

UC San Diego

UC San Diego Previously Published Works

Title

Polysomnographic Assessment of Sleep Disturbances in Cancer Development A Historical Multicenter Clinical Cohort Study

Permalink

<https://escholarship.org/uc/item/3nr427t4>

Journal

CHEST Journal, 164(2)

ISSN

0012-3692

Authors

Kendzerska, Tetyana

Murray, Brian J

Gershon, Andrea S

et al.

Publication Date

2023-08-01

DOI

10.1016/j.chest.2023.03.006

Peer reviewed

Polysomnographic Assessment of Sleep Disturbances in Cancer Development

A Historical Multicenter Clinical Cohort Study



Tetyana Kendzerska, MD, PhD; Brian J. Murray, MD; Andrea S. Gershon, MD; Marcus Povitz, MD; Daniel I. McIsaac, MD, MPH; Gregory L. Bryson, MD; Robert Talarico, MSc; John Hilton, MD; Atul Malhotra, MD; Richard S. Leung, MD, PhD; and Mark I. Boullos, MD



BACKGROUND: Many cellular processes are controlled by sleep. Therefore, alterations in sleep might be expected to stress biological systems that could influence malignancy risk.

RESEARCH QUESTION: What is the association between polysomnographic measures of sleep disturbances and incident cancer, and what is the validity of cluster analysis in identifying polysomnography phenotypes?

STUDY DESIGN AND METHODS: We conducted a retrospective multicenter cohort study using linked clinical and provincial health administrative data on consecutive adults free of cancer at baseline with polysomnography data collected between 1994 and 2017 in four academic hospitals in Ontario, Canada. Cancer status was derived from registry records. Polysomnography phenotypes were identified by k-means cluster analysis. A combination of validation statistics and distinguishing polysomnographic features was used to select clusters. Cox cause-specific regressions were used to assess the relationship between identified clusters and incident cancer.

RESULTS: Among 29,907 individuals, 2,514 (8.4%) received a diagnosis of cancer over a median of 8.0 years (interquartile range, 4.2-13.5 years). Five clusters were identified: mild (mildly abnormal polysomnography findings), poor sleep, severe OSA or sleep fragmentation, severe desaturations, and periodic limb movements of sleep (PLMS). The associations between cancer and all clusters compared with the mild cluster were significant while controlling for clinic and year of polysomnography. When additionally controlling for age and sex, the effect remained significant only for PLMS (adjusted hazard ratio [aHR], 1.26; 95% CI, 1.06-1.50) and severe desaturations (aHR, 1.32; 95% CI, 1.04-1.66). Further controlling for confounders, the effect remained significant for PLMS, but was attenuated for severe desaturations.

INTERPRETATION: In a large cohort, we confirmed the importance of polysomnographic phenotypes and highlighted the role that PLMS and oxygenation desaturation may play in cancer. Using this study's findings, we also developed an Excel (Microsoft) spreadsheet (polysomnography cluster classifier) that can be used to validate the identified clusters on new data or to identify which cluster a patient belongs to.

TRIAL REGISTRY: ClinicalTrials.gov; Nos.: NCT03383354 and NCT03834792; URL: www.clinicaltrials.gov
CHEST 2023; 164(2):517-530

KEY WORDS: cluster analysis; incident cancer; sleep disturbances

FOR EDITORIAL COMMENT, SEE PAGE 287

ABBREVIATIONS: AHI = apnea-hypopnea index; aHR = adjusted hazard ratio; CHF = chronic heart failure; PLMS = periodic limb movements of sleep; REM = rapid eye movement; RLS = restless leg syndrome; SaO₂ = oxygen saturation; WASO = wakefulness after sleep onset

AFFILIATIONS: From the Department of Medicine, Faculty of Medicine (T. K. and J. H.), the Departments of Anesthesiology and Pain Medicine (D. I. M. and G. L. B.), The Ottawa Hospital/University of Ottawa, the Clinical Epidemiology Program (T. K., D. I. M., and

Take-home Points

Study Question: What is the association between polysomnographic measures of sleep disturbances and incident cancer, and what is the validity of cluster analysis in identifying polysomnographic phenotypes?

Results: Five clusters were identified: mild (mildly abnormal polysomnography findings), poor sleep, severe OSA or sleep fragmentation, severe desaturations, and periodic limb movements of sleep (PLMS). The associations between cancer and only PLMS and severe desaturations clusters compared with the mild cluster were significant while controlling for clinic and year of the polysomnography and demographics.

Interpretation: In a large cohort, we confirmed the importance of polysomnographic phenotypes and highlighted the role that PLMS and oxygenation desaturation may play in cancer. Using this study's findings, we also developed an Excel (Microsoft) spreadsheet (polysomnography cluster classifier) that can be used to validate the identified clusters on new data or to identify which cluster a patient belongs to.

Sleep is a fundamental health phenomenon with distinct and quantifiable biological, perceptual, behavioral, and temporal attributes.¹ Sleep impacts many behavioral and physiological processes that may influence cancer development, progression, recovery, quality of life, and morbidity and mortality.¹ Emerging evidence suggests that sleep-related disturbances such as sleep-disordered breathing, sleep fragmentation, abnormal sleep architecture, and disruption of circadian rhythms may be linked directly to cancer development and progression.² However, whereas most prior studies have

focused on measures of sleep-disordered breathing and nocturnal hypoxemia,³ potential evidence of an association between cancer risk with abnormal sleep architecture and sleep fragmentation has been inconsistent and limited.⁴ A meta-analysis suggested that although short sleep duration may increase cancer risk in Asians and long sleep duration may be associated with an increased risk of colorectal cancer, these findings were not consistent in a dose-response meta-analysis.⁴ Other common sleep disorders that have not been explored as extensively include restless legs syndrome, which is associated frequently with periodic limb movements of sleep (PLMS). These limb movements can reflect the various causes and effects of malignancy. For example, malignancy may be associated with iron loss and consequent restless legs syndrome with PLMS.⁵ Alternatively, periodic limb movements can occur with injury to a peripheral nerve,⁶ the spinal cord,⁷ or brain,⁸ all of which may be primary or metastatic sites for various forms of cancer. Despite the importance of sleep to cancer and the availability of numerous tools for measuring sleep quality and quantity, objective sleep measurements usually are underused in cancer studies. A need exists for robust and generalizable data regarding the association between sleep disturbances and cancer.

Polysomnography provides a rich set of sleep-related physiologic information obtained in routine clinical practice to inform clinical decision-making and to support research.⁹ Cluster analysis using polysomnography variables may help to identify associations between health outcomes and unique polysomnographic features or phenotypes beyond considering each variable separately in a prognostic model.¹⁰ The expected benefit of polysomnography phenotyping would be to allow for more directed diagnostic and management strategies, which may lead to improved health outcomes and better-designed clinical trials.¹¹ Prior work has demonstrated that cluster analysis may identify relationships not seen in traditional modeling; for example, it has been shown that phenotypic clusters identified using routine polysomnography data can capture the risk of cardiovascular outcomes otherwise missed by conventional OSA severity classifications.¹² However, only a few studies have examined the association between polysomnographic clusters and outcomes,^{10,12} suggesting limited reproducibility and stability of phenotypic clusters.¹¹ Thus, before using polysomnographic phenotypes in clinical practice, a need exists to develop reproducible clusters.¹¹

G. L. B.), Ottawa Hospital Research Institute, ICES (T. K., D. I. M., and R. T.), Ottawa; ICES, Toronto (A. S. G.); the Department of Medicine (B. J. M., A. S. G., R. S. L., and M. I. B.), the Division of Neurology (B. J. M. and M. I. B.), University of Toronto, St. Michael's Hospital (R. S. L.), the Hurvitz Brain Sciences Research Program (B. J. M. and M. I. B.), the Division of Respiriology (A. S. G.), Sunnybrook Health Sciences Centre, Toronto; the Department of Medicine (M. P.), Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, the Department of Medicine (M. P.), Cumming School of Medicine, University of Calgary, Calgary, AB, Canada; and the Department of Medicine (A. M.), University of California, San Diego, La Jolla, CA.

Drs Kendzerska and Murray contributed equally to this manuscript as co-first authors.

CORRESPONDENCE TO: Tetyana Kendzerska, MD, PhD; email: tkendzerska@toh.ca

Copyright © 2023 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2023.03.006>

In this study, we examined the association between polysomnographic measures of sleep disturbance and incident cancer controlling for known cancer risk factors and using four large clinical cohorts of patients with

suspected OSA who were free of any cancer at study entry. We also assessed the validity and reproducibility of a cluster analysis in identifying distinct polysomnographic phenotypes.

Study Design and Methods

Study Design

A historical cohort study was conducted using already linked provincial health administrative data and polysomnographic (sleep study) data from four large independent Ontario clinical sleep cohorts from 1991 through 2017. Participants were followed up from the diagnostic sleep study (index date) to the end of March 2020. The ethics committees of all institutions involved approved this study. ICES is an independent, nonprofit research institute whose legal status under Ontario's health information privacy law allows it to collect and analyze health-care and demographic data for health system evaluation and improvement.

Data Sources

Provincial Health Administrative Data: ICES has housed high-quality individual-level databases on publicly funded services in Ontario since 1988 (<https://datadictionary.ices.on.ca>).¹³⁻¹⁵ The databases used in this study are held securely in a de-identified form at ICES and were linked using unique encoded identifiers at ICES. Information on cancer status was derived from the Ontario Cancer Registry, which is based on multiple combined data sources to provide good-quality incidence data since 1964.¹⁶⁻¹⁹ Data regarding claims for positive airway pressure therapy was obtained through the Ontario Assistive Devices Program database.²⁰ The Registered Persons Database was used to derive vital status information.

Clinical Databases: Clinical databases included clinical and detailed polysomnography data and are described elsewhere²¹⁻²⁴: (1) St. Michael's Hospital Sleep Laboratory database, dating from 1994 through 2010 (Toronto); (2) the Sunnybrook Health Sciences Centre sleep database, dating from 2010 through 2015 (Toronto; [ClinicalTrials.gov](https://clinicaltrials.gov) Identifier: NCT03383354); (3) the London Health Sciences Centre Sleep and Apnea Assessment Unit polysomnography database, dating from 2007 through 2015 (London); and (4) The Ottawa Hospital Sleep Database, dating from 2015 through 2017 (Ottawa; [ClinicalTrials.gov](https://clinicaltrials.gov) Identifier: NCT03834792). Details are also provided in e-Table 1.

Populations of Interest

All consecutive adults who underwent a diagnostic sleep study (level 1 polysomnography) between 1994 and 2017 at four large academic hospitals (Ontario, Canada) listed previously with available detailed information from polysomnography and who were free of any cancer diagnosis at baseline were considered for inclusion. Details on cohort creation are provided in e-Figure 1. We previously used this cohort to assess the relationship between measures of OSA severity only and incident cancer.²⁴ For this study, we focused only on individuals with detailed information from polysomnography (not only on those with available information on OSA severity as assessed by the apnea-hypopnea index [AHI], as reported previously²⁴). This reduced the sample size of the cohort, but the median follow-up was extended from 7 to 8 years.

Outcomes

The primary outcome was the time from the index date to the incident cancer diagnosis. The following subgroups of suspected cancer causes were considered using previously used definitions²⁴: tobacco use

related, alcohol related, virus or immune related, and hormone related, as well as detectable vs nondetectable by screening. Details are provided in e-Table 2. Each individual was followed up until cancer diagnosis, death, emigration from Ontario, or the end of the follow-up period (March 31, 2020), whichever occurred first.

Exposures

All available polysomnography variables collected from at least three of the four clinical centers were considered as potential exposures (Table 1). Given that we previously found that the severity of OSA as measured by AHI and nocturnal hypoxemia (mean oxygen saturation [SaO₂] in sleep and percent of sleep time spent with SaO₂ < 90%) were associated independently with incident cancer, controlling for confounders,²⁴ we did not focus on these measures separately in this study. However, these measures were considered in the cluster analysis (details provided subsequently in the "Cluster Analysis" section).

Confounders and Risk Factors

A large number of potential confounders and risk factors available from clinical and health administrative databases were considered using a previously developed theoretical framework (also in e-Fig 2)²⁴: (1) from clinical databases at the date of the sleep study—age, sex, BMI,²⁶⁻²⁸ and self-reported tobacco use status^{29,30}; (2) from health administrative databases—separate prior comorbidities (chronic heart failure [CHF], COPD, hypertension, diabetes, obesity, and alcohol use disorder), outpatient or inpatient visits for depression or anxiety within 5 years before the index date, Charlson comorbidity index (to adjust for multiple comorbidities),³¹ demographics (neighborhood income, rural and immigrant status) at the date of the sleep study, prior health-care exposure to adjust for a potential detection bias (number of primary care office visits within 1 year before the index date), the date of the initial cancer diagnosis, cancer type, cancer stage, and OSA-related treatment in follow-up, such as initiation of positive airway pressure treatment or maxillomandibular advancement, uvulopalatopharyngoplasty, or bariatric surgical interventions (e-Table 3).³¹⁻³⁷

Analyses

Descriptive Statistics: Descriptive statistics were calculated to characterize the study population by cancer status and by clusters. We calculated standardized differences while comparing baseline characteristics between groups; a threshold of > 0.1 was used as an indicator of a meaningful difference between groups. Survival across clusters was estimated using the cumulative incidence function³⁸ for incident cancer and the Kaplan-Meier method for all-cause mortality and was compared between categories with Gray's test³⁹ and the log-rank test, respectively.

Cluster Analysis: For the primary analysis, instead of focusing on each polysomnographic measure separately, we used cluster analyses to identify clinically meaningful polysomnographic subgroups or phenotypes (e-Fig 3). Of 15 available polysomnography variables (Table 1), 13 polysomnography variables were considered in the cluster analysis because they were available in all four clinical centers. Rapid eye movement (REM) sleep latency and wakefulness after sleep onset (WASO) were available in only three clinical cohorts (e-Table 4); thus, they were excluded from the cluster analysis, but were considered in the secondary analysis as separate polysomnography indexes (see details in "Secondary Analyses")

TABLE 1 | Baseline Cohort Characteristics by the Total Sample and Polysomnographic Clusters

Cohort Characteristics ^a	Cluster 1: Severe OSA or Sleep Fragmentation (n = 1,824)	Cluster 2: PLMS (n = 1,010)	Cluster 3: Mild (n = 10,713)	Cluster 4: Severe Desaturation (n = 552)	Cluster 5: Poor Sleep (n = 2,415)	Total (N = 16,514)
Polysomnography variables						
TST, h	5.13 (4.24-5.92)	5.29 (4.48-6.02)	5.93 (5.33-6.52)	5.00 (3.85-5.86)	3.78 (2.88-4.53)	5.59 (4.71-6.30)
Sleep efficiency, %	73.5 (61.1-82.2)	76.0 (64.8-85.0)	85.1 (77.7-90.6)	72.3 (58.4-83.9)	54.9 (43.1-64.1)	80.5 (68.4-88.5)
No. of awakenings in TST	40.0 (29.0-56.0)	27.0 (19.0-36.0)	23.0 (17.0-30.0)	30.0 (20.0-42.0)	22.0 (15.0-31.0)	24.0 (17.0-33.0)
Arousal index/h	52.3 (39.4-68.5)	28.9 (15.1-44.7)	12.7 (6.9-20.8)	43.6 (19.2-75.8)	13.6 (6.8-24.4)	15.4 (7.9-28.6)
PLMI/h	0.0 (0.0-5.7)	60.5 (47.9-83.8)	0.0 (0.0-3.1)	0.0 (0.0-6.7)	0.0 (0.0-2.3)	0.0 (0.0-5.6)
PLMI ≥ 15	252 (13.8)	1,010 (100.0)	1,013 (9.5)	103 (18.7)	232 (9.6)	2,610 (15.8)
REM sleep, %	11.0 (6.0-15.6)	14.8 (9.8-19.6)	17.4 (13.2-21.4)	11.3 (5.4-16.7)	8.1 (1.8-13.2)	15.4 (10.1-20.1)
Stage 1, %	23.0 (13.8-38.6)	10.4 (6.6-16.2)	8.5 (5.3-13.1)	16.5 (7.9-31.0)	17.5 (10.7-26.2)	10.5 (6.2-17.6)
Sleep onset latency, min	11.0 (5.5-21.7)	13.5 (6.3-26.2)	10.0 (5.0-18.1)	11.6 (4.9-24.5)	45.5 (25.0-73.5)	12.2 (5.9-24.8)
REM latency, min	163.0 (101.8-249.8)	150.8 (98.8-233.5)	111.8 (79.5-169.0)	128.5 (87.0-220.5)	173.5 (108.8-255.5)	124.0 (83.0-191.5)
WASO, min	94.3 (59.7-141.51)	79.9 (50.6-120.5)	48.5 (28.0-77.4)	92.8 (51.2-149.6)	140.9 (93.5-190.2)	63.0 (34.5-107.0)
Apnea index, events/h	17.6 (5.2-36.7)	0.8 (0.0-4.1)	0.40 (0.0-2.5)	4.90 (0.3-35.8)	0.40 (0.0-2.7)	0.7 (0.0-4.5)
Hypopnea index, events/h	23.7 (11.2-38.6)	7.1 (1.5-15.8)	3.6 (0.5-10.6)	15.9 (5.0-36.7)	2.2 (0.2-9.3)	4.8 (0.6-14.2)
AHI, events/h	57.9 (43.6-74.6)	13.8 (5.7-25.6)	8.6 (2.8-18.3)	50.0 (16.6-83.7)	9.1 (2.4-20.5)	11.4 (3.6-26.8)
AHI categories²⁵						
Normal, < 5	11 (0.6)	222 (22.0)	3,836 (35.8)	40 (7.2)	888 (36.8)	4,997 (30.3)
Mild, 5-14.9	29 (1.6)	314 (31.1)	3,485 (32.5)	81 (14.7)	681 (28.2)	4,590 (27.8)
Moderate, 15-30	122 (6.7)	280 (27.7)	2,312 (21.6)	76 (13.8)	532 (22.0)	3,322 (20.1)
Severe, > 30/h	1,662 (91.1)	194 (19.2)	1,080 (10.1)	355 (64.3)	314 (13.0)	3,605 (21.8)
Time with SaO ₂ < 90%, %	4.7 (0.8-13.6)	0.10 (0.0-1.9)	0.0 (0.00-0.4)	67.8 (51.5-87.1)	0.0 (0.0-0.9)	0.1 (0.0-1.5)
Mean SaO ₂ , %	93.3 (92.0-94.7)	94.7 (93.3-95.8)	95.30 (94.0-96.3)	87.1 (84.7-89.0)	94.6 (93.0-95.9)	95.0 (93.4-96.1)
Self-reported symptoms						
Restless legs	504 (27.6)	491 (48.6)	2,864 (26.7)	169 (30.6)	699 (28.9)	4,727 (28.6)
Epworth sleepiness score/24	8.0 (5.0-12.0)	8.0 (4.0-11.0)	8.0 (5.0-12.0)	10.0 (5.0-13.0)	7.0 (4.0-11.0)	8.0 (5.0-12.0)
Demographics						
Age, y	55.0 (45.0-65.0)	59.0 (50.0-69.0)	46.0 (36.0-56.0)	58.0 (48.0-69.0)	55.0 (43.0-66.0)	50.0 (39.0-60.0)
Male sex	1,405 (77.0)	690 (68.3)	5,764 (53.8)	346 (62.7)	1,237 (51.2)	9,442 (57.2)

(Continued)

TABLE 1] (Continued)

Cohort Characteristics ^a	Cluster 1: Severe OSA or Sleep Fragmentation (n = 1,824)	Cluster 2: PLMS (n = 1,010)	Cluster 3: Mild (n = 10,713)	Cluster 4: Severe Desaturation (n = 552)	Cluster 5: Poor Sleep (n = 2,415)	Total (N = 16,514)
Rural location	42 (2.3)	51 (5.0)	339 (3.2)	18 (3.3)	106 (4.4)	556 (3.4)
BMI, kg/m ²	31.7 (27.9-36.6)	28.9 (25.4-32.4)	28.0 (24.8-32.1)	35.4 (30.6-41.7)	28.5 (25.0-33.4)	28.7 (25.2-33.3)
Current tobacco users (self-reported)	287 (15.7)	174 (17.2)	1,595 (14.9)	100 (18.1)	343 (14.2)	2,499 (15.1)
Neighborhood income quintile						
1	375 (20.6)	175 (17.3)	1,726 (16.1)	142 (25.7)	409 (16.9)	2,827 (17.1)
2	311 (17.1)	174 (17.2)	1,917 (17.9)	112 (20.3)	487 (20.2)	3,001 (18.2)
3	297 (16.3)	142 (14.1)	1,849 (17.3)	101 (18.3)	423 (17.5)	2,812 (17.0)
4	318 (17.4)	196 (19.4)	2,160 (20.2)	84 (15.2)	477 (19.8)	3,235 (19.6)
5	512 (28.1)	316 (31.3)	2,999 (28.0)	111 (20.1)	613 (25.4)	4,551 (27.6)
Prevalent comorbidities						
Alcohol misuse disorder	144 (7.9)	90 (8.9)	540 (5.0)	48 (8.7)	160 (6.6)	982 (5.9)
CHF	304 (16.7)	202 (20.0)	534 (5.0)	201 (36.4)	225 (9.3)	1,466 (8.9)
COPD	443 (24.3)	308 (30.5)	1,374 (12.8)	289 (52.4)	436 (18.1)	2,850 (17.3)
Diabetes	758 (41.6)	356 (35.2)	2,060 (19.2)	286 (51.8)	605 (25.1)	4,065 (24.6)
Hypertension	1,262 (69.2)	646 (64.0)	4,070 (38.0)	425 (77.0)	1,229 (50.9)	7,632 (46.2)
Outpatient or inpatient visits for depression or anxiety in the previous 5 y	901 (49.4)	580 (57.4)	5,958 (55.6)	294 (53.3)	1,461 (60.5)	9,194 (55.7)
Primary care visits in the previous year	6.0 (3.0-10.0)	6.0 (3.0-10.0)	4.00 (2.0-8.0)	7.0 (4.0-12.0)	5.0 (2.0-9.0)	5.0 (2.0-9.0)

Data are presented as No. (%) or median (interquartile range). AHI = apnea-hypopnea index; CHF = chronic heart failure; PLMI = periodic leg movement index; REM = rapid eye movement; SaO₂ = oxygen saturation; TST = total sleep time; WASO = wakefulness after sleep onset.

^aAll $P < .001$ (overall comparison χ^2 or one-way analysis of variance, as applicable, for every variable).

section). We used principal component analysis with a varimax rotation⁴⁰ to guide our decision on variable reduction (e-Table 5). We considered factors (ie, principal components) that explained at least 75% of the variance as important.¹² We reduced the number of variables for those redundant based on correlation analysis and clinical justifications. Specifically, because of the clinical and statistical redundancy between total sleep time and sleep efficiency (second principal component; $r > 0.88$), we decided to keep sleep efficiency only in the cluster analysis. AHI was chosen over the separate apnea and hypopnea indexes because it is the most used metric in clinical practice (first principal component; $r \geq 0.70$). Then, using the 10 selected variables, we used the k-means clustering algorithm to identify clusters.^{41,42} We computed and plotted select cluster validation statistics, such as the sum of squared error, the silhouette index, and the gap statistic with 500 bootstrap resamples, across a minimum cluster size of three and a maximum of nine (e-Fig 4).^{41,42} These were chosen to capture complexity beyond a binary variable without having an extreme number of groups. Because no consensus exists regarding which cluster validation statistics are most appropriate, we relied on clinical expertise to guide the final number of selected clusters and used the validation statistics to confirm the decision, rather than to derive the number of clusters independently. From the three to seven clusters identified using the 10 polysomnography variables tested, five were considered for the analysis: mild (mildly abnormal polysomnography findings; reference group for the statistical analysis), poor sleep, severe OSA or sleep fragmentation, severe desaturations, and PLMS (e-Figs 4, 5). We selected five over four clusters to differentiate between severe OSA without significant oxygen desaturations and severe desaturations groups, given different associated risks on adverse health outcomes.¹²

After clusters were identified, we evaluated their construct validity (how well the clusters measure a specific construct) and predictive validity (how well the clusters predict future relevant outcomes).⁴³ Specifically, to understand construct validity, we explored the distribution of clusters by polysomnography variables (e-Fig 5), sex, and age. To test predictive validity, we assessed the association between clusters and all-cause mortality. We selected mortality for validation as an independent outcome because we anticipated or hypothesized the directions of the associations between polysomnography measures,^{21,44,45} including polysomnographic phenotypes¹⁰ and all-cause mortality based on the current evidence. Finally, given that the k-means cluster algorithm is entirely data driven based on Euclidean distances between observations, the clusters are not directly reproducible and can not be derived or validated on new data. To

overcome this limitation, we fit a polytomous (multinomial) regression model with the outcome being the identified clusters and the exposures being the standardized polysomnography parameters that formed the cluster analysis. The coefficients from this model then were used to assign a given individual to a cluster based on the polysomnography parameter values. To reiterate, this means the five-cluster solution can be reproduced on new datasets or even in real time in a clinic setting using the model coefficients and input polysomnography parameter values (details are presented in the Data Supplement: Polysomnography Cluster Classifier).

Adjustment for Covariates: We used Cox cause-specific regressions, with death considered as a competing risk in incident cancer, to assess the relationship between identified clusters and incident cancer, controlling for the covariates described previously. For a variable selection process, we developed a theoretical framework published previously²⁴ (e-Fig 2).

Consistent with our previous study,²⁴ we entered covariates into the statistical models sequentially as follows: model 1, adjusted for the clinic site and year of sleep study, given potential differences in scoring criteria between sites and over time; model 2, additionally adjusted for age and sex; model 3, additionally adjusted for alcohol use disorder, CHF, COPD, depression or anxiety, hypertension, and diabetes at baseline; and model 4, additionally adjusted for OSA treatment considered as a time-varying covariate. In the sensitivity analyses, given missing values for BMI ($n = 363$ missing [1.2%]) and tobacco use status ($n = 14,613$ missing [48.9%]), we additionally adjusted for Charlson comorbidity index and BMI (model 5), as well as tobacco use status (model 6). Analyses were performed only on the clinical cohorts combined.

Secondary Analyses

In the secondary analyses, while investigating the effect of each polysomnography index separately using Cox cause-specific regressions, we additionally adjusted the statistical model for OSA severity as measured by AHI and time spent with $\text{SaO}_2 < 90\%$ (model 1A).

Complete case analysis was considered. The numbers of missing values for polysomnography variables are presented in e-Table 4. All statistical analyses were performed in the secure environment at ICES following Ontario privacy standards using SAS Enterprise Guide version 7.1., SAS version 9.4 software (SAS Institute, Inc.).

Results

Among 29,907 individuals considered for inclusion (median age, 50 years; 58.3% male), 2,514 individuals (8.4%) demonstrated any type of cancer over a median follow-up of 8.0 years (interquartile range, 4.2-13.5 years) (Table 1). Of these 2,514 individuals, 516 individuals (20.5%) demonstrated tobacco use-related cancer, 314 individuals (12.5%) demonstrated virus- or immune-related cancer, 103 individuals (4.1%) demonstrated alcohol-related cancer, and 787 individuals (31.3%) demonstrated hormone-related cancer (Table 2).

Individuals who received a cancer diagnosis were more likely to be older with higher BMI and prevalent CHF, COPD, diabetes, and hypertension; they were also more

likely to have reduced total sleep time, sleep efficiency, and mean SaO_2 and an increased number of awakenings, periodic leg movement index, WASO, AHI, and time spent with SaO_2 of $< 90\%$ (e-Table 6).

Cluster Analysis

Cohort characteristics of the final five clusters are presented in Table 1. Identified clusters demonstrated strong predictive validity: controlling for confounders, all clusters compared with the mild cluster were associated strongly with all-cause mortality (adjusted hazard ratio [aHR], > 1.40 in all models), with the greatest effect noted for the severe desaturations cluster (aHR, 2.34 in the final model; 95% CI, 1.86-2.94) (Fig 1, Table 3).

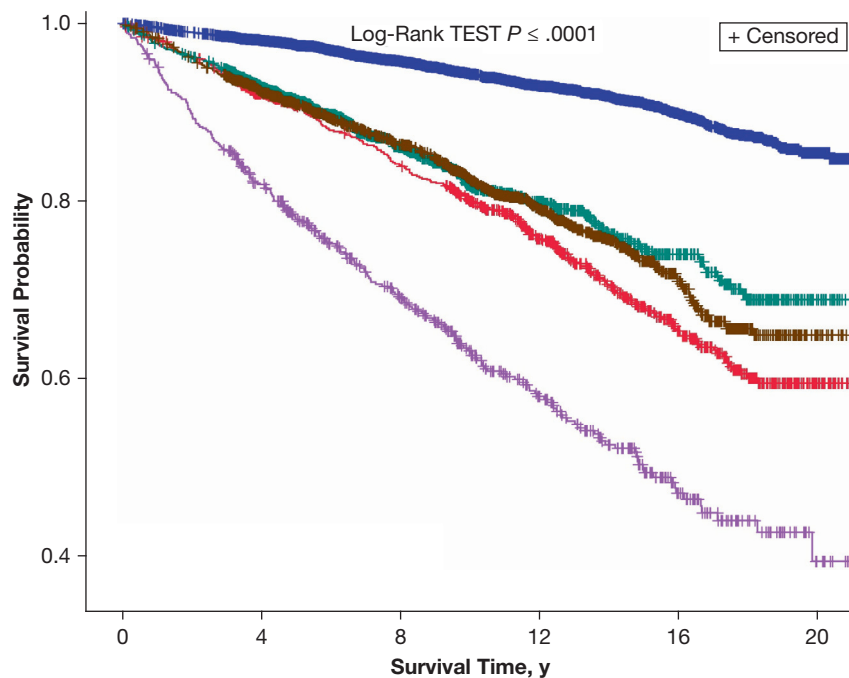
TABLE 2] Outcome Characteristics by the Total Sample and Polysomnographic Clusters

Outcome Characteristic	Cluster 1: Severe OSA or Sleep Fragmentation (n = 1,824)	Cluster 2: PLMS (n = 1,010)	Cluster 3: Mild (n = 10,713)	Cluster 4: Severe Desaturation (n = 552)	Cluster 5: Poor Sleep (n = 2,415)	Total (n = 16,514)
Cancer, total	221 (12.1)	175 (17.3)	788 (7.4)	80 (14.5)	184 (7.6)	1,448 (8.8)
Follow-up time, y ^a	9.91 (4.94-13.59)	10.16 (4.16-15.01)	8.49 (4.28-14.04)	7.96 (4.13-12.95)	4.82 (3.62-9.30)	7.96 (4.19-13.49)
Mean follow-up time, y ^a	9.64 ± 5.07	9.94 ± 5.91	9.31 ± 5.56	8.53 ± 5.54	6.82 ± 4.73	8.99 ± 5.49
Rate per 100 person-y (95% CI) ^b	1.26 (1.10-1.43)	1.74 (1.50-2.02)	0.79 (0.74-0.85)	1.70 (1.37-2.12)	1.12 (0.97-1.29)	0.98 (0.93-1.03)
Cancer subtypes						
Tobacco use related	43 (2.4)	39 (3.9)	145 (1.4)	30 (5.4)	37 (1.5)	294 (1.8)
Immune related	24 (1.3)	29 (2.9)	106 (1.0)	7 (1.3)	29 (1.2)	195 (1.2)
Alcohol related	10 (0.5)	8 (0.8)	31 (0.3)	6 (1.1)	12 (0.5)	67 (0.4)
Hormone related	66 (3.6)	52 (5.1)	282 (2.6)	14 (2.5)	51 (2.1)	465 (2.8)
Screening detected	88 (4.8)	71 (7.0)	302 (2.8)	19 (3.4)	65 (2.7)	545 (3.3)
All-cause mortality	280 (15.4)	181 (17.9)	450 (4.2)	176 (31.9)	266 (11.0)	1,353 (8.2)
Time from the index date to death, emigration from Ontario, or the end of the follow-up, y	10.33 (5.57-14.02)	11.46 (4.60-15.82)	9.28 (4.38-14.51)	8.63 (4.57-13.70)	4.91 (3.76-9.60)	8.85 (4.34-14.06)
Time from the index date to death, emigration from Ontario, or the end of the follow-up, y	10.14 ± 5.02	10.73 ± 5.83	9.65 ± 5.60	9.02 ± 5.56	7.13 ± 4.79	9.38 ± 5.53

Data are presented as No. (%), mean ± SD, or median (interquartile range), unless otherwise indicated. PLMS = periodic limb movements of sleep.

^aTime from the index date (the date of the sleep study) to cancer diagnosis, death, emigration from Ontario, or the end of the follow-up period (March 31, 2020), whichever occurred first.

^b $P < .0001$ (between-clusters difference).



	No. at patients at risk					
	0	4	8	12	16	20
— Mild/Normal Sleep	10,713	8,652	5,777	3,819	1,989	203
— PLMS	1,010	814	653	465	237	30
— Poor Sleep	2,414	1,701	778	423	209	23
— Severe OSA/Sleep Fragmentation	1,824	1,575	1,198	696	252	20
— Severe Oxygen Desaturation	552	432	293	175	75	10

Figure 1 – Unadjusted Kaplan-Meier survival curves for all-cause mortality stratified by polysomnographic clusters: cluster 1, severe OSA or sleep fragmentation; cluster 2, periodic limb movements of sleep; cluster 3, mild, representing a mildly abnormal polysomnography findings; cluster 4, severe desaturations; and cluster 5, poor sleep. PLMS = periodic limb movements of sleep.

Primary Analyses

The associations between incident cancer and all clusters compared with the mild cluster were significant in the univariate analyses and when controlling for the clinic site and year of the sleep study (Fig 2, Table 3).

However, when additionally controlling for age and sex, the effect remained significant only for the PLMS (aHR, 1.26; 95% CI, 1.06-1.50) and severe desaturations (aHR, 1.32; 95% CI, 1.04-1.66) clusters. The effect remained significant for the PLMS cluster, but was attenuated further for the severe desaturations cluster when controlling further for confounders.

Secondary Analyses

While investigating the association of each of the 15 polysomnographic measures separately in the statistical model (e-Figs 6-8, e-Table 7)—controlling for the clinic site, year of sleep study, age, and sex—lower sleep efficiency, reduced percentage of REM sleep and mean

SaO₂, increased AHI, time spent with SaO₂ < 90%, periodic leg movement index, and WASO were associated significantly with incident cancer. After controlling for other covariates, the effect was attenuated, but remained significant, in at least four models for a reduced percentage of REM sleep and increased WASO, AHI, periodic leg movement index, and time spent with SaO₂ < 90%.

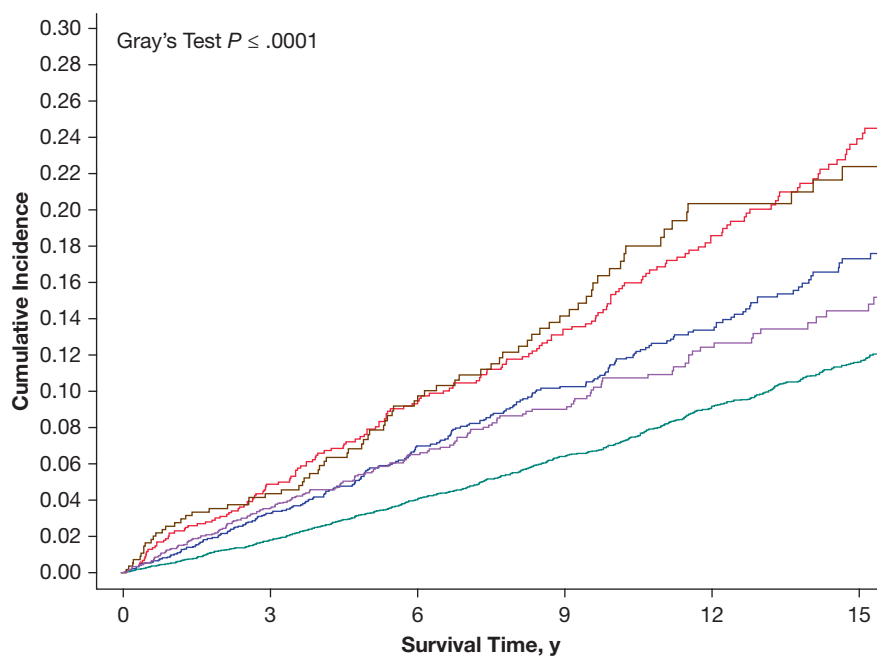
Estimates from the polytomous (multinomial) regression model, with the outcome being the identified clusters and the exposures being standardized polysomnographic parameters, are presented in e-Table 8. We also provided an Excel (Microsoft) spreadsheet (polysomnography cluster classifier) in which a user can input the polysomnography parameters of a given patient and receive an output that identifies which cluster the given patient belongs to. This model can be used to validate the clusters on new data or to identify which cluster a patient belongs to in real time.

TABLE 3] Association Between Each Cluster Considered Separately and Incident Cancer, Controlling for Confounders^a

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Incident cancer						
Cluster 1: severe OSA or sleep fragmentation	1.59 (1.37-1.85)	1.11 (0.95-1.30)	1.06 (0.90-1.24)	1.04 (0.89-1.22)	0.99 (0.84-1.16)	1.00 (0.84-1.19)
Cluster 2: PLMS	2.26 (1.91-2.66)	1.26 (1.06-1.50)	1.22 (1.03-1.45)	1.22 (1.03-1.45)	1.20 (1.01-1.43)	1.20 (1.00-1.44)
Cluster 4: severe desaturation	2.15 (1.70-2.70)	1.32 (1.04-1.66)	1.14 (0.90-1.45)	1.12 (0.88-1.43)	1.01 (0.78-1.29)	1.07 (0.82-1.41)
Cluster 5: poor sleep	1.42 (1.21-1.67)	1.04 (0.88-1.22)	1.01 (0.85-1.20)	1.01 (0.85-1.19)	0.99 (0.84-1.17)	1.00 (0.83-1.20)
Cluster 3: mild	Reference	Reference	Reference	Reference	Reference	Reference
All-cause mortality						
Cluster 1: severe OSA or sleep fragmentation	3.56 (3.06-4.13)	1.95 (1.67-2.29)	1.70 (1.45-1.99)	1.65 (1.40-1.93)	1.68 (1.43-1.98)	1.68 (1.40-2.00)
Cluster 2: PLMS	4.28 (3.59-5.10)	1.73 (1.44-2.07)	1.56 (1.30-1.87)	1.54 (1.28-1.85)	1.39 (1.16-1.68)	1.44 (1.18-1.75)
Cluster 4: severe desaturation	7.87 (6.60-9.37)	3.79 (3.16-4.53)	2.23 (1.85-2.70)	2.17 (1.80-2.63)	2.43 (1.99-2.98)	2.35 (1.87-2.96)
Cluster 5: poor sleep	3.24 (2.78-3.79)	1.96 (1.67-2.30)	1.81 (1.54-2.12)	1.81 (1.54-2.12)	1.85 (1.57-2.17)	1.74 (1.45-2.08)
Cluster 3: mild	Reference	Reference	Reference	Reference	Reference	Reference

Estimates are presented as hazard ratios (95% CI). Boldface values are those statistically significant with $P < .05$. PLMS = periodic limb movements of sleep.

^aWe entered covariates into the statistical models sequentially: model 1, adjusted for the clinic site and year of sleep study, given potential differences in scoring criteria between sites and over time; model 2, additionally adjusted for age and sex; model 3, additionally adjusted for alcohol misuse disorder, chronic heart failure, COPD, depression or anxiety, hypertension, and diabetes at baseline; model 4, additionally adjusted for OSA treatment considered as a time-varying covariate; model 5, additionally adjusted for Charlson comorbidity index and BMI; and model 6, additionally adjusted for tobacco use status.



Currently Event-Free	0	3	6	9	12	15
— Severe OSA/Sleep Fragmentation	1,824	1,654	1,267	1,055	636	310
— PLMS	1,010	902	638	580	407	254
— Mild/Normal Sleep	10,713	10,192	6,251	5,175	3,560	2,254
— Severe Oxygen Desaturation	552	457	322	246	159	94
— Poor Sleep	2,415	2,166	945	645	388	234

Figure 2 – Estimated cumulative incidence of cancer stratified by polysomnographic clusters: cluster 1, severe OSA or sleep fragmentation; cluster 2, periodic limb movements of sleep; cluster 3, mild, representing mildly abnormal polysomnography findings; cluster 4, severe desaturations; and cluster 5, poor sleep. PLMS = periodic limb movements of sleep.

Discussion

In a multicenter clinical cohort of adults with suspected OSA free of cancer diagnosis at baseline, our findings support the importance of identifying polysomnographic clusters and highlight the important role that PLMS and oxygenation desaturation may play in cancer-related outcomes. Poor-quality sleep was associated with increased malignancy risk in our study across a variety of measures, such as decreased sleep efficiency, increased percentage of stage N1 sleep, reduction in REM sleep, and increased WASO. Intrinsic sleep disorders such as sleep apnea may contribute to these effects because we saw associations with sleep apnea and oxygen desaturation. A novel finding was the potential association of periodic limb movements with malignancy.

Similar to two previous studies that used cluster analyses to investigate an association between polysomnographic features and health outcomes, we identified mild, poor sleep, severe OSA, hypoxemia, and PLMS clusters.^{10,12} In adjusted analyses from both previous studies,

hypoxemia and PLMS clusters were associated with the risk of cardiovascular events and mortality, supporting the potential clinical importance of these clusters.¹² Reproducibility and stability of the cluster analysis and clusters is an important limitation of previously published studies. Although in our study we neither found considerable benefits of the cluster analysis vs separate polysomnography variables in better prediction nor identified relationships not seen in traditional modeling compared with other studies,¹² the regression models presented in our study allow for external validation of the identified clusters on new data and facilitate calculating a patient's cluster in real time (including for individuals with no clear polysomnography pattern), as well as contribute to developing a reproducible approach for the cluster analysis.

A potential association of PLMS with cancer was noted. Although emerging evidence suggests that PLMS may be a prognostic factor for incident cardiovascular events and mortality,⁴⁵ links between PLMS and cancer have

not been established. One interpretation of our results is that PLMS may contribute to the pathogenesis of cancer. Inflammation is one of the mechanisms that may link PLMS with cancer.^{46,47} In a study of patients with restless leg syndrome (RLS), those with elevated PLM indexes (≥ 45 PLMS/h of sleep) showed more than a threefold chance of having an elevated C-reactive protein level, a marker of systemic inflammation, compared with patients with RLS with lower PLM indexes. In that study, the presence of RLS was not reported to be associated with increased levels of serum C-reactive protein, suggesting that the presence of PLMS was the main modulator of the elevated C-reactive protein levels.⁴⁶ Other than inflammation mechanisms such as oxidative stress,⁴⁸ sleep disturbance,⁴⁹ iron loss, and metastatic deposits⁵⁻⁸ also may play a role in linking PLMS with cancer. Furthermore, emerging evidence suggests a link between dopamine signalling and cancer,⁵⁰ which is important because PLMS are also known to be mediated strongly by dopaminergic factors.⁵¹

Alternatively, PLMS may be a marker of comorbidities that frequently co-occur in patients with cancer. PLMS are associated with sleep disorders such as RLS, which are associated with low iron levels that are common in some forms of cancer.⁵² PLMS also are linked with numerous medical conditions such as CHF and hypertension (controlled for in our analyses), as well as end-stage renal disease and liver disease⁵³ that may be present in some forms of cancer.

Although this work raised many further questions, it also identified practical testable sleep interventions that can be explored to reduce malignancy risk.^{54,55} It is hoped that attention to sleep and various polysomnography metrics (such as PLMS and oxygen levels) on a population level may help to mitigate cancer incidence. Given the growing recognition of PLMS as a separate polysomnographic phenotype^{10,12} and potential associations between PLMS and a higher risk of CVD, mortality,⁴⁵ and now cancer, a need exists for ambulatory approaches to detect PLMS⁵⁶ to understand better the pathophysiologic mechanisms through which PLMS may contribute to adverse health consequences. Future work will need to delineate which patient populations with PLMS are at higher risk of adverse health outcomes.

Our study had many strengths, including (1) long follow-up through the provincial health administrative data; (2) a wide range of polysomnography indexes from

different sleep centers across Ontario, which represented diverse populations contributing to generalizability of our findings; (3) access to a validated high-quality cancer registry to define outcomes; and (4) robust confounder control data.

The main limitation of our study was its retrospective and observational study design using data from clinical academic centers only, over different periods, which may limit its generalizability and may increase the potential for unmeasured confounding, although we have been able to replicate findings from other studies. Although our study was limited by the number of variables available in some clinical centers, this represents a real-world situation in which a different set of polysomnography variables are available.⁵⁷ Because no gold standard exists, some bias always accompanies selecting parameters for validation statistics; however, all-cause mortality has minimal misclassification bias and has been well studied in the sleep medicine literature. Finally, although the sample size was robust, we recognize that we still were underpowered for some subgroups, for example, specific tumor types. In addition, a likelihood exists that focusing only on those with detailed polysomnography data could bias our results through selection. However, baseline characteristics of the initial cohort,²⁴ of the cohort used in the study, and of included vs excluded individuals from the cluster analysis were comparable (e-Table 9), indicating that, if anything, our results are biased toward the null because of reduced sample size. Given the study objective to identify polysomnographic phenotypes, self-reported symptoms-related variables were not considered, but should be evaluated in future work.

Interpretation

In a large clinical cohort of adults with suspected OSA free of cancer at baseline, we confirmed the importance of identifying polysomnographic phenotypes and highlighted the important role that PLMS and oxygenation desaturation may play in cancer-related outcomes. An Excel spreadsheet (polysomnography cluster classifier) developed via our study allows for external validation of the identified clusters on new data and facilitates calculating a patient's cluster in real time (including for individuals with no clear polysomnography pattern) as well as contributes to developing a reproducible approach for future cluster analyses.

Funding/Support

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. This study also received funding from the Lung Health Foundation (also known as the Lung Association, Ontario), Grant-in-Aid, the American Thoracic Society Foundation [Unrestricted Grant], 2020 CHEST Foundation Research Grant in Sleep Medicine, the Ottawa Hospital Sleep Walk, and the University of Ottawa Department of Medicine Developmental Research Grant. Parts of this material are based on data and information compiled and provided by Ontario Ministry of Health, Ontario Health, the Canadian Institute for Health Information, Cancer Care Ontario (now known as Ontario Health), and Immigration, Refugees and Citizenship Canada's Permanent Resident Database. The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Financial/Nonfinancial Disclosures

The authors have reported to *CHEST* the following: M. P. was supported by the Academic Medical Association of Southwestern Ontario. M. P. also reports financial support for his research projects outside of the submitted work from Zennea Technologies, Jazz Pharmaceuticals, and Paladin Labs, and is a clinical consultant at the Medical Advisory

Board of Paladin Labs and Jazz Pharmaceuticals. A. M. reports income related to medical education from Livanova, Corvus, Jazz and Equillum. A. M. was supported by the [National Heart, Lung, and Blood Institute](#); ResMed provided a philanthropic donation to University of California, San Diego. T. K. was supported by the 2020 PSI Graham Farquharson Knowledge Translation Fellowship Award, received a speaker honorarium from AstraZeneca Canada, Inc., and is a clinical consultant at Pitolisant Medical Advisory Board (Paladin Labs, Inc.). T. K. also reports funding from the Canadian Institutes of Health Research outside of the submitted work. D. I. M. receives support from The Ottawa Hospital Anesthesia Alternate Funds Association and a Research Chair from the Faculty of Medicine, University of Ottawa. Outside of the submitted work, M. I. B. reports funding from the Canadian Institutes of Health Research, Canadian Partnership for Stroke Recovery, Alternative Funding Plan from the Academic Health Sciences Centres of Ontario, Ontario Genomics, and McLaughlin Centre for Molecular Medicine. He also reports consulting fees and honoraria from Jazz Pharmaceuticals, Paladin Labs, Eisai, and the Ontario MD Peer Leader Program; travel support from McGill University; and receipt of sleep equipment or research support from Braebon Medical Corporation, The Mahaffy Family Research Fund, the Dr. Robert Maggisano Research Fund, and Green Mountain. None declared (B. J. M., A. S. G., G. L. B., R. T., J. H., R. S. L.).

Acknowledgments

Author contributions: All authors were involved in the following: study conception and design, interpretation of data, revising the manuscript critically for the accuracy and important intellectual content, and final approval of the version to be published. M. P. is a custodian of the London Health Sciences Centre and Sleep Apnea Assessment Unit polysomnography database. R. S. L. is a custodian of the St. Michael's Hospital Sleep Laboratory database. M. I. B. and B. J. M. are custodians of the Sunnybrook Health Sciences Centre sleep database. T. K. is a custodian of The Ottawa Hospital Sleep Database. D. I. M. and G. L. B. are custodians of the Ottawa Hospital Surgical Sleep Database. T. K. additionally was involved in the literature search, dataset creation plan, data analyses, and drafting of the manuscript. R. T. and T. K. additionally were involved in data analyses. T. K. and R. T. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Role of sponsors: The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Research ethics and patient consent: The ICES (formerly Institute for Clinical Evaluative Sciences) is a prescribed entity under section 45 of Ontario's Personal Health Information Protection Act. Section 45 is the provision that enables analysis and compilation of statistical information related to the management, evaluation, and monitoring of, allocation of resources to, and planning for the health system. Section 45 authorizes health information custodians to disclose personal health information to a prescribed entity, like ICES, without consent for such purposes. Projects conducted wholly under section 45, by definition, do not require review by a Research Ethics Board. A confirming letter from the REB of Sunnybrook Health Sciences Centre is available on request. As a prescribed entity, ICES must submit to triannual review and

approval of its privacy and security policies, procedures, and practices by Ontario's Information and Privacy Commissioner. These include policies, practices, and procedures that require internal review and approval of every project by ICES' Privacy and Compliance Office. ICES was approved by the Commissioner for a fifth time in 2017.

Data availability: The dataset from this study is held securely in coded form at ICES. Although data sharing agreements prohibit ICES from making the dataset publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at www.ices.on.ca/DAS. The full dataset creation plan and underlying analytic code are available from the authors on request, understanding that the computer programs may rely on coding templates or macros that are unique to ICES and therefore are either inaccessible or may require modification.

Other contributions: The authors thank Victor Hoffstein, MD, for creating and

maintaining the St. Michael's Hospital Sleep Laboratory database and the research assistants at the London Health Sciences Centre, Nariko Kuwahara, MSc, and Sherry Coulson, MA, and at the Ottawa Hospital Research Institute, Moussa Meteb, MD, Aseel Ahmad, BSc, and Randa Ahmad, BSc, for their roles in databases preparation.

Additional information: The e-Figures and e-Tables are available online under "Supplementary Data."

References

- Redeker NS, Pigeon WR, Boudreau EA. Incorporating measures of sleep quality into cancer studies. *Supportive Care Cancer*. 2015;23(4):1145-1155.
- Owens RL, Gold KA, Gozal D, et al. Sleep and breathing . . . and cancer? *Cancer Prev Res (Phila)*. 2016;9(11):821-827.
- Tan BKJ, Teo YH, Tan NKW, et al. Association of obstructive sleep apnea and nocturnal hypoxemia with all-cancer incidence and mortality: a systematic review and meta-analysis. *J Clin Sleep Med*. 2022;18(5):1427-1440.
- Chen Y, Tan F, Wei L, et al. Sleep duration and the risk of cancer: a systematic review and meta-analysis including dose-response relationship. *BMC Cancer*. 2018;18(1):1149.
- Trenkwalder C, Allen R, Hogl B, et al. Comorbidities, treatment, and pathophysiology in restless legs syndrome. *Lancet Neurol*. 2018;17(11):994-1005.
- Esen R, Ediz L, Gulcu E, et al. Restless legs syndrome in multiple myeloma patients. *J Clin Med Res*. 2012;4(5):318-322.
- Lee MS, Choi YC, Lee SH, Lee SB. Sleep-related periodic leg movements associated with spinal cord lesions. *Mov Disord*. 1996;11(6):719-722.
- Ruppert E, Hacquard A, Tatu L, et al. Stroke-related restless legs syndrome: clinical and anatomo-functional characterization of an emerging entity. *Eur J Neurol*. 2022;29(4):1011-1016.
- Bianchi MT, Russo K, Gabbidon H, Smith T, Goparaju B, Westover MB. Big data in sleep medicine: prospects and pitfalls in phenotyping. *Nat Sci Sleep*. 2017;9:11-29.
- Kim JW, Won TB, Rhee CS, Park YM, Yoon IY, Cho SW. Polysomnographic phenotyping of obstructive sleep apnea and its implications in mortality in Korea. *Sci Rep*. 2020;10(1):13207.
- Zinchuk A, Yaggi HK. Phenotypic subtypes of OSA: a challenge and opportunity for precision medicine. *Chest*. 2020;157(2):403-420.
- Zinchuk AV, Jeon S, Koo BB, et al. Polysomnographic phenotypes and their cardiovascular implications in obstructive sleep apnoea. *Thorax*. 2018;73(5):472-480.
- Institute for Clinical Evaluative Sciences. Improving Health Care Data in Ontario. ICES Investigative Report. Institute for Clinical Evaluative Sciences; 2005.
- Juurlink DPC, Croxford R, Chong A, Austin P, Tu J, Laupacis A. Canadian Institute for Health Information Discharge Abstract Database: A Validation Study. Institute for Clinical Evaluative Sciences; 2006.
- Goel V; Canadian Medical Association, Institute for Clinical Evaluative Sciences in Ontario. *Patterns of Health Care in Ontario*. 2nd ed. Canadian Medical Association [for] the Institute for Clinical Evaluative Sciences in Ontario; 1996.
- Robles SC, Marrett LD, Clarke EA, Risch HA. An application of capture-recapture methods to the estimation of completeness of cancer registration. *J Clin Epidemiol*. 1988;41(5):495-501.
- McLaughlin JR, Kreiger N, Marrett LD, Holowaty EJ. Cancer incidence registration and trends in Ontario. *Eur J Cancer*. 1991;27(11):1520-1524.
- Hall S, Schulze K, Groome P, Mackillop W, Holowaty E. Using cancer registry data for survival studies: the example of the Ontario Cancer Registry. *J Clin Epidemiol*. 2006;59(1):67-76.
- Brenner DR, Tammemagi MC, Bull SB, Pinnaduwa D, Andrulis IL. Using cancer registry data: agreement in cause-of-death data between the Ontario Cancer Registry and a longitudinal study of breast cancer patients. *Chron Dis Canada*. 2009;30(1):16-19.
- Ministry of Health and Long-Term Care, Policies and Procedures Manual for the Assistive Devices Program, January, 2010, 1-121. Accessed March 22, 2023. https://www.health.gov.on.ca/en/pro/programs/adp/policies_procedures_manuals/docs/pp_adp_manual.pdf.
- Kendzerska T, Gershon AS, Hawker G, Leung RS, Tomlinson G. Obstructive sleep apnea and risk of cardiovascular events and all-cause mortality: a decade-long historical cohort study. *PLoS Med*. 2014;11(2):e1001599.
- Mclsaac DI, Gershon A, Wijesundera D, Bryson GL, Badner N, van Walraven C. Identifying obstructive sleep apnea in administrative data: a study of diagnostic accuracy. *Anesthesiology*. 2015;123(2):253-263.
- Kendzerska T, Gershon AS, Povitz M, et al. Polysomnographic markers of obstructive sleep apnea severity and cancer-related mortality: a large retrospective multicenter clinical cohort study. *Ann Am Thorac Soc*. 2022;19(5):807-818.
- Kendzerska T, Povitz M, Leung RS, et al. Obstructive sleep apnea and incident cancer: a large retrospective multicenter clinical cohort study. *Cancer Epidemiol Biomarkers Prev*. 2021;30(2):295-304.
- Fleetham J, Ayas N, Bradley D, et al. Canadian Thoracic Society guidelines: diagnosis and treatment of sleep disordered breathing in adults. *Can Respir J*. 2006;13(7):387-392.
- Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. *Nat Rev Endocrinol*. 2014;10(8):455-465.
- Taghizadeh N, Boezen HM, Schouten JP, Schroder CP, Elisabeth de Vries EG, Vonk JM. BMI and lifetime changes in BMI and cancer mortality risk. *PLoS One*. 2015;10(4):e0125261.
- Stone TW, McPherson M, Gail Darlington L. Obesity and cancer: existing and new hypotheses for a causal connection. *EBioMedicine*. 2018;30:14-28.
- Abrams JA, Lee PC, Port JL, Altorki NK, Neugut AI. Cigarette smoking and risk of lung metastasis from esophageal cancer. *Cancer Epidemiol Biomarkers Prev*. 2008;17(10):2707-2713.
- Zhang Q, Tang X, Zhang ZF, Velikina R, Shi S, Le AD. Nicotine induces hypoxia-inducible factor-1alpha expression in human lung cancer cells via nicotinic acetylcholine receptor-mediated signaling pathways. *Clin Cancer Res*. 2007;13(16):4686-4694.
- Sogaard M, Thomsen RW, Bossen KS, Sorensen HT, Norgaard M. The impact of comorbidity on cancer survival: a review. *Clin Epidemiol*. 2013;5(suppl 1):3-29.
- North CM, Christiani DC. Women and lung cancer: what is new? *Semin Thorac Cardiovasc Surg*. 2013;25(2):87-94.
- Parkin DM. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br J Cancer*. 2011;105(suppl 2):S2-S5.
- Parkin DM, Boyd L, Walker LC. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br J Cancer*. 2011;105(suppl 2):S77-S81.
- Stephenson GD, Rose DP. Breast cancer and obesity: an update. *Nutr Cancer*. 2003;45(1):1-16.
- Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med*. 2003;348(17):1625-1638.
- Hystad P, Carpiano RM, Demers PA, Johnson KC, Brauer M. Neighbourhood socioeconomic status and individual lung cancer risk: evaluating long-term exposure measures and mediating mechanisms. *Soc Sci Med*. 2013;97:95-103.
- Wolbers M, Koller MT, Witteman JC, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiology*. 2009;20(4):555-561.
- 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.
- Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374(2065):20150202.
- Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Series B*. 2001;63(2):411-423.
- Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for

- determining the relevant number of clusters in a data set. *J Stat Softw.* 2014;61(6):1-36.
43. de Vet HCW. *Measurement in Medicine: A Practical Guide.* Cambridge University Press; 2011.
 44. Kendzerska T, Mollayeva T, Gershon AS, Leung RS, Hawker G, Tomlinson G. Untreated obstructive sleep apnea and the risk for serious long-term adverse outcomes: a systematic review. *Sleep Med Rev.* 2014;18(1):49-59.
 45. Kendzerska T, Kamra M, Murray BJ, Boulos MI. Incident cardiovascular events and death in individuals with restless legs syndrome or periodic limb movements in sleep: a systematic review. *Sleep.* 2017;40(3).
 46. Trotti LM, Rye DB, De Staercke C, Hooper WC, Quyyumi A, Bliwise DL. Elevated C-reactive protein is associated with severe periodic leg movements of sleep in patients with restless legs syndrome. *Brain Behav Immun.* 2012;26(8):1239-1243.
 47. Murase K, Hitomi T, Hamada S, et al. The additive impact of periodic limb movements during sleep on inflammation in patients with obstructive sleep apnea. *Ann Am Thorac Soc.* 2014;11(3):375-382.
 48. Chen CY, Yu CC, Chen CL. Nocturnal periodic limb movements decrease antioxidant capacity in post-stroke women. *Acta Neurol Scand.* 2016;133(4):245-252.
 49. Hakim F, Wang Y, Zhang SX, et al. Fragmented sleep accelerates tumor growth and progression through recruitment of tumor-associated macrophages and TLR4 signaling. *Cancer Res.* 2014;74(5):1329-1337.
 50. Grant CE, Flis AL, Ryan BM. Understanding the role of dopamine in cancer: past, present and future. *Carcinogenesis.* 2022;43(6):517-527.
 51. Walters AS, Rye DB. Review of the relationship of restless legs syndrome and periodic limb movements in sleep to hypertension, heart disease, and stroke. *Sleep.* 2009;32(5):589-597.
 52. Ploug M, Kroijer R, Qvist N, Lindahl CH, Knudsen T. Iron deficiency in colorectal cancer patients: a cohort study on prevalence and associations. *Colorectal Dis.* 2021;23(4):853-859.
 53. Moretti R, Caruso P, Tecchiolli M, Gazzin S, Tiribelli C. Management of restless legs syndrome in chronic liver disease: a challenge for the correct diagnosis and therapy. *World J Hepatol.* 2018;10(3):379-387.
 54. Balachandran DD, Miller MA, Faiz SA, Yennurajalingam S, Innominato PF. Evaluation and management of sleep and circadian rhythm disturbance in cancer. *Curr Treat Options Oncol.* 2021;22(9):81.
 55. Johnson JA, Rash JA, Campbell TS, et al. A systematic review and meta-analysis of randomized controlled trials of cognitive behavior therapy for insomnia (CBT-I) in cancer survivors. *Sleep Med Rev.* 2016;27:20-28.
 56. Athavale Y, Krishnan S, Raissi A, et al. Actigraphic detection of periodic limb movements: development and validation of a potential device-independent algorithm. A proof of concept study. *Sleep.* 2019;42(9):zsz117.
 57. Boulos MI, Jairam T, Kendzerska T, Im J, Mekhael A, Murray BJ. Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. *Lancet Respir Med.* 2019;7(6):533-543.