	Tertile of daily cumulative average total hydrocarbons[a], n (%)			P value[b]	Total (N = 589,135)
	T1 (lowest) (n = 193,675)	T2 (n = 189,890)	T3 (highest) (n = 205,570)		
Urinary bladder cancer	118	420	1340	<0.001	1878
Incident UBC per 100,000	60.9	221.2	651.8		–
Odds ratio (95% CI)	Reference	3.64 (2.96–4.46)	10.76 (8.92–12.99)		–
Attributable risk proportion % (95% CI)	Reference	56.62 (49.70–63.54)	83.41 (80.53–86.28)		–
Age, years				<0.001	
20–29	45,209 (23.3)	47,935 (25.2)	48,808 (23.7)		141,952 (24.1)
30–39	50,022 (25.8)	45,760 (24.1)	49,580 (24.1)		145,362 (24.7)
40–49	46,564 (24.0)	42,483 (22.4)	41,402 (20.1)		130,449 (22.1)
50–59	25,263 (13.0)	26,173 (13.8)	22,656 (11.0)		74,092 (12.6)
60–69	17,298 (8.9)	16,466 (8.7)	19,758 (9.6)		53,522 (9.1)
≥70	9319 (4.8)	11,073 (5.8)	23,366 (11.4)		43,758 (7.4)
Mean ± SD	41.7 ± 14.5	41.8 ± 15.2	43.9 ± 17.2	<0.001	42.5 ± 15.7
Sex				<0.001	
Male	95,637 (49.4)	94,736 (49.9)	106,860 (52.0)		297,233 (50.5)
Comorbidities[c]					
Essential hypertension	48,441 (25.0)	45,755 (24.1)	47,055 (22.9)	<0.001	141,251 (24.0)
Chronic cystitis	1520 (0.8)	1744 (0.9)	1568 (0.8)	<0.001	4832 (0.8)
Smoking-related diagnoses	22,684 (11.7)	23,613 (12.4)	25,981 (12.6)	<0.001	72,278 (12.3)
Alcohol use disorders	8235 (4.3)	7518 (4.0)	8453 (4.1)	<0.001	24,206 (4.1)
Morbid obesity	4060 (2.1)	3962 (2.1)	4266 (2.1)	0.887	12,288 (2.1)
Spinal cord injury	2656 (1.4)	2272 (1.2)	2407 (1.2)	<0.001	7335 (1.2)
Diabetes mellitus	44,653 (23.1)	43,798 (23.1)	49,752 (24.2)	<0.001	138,203 (23.5)
Chronic liver disease	41,662 (21.5)	41,122 (21.7)	44,706 (21.7)	0.191	127,490 (21.6)
Gout	36,762 (19.0)	35,001 (18.4)	38,368 (18.7)	<0.001	110,131 (18.7)
Chronic kidney disease	9603 (5.0)	8945 (4.7)	9544 (4.6)	<0.001	28,092 (4.8)
Pesticide exposures	766 (0.4)	575 (0.3)	426 (0.2)	<0.001	1767 (0.3)
Dyslipidemia	63,528 (32.8)	62,560 (32.9)	65,176 (31.7)	<0.001	191,264 (32.5)
Level of urbanization[d]				<0.001	
High	101,049 (52.2)	100,212 (52.8)	118,698 (57.7)		319,959 (54.3)
Medium-high	68,932 (35.6)	62,200 (32.8)	53,801 (26.2)		184,933 (31.4)
Medium	10,629 (5.5)	13,159 (6.9)	13,849 (6.7)		37,637 (6.4)
Low	224 (0.1)	2185 (1.15)	2743 (1.3)		5152 (0.9)
Unknown	12,841 (6.6)	12,134 (6.4)	16,479 (8.0)		41,454 (7.0)
Black-foot disease endemic region[e]				<0.001	
No	193,213 (99.8)	186,302 (98.1)	204,204 (99.3)		583,719 (99.1)
Yes	462 (0.2)	3588 (1.9)	1366 (0.7)		5416 (0.9)

Table 1. Odds ratios of incident urinary bladder cancer by tertile of the total hydrocarbons' exposure and characteristics of the cohorts. *CI* confidence interval, *SD* standard deviation, *UBC* urinary bladder cancer. [a]The tertile values, in ppm, are as follows: T1: < 2.17, T2: ≥ 2.17 and < 2.32, and T3: ≥ 2.32. [b]The Chi-squared test or one-way analysis of variance among tertiles of total hydrocarbons. [c]Comorbidities were before the survival date. [d]Level of urbanization was defined at the beginning of the follow-up period. [e]The historic black-foot endemic region included six southwestern coastal townships where people drank arsenic-contaminated well water before tap water installation in 1979–2003.

The distributions of daily average concentrations of air pollutants over the 10-year exposure period are shown in Supplementary Figs. S1 and S2 (SO_2, CO_2, CO, O_3, PM_{10}, and $PM_{2.5}$, are shown in Fig. S1; NOx, NO, NO_2, THC, NMHC, and CH_4 in Fig. S2). Both ambient pollutants showed a long-term monotonic downward trend over time by the Mann–Kendall test. Sen's method used to estimate the slope of the trend over the study period revealed a downward trend of − 0.03 units per year for THC (95% CI, − 0.04 to − 0.02; p < 0.001) and − 0.01 units/ year for NMHC (95% CI, − 0.01 to − 0.01; p < 0.001) (Supplementary Fig. S3).

Ambient air THC, NMHC exposure, and incident UBC. Among the target pollutant concentrations to which all cohort participants were exposed, pollutant levels were categorized into tertiles, with T1 being the lowest and T3 being the highest. At the end of the follow-up period, prolonged exposure to THC increased the number of newly diagnosed UBC in a dose-dependent manner: 60.9 cases per 100,000 study individuals under

| Characteristics | Tertile of daily cumulative average NMHC[a], n (%) | | | P value[b] | Total (N = 589,135) |
	T1 (lowest) (n = 185,291)	T2 (n = 207,464)	T3 (highest) (n = 196,380)		
Urinary bladder cancer	315	725	838		1878
Incident UBC per 100,000	170.0	349.5	426.7	< 0.001	–
Odds ratio (95% CI)	Reference	2.06 (1.80–2.35)	2.52 (2.21–2.87)		–
Attributable risk proportion % (95% CI)	Reference	35.88 (29.96–41.80)	43.83 (38.54–49.12)		
Age, years				< 0.001	
20–29	39,467 (21.3)	50,316 (24.3)	52,169 (26.6)		141,952 (24.1)
30–39	45,261 (24.4)	49,483 (23.9)	50,618 (25.8)		145,362 (24.7)
40–49	42,186 (22.8)	47,122 (22.7)	41,141 (20.9)		130,449 (22.1)
50–59	24,477 (13.2)	28,400 (13.7)	21,215 (10.8)		74,092 (12.6)
60–69	20,474 (11.0)	17,609 (8.5)	15,439 (7.9)		53,522 (9.1)
≥ 70	13,426 (7.2)	14,534 (7.0)	15,798 (8.0)		43,758 (7.4)
Mean ± SD	43.5 ± 15.5	42.4 ± 15.5	41.7 ± 16.1	< 0.001	42.5 ± 15.7
Sex				< 0.001	
Male	92,719 (50.0)	104,005 (50.1)	100,509 (51.2)		297,233 (50.5)
Comorbidities[c]					
Essential hypertension	47,351 (25.6)	49,796 (24.0)	44,104 (22.5)	< 0.001	141,251 (24.0)
Chronic cystitis	2142 (1.2)	1411 (0.7)	1279 (0.7)	< 0.001	4832 (0.8)
Smoking-related diagnoses	25,284 (13.6)	24,318 (11.7)	22,676 (11.5)	< 0.001	72,278 (12.3)
Alcohol use disorders	8045 (4.3)	8823 (4.3)	7338 (3.7)	< 0.001	24,206 (4.1)
Morbid obesity	3700 (2.0)	4183 (2.0)	4405 (2.2)	< 0.001	12,288 (2.1)
Spinal cord injury	2763 (1.5)	2669 (1.3)	1903 (1.0)	< 0.001	7335 (1.2)
Diabetes mellitus	46,637 (25.2)	48,184 (23.2)	43,382 (22.1)	< 0.001	138,203 (23.5)
Chronic liver disease	42,373 (22.9)	44,262 (21.3)	40,855 (20.8)	< 0.001	127,490 (21.6)
Gout	37,713 (20.4)	37,634 (18.1)	34,784 (17.7)	< 0.001	110,131 (18.7)
Chronic kidney disease	9655 (5.2)	10,264 (4.9)	8173 (4.2)	< 0.001	28,092 (4.8)
Pesticide exposures	960 (0.5)	601 (0.3)	206 (0.1)	< 0.001	1767 (0.3)
Dyslipidemia	64,086 (34.6)	66,012 (31.8)	61,166 (31.1)	< 0.001	191,264 (32.5)
Level of urbanization[d]				< 0.001	
High	98,404 (53.1)	103,259 (49.8)	118,296 (60.2)		319,959 (54.3)
Medium-high	60,327 (32.6)	76,580 (36.9)	48,026 (24.5)		184,933 (31.4)
Medium	14,464 (7.8)	10,070 (4.9)	13,103 (6.7)		37,637 (6.4)
Low	207 (0.1)	2244 (1.1)	2701 (1.4)		5152 (0.9)
Unknown	11,889 (6.4)	15,311 (7.4)	14,254 (7.3)		41,454 (7.0)
Black-foot disease endemic region[e]				< 0.001	
No	180,312 (97.3)	207,130 (99.8)	196,277 (99.9)		583,719 (99.1)
Yes	4979 (2.7)	334 (0.2)	103 (0.1)		5416 (0.9)

Table 2. Odds ratios of incident urinary bladder cancer by each tertile of nonmethane hydrocarbons exposure and characteristics of the cohorts. *CI* confidence interval, *NMHC* nonmethane hydrocarbon, *SD* standard deviation, *UBC* urinary bladder cancer. [a]The tertile values, in ppm, are as follows: T1: < 0.25, T2: ≥ 0.25 and < 0.33, and T3: ≥ 0.33. [b]The Chi-squared test or one-way analysis of variance among tertiles of nonmethane hydrocarbons. [c]Comorbidities were before the survival date. [d]Level of urbanization was defined at the beginning of the follow-up period. [e]The historic black-foot endemic region included six southwestern coastal townships where people drank arsenic-contaminated well water before tap water installation in 1979–2003.

T1 concentration exposure, 221.2 under T2 exposure, and 651.8 under T3 exposure (Table 1). For T2 exposure, the odds ratio (OR) was 3.64 (95% CI 2.96–4.46) when compared to that of T1 exposure. Moreover, when comparing T3 to T1, the OR was 10.76 (95% CI 8.92–12.99) (Table 1). Ambient air NMHC exposure was associated with 170.0/100,000 enrollees over the entire study period for the lowest T1 exposure; 349.5 for T2; and 426.7 for T3, indicating a dose-dependent effect (Table 2). The OR was 2.06 (95% CI 1.80–2.35) for T2 exposure when compared to that of T1 exposure. The OR was 2.52 (95% CI 2.21–2.87) for T3 compared to T1 (Table 2).

Tables 3, 4, and Supplementary Table S3 present the single- and multiple-pollutant models for per 0.13 ppm or 0.09 ppm increase in THC or NMHC on different stratifications. We fitted two-pollutant models for THC by controlling for the concomitant exposure to SO_2, PM_{10}, or $PM_{2.5}$. Two-pollutant models for NMHC were used to control for the concomitant exposure to SO_2 or CH_4, and the three-pollutant model for NMHC while controlling for SO_2 and CH_4.

Pollutant category	Controlling pollutant[a]	Crude HR (95% CI)	Adjusted HR[b] (95% CI)
THC (0.13 ppm increase)	–	2.87 (2.79, 2.94)[‡]	1.83 (1.75, 1.91)[‡]
	SO_2	2.97 (2.89, 3.06)[‡]	1.82 (1.74, 1.90)[‡]
	PM_{10}	3.41 (3.31, 3.51)[‡]	2.01 (1.92, 2.11)[‡]
	$PM_{2.5}$	3.41 (3.32, 3.50)[‡]	2.09 (1.99, 2.19)[‡]
NMHC (0.09 ppm increase)	–	2.04 (1.96, 2.12)[‡]	1.37 (1.32, 1.43)[‡]
	SO_2	2.06 (1.98, 2.14)[‡]	1.36 (1.30, 1.42)[‡]
	CH_4	1.45 (1.41, 1.49)[‡]	1.14 (1.09, 1.18)[‡]
	SO_2, CH_4	1.45 (1.40,1.49)[‡]	1.10 (1.06, 1.15)[‡]

Table 3. Crude and adjusted hazard ratios of developing urinary bladder cancer during long-term THC or NMHC air pollutants exposure at a standard deviation (SD) increment controlled for $PM_{2.5}$ and other air pollutants. *HR* hazard ratio, *CI* confidence interval, SO_2 sulfur dioxide, PM_{10} particulate matter < 10 μm in size, $PM_{2.5}$ particulate matter < 2.5 μm in size, CH_4 methane. [a]Additional pollutants were added into the pollutant models for multiple analysis only when Pearson's correlation coefficient was < 0.3. [b]Cox regression models were adjusted for age, sex, level of urbanization, the black-foot disease endemic region, essential hypertension, chronic cystitis, smoking-related diagnoses, alcohol use disorders, morbid obesity, spinal cord injury, chronic liver disease, diabetes mellitus, gout, chronic kidney disease, pesticide exposures, dyslipidemia, lag0–2, season, and ambient temperature, controlled pollutants (weak correlation with THC or NMHC). [‡]p < 0.001.

Pollutant category	Tertiles of average daily pollutant[a]	Population	Urinary bladder cancer	Person-year	Adjusted HR[b] (95% CI)
THC	T1 (lowest)	Total (N = 589,135)	118	2,472,031	Reference
	T2		420	2,381,570	1.23 (0.98, 1.54)
	T3 (highest)		1340	2,407,071	3.16 (2.54, 3.94)[‡]
	p for trend				< 0.001
	T1 (lowest)	Male (n = 297,233)	80	1,233,313	Reference
	T2		277	1,199,760	1.26 (0.95, 1.68)
	T3 (highest)		861	1,236,185	3.16 (2.40, 4.16)[‡]
	p for trend				< 0.001
	T1 (lowest)	Female (n = 291,902)	38	1,238,718	Reference
	T2		143	1,181,810	1.15 (0.78, 1.68)
	T3 (highest)		479	1,170,885	3.11 (2.15, 4.49)[‡]
	p for trend				< 0.001
NMHC	T1 (lowest)	Total (N = 589,135)	315	2,334,352	Reference
	T2		725	2,543,067	1.67 (1.41, 1.97)[‡]
	T3 (highest)		838	2,383,252	1.94 (1.61, 2.34)[‡]
	p for trend				< 0.001
	T1 (lowest)	Male (n = 297,233)	208	1,174,828	Reference
	T2		469	1,280,054	1.65 (1.35, 2.03)[‡]
	T3 (highest)		541	1,214,377	1.98 (1.57, 2.49)[‡]
	p for trend				< 0.001
	T1 (lowest)	Female (n = 291,902)	107	1,159,525	Reference
	T2		256	1,263,013	1.62 (1.22, 2.15)[‡]
	T3 (highest)		297	1,168,875	1.83 (1.33, 2.51)[‡]
	p for trend				< 0.001

Table 4. Cox regression-derived adjusted hazard ratios for incident urinary bladder cancer associated with each tertile of ambient THC or NMHC exposure stratified by sex. *HR* hazard ratio, *CI* confidence interval, *THC* total hydrocarbons, *NMHC* nonmethane hydrocarbons. [a]The tertile values, in ppm (THC, NMHC), are as follows: THC (T1: < 2.17, T2: ≥ 2.17 and < 2.32, and T3: ≥ 2.32); NMHC (T1: < 0.25, T2: ≥ 0.25 and < 0.33, and T3: ≥ 0.33). [b]Cox regression models were adjusted for age, sex, level of urbanization, the black-foot disease endemic region, essential hypertension, chronic cystitis, smoking-related diagnoses, alcohol use disorders, morbid obesity, spinal cord injury, chronic liver disease, diabetes mellitus, gout, chronic kidney disease, pesticide exposures, dyslipidemia, lag0–2, season, and ambient temperature. [‡]p < 0.001.

Without controlling for confounding air pollutants, the adjusted hazard ratio (HR) for UBC development was 1.83 (95% CI 1.75–1.91; $p < 0.001$) per 0.13 ppm increase of THC; after controlling for $PM_{2.5}$, adjusted HR was even higher at 2.09 (95% CI 1.99–2.19). The adjusted HR was 1.37 (95% CI 1.32–1.43; $p < 0.001$) per 0.09 ppm increase in ambient NMHC concentration. After controlling for SO_2 and CH_4, the adjusted HR was 1.10 (95% CI 1.06–1.15). Ambient NMHC controlling for the most important confounder of air pollutant SO_2 resulted in an adjusted HR of 1.36 (1.30–1.42, $p < 0.001$) (Table 3).

Table 4 presents the Cox proportional hazards regression analysis of the two targeted pollutant categories divided into three tertiles. The lowest tertile was used as the reference in each case, and the estimated HRs were adjusted for age, sex, lag 0–2, season, ambient temperature, level of urbanization, black-foot endemic region, and comorbidities. These results were consistent with those obtained from earlier multivariate analyses. Our research shows that the overall population exposure to the highest tertile (T3) of THC significantly increased the risk of UBC, with an adjusted HR (95% CI) of 3.16 (2.54–3.94; $p < 0.001$) (Table 4). The following results shows adjusted HRs associated with each tertile of airborne THC or NMHC exposure stratified by sex. For THC, the adjusted HRs of UBC for T3 were 3.16 (2.40–4.16; $p < 0.001$) for males and 3.11 (2.15–4.49; $p < 0.001$) for females. For NMHC, the adjusted HRs for T3 were 1.98 (1.57–2.49; $p < 0.001$) for males and 1.83 (1.33–2.51; $p < 0.001$) for females. When data regarding sex were stratified or merged for analysis, statistically significant correlations of adjusted HRs were measured for T2 and T3 compared with T1. The analysis demonstrated an association between the targeted pollutants and UBC risk in a dose-dependent manner but had no sex difference in the magnitude of risk.

Sensitivity analyses. To assess the sex-specific differences and the confounding effect of diabetes mellitus status on the association between the targeted pollutants and UBC cancer development, and to reveal any unexpected hidden relationships, we performed sensitivity analyses to compute causal effects only within the underlying strata. The results demonstrated no unexpected relationship, sex-specific differences, or ameliorating impact in non-diabetics (Supplementary Tables S3 and S4).

Before controlling for other pollutants, newly diagnosed UBC in males was significantly positively associated with the daily average concentration over the 10-year period for THC and NMHC with adjusted HR (95% CI) of 1.78 (1.68–1.88; $p < 0.001$) and 1.34 (1.27–1.41; $p < 0.001$) (Supplementary Table S3). Moreover, the hazards were numerically larger in females showing that the adjusted HR (95% CI) for developing incident UBC by THC and NMHC was 1.91 (1.78–2.06; $p < 0.001$) and 1.43 (1.33–1.53; $p < 0.001$). Among them, THC controlling for $PM_{2.5}$ resulted in adjusted HRs for males and females of 2.04 (1.92–2.17; $p < 0.001$) and 2.19 (2.02–2.37; $p < 0.001$), respectively. NMHC controlling for SO_2 resulted in adjusted HRs for males and females of 1.31 (1.24–1.38; $p < 0.001$) and 1.43 (1.33–1.54; $p < 0.001$), respectively.

Before controlling for other pollutants, newly diagnosed UBC in people with diabetes mellitus was significantly positively associated with the daily average concentration over the 10-year period for THC and NMHC, with adjusted HRs (95% CI of 1.99 [1.84–2.16; $p < 0.001$]) and 1.48 (1.36–1.60; $p < 0.001$), respectively (Supplementary Table S4). However, the magnitude of these risks was not significantly reduced in non-diabetics. Namely, the newly diagnosed UBC without diabetes mellitus for THC and NMHC with adjusted HR was 1.76 (1.67–1.86; $p < 0.001$) and 1.33 (1.27–1.40; $p < 0.001$), respectively. Among them, THC controlling for $PM_{2.5}$ resulted in adjusted HRs for people with and without diabetes mellitus of 2.34 (2.14–2.55; $p < 0.001$) and 2.00 (1.88–2.11; $p < 0.001$), respectively. NMHC controlling for SO_2 resulted in adjusted HRs for people with and without diabetes mellitus of 1.48 (1.37–1.61; $p < 0.001$) and 1.31 (1.24–1.38; $p < 0.001$), respectively.

Cumulative incidences of UBC compared between different tertiles. We observed slight changes in the effects of THC and NMHC after controlling for other pollutants, and the directions of the effect estimates did not change, suggesting that our findings were robust against potential confounders. The appropriateness of the Cox proportional hazards model is supported by the plot in the upper panel of Fig. 2, showing the log (−log [survival function]) versus survival time. Cumulative UBC incidence for the targeted pollutants was assessed using the Kaplan–Meier method (Fig. 2, lower panel), presenting a clear trend of increased UBC risk with increased exposure to each targeted pollutant. Over the entire follow-up period, when T1 exposure was set as a reference for comparison, the adjusted HRs by T2 and T3 exposure to THC was 1.23 (0.98–1.54) and 3.16 (2.54–3.94), respectively; it was 1.67 (1.41–1.97) and 1.94 (1.61–2.34) corresponding to NMHC T2/T3 exposure, respectively (Table 4). In addition, statistically significant differences in UBC occurrence were observed among the tertiles of the targeted pollutant categories (log-rank test; $p < 0.001$). A visualization summary to display the dose–response effect by escalated exposure in tertile is displayed in Fig. 3.

Discussion

This nationwide population-based cohort study linked national insurance claims data to open government data to investigate the association between long-term exposure to ambient air pollutants, THC or NMHC in Taiwan, and UBC risk. Our novel analysis shows a positive correlation between exposure to HC (THC or NMHC) in the ambient air for 10 years and UBC risk in people aged ≥ 20 years. Each additional unit of SD (0.13 ppm; 0.09 ppm) concentration of THC and NMHC increases the risk of bladder cancer by 83% and 37%, respectively. Furthermore, sensitivity analyses showed that these relationships did not change according to sex or presence or absence of diabetes. A large body of epidemiological evidence has indicated that diabetes is an independent risk factor for increased rates of heterogeneous types of cancer occurrence and death. The incidence and mortality of various types of cancer, including UBC, have a modest increase in patients with diabetes[25]. The magnitude of the risk of UBC mortality in people with diabetes in terms of the HR, after being adjusted for baseline age, smoking status, and body mass index, was increased by 40% when compared to non-diabetics[25]. We therefore performed
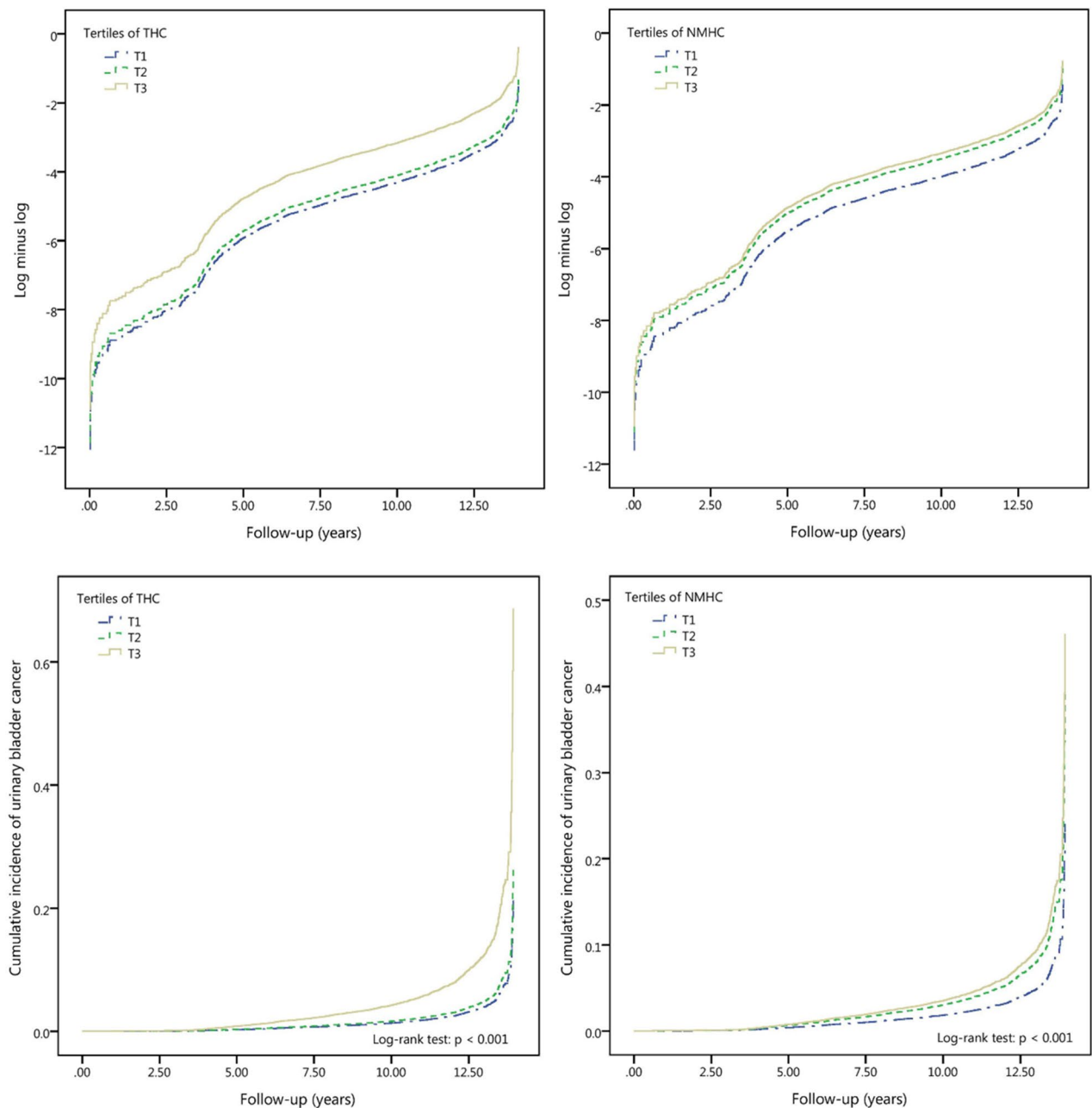
Figure 2. The log minus log survival plot by ambient air THC and NMHC pollutants, and the cumulative incidence of urinary bladder cancer in individuals among THC and NMHC pollutants' tertile. Upper panel: To evaluate the proportional hazards (PH) assumption involving the comparison of estimated $-\ln(-\ln)$ survival curves over different THC and NMHC tertiles, the plot of log ($-$log [survival function]) versus survival time in THC and NMHC air pollutants was constructed. The graphical approach showing parallel curves over time provides support of the PH assumption. Lower panel: Cumulative incidence of urinary bladder cancer for individuals among tertiles of THC and NMHC pollutants. The tertile values, in ppm (THC, NMHC), are as follows: THC (lowest tertile, T1: < 2.10; medium tertile, T2: ≥ 2.10 and < 2.25; and highest tertile, T3: ≥ 2.25); NMHC (lowest tertile, T1: < 0.26; medium tertile, T2: ≥ 0.26 and < 0.34; and highest tertile, T3: ≥ 0.34).

sensitivity analyses and observed a consistent effect on patients without diabetes: the association in our study remained after controlling for simultaneous exposure to other pollutants, particularly $PM_{2.5}$.

Many studies have demonstrated that air pollution adversely impacts health, and may result in problems including cancer; however, few have focused on specific pollutants in ambient air[2,14,15,20,23,26–28]. Most studies investigated the role of $PM_{2.5}$, NO_2, NOx, SO_2, and respirable elemental carbon as a proxy for diesel exhaust in the overall incidence of cancers, including UBC[2,14,23]. By contrast, our study presents a novel analysis of the association between the long-term exposure to ambient NMHC and THC and the risk of UBC. VOCs, including benzene and formaldehyde, in diesel engine exhaust emissions can be positively correlated with THC emissions, contributing to aggravated ground-level O_3 pollution when intense solar radiation and high temperatures and

8
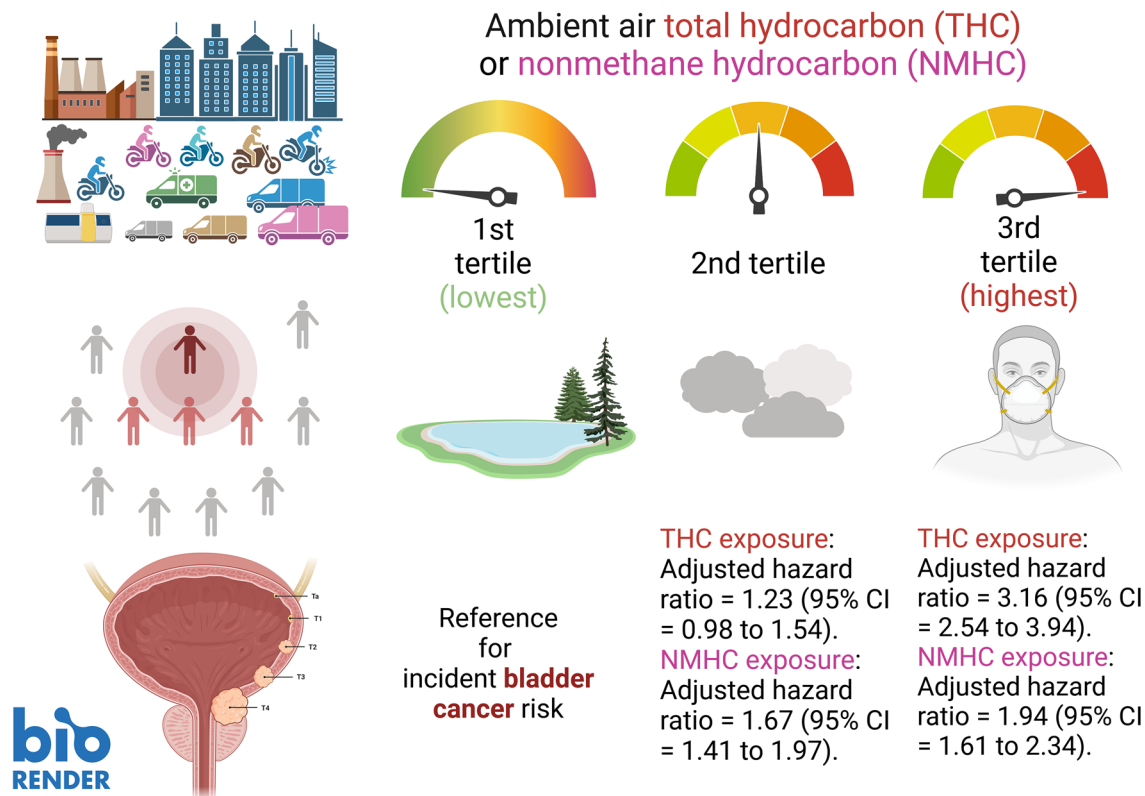
Figure 3. A schematic diagram summarizing the study results of the risk of developing urinary bladder cancer associated with long-term exposure to different tertiles of airborne THC or NMHC concentration. The magnitudes of risk show a significant exposure–response relationship.

low humidity occur. Our prespecified models demonstrated that after controlling for $PM_{2.5}$; health comorbidities; levels of urbanization, black-foot disease endemic region, season and ambient temperatures, each 0.13 ppm increase in THC concentrations would lead to a two-fold increase in the risk of incident UBC even in people without diabetes mellitus (adjusted HR 2.00; 95% CI 1.88–2.11). When replacing $PM_{2.5}$ with SO_2, an increase of 74% in the risk of developing UBC was still observed (adjusted HR 1.74; 95% CI 1.65–1.84).

A recent systematic review and meta-analysis study pointed out that petroleum industry work was associated with a modest increased risk of various cancers, including UBC (effect size = 1.25, 95% CI 1.09–1.43)[29]. A Spanish case–control study indicated that living more than 40 years in a city with more than 100,000 inhabitants was associated with an increased risk for UBC (OR = 1.30, 95% CI 1.04–1.63)[30]. Emissions of PAHs and diesel from industries near the residence, as evaluated by experts, were associated with an increased risk (OR = 1.29, 95% CI 0.85–1.98)[30]. In addition, previous research has shown that, for urban bus drivers and tramway employees who were employed for > 3 months, the risk of bladder cancer (standardized incidence ratio [SIR] = 1.4, 95% CI 1.2–1.6) was significantly increased[22]. The SIRs and 95% CIs for bladder cancer in road transportation workers compared with those in the whole population were 1.26 and 1.03–1.52, respectively[15]. Nevertheless, assuming occupational exposure, such as among road transportation workers, motor vehicle mechanics and repairers, garage mechanics, underground mines workers, and petroleum industry workers, as a proxy to ambient air pollution may be invalid since it is possible that the route of entry into the human body might have been via skin contact or even through the mucosal organs and not from inhalational hazard sources. Similarly, using petrol station density or annual industrial waste gas emissions to represent exposure to air pollution may also be groundless due to poor representation of specific pollutants. Thus, our study specifically targeted air pollutants such as THC and NMHC to offer new evidence on the impact of these precise components on the development of UBC.

In addition, we used a multiple-pollutant model to control the association between HC and the incidence of bladder cancer. We found that compared with the THC single-pollutant model, THC controlling for $PM_{2.5}$ has a higher risk of bladder cancer. Although previous studies used mortality rather than new cancer incidence as endpoints, they still showed that $PM_{2.5}$ has a significant deleterious effect on bladder cancer[17,18]. In an analysis of 623,048 ACS CPS-II participants in the United States, there was a significant adverse association between $PM_{2.5}$ and bladder cancer mortality (HR per 4.4 μg/m³, 1.13; 95% CI 1.03–1.23; N = 1324)[17]. Thus, it can be seen that in the multiple-pollutant model, $PM_{2.5}$ may increase the risk of target pollutants in bladder cancer.

Sufficient or convincing evidence about UBC carcinogenesis is associated with genetic susceptibility, cigarette smoking, diet, endemic *Schistosoma haematobium* infestation, and environmental pollution exposure[31]. However, the biological mechanism of the urinary tract carcinogenesis related to ambient air pollution remains unclear. Air pollution contains various mutagens and carcinogens, which may play a role in chronic systemic inflammation,

oxidative stress, and DNA damage in tissues other than the lungs[12,14,24,32,33]. Multiple lines of indirect evidence have shed light on the mechanistic explanation for ambient air THC and NMHC-related UBC carcinogenesis. The toxic potential of environmental pollutants can induce oxidative stress and inflammatory potential[32–35]. However, there is no evidence in animal studies for such a relationship between ambient hydrocarbon exposure and UBC, even in the case of the pollutant PAHs. Fundamentally, PAHs are local carcinogens, i.e., they trigger cancer development at the site of exposure, such as the lung with inhalation, skin with dermal application, or stomach and duodenum with oral ingestion, in animal models[36–38]. The biological plausibility of hydrocarbon pollutant-induced lower urinary tract carcinogenesis remains elusive in light of the research findings from studies with other sources of PAHs. Before we have more convincing and direct evidence for UBC's ambient air THC and NMHC-induced carcinogenesis, the canonical oxidative and inflammatory mechanisms and other mechanisms could be investigated in THC and NMHC carcinogenesis models in future research.

The strengths of this study are as follows. First, this is a nationwide study using a large population derived from the NHIRD, which contains the medical care data of 2296 million people (99% of Taiwan's population) under the National Health Insurance Plan. Second, this study is based on a 10-year long-term follow-up, which can provide adequate follow-up time to assess UBC development. Third, few epidemiological studies have assessed the association between airborne HC and UBC in Asia. Many studies have dealt with an increase in mortality, rather than cancer incidence. These endpoints are different in that the increase in mortality from UBC is due to the promoting effect of pollutants on cancer progression, whereas an increase in UBC incidence indicates the presence of the carcinogenic effect of air pollutants. Fourth, we considered and included the most important known risk factors for UBC in our calculation models and examined the possible synergistic effects between air pollutants. Finally, this study explored the association between HC and UBC through sex and diabetes stratification.

This linkage dataset cohort study has several limitations. First, because study subjects might relocate to a new place, exposure to air pollutants may change accordingly if the geographic location has a different level of pollutant concentration. The impact of this transition was difficult to measure. However, for most subjects in the cohort, a fixed postal address for one exposure time was reliable. Furthermore, the postal code-based study methodology does not consider that the individuals would be away from their residences, and their exposures at these times could be very different. Second, we could not measure occupational exposure to other hazardous chemicals in the workplace. Although we utilized the smoking-related diagnoses as a proxy for heavy smokers, this adjustment is considered limited in its ability to evaluate confounding by cigarette smoking. Thus, the impact of this likely confounding from cigarette smoking needs to be considered, a potential in this type of analysis that is not easily evaluated. Lastly, individual dietary habit records were not available to us. We were therefore unable to measure the possible exposure to food chain carcinogens in the current study. Although we do not believe that the lack of individual dietary habits and other lifestyles data had a large enough effect to significantly alter the risk of developing UBC in this large population-based cohort study, it may have caused possible biases, thereby affecting the association's estimates.

Conclusions

This ambispective cohort study offered new evidence to suggest that long-term exposure to THC and NMHC may be a risk factor for UBC. The results indicate a possible link between HC and UBC risk. Further, in a stratified analysis of the population in Taiwan by sex or diabetes status, long-term exposure to the two target pollutant categories was associated with an increased risk of UBC.

The leading cause and mechanism of the disease remain unclear. Therefore, the best prevention method currently available is to avoid dangerous factors as much as possible or reduce exposure to hazardous environments. Early detection and treatment are also important for prevention and treatment. With the global aging trend, the prevalence and burden of UBC may also increase. Acknowledging pollutant sources that harmful to health can provide information for risk management strategies and help decision-makers formulate more targeted air pollution regulations. These findings may have important public health implications for preventing UBC.

Methods

Ethics approval and consent to participate.　This study was approved after a full ethical review by the Institutional Review Board (IRB) of the China Medical University, Taichung, Taiwan (approval number: CMUH104-REC2-115 [CR-6]). In addition, because de-identified/anonymized data were used from the NHIRD, the IRB waived the requirement to obtain informed consent from the study participants. All experiments were performed according to confidentiality guidelines set forth by the Taiwan Personal Information Protection Act regulations. The entire study was conducted in accordance with the Declaration of Helsinki.

Data sources for linkage dataset ambispective cohort research.　Health data were obtained from the Longitudinal Health Insurance Database 2000 (LHID2000) within the NHIRD, including claims data for 1 million randomly selected individuals, from 1996 to 2013[39]. The NHIRD, established in 1996 in Taiwan, contains healthcare data of 22.96 million people (99% of Taiwan's population) under a universal health insurance program, including all claims data (ambulatory care claims and inpatient claims) and prescriptions dispensed at pharmacies, the registry for beneficiaries, registry for medical facilities, and registry for medical specialists. To establish demographic characteristics for research, patient-level information is gathered by linking these data files using the identification number of insured individuals. As recorded in the database, each individual's health and disease status was assigned an International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) until recently, when ICD-10-CM was implemented. To enhance the reliability of the NHIRD data, the observation period was set as 2000–2013.

In addition, the Environment Resource Datasets[40] are publicly available from open government data. This dataset was obtained by the Environmental Protection Administration of Taiwan, which determined ambient pollutants and temperatures at 76 monitoring stations across Taiwan, from 1993 to 2013. In this linkage database research, we used the postal code location as a proxy for the residence location from the NHIRD dataset and matched the postal code locations to the corresponding air quality monitoring stations in the Environmental Protection Administration (EPA) Open Dataset.

Study design and study population.   A nationwide linkage database ambispective cohort design was used for this study from January 1, 2000, to December 31, 2013. The selection of the study subjects is depicted in Fig. 1. Among the 1 million subjects in the LHID2000 database, individuals aged 20 years and above were enrolled on January 1, 2000 (n = 594,297). Those with missing or unknown records for sex and birth were excluded. Those with UBC (n = 351) and cancer from any other primary site (n = 10,415) diagnosed before the beginning of the study period; those with only one claim record during the study period (n = 304); and to avoid the reverse causation bias, those with outcome diagnosis made before July 2003 (n = 14,696) were excluded. Ultimately, 594,297 subjects were selected for further merging with the EPA dataset through postal codes linked to the location of the air quality monitoring stations. 5162 enrollees did not have postal code information or EPA monitoring station data, thus were excluded. After the linkage dataset merge, we finally tracked 589,135 study subjects for the research.

Exposure modeling.   We established from our research hypothesis to measure exposure to the targeted pollutants. Our devised exposure model for this study incorporated ambient concentrations of targeted pollutants over time while simultaneously addressing personal exposures tracked with the residential information and the duration of contact as input variables to estimate the cumulative individual exposure from inhalation. We have previously reported a similar exposure modeling approach which has drawn acceptance from the exposure and outcomes research community[41]. We determined the concentrations of 12 ambient air pollutants monitored by the EPA in Taiwan over a prespecified study period. The study targets were THCs and NMHCs. To examine the association between long-term exposure to targeted air pollutants and the development of newly diagnosed UBC, we measured the risk magnitude after controlling for other non-targeted pollutants over the exposure period. Non-targeted pollutants were included in the subsequent multiple-pollutant analyses. These were selected based on weak correlations (Pearson's correlation coefficients < 0.3) of target pollutants with 10 other monitored air pollutants: sulfur dioxide (SO_2); ozone (O_3); carbon monoxide (CO); carbon dioxide (CO_2); nitrogen oxides (NO_X); nitrogen monoxide (NO); nitrogen dioxide (NO_2); particulate matter < 10 μm in size (PM_{10}); particulate matter < 2.5 μm in size ($PM_{2.5}$); and methane (CH_4) (Supplementary Table S1). Daily air quality data were collected at 76 monitoring stations from July 1, 1993, to December 31, 2013, and maintained by the EPA[40]. The locations where air pollutants were recorded were selected to form an integrated geographic information system. Using this system, each study patient was linked to the appropriate monitoring region by postal code, and the change in residence was considered through insurance registration during the study period. A patient's long-term exposure to each air pollutant was defined as the cumulative concentration during the measurement period (i.e., 10 years before the survival date) averaged per day. Therefore, the long-term exposure to each air pollutant ($LEAP_{ij}$) (i = SO_2, O_3, CO, CO_2, NO_X, NO, NO_2, PM_{10}, $PM_{2.5}$, THC, NMHC, and CH_4) for a patient living in the region served by the air quality monitoring station j was calculated as follows[41,42]:

$$LEAP_{ij} = \frac{\sum_{t=m}^{n} AP_{ijt}}{d}$$

where AP_i is the ambient air pollution level for pollutant category i, m is the start date of the measurement period (10 years before the survival date), n is the end date of the measurement period (survival date), and d is the number of days in the measurement period.

This research also investigated the long-term trends of the airborne pollutants, THC and NMHC, employing the Mann–Kendall test to statistically assess whether there is a monotonic upward or downward trend of THC and NMHC over time. The test investigates the null hypothesis, H0, of no trend, indicating the observations are arbitrarily ordered in time, or the alternative hypothesis, H1, where either an increasing or decreasing monotonic trend is present. Sen's method (the Thell–Sen estimator) was used to estimate the slope of these trends[43].

Study outcomes.   From the included population, we identified people who received a first-time diagnosis of either invasive or in situ UBC during the study period, based on ICD-9-CM codes 188 for invasive carcinoma and 233.7 for urinary bladder carcinoma-in-situ, respectively. Individuals were considered to have UBC if they visited an outpatient clinic ≥ 3 times with a UBC diagnosis or had been hospitalized because of UBC[44]. The earliest hospitalization or outpatient visit with UBC diagnosis was assigned as the diagnosis date and served as the newly diagnosed date of UBC for all subsequent analyses. We defined survival (the expected duration of time until the outcome event) with an endpoint date of either UBC diagnosis, death, or December 31, 2013, the final observation date, whichever occurred first.

Comorbidities as confounding factors for UBC outcome were collected.   Information on comorbid conditions of patients was determined from the LHID2000 based on ICD-9-CM codes. The following comorbidities were considered essential: hypertension (401–405); chronic cystitis (595.1, 595.2); smoking-related diagnosis (305.1, 491.0, 491.2, 492.8, 496, 523.6, 989.84, V15.82, 649.0); alcohol use disorders (265.2, 291, 303, 305.0, 357.5, 425.5, 535.3, 571.0, 571.1, 571.2, 571.3, 980.0, V11.3); morbid obesity (278, 646.1, 649.1, 649.2,

V45.86, V65.3, V77.8)[47]; spinal cord injury (806, 952, 336.1)[47]; chronic liver disease (571, 572.2–572.9); diabetes mellitus (249, 250, 648.8, 648.0); gout (274); chronic kidney disease (403, 404, 582.9, 585, 646.2, 792.5, 996.1, 999); pesticide exposures (989.1, 989.2, 989.3, 989.4); and dyslipidemia (272). These were identified and defined according to the diagnostic history collected from at least three outpatient visits or a single hospital admission before the survival date.

Levels of urbanization and the historic black-foot disease endemic regions as confounders. Seven clusters of urbanization stratification were grouped into four levels: high, medium–high, medium, and low urbanization, according to the previously published consensus methodology[45]. Clusters of high bladder cancer incidence rate in the black-foot disease endemic regions included six southwestern coastal townships where people unknowingly drank arsenic-contaminated well water before tap water installation in 1979–2003[46–48]. The study subjects originating from these regions were identified and regarded as a confounder for adjustment in the multivariate Cox models.

Statistical analysis. The Chi-squared test (for categorical variables) and one-way analysis of variance (for continuous variables) were used to test for differences in demographic characteristics and distribution of comorbidities among tertiles of the targeted pollutant concentrations. UBC risk associated with each targeted pollutant category, expressed as hazard ratios (HRs) with 95% confidence intervals (CIs), was examined using Cox proportional hazards regression, considering potential confounders. To control the confounding effects of other pollutants, the possible link between air pollutants was used to assess the effects of multiple pollutants, by controlling others that were based on the selection of weak correlations with other air pollutants (i.e., the absolute value of the correlation coefficients between each of the two air pollutants was lower than 0.3; Supplementary Table S1). To avoid potential collinearity problems, we did not include pollutants with high correlations in the same regression model. The effect of each targeted pollutant on the risk of newly diagnosed UBC was estimated as the adjusted HR for the change in standard deviation (SD) over the follow-up period.

Local research has identified a V/U-shaped relationship between air pollutants and ambient temperature, showing significant effects at both ends of extreme temperatures in the region[49]. Therefore, to control the impact of weather conditions on air pollution and UBC, the ambient temperature should be one of the confounding factors in the pollutant models. Additionally, to control for short-term pollutant exposure effects, we used a lag of 0–2 days (average concentration levels on the same day of the UBC diagnosis, and one and two days before) for all air pollutants as one of the adjusting factors. Because air pollutant levels vary depending on the weather conditions, adjustment for the season is usually considered an important modifier in ambient air pollution-related biological effects in East Asia[19]. In the present study, multiple-pollutant models for two targeted pollutants were fitted, the independent effects of each targeted pollutant were adjusted for age, sex, comorbidities, level of urbanization, lag of 0–2 days, season (seasonal trends in UBC onset), and ambient temperature were estimated, and other pollutants that showed weak correlations were controlled. The concentration data of the targeted pollutants were divided into three tertiles, T1, T2, and T3, and adjusted HRs with 95% CIs were re-calculated.

Attributable risk proportion (ARP), as a percentage, estimates the proportion of UBC in the study population that is attributable to ambient air pollutants exposure. The incidence of exposure in the study population (not the entire general population) was estimated to calculate ARP. The calculation of ARP is as follows: From the exposure and outcome's 2 by 2 table, odds ratio (OR) = (a × d)/(b × c); study population exposure (SPe) = c/(c + d); finally, ARP (%) = 100 × (SPe × (OR-1))/(1 + (SPe × (OR-1))).

Sensitivity analyses examined whether the effects of pollutant categories differed between males and females. In addition, studies have pointed out that diabetes is related to a higher risk of UBC[50]; we decided to use diabetes stratification to explore whether the pollutant category would have a significant impact on the non-diabetic population. Kaplan–Meier analysis was used to determine the cumulative incidence of UBC, and the log-rank test was used to evaluate the difference among tertiles of concentrations of the target pollutants. The analyses were performed using the MetaTrial Platform and Statistical Product and Service Solutions (Version 22). All statistical tests were two-sided; p values of 0.05 were considered statistically significant.

Data availability
All data and related metadata underlying reported findings have been deposited in the public data repository: Mendeley Data (https://data.mendeley.com) with a digital object identifier (DOI) as https://doi.org/10.17632/gtpf8t5r9w.6 (direct link).

References
1. Turner, M. C. et al. Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. *CA Cancer J. Clin.* https://doi.org/10.3322/caac.21632 (2020).
2. Cohen, G. et al. Chronic exposure to traffic-related air pollution and cancer incidence among 10,000 patients undergoing percutaneous coronary interventions: A historical prospective study. *Eur. J. Prev. Cardiol.* **25**, 659–670. https://doi.org/10.1177/2047487318760892 (2018).
3. Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. & Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **525**, 367–371. https://doi.org/10.1038/nature15371 (2015).
4. Espina, C. et al. European Code against Cancer 4th Edition: Environment, occupation and cancer. *Cancer Epidemiol.* **39**(Suppl 1), S84–S92. https://doi.org/10.1016/j.canep.2015.03.017 (2015).

5. Guo, S., Yang, F., Tan, J. & Duan, J. Nonmethane hydrocarbons in ambient air of hazy and normal days in Foshan, South China. *Environ. Eng. Sci.* **29**, 262–269. https://doi.org/10.1089/ees.2010.0421 (2012).

6. Stewart, G. J. *et al.* Sources of non-methane hydrocarbons in surface air in Delhi, India. *Faraday Discuss.* **226**, 409–431. https://doi.org/10.1039/d0fd00087f (2021).

7. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. https://doi.org/10.3322/caac.21660 (2021).

8. Jhuang, J. R., Chiang, C. J., Su, S. Y., Yang, Y. W. & Lee, W. C. Reduction in the incidence of urological cancers after the ban on Chinese herbal products containing aristolochic acid: An interrupted time-series analysis. *Sci. Rep.* **9**, 19860. https://doi.org/10.1038/s41598-019-56394-y (2019).

9. Chang, H. C., Wang, M. C., Liao, H. C., Cheng, S. F. & Wang, Y. H. Hazard prevention regarding occupational accidents involving blue-collar foreign workers: A perspective of Taiwanese manpower agencies. *Int. J. Environ. Res. Public Health* **13**, 706. https://doi.org/10.3390/ijerph13070706 (2016).

10. Chen, T. H., Huang, J. J., Chang, F. C., Chang, Y. T. & Chuang, H. Y. Effect of workplace counseling interventions launched by workplace health promotion and tobacco control centers in Taiwan: An evaluation based on the Ottawa charter. *PLoS One* **11**, e0150710. https://doi.org/10.1371/journal.pone.0150710 (2016).

11. Hsiao, B. Y. *et al.* Ensemble forecasting of a continuously decreasing trend in bladder cancer incidence in Taiwan. *Sci. Rep.* **11**, 8373. https://doi.org/10.1038/s41598-021-87770-2 (2021).

12. Cohen, G. *et al.* Long-term exposure to traffic-related air pollution and cancer among survivors of myocardial infarction: A 20-year follow-up study. *Eur. J. Prev. Cardiol.* **24**, 92–102. https://doi.org/10.1177/2047487316669415 (2017).

13. Collarile, P. *et al.* Residence in proximity of a coal-oil-fired thermal power plant and risk of lung and bladder cancer in North-Eastern Italy. A population-based study: 1995–2009. *Int. J. Environ. Res. Public Health* **14**, 860. https://doi.org/10.3390/ijerph14080860 (2017).

14. Cong, X. Air pollution from industrial waste gas emissions is associated with cancer incidences in Shanghai, China. *Environ. Sci. Pollut. Res. Int.* **25**, 13067–13078. https://doi.org/10.1007/s11356-018-1538-9 (2018).

15. Lee, W., Kang, M. Y., Kim, J., Lim, S. S. & Yoon, J. H. Cancer risk in road transportation workers: A national representative cohort study with 600,000 person-years of follow-up. *Sci. Rep.* **10**, 11331. https://doi.org/10.1038/s41598-020-68242-5 (2020).

16. Tsai, S., Tiao, M. M., Kuo, H. W., Wu, T. N. & Yang, C. Y. Association of bladder cancer with residential exposure to petrochemical air pollutant emissions in Taiwan. *J. Toxicol. Environ. Health A* **72**, 53–59. https://doi.org/10.1080/15287390802476934 (2009).

17. Turner, M. C. *et al.* Ambient air pollution and cancer mortality in the Cancer Prevention Study II. *Environ. Health Perspect.* **125**, 087013. https://doi.org/10.1289/ehp1249 (2017).

18. Yeh, H. L. *et al.* Spatial analysis of ambient PM(2.5) exposure and bladder cancer mortality in Taiwan. *Int. J. Environ. Res. Public Health* **14**, 508. https://doi.org/10.3390/ijerph14050508 (2017).

19. Yoon, J. Y., Kwak, M. S., Kim, H. I. & Cha, J. M. Seasonal variations in the diagnosis of the top 10 cancers in Korea: A nationwide population-based study using a common data model. *J. Gastroenterol. Hepatol.* **36**, 3371–3380. https://doi.org/10.1111/jgh.15634 (2021).

20. Ho, C. K., Peng, C. Y. & Yang, C. Y. Traffic air pollution and risk of death from bladder cancer in Taiwan using petrol station density as a pollutant indicator. *J. Toxicol. Environ. Health A* **73**, 23–32. https://doi.org/10.1080/15287390903248869 (2010).

21. Pedersen, M. *et al.* Is there an association between ambient air pollution and bladder cancer incidence? Analysis of 15 European cohorts. *Eur. Urol. Focus* **4**, 113–120. https://doi.org/10.1016/j.euf.2016.11.008 (2018).

22. Soll-Johanning, H., Bach, E. & Jensen, S. S. Lung and bladder cancer among Danish urban bus drivers and tramway employees: A nested case–control study. *Occup. Med. (Lond)* **53**, 25–33. https://doi.org/10.1093/occmed/kqg004 (2003).

23. Turner, M. C. *et al.* Ambient air pollution and incident bladder cancer risk: Updated analysis of the Spanish Bladder Cancer Study. *Int. J. Cancer* **145**, 894–900. https://doi.org/10.1002/ijc.32136 (2019).

24. Castaño-Vinyals, G. *et al.* Bulky DNA adduct formation and risk of bladder cancer. *Cancer Epidemiol. Biomark. Prev.* **16**, 2155–2159. https://doi.org/10.1158/1055-9965.Epi-07-0184 (2007).

25. Rao Kondapally Seshasai, S. *et al.* Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N. Engl. J. Med.* **364**, 829–841. https://doi.org/10.1056/NEJMoa1008862 (2011).

26. Koutros, S. *et al.* Diesel exhaust and bladder cancer risk by pathologic stage and grade subtypes. *Environ. Int.* **135**, 105346. https://doi.org/10.1016/j.envint.2019.105346 (2020).

27. Latifovic, L. *et al.* Bladder cancer and occupational exposure to diesel and gasoline engine emissions among Canadian men. *Cancer Med.* **4**, 1948–1962. https://doi.org/10.1002/cam4.544 (2015).

28. Zare Sakhvidi, M. J., Lequy, E., Goldberg, M. & Jacquemin, B. Air pollution exposure and bladder, kidney and urinary tract cancer risk: A systematic review. *Environ. Pollut.* **267**, 115328. https://doi.org/10.1016/j.envpol.2020.115328 (2020).

29. Onyije, F. M., Hosseini, B., Togawa, K., Schüz, J. & Olsson, A. Cancer incidence and mortality among petroleum industry workers and residents living in oil producing communities: A systematic review and meta-analysis. *Int. J. Environ. Res. Public Health* **18**, 4343. https://doi.org/10.3390/ijerph18084343 (2021).

30. Castaño-Vinyals, G. *et al.* Air pollution and risk of urinary bladder cancer in a case–control study in Spain. *Occup. Environ. Med.* **65**, 56–60. https://doi.org/10.1136/oem.2007.034348 (2008).

31. Jahrreiss, V., Pradere, B., Laukhtina, E., Mori, K. & Shariat, S. F. Catalog of exogenous risk factors for bladder carcinogenesis. *Curr. Opin. Urol.* **30**, 449–456. https://doi.org/10.1097/mou.0000000000000753 (2020).

32. Boada, L. D. *et al.* Exposure to polycyclic aromatic hydrocarbons (PAHs) and bladder cancer: Evaluation from a gene-environment perspective in a hospital-based case–control study in the Canary Islands (Spain). *Int. J. Occup. Environ. Health* **21**, 23–30. https://doi.org/10.1179/2049396714y.0000000085 (2015).

33. de Oliveira Alves, N. *et al.* Inflammation response, oxidative stress and DNA damage caused by urban air pollution exposure increase in the lack of DNA repair XPC protein. *Environ. Int.* **145**, 106150. https://doi.org/10.1016/j.envint.2020.106150 (2020).

34. Borza, A. *et al.* Synergism of aromatic amines and benzo[a]pyrene in induction of Ah receptor-dependent genes. *Arch. Toxicol.* **82**, 973–980. https://doi.org/10.1007/s00204-008-0381-z (2008).

35. le Tuyen, H. *et al.* Aryl hydrocarbon receptor mediated activities in road dust from a metropolitan area, Hanoi-Vietnam: Contribution of polycyclic aromatic hydrocarbons (PAHs) and human risk assessment. *Sci. Total Environ.* **491–492**, 246–254. https://doi.org/10.1016/j.scitotenv.2014.01.086 (2014).

36. Kallistratos, G. & Fasske, E. Inhibition of benzo(a)pyrene carcinogenesis in rats with vitamin C. *J. Cancer Res. Clin. Oncol.* **97**, 91–96. https://doi.org/10.1007/bf00411283 (1980).

37. Cui, X. S., Torndal, U. B., Eriksson, L. C. & Möller, L. Early formation of DNA adducts compared with tumor formation in a long-term tumor study in rats after administration of 2-nitrofluorene. *Carcinogenesis* **16**, 2135–2141. https://doi.org/10.1093/carcin/16.9.2135 (1995).

38. Cavalieri, E. & Rogan, E. Role of radical cations in aromatic hydrocarbon carcinogenesis. *Environ. Health Perspect.* **64**, 69–84. https://doi.org/10.1289/ehp.856469 (1985).

39. Hsieh, C. Y. *et al.* Taiwan's National Health Insurance Research Database: Past and future. *Clin. Epidemiol.* **11**, 349–358. https://doi.org/10.2147/clep.S196293 (2019).

40. Wong, P. Y. *et al.* Using a land use regression model with machine learning to estimate ground level PM(2.5). *Environ. Pollut.* **277**, 116846. https://doi.org/10.1016/j.envpol.2021.116846 (2021).

41. Zhang, H. W. *et al.* Long-term exposure to ambient hydrocarbons increases dementia risk in people aged 50 years and above in Taiwan. *Curr. Alzheimer Res.* **16**, 1276–1289. https://doi.org/10.2174/1567205017666200103112443 (2019).
42. Zhang, H. W. *et al.* Long-term ambient hydrocarbons exposure and incidence of ischemic stroke. *PLoS One* **14**, e0225363. https://doi.org/10.1371/journal.pone.0225363 (2019).
43. Lorelei de Jesus, A. *et al.* Long-term trends in PM(2.5) mass and particle number concentrations in urban air: The impacts of mitigation measures and extreme events due to changing climates. *Environ. Pollut.* **263**, 114500. https://doi.org/10.1016/j.envpol.2020.114500 (2020).
44. Kok, V. C., Zhang, H. W., Lin, C. T., Huang, S. C. & Wu, M. F. Positive association between hypertension and urinary bladder cancer: Epidemiologic evidence involving 79,236 propensity score-matched individuals. *Ups J. Med. Sci.* **123**, 109–115. https://doi.org/10.1080/03009734.2018.1473534 (2018).
45. Chen, P. C. *et al.* A cohort study evaluating the risk of stroke associated with long-term exposure to ambient fine particulate matter in Taiwan. *Environ. Health* **21**, 43. https://doi.org/10.1186/s12940-022-00854-y (2022).
46. Chen, T. C., Hseu, Z. Y., Jean, J. S. & Chou, M. L. Association between arsenic and different-sized dissolved organic matter in the groundwater of black-foot disease area, Taiwan. *Chemosphere* **159**, 214–220. https://doi.org/10.1016/j.chemosphere.2016.06.007 (2016).
47. Chang, C. W. *et al.* Comparative analysis of patients with upper urinary tract urothelial carcinoma in black-foot disease endemic and non-endemic area. *BMC Cancer* **21**, 80. https://doi.org/10.1186/s12885-021-07799-4 (2021).
48. Lin, M. Y. *et al.* Incidence and survival variations of upper tract urothelial cancer in Taiwan (2001–2010). *Int. J. Urol.* **29**, 121–127. https://doi.org/10.1111/iju.14731 (2022).
49. Wu, D. W. *et al.* The impact of the synergistic effect of temperature and air pollutants on chronic lung diseases in subtropical Taiwan. *J. Pers. Med.* **11**, 819. https://doi.org/10.3390/jpm11080819 (2021).
50. Turati, F. *et al.* Diabetes mellitus and the risk of bladder cancer: An Italian case–control study. *Br. J. Cancer* **113**, 127–130. https://doi.org/10.1038/bjc.2015.178 (2015).

Competing interests
Han-Wei Zhang, Hsiao-Ching Peng reports equipment and statistical analysis were provided by Biomedica Corporation. Han-Wei Zhang reports a relationship with Biomedica Corporation in the form of board membership. Hsiao-Ching Peng reports a relationship with Biomedica Corporation in the form of employment. Both received support for using the product 'MetaTrial Platform' to conduct the present research. Both declare that they had full access to all the data in this study and take complete responsibility for the integrity of the data and the accuracy of the data analysis. The other authors declare that they have no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-25425-6.

Correspondence and requests for materials should be addressed to V.C.K.

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.