

## Identifying Cases of Sleep Disorders through International Classification of Diseases (ICD) Codes in Administrative Data

Jolley, RJ<sup>1\*</sup>, Liang, Z<sup>1</sup>, Peng, M<sup>1</sup>, Pendharkar, SR<sup>1,2</sup>, Tsai, W<sup>1,2</sup>, Chen, G<sup>1</sup>, Eastwood, CA<sup>1</sup>, Quan, H<sup>1</sup>, and Ronksley, PE<sup>1</sup>

### Submission History

Submitted:	29/12/2017
Accepted:	21/03/2018
Published:	10/07/2018

<sup>1</sup>Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada  
<sup>2</sup>Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

### Abstract

#### Objectives

Prevalence, and associated morbidity and mortality of chronic sleep disorders have been limited to small cohort studies, however, administrative data may be used to provide representation of larger population estimates of disease. With no guidelines to inform the identification of cases of sleep disorders in administrative data, the objective of this study was to develop and validate a set of ICD-codes used to define sleep disorders including narcolepsy, insomnia, and obstructive sleep apnea (OSA) in administrative data.

#### Methods

A cohort of adult patients, with medical records reviewed by two independent board-certified sleep physicians from a sleep clinic in Calgary, Alberta between January 1, 2009 and December 31, 2011, was used as the reference standard. We developed a general ICD-coded case definition for sleep disorders which included conditions of narcolepsy, insomnia, and OSA using: 1) physician claims data, 2) inpatient visit data, 3) emergency department (ED) and ambulatory care data. We linked the reference standard data and administrative data to examine the validity of different case definitions, calculating estimates of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

#### Results

From a total of 1186 patients from the sleep clinic, 1045 (88.1%) were classified as sleep disorder positive, with 606 (51.1%) diagnosed with OSA, 407 (34.4%) with insomnia, and 59 (5.0%) with narcolepsy. The most frequently used ICD-9 codes were general codes of 307.4 (Nonorganic sleep disorder, unspecified), 780.5 (unspecified sleep disturbance) and ICD-10 codes of G47.8 (other sleep disorders), G47.9 (sleep disorder, unspecified). The best definition for identifying a sleep disorder was an ICD code (from physician claims) 2 years prior and 1 year post sleep clinic visit: sensitivity 79.2%, specificity 28.4%, PPV 89.1%, and NPV 15.6%. ICD codes from ED/ambulatory care data provided similar diagnostic performance when at least 2 codes appeared in a time period of 2 years prior and 1 year post sleep clinic visit: sensitivity 71.9%, specificity 54.6%, PPV 92.1%, and NPV 20.8%. The inpatient data yielded poor results in all tested ICD code combinations.

#### Conclusion

Sleep disorders in administrative data can be identified mainly through physician claims data and with some being determined through outpatient/ambulatory care data ICD codes, however these are poorly coded within inpatient data sources. This may be a function of how sleep disorders are diagnosed and/or reported by physicians in inpatient and outpatient settings within medical records. Future work to optimize administrative data case definitions through data linkage are needed.

## Background

Sleep disorders are common; however, prevalence estimates of different sleep disorders vary. It has been reported that 35 to 40% of the US adult population annually are affected by problems falling asleep or daytime sleepiness (1). The International Classification of Sleep Disorders 3rd Edition (ICSD-3)(2) has

identified over 80 different types of sleep disorders, the most common of which include obstructive sleep apnea (OSA), insomnia and narcolepsy. Despite the prevalence of sleep disorders, methodologies used for identifying such patients from administrative data are limited.

Sleep disorders can both be affected by disease and impact disease, with sleep deprivation and associated disorders hav-

\*Corresponding Author:

Email Address: [rjolley@ucalgary.ca](mailto:rjolley@ucalgary.ca) (RJ Jolley)

ing been shown to have large impacts on health (3). People with daytime sleepiness have a tendency to fall asleep in inappropriate places at inappropriate times, and have poor concentration which impacts daily functioning, work productivity and can even cause motor-vehicle accidents (4). Specifically, OSA has been associated with hypertension, heart failure, ischemic heart disease, arrhythmias, metabolic syndrome, pulmonary hypertension, stroke, and depression (5). Other sleep disorders such as insomnia have a high co-occurrence with psychiatric disorders (6). Recently the possible relationship between vaccinations and narcolepsy has been explored, with a recent reported increase in narcolepsy cases in Europe associated with large-scale population immunization campaigns such as the adjuvanted 2009 influenza A (H1N1) pandemic vaccines (7). In Alberta, Canada, the prevalence of sleep disorders, in particular OSA, is increasing mainly due to the aging population and rising rates of chronic obstructive pulmonary disease and obesity (8).

Population studies have revealed that sleep disorders are affecting many more people worldwide than previously understood. However, these population estimates vary. For instance, narcolepsy has been estimated in one study at 25 to 50 per 100000 when cataplexy is required and higher at 56 per 100000 when cataplexy is not required (9). Estimates of the prevalence of chronic insomnia vary depending on the criteria used and population studied but have been reported to be from 10% to as high as 40%(10-15). A Canadian study surveyed patients and found 29.9% reported symptoms of insomnia with 9.5% meeting criteria for insomnia syndrome (16). OSA has been estimated to affect at least 24% of men and 9% of women, however some studies have reported the prevalence of OSA associated with accompanying daytime sleepiness to be approximately 3 to 7% for adult men and 2 to 5% for adult women in the general population (17).

The differences in prevalence estimates across studies have been attributed to differences in sampling populations, varied techniques used for monitoring sleep and breathing, and variability in definitions used to determine cases (18). Specifically, many studies lacked sample size or have employed self-reported survey measures when defining sleep disorders (12). Population-based studies utilizing administrative health data to estimate incidence and prevalence are a less expensive and more practical means to study disease exposure, processes of care, and health outcomes at a population level. Having a clear understanding of population prevalence and incidence estimates of sleep disorders for instance is highly important particularly for health services planning and financing for the provision of local and national sleep-medicine services including timely access to specialist referral, diagnosis, treatment, and monitoring (8).

The World Health Organization created the International Classification of Diseases (ICD)(19) classification system, now one of the main coding systems used in administrative health data worldwide. However, when using administrative data, consistent and accurate coding algorithms or combinations of ICD codes to identify diseases are necessary as different coding algorithms can markedly affect the reported prevalence and incidence for a given disease (20). Studies have shown that case identification of chronic conditions (e.g. hypertension and diabetes) are valid and stable over time in Canadian administrative data when compared with clinical data sources (21-24).

Clinical documentation is considered the best available 'truth' or 'gold standard' for validating administrative data, especially when diagnoses are based on condition-specific testing, like sleep studies (25).

The objective of this study was to determine sleep disorders could be identified using ICD-coded administrative data. We developed and tested the accuracy of ICD codes for defining specific sleep disorders (including narcolepsy, insomnia, and OSA) in Alberta, Canada in a cohort of patients referred to a specialty sleep clinic, using ICD-coded physician claims, discharge abstract administrative data and ambulatory care data compared against a medical record review.

## Methods

### Study population and data sources

This study used chart review and survey data previously collected for a larger study examining the identification of OSA (26) as well as insomnia in sleep centre patients (27). The study population included all adults over the age of 18 years who completed an online questionnaire and underwent clinical assessment and/or sleep diagnostic testing at the Foothills Medical Centre Sleep Centre (FMC-SC) between January 1, 2009 and January 1, 2011. The FMC-SC is a publicly funded academic sleep centre that provides diagnostic testing and clinical services and is the only publicly funded tertiary referral sleep centre in Calgary, Alberta. They receive approximately 2500 referrals annually and over 5000 patient visits occur each year for diagnostic services including approximately 1000 polysomnography tests and over 2000 ambulatory tests. All newly referred patients filled out a questionnaire containing questions regarding occupation, sleep patterns, sleep aids, previous medical history and relevant clinical questions.

### Referral online questionnaires

Each patient completed an online questionnaire upon referral to the sleep clinic. The online questionnaire was composed of 108 questions that provide a comprehensive overview of the patient's demographics, anthropometrics, snoring history, daytime function, and medical history, as well as sleep schedule, behavior, and complaints. The questionnaire also included the Epworth Sleepiness Scale (ESS), a self-administered questionnaire used to determine daytime sleepiness (28), the Insomnia Severity Index (ISI) questionnaire (29), designed to assess the nature, severity and impact of insomnia, and a Patient Health Questionnaire (PHQ)(30) used to screen, diagnose, monitor and measure the severity of depression.

### Medical record review and determination of reference standard diagnosis (clinical sleep disorder diagnosis)

The reference standard diagnosis for each condition was determined from the identified cohort by two American Academy of Sleep Medicine board-certified physicians who independently reviewed all patient medical charts, assigning both a primary and a secondary diagnosis to each patient record in the cohort. Diagnoses were assigned and determined through a medical

record review which included collection of all clinical information including previous medical history, Sleep Centre testing results, as well as the patient-reported survey questionnaires as described above. The two physicians assigned a sleep disorder diagnosis based on the ICSD-2 criteria (most current version at the time of data collection). If there was evidence of a co-existing sleep disorder, a secondary diagnosis was also assigned by the reviewing physicians. If there were any disagreements between the physicians in assigning a diagnosis, the patient was not included in the analysis. For the reference standard diagnosis, no ICD-codes were assigned, only a clinical diagnostic category.

Diagnostic categories were selected prior to the chart review based on ICSD-2 criteria, and included insomnia, central nervous system (CNS) hypersomnolence (narcolepsy), OSA, and other (a category including all less common sleep disorders). A sleep clinic index visit date was defined as the date of the first sleep clinic visit as recorded by the medical record reviewer in the original data set, denoted as the index date.

Identification of insomnia was determined by taking into account the clinical history, the Insomnia Severity Index (ISI) questionnaire and any other testing that may have been performed (e.g. polysomnography). A diagnosis of CNS/hypersomnolence (narcolepsy) required a multiple sleep latency test (MSLT) with a mean sleep latency of < 8 minutes. When other causes for sleepiness were excluded, the definition of narcolepsy was based on a mean sleep latency of < 8 minutes and either: two sleep onset rapid eye movement episodes on the multiple sleep latency or a history of cataplexy (documented in the physician notes), along with excluding any other cause of sleepiness. An RDI > 5 events/hr was used to define OSA, while OSA syndrome required a diagnosis of OSA with symptoms. For the purpose of this study, we grouped cases of OSA for analysis. We considered a patient to have a diagnosis of a sleep disorder if the patient was assigned a primary or secondary diagnosis in the chart review for one of the following: insomnia, CNS/hypersomnolence (narcolepsy), OSA, and other (including diagnoses such as central sleep apnea, restless leg syndrome, upper airway resistance syndrome, and parasomnia).

### Administrative data sources

Three Alberta health administrative databases were linked with the chart review and survey data using the provincial personal healthcare number (PHN) as the unique identifier; 1) Practitioners' Claims (PC) database: The Alberta government is the sole payer of the submitted physician claims and maintains an electronic database for this purpose. Each claim includes the PHN, date of service, up to three diagnostic codes (*International Classification of Diseases, Ninth Revision, ICD-9*) and one service tariff code; 2) Hospital Discharge Abstract Database (DAD): The DAD includes all inpatient visits and captures information including administrative, clinical, and demographic information from every inpatient medical record for that visit at the time of discharge (31). Up to 25 diagnosis codes are assigned using the ICD, *Tenth Revision Canadian Edition* (ICD-10-CA) (32). The ICD-10-CA version expands on the ICD-9 codes by providing more specific sub-types of each condition listed (see supplemental tables 1 and 2 for comparison); 3) The National Ambulatory Care Reporting System

(NACRS) (formally known as the Ambulatory Care Classification System): The NACRS Database includes all data collected at the time of visit for hospital-based and community-based ambulatory care including day surgery, emergency department (ED) visits, and other community-based and outpatient clinics. NACRS also utilizes the ICD-10-CA.

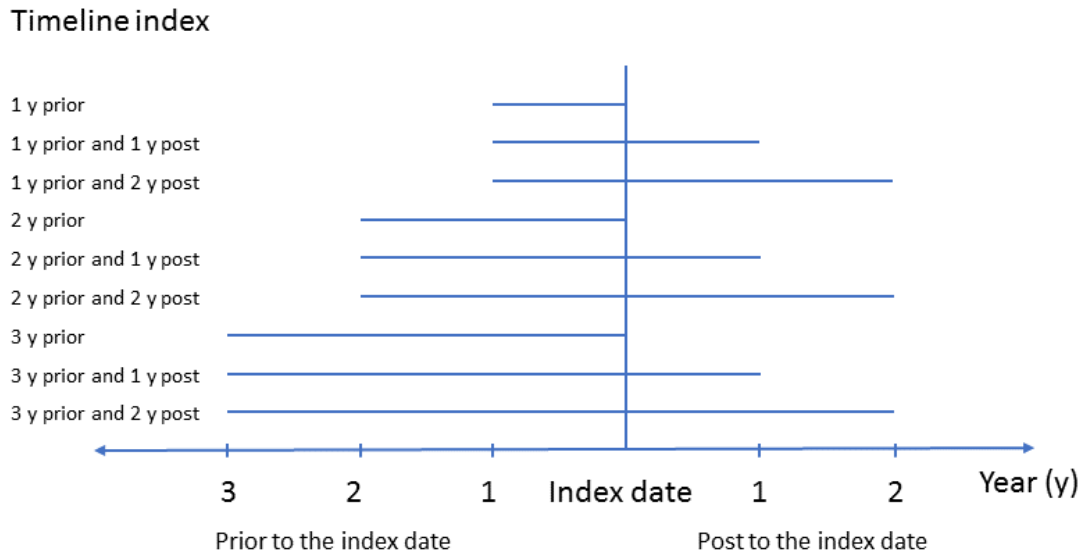
These data sources were selected to capture all sleep disorder diagnosis codes, as the majority of sleep disorder patients are managed or diagnosed at outpatient primary care or specialty sleep centre clinics. However, the diagnosis of a sleep disorder may also be recorded in an inpatient visit or outpatient setting. Therefore, the physician claims, inpatient discharge data, and any outpatient visits were used to validate an ICD coded case definition for sleep disorder identification. All data was accessed with local health research ethics board approval through the data holder provincial health authority -Alberta Health Services and with special permission to access physician claims data through the provincial government Ministry of Health – Alberta Health.

### Case definitions for sleep disorder using ICD coded administrative data

Using the PHN, all patients were linked to the administrative data sources as described above using data from a 7-year period (between 2006 and 2013). In order to capture the majority of potential coded cases, data were examined 3 years prior to and up to 2 years after the index visit as determined in the original study chart review and survey data collection occurring between January 1, 2009 and December 31, 2011. Patients were identified in administrative data as having a positive sleep disorder diagnosis if they had any one of the relevant ICD diagnostic codes (Supplementary Appendix 1 and 2) within any field of the administrative databases from PC (contains up to 3 diagnostic coding fields per visit; ICD-9 coded) DAD (contains up to 25 diagnosis coding fields; ICD-10-CA coded), and NACRS (contains up to 16 diagnosis fields; coded using ICD-10-CA). In addition to the diagnostic codes, we also included any procedure codes such as 8917 (polysomnogram); 8918 (other sleep disorder function tests); and a health service function code 03.19D (other non-operative measurements and examinations of nervous system and sense organs NEC Sleep polygraph studies for apnea and SIDS, interpretation). A diagnosis for selected sleep disorders (subtyped) included any relevant codes that differentiated that specific condition (Supplementary Appendix 1 and 2).

Case definitions for general sleep disorders (not subtyped) and subtyped sleep disorders including narcolepsy (CNS/hypersomnolence), OSA, and insomnia were constructed initially using varying combinations of (PC, NACRS and DAD data, as well as different time periods from up to 3 years prior to the recorded medical record review clinic date to 2 years following the index date, generating a total of 9 time points tested (Figure 1). The case definitions included detecting a code at just the index date, or at any time point prior or post index date.

Figure 1: Nine time points tested to identify the sleep disorder cases using ICD codes



## Statistical Analysis

Descriptive statistics were used to compare the clinical and demographic characteristics of patients stratified according to the presence or absence of clinically significant sleep disorder. To estimate the validity of the administrative case definitions, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each combination of ICD codes were calculated against the sleep-specific physician-reviewed database as the reference standard. All analyses were performed using SAS 9.4 (SAS Institute). Results are presented as mean ± standard deviation unless otherwise stated. For all statistical tests,  $P < 0.05$  was considered to be statistically significant.

## Results

A total of 1223 patients were included in analysis from the sleep study cohort with completed consent. From this cohort, 1200 patients had both physicians agree on the primary diagnosis, and 14 patients did not have an available PHN to link with administrative data. The final cohort included 1186 patients (Figure 2).

There was a total of 1045 (88.1%) patients diagnosed with a sleep disorder, while 141 (11.9%) did not have a diagnosis of a sleep disorder (Table 1). The mean age of all patients included was  $45.4 \pm 12.1$  years, 56.9% were men, and mean body mass index (BMI) was  $30.7 \pm 7.7$  (kg/m<sup>2</sup>), with significantly more men, and patients with elevated BMI having sleep disorders. The self-reported co-morbidities of asthma, chronic obstructive pulmonary disorder/ emphysema, depression, diabetes, kidney disease, myocardial infarction, previous heart failure and hypertension were similar across groups. The ESS mean total score was higher in those with a sleep disorder ( $11.3 \pm 5.7$ ) (ESS >10 is abnormal) versus those without a sleep disorder ( $9.9 \pm 5.1$ ) ( $p < 0.01$ ). The ISI mean total score was also higher in those with a sleep disorder ( $3.3 \pm 7.0$ ) versus those without ( $1.6 \pm 4.7$ ) ( $p < 0.01$ ). Patients without

a sleep disorder reported a higher total mean sleep time of  $7.3 \pm 1.4$  hours versus  $6.7 \pm 1.6$  hours ( $p < 0.0001$ ).

Table 2 summarizes the distribution of the diagnoses of sleep disorders within the cohort. Most of the patients were diagnosed with OSA (primary diagnosis 46.0%), followed by insomnia (primary diagnosis 28.2%), followed by CNS/hypersomnolence (narcolepsy) (primary diagnosis 4.8%) and restless leg syndrome (primary diagnosis 2.3%). Three ICD-9 codes were used for all sleep disorder diagnoses: 307.4 (non-organic sleep disorder 36.7%), 780.5 (unspecified sleep disturbance 61.0%) and 347.0 (narcolepsy 2.3%). Eight ICD-10-CA codes were primarily used including two general sleep disorder codes: G47.8 (other sleep disorders, includes: Kleine-Levin syndrome 91%) and G47.9 (sleep disorder, unspecified 7.1%). Six other more specific ICD-10-CA codes were used at varying frequencies: G47.0 (disorders of initiating and maintaining sleep insomnias 0.5%), G47.2 (disorders of the sleep-wake schedule 0.7%), G47.30 (sleep apnea, unspecified 0.6%), G47.4 (narcolepsy and cataplexy 0.1%), G47.31 (primary central sleep apnea 0.02%) and G25.8 (other specified extrapyramidal and movement disorders, including: restless leg syndrome and stiff-man syndrome 0.02%). No procedure codes (8917, 8918), or health service function code (03.19D) were recorded.

The performance of the ICD-coded case definitions are shown in Table 3 for the most valid ICD codes tested. The case definitions tested included any one of the ICD codes that indicated a sleep disorder (Supplemental Table 1 and 2). As there were so few codes used for each sleep condition, to facilitate analysis, we did not examine each type of sleep disorder individually. Out of a total of nine different administrative data case definition timing points tested (Figure 1), three of these gave the most optimal performance results (Table 3). The case definitions with the timing of 2 years prior and 1 year post index date gave the most optimal definitions versus the definitions with only 2 years prior but without 1 year post clinic visit. With at least 1 code appearing in PC data, captured patients ( $n=1186$ ) that could be linked, had sensi-

Figure 2: Study population inclusion flow diagram

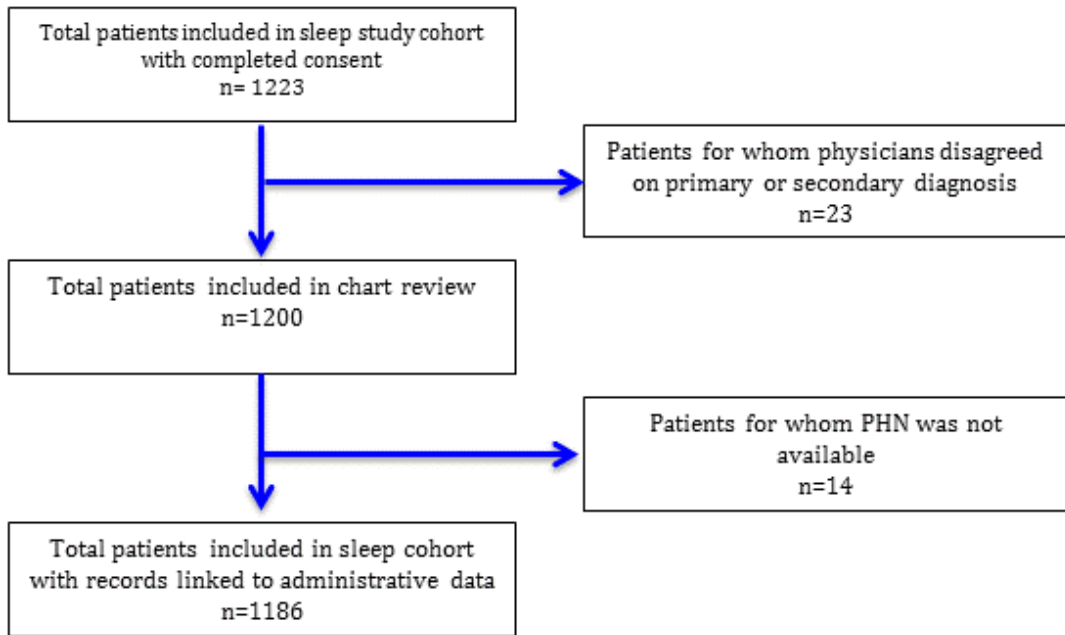


Table 1: Patient characteristics from the chart review

Characteristic	All Patients	Diagnoses of sleep disorder		P-value
	(n=1186)	Yes (n=1045)	No (n=141)	
Age (years)	45.4 ± 12.1	45.4 ± 12.0	45.6 ± 12.8	0.821
Sex (male), %	56.9	58.3	46.1	0.006
Body Mass Index (kg/m <sup>2</sup> )	30.7 ± 7.7	30.9 ± 7.6	29.5 ± 8.0	0.048
Co-morbidities:				
Arrhythmia/Atrial Fibrillation, %	3.3	3.2	4.3	0.493
Asthma, %	3.5	3.6	2.8	0.630
COPD/Emphysema, %	0.7	0.6	1.4	0.250
Depression, %	13.6	13.6	13.5	0.971
Diabetes, %	6.4	6.1	8.5	0.278
Kidney Disease, %	4.2	4.1	5.0	0.637
Myocardial Infarction (MI), %	0.8	0.9	0.7	0.853
Previous Heart Failure, %	0.7	0.7	0.7	0.957
Hypertension, %	7.0	7.1	6.4	0.760
Epworth Sleepiness Scale (ESS)	11.1 ± 5.7	11.3 ± 5.7	9.9 ± 5.1	0.006
Insomnia Severity Index (ISI) Total	3.1 ± 6.8	3.3 ± 7.0	1.6 ± 4.7	0.006
Sleep Latency (hours)	0.5 ± 0.7	0.5 ± 0.8	0.4 ± 0.4	0.112
Total Sleep Time (hours)	6.8 ± 1.6	6.7 ± 1.6	7.3 ± 1.4	<.0001

Data presented as mean ± standard deviation



Table 2: Frequency for selected sleep disorder diagnoses as either a primary or secondary diagnosis based on physician medical chart review

Condition	Diagnosis*		
	Total	Primary	Secondary
OSA	606 (51.1)	546 (46.0)	60 (5.1)
OSA/hypoventilation	28 (12.4)	27 (2.3)	1 (0.1)
Insomnia	407 (34.4)	334 (28.2)	73 (6.2)
CNS hypersomnolence	59 (5.0)	57 (4.8)	2 (0.2)
Restless leg syndrome	48 (4.1)	27 (2.3)	21 (1.8)
UARS	28 (2.4)	21 (1.8)	7 (0.6)
Parasomnia	14 (1.2)	7 (0.6)	7 (0.6)
CSA	8 (0.6)	4 (0.3)	4 (0.3)
Snoring	59 (5.0)	41 (3.5)	18 (1.5)
Depression	13 (1.3)	11 (0.9)	5 (0.4)
Fatigues	15 (1.3)	9 (0.8)	6 (0.5)
Fibromyalgia	10 (0.8)	6 (0.5)	4 (0.3)
Normal	49 (4.1)	49 (4.1)	NA
Other	115 (9.7)	47 (4.0)	68 (5.7)

Data presented as n (%). \*Primary diagnosis refers to the main diagnosis assigned by the physicians performing the medical record; Secondary diagnosis refers to the alternate diagnosis assigned by the physician if more than one was present; UARS - upper airway resistance syndrome; OSA - obstructive sleep apnea; CNS - central nervous system; CSA- central sleep apnea

tivity of 79.2%, specificity of 28.4%, PPV of 89.1% and NPV of 15.6% versus the 2 years prior with no codes appearing post index date with a lower sensitivity of 36.0%, specificity of 71.6%, PPV of 90.4% and NPV of 13.1% (Table 3).

The NACRS data also had optimal performance of a higher sensitivity and PPV when at least one ICD-code appeared in the 2 years prior and 1 year post index date with sensitivity of 85.4%, and PPV of 88.8%, however, lowered the specificity to 20.6%. When there was at least one code appearing in PC or one code appearing in NACRS, the Sn of the definition increased significantly to 90.8% and PPV of 88.9%, but again lowered the to specificity 15.6%, with a similar value of NPV of 18.6%. When adding at least 1 DAD code to that same definition, the increase was minimal, sensitivity increased by 0.1%, and the NPV increased by 0.2% while specificity and PPV remained the same.

## Discussion

Our study examined whether administrative data could be used to identify different sleep disorders. Overall, the quality of administrative data recording sleep disorders in our sample is moderate and varied by the data type and length of observation. While sleep disorders can be identified from a sleep related ICD code 2 years prior and 1 year post diagnosis; the ability to differentiate between specific sleep disorders was poor. The sensitivity was 78.7% for ICD definition of 1 ICD code recorded in PC or 1 ICD code recorded in DAD and increased to 90.9% when 1 ED/ambulatory care data record with a relevant code was added (definition of 1 PC or 1 DAD or 1 NACRS). The subtypes of sleep disorders are not captured well in these data.

Patients with sleep disorders are mainly treated at outpatient clinics and some severe patients are admitted to hospi-

tals. While defining cohorts of sleep disorders, both out and inpatient data are usually used. Our study found that the physician claims and ambulatory care data captured the majority of patients, while the inpatient hospital discharge data did not offer added value to the ICD definitions. Few cases were coded in the DAD (sensitivity and NPV were low). This is most likely due to the diagnosis of a sleep disorder occurring in the outpatient setting such as in a specialty sleep clinic.

Length of observation is related to ICD case definition validity. We found that both physician claims and ambulatory care data over the period of 2 years prior and 1 year post index date perform better than other time periods. The inclusion of a year following the initial sleep clinic visit allowed for an increased capture of patients with a recorded sleep diagnosis. This is most likely due to the fact that patients we studied were specifically referred to a sleep clinic and that the sleep clinic referral would confirm a diagnosis of a sleep disorder. As well, long wait times at the FMC Sleep Centre could lead to a long lag time between the time the referring physician coded a sleep disorder and when the sleep diagnosis was confirmed by the FMC sleep physicians. This along with the 6-month lag in time it takes for physician billing codes and professional coders to input data into the system could account for this lag in times of recordings and hence not being captured.

Most cases in this population had OSA and insomnia, with narcolepsy accounting for only approximately 5% of the total. However, most patients were coded with only one of four generic ICD-codes depending on the dataset: either ICD-9 codes 307.4 (nonorganic sleep disorder, unspecified), 780.5 (unspecified sleep disturbance), or ICD-10 codes of G47.8 (other sleep disorders; includes: Kleine-Levin syndrome), or G47.9 (sleep disorder, unspecified). Since the individual sleep disorders (i.e. OSA, insomnia and narcolepsy) could not be estimated separately, as there were so few specific ICD codes used by physicians, the discriminating value of the administra-

Table 3: Sensitivity, specificity, and predictive values for claims-based administrative ICD code combinations to identify sleep disorders (includes any ICD code identifying a sleep disorder)

Administrative data diagnosis definition	Timing of ICD code appearance											
	2 y prior to ID				1 y prior to ID				2 y prior and 1 y post ID			
	Sn	Sp	PPV	NPV	Sn	Sp	PPV	NPV	Sn	Sp	PPV	NPV
1 PC	36.0	71.6	90.4	13.1	33.3	73.0	90.2	12.9	79.2	28.4	89.1	15.6
1 DAD	1.7	98.6	90.0	11.9	1.4	98.6	88.2	11.9	2.3	98.6	92.3	12.0
1 PC or 1 DAD	36.6	71.2	89.3	14.5	34.0	71.8	88.8	14.1	79.7	28.8	88.1	17.7
2 PC or 1 DAD	15.3	85.9	87.8	13.3	13.1	90.4	90.0	13.6	43.0	66.7	89.5	15.1
3 PC or 1 DAD	8.6	92.3	88.1	13.3	6.2	94.2	87.7	13.2	23.8	82.1	89.7	14.0
1 NACRS	12.7	90.8	91.1	12.3	7.9	96.5	90.8	12.4	85.4	20.6	88.8	15.9
2 NACRS	8.2	95.7	93.5	12.3	7.9	96.5	94.3	12.4	71.9	54.6	92.1	20.8
3 NACRS	4.6	99.3	98.0	12.3	4.4	99.3	97.9	12.3	45.6	89.4	97.0	18.2
1 PC or 1 NACRS	38.3	68.8	90.1	13.1	35.7	70.5	88.9	14.2	90.8	15.6	88.9	18.6
1 PC or 1 NACRS or 1 DAD	38.9	68.9	89.1	14.5	36.3	69.2	88.6	14.1	91.0	15.4	87.7	20.5

Data presented as %. \*At least noted number of the relevant ICD diagnostic code appearance in the noted database; ID - Index date (Sleep clinic visit date, the date recorded in the medical record review as the date patient was consulted in the sleep clinic); y - year(s); PC - Practitioners' Claims Database (Physician Claims); DAD - Discharge Abstract Database (inpatient hospital discharge); NACRS - National Ambulatory Care Reporting System (emergency department/ambulatory care); Sn - Sensitivity; Sp - Specificity; PPV - Positive Predictive Value; NPV - Negative Predictive Value

tive data to identify sleep disorders is limited. The generic nature of the codes that are being used by physicians and coders is part of the challenge, therefore physician billing practices and the choice of alternative generic codes have an impact on capturing cases accurately. This may explain the sub-optimal sensitivity of specific sleep disorders.

This study has limitations. First, our sample is of patients who were referred to only one of a few sleep specialty clinics in the region. In limiting this to a single-centre, the sleep disorder prevalence is relatively high among this population, as only those patients with a suspected sleep disorder referral would have been included, which would impact the NPV and PPV as seen in our results. As well, the chart review involved self-reported questionnaire data, which may lead to potential self-report bias and may have overestimated the true prevalence of conditions such as OSA and insomnia. Although the extent of this bias cannot be assessed, those patients referred to the clinic not only completed the self-assessment questionnaire but also a clinical assessment with clinical sleep disorder testing, and therefore the medical record review would have mitigated any potential bias. Sleep disorders may also be potentially diagnosed and/or managed by primary care physicians. In this study we did not include primary care clinic patients, and therefore if applying these results to the wider population, generalizing the results to primary care clinic patients should be done with caution. As well, because our study utilized medical record review data from a previous study that examined specifically patients having a diagnosis of a sleep disorder as the gold standard, testing the ICD codes determined through this initial agreement study on another undefined cohort was not feasible, as the prevalence of the disease is low in the general population. Finally, due to the low prevalence of certain sleep disorders and physician ICD code submission patterns (only general codes used), we were unable to validate the subtypes of sleep disorders against particular ICD codes that

would match these patients.

Overall, this study determined that sleep disorders could be identified through physician billing codes and outpatient/ambulatory care data, yet generally were poorly coded in the inpatient database. Due to the generic nature of the codes, individual sleep disorders such as narcolepsy or insomnia could not be specifically identified. To study specific sleep disorders at the population level using administrative data, caution must be taken in only applying a ICD codes specific for each disorder.

## Statement on conflicts of interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## References

1. Hossain JL, Shapiro CM. The prevalence, cost implications, and management of sleep disorders: an overview. *Sleep Breath* 2002;6(2):85-102. <https://doi.org/10.1007/s11325-002-0085-1>
2. American Sleep Disorders Association. *The International Classification of Sleep Disorders- revised edition*. Rochester, MN: Davies; 1997.

3. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Chapter 3: Extent and Health Consequences of Chronic Sleep Loss and Sleep Disorders. Harvey R Colten and Bruce M Altevogt. Institute of Medicine (US) Committee on Sleep Medicine and Research. Washington (DC): National Academies Press (US); 2006. <https://doi.org/10.17226/11617>
4. Aldrich MS. Automobile sleep related accidents in patients with sleep disorders. *Sleep* 1989; 12:487-94.
5. Heinzer R, Vat S, Marques-Vidal P et al. Prevalence of Sleep disordered breathing in the general population: the HypnoLaus study. *Lancet Resp Med* 2015; 3:310-8. [https://doi.org/10.1016/s2213-2600\(15\)00043-0](https://doi.org/10.1016/s2213-2600(15)00043-0)
6. Ohayon MM. Prevalence of DSM-IV diagnostic criteria of insomnia: distinguishing insomnia related to mental disorders from sleep disorders. *J Psychiatr Res* 1997;31:333-46.
7. Johansen K, Brasseur D, MacDonald N, Nohynek H, Vandeputte J, Wood D, Neels P, Where are we in our understanding of the association between narcolepsy and one of the 2009 adjuvanted influenza A (H1N1) vaccines? *Biologicals* 2016;44(4):276-80. <https://doi.org/10.1016/j.biologicals.2016.04.007>
8. Level I and Level III sleep studies for the diagnosis of sleep disordered breathing (SDB) in adults. Final Report. Edmonton (AB): Prepared for Alberta Health and Wellness by the Health Technology & Policy Unit, School of Public Health, University of Alberta; 2013. Available: <https://open.alberta.ca/dataset/34ab89d3-320a-4951-af89-2aba5cc34e52/resource/700b3b71-011b-4553-9f17-9d9f375b6f50/download/AHTDP-Sleep-UofA-2013.pdf>.
9. Longstreth WT, The Epidemiology of narcolepsy. *Sleep* 2007;30(1):13-26.
10. Hatoum HT, Kania CM, Kong SX, Wong JM, Mendelson WB. Prevalence of insomnia: a survey of the enrollees at five managed care organizations. *Am J Manag Care* 1998;4:79-86.
11. Blais FC, Morin CM, Boisclair A, Grenier V, Guay B. Insomnia. Prevalence and treatment of patients in general practice [article in French]. *Can Fam Physician* 2001;47:759-67.
12. Hohagen F, Rink K, Kappler C, et al. Prevalence and treatment of insomnia in general practice. A longitudinal study. *Eur Arch Psychiatry Clin Neurosci* 1993;242:329-36.
13. Hohagen F, Kappler C, Schramm E, et al. Prevalence of insomnia in elderly general practice attenders and the current treatment modalities. *Acta Psychiatr Scand* 1994;90:102-8.
14. Simon GE, VonKorff M. Prevalence, burden, and treatment of insomnia in primary care. *Am J Psychiatr* 1997;154:1417-23. <https://doi.org/10.1176/ajp.154.10.1417>
15. Leger D, Partinen M, Hirshkowitz M, Chokroverty S, Hedner J. Characteristics of insomnia in a primary care setting: EQUINOX survey of 5293 insomniacs from 10 countries. *Sleep Med* 2010;11:987-98. <https://doi.org/10.1016/j.sleep.2010.04.019>
16. Morin CM, LeBlanc M, Bélanger L, Ivers H, Mérette C, Savard J. Prevalence of Insomnia and Its Treatment in Canada. *Can J Psych* 2011; 56(9):540-8. <https://doi.org/10.1177/070674371105600905>
17. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM; Increased Prevalence of Sleep-Disordered Breathing in Adults. *Am J Epidemiol* 2013 May 1;177(9):1006-14. <https://doi.org/10.1093/aje/kws342>
18. Punjabi NM. The Epidemiology of Adult Obstructive Sleep Apnea. *Proceedings of the American Thoracic Society* 2008;5(2):136-43. <https://doi.org/10.1513/pats.200709-155MG>
19. History of the Development of the ICD Taken from the World Wide Web. Available from: <http://www.who.int/classifications/icd/en/HistoryOfICD.pdf>.
20. O'Malley KJ, Cook KF, Price MD, Raiford Wildes K, Hurdle JF, Ashton CM. Measuring Diagnosis: ICD Code Accuracy. *Health Research Methods* 2005;40:1620-39. <https://doi.org/10.1111/j.1475-6773.2005.00444.x>
21. Robitaille C, Dai S, Waters, C, Loukine L, Bancej C, Quach S, Ellison J, Campbell N, Tu K, Reimer K, Walker R, Smith M, Blais C, Quan H. Diagnosed hypertension in Canada: Incidence, prevalence, and associated mortality. *CMAJ*. 2012 Jan 10; 184(1):E49-56.
22. Quan H, Khan N, Hemmelgarn BR, Tu K, Chen G, Campbell N, Hill MD, Ghali WA, McAllister FA. Hypertension Outcome and Surveillance Team of the Canadian a Education Programs. *Hypertension*. 2009 Dec;54(6):1423-8.
23. Jiang J, Southern DA, Beck C, James M, Lu M, Quan H. Validity of Canadian discharge abstract data for hypertension and diabetes from 2002 to 2013. *CMAJ Open*. 2016;4(4):E646-E653. <https://doi.org/10.9778/cmajo.20160128>.
24. Januel JM, Luthi JC, Quan H, Borst F, Taffe P, Ghali WA, Burnard B. Improved accuracy of co-morbidity coding over time after the introduction of ICD-10 administrative data. *BMC Health Services Res* 2011, 11:194.
25. Gordis L. *Epidemiology* (5<sup>th</sup> ed.). 2014. Saunders, and imprint of Elsevier. Baltimore, MD.
26. Severson CA, Pendharkar SR, Ronksley PE, Tsai WH. Use of electronic data and existing screening tools to identify clinically significant obstructive sleep apnea. *Can Respir J* 2015 Jul-Aug;22(4):215-20.



27. Severson CA, Tsai WH, Ronksley PE, Pendharkar SR. Identification of Insomnia in a Sleep Center Population Using Electronic Health Data Sources and the Insomnia Severity Index. *J Clin Sleep Med* 2013 Jul 15;9(7): 655–60. <https://doi.org/10.5664/jcsm.2830>
28. Johns MW. Daytime Sleepiness, Snoring and Obstructive Sleep Apnea. Epworth Sleepiness Scale. *Chest* 1993;103:30-6.
29. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. 2011 May 1;34(5):601-8.
30. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ*. 2012 Feb 21;184(3):E191-6. <https://doi.org/10.1503/cmaj.110829>
31. CIHI Discharge Abstract Database (DAD) Metadata. Available from: <https://www.cihi.ca/en/discharge-abstract-database-metadata>.
32. World Health Organization 1992 International Statistical Classification of Disease and Health Related Problems, 10th Revision (ICD-10). Geneva: World Health Organization.



# Supplementary Appendices

## Supplementary Appendix 1: ICD-9-CM codes used to define sleep disorders

ICD-9-CM	CODE DESCRIPTION
291.82	Alcohol induced sleep disorders
292.85	Drug induced sleep disorders
307.4	Nonorganic Sleep Disorder Unspecified
307.41	Transient Disorder of Initiating or Maintaining Sleep
307.42	Persistent Disorder of Initiating or Maintaining Sleep
307.45	Circadian Rhythm Sleep Disorder of Nonorganic Organic
307.46	Sleep Arousal Disorder
307.47	Other Dysfunctions of Sleep Stages or Arousal from Sleep
307.48	Repetitive intrusions of Sleep
307.49	Other Specific Disorders of sleep nonorganic origin
327.11	Idiopathic hypersomnia with long sleep time
327.12	Idiopathic hypersomnia without long sleep time
327.2	Organic Sleep apnea, Unspecified
327.21	Primary central sleep apnea
327.23	Obstructive sleep apnea (adult) (pediatric)
327.24	Idiopathic sleep related non-obstructive alveolar hypoventilation
327.26	Sleep related hypoventilation/hypoxemia in conditions classified elsewhere
327.27	Central sleep apnea in conditions classified elsewhere
327.29	Other organic sleep apnea
327.3	Circadian rhythm sleep disorder, unspecified
327.31	Circadian rhythm sleep disorder, delayed sleep phase type
327.32	Circadian rhythm sleep disorder, advanced sleep type
327.33	Circadian rhythm sleep disorder, Irregular sleep-wake type
327.34	Circadian rhythm sleep disorder, free running type
327.35	Circadian rhythm sleep disorder, jet lag type
327.36	Circadian rhythm sleep disorder, Shift work type
327.37	Circadian rhythm sleep disorder in conditions classified elsewhere
327.39	Other Circadian rhythm sleep disorder
327.42	REM Sleep behaviour disorder
327.43	Recurrent Isolated Sleep Paralysis
327.52	Sleep Related Leg Cramps
327.53	Sleep Related Bruxism
327.59	Other Organic Sleep Related Movement Disorders
327.8	Other organic sleep disorders
333.94	Restless Leg Syndrome
347.0	Narcolepsy
347.00	Narcolepsy without Cataplexy
347.01	Narcolepsy with Cataplexy
347.10	Narcolepsy without Cataplexy and other disorder
347.11	Narcolepsy with Cataplexy and other disorder
780.5	Unspecified sleep disturbance
780.51	Insomnia with sleep apnea, unspecified
780.53	Hypersomnia with sleep apnea, unspecified
780.55	Disruption of 24 hour sleep wake cycle, unspecified
780.56	Dysfunctions associated with sleep stages or arousal from sleep
780.57	Unspecified sleep apnea
780.58	Sleep related movement disorder, unspecified
780.59	Other sleep disturbances
8917	Polysomnogram (PSG)
8918	Other Sleep Disorder Function Tests
V69.4	Lack of adequate sleep



## Supplementary Appendix 2: ICD-10-CA codes used to define sleep disorders

ICD-10-CA	CODE DESCRIPTION
F51	Nonorganic sleep disorders
F51.0	Nonorganic insomnia
F51.1	Nonorganic hypersomnia
F51.11	Primary hypersomnia
F51.13	Hypersomnia due to other mental disorder
F51.19	Other hypersomnia not due to a substance or known physiological condition
F51.2	Nonorganic disorder of the sleep-wake schedule
F51.3	Sleepwalking [somnambulism]
F51.4	Sleep terrors (night terrors)
F51.5	Nightmares
F51.8	Other nonorganic sleep disorders
F51.9	Nonorganic sleep disorder, unspecified
G25.8	Other specified extrapyramidal and movement disorders – Includes: Restless legs syndrome; Stiff-man syndrome
G47	Other sleep disorders
G47.0	Disorders of initiating and maintaining sleep (insomnias)
G47.1	Disorders of excessive somnolence (hypersomnias)
G47.10	Hypersomnia, unspecified
G47.11	Idiopathic hypersomnia with long sleep time
G47.11	Idiopathic hypersomnia with long sleep time
G47.12	Idiopathic hypersomnia without long sleep time
G47.12	Idiopathic hypersomnia without long sleep time
G47.13	Recurrent hypersomnia
G47.14	Hypersomnia due to medical condition
G47.19	Other hypersomnia
G47.2	Disorders of the sleep-wake schedule
G47.3	Sleep apnea
G47.30	Sleep apnea, unspecified
G47.30o	Sleep apnea, obstructed
G47.31	Primary central sleep apnea
G47.310	Sleep apnea, central
G47.33	Obstructive sleep apnea (adult) (pediatric)
G47.34	Idiopathic sleep related non-obstructive alveolar hypoventilation
G47.37	Central sleep apnea in conditions classified elsewhere
G47.38o	Other sleep apnea
G47.39	Other sleep apnea
G47.4	Narcolepsy and cataplexy
G47.411	Narcolepsy with cataplexy
G47.419	Narcolepsy without cataplexy
G47.421	Narcolepsy in conditions classified elsewhere with cataplexy
G47.429	Narcolepsy in conditions classified elsewhere without cataplexy
G47.8	Other sleep disorders; Includes: Kleine-Levin syndrome
G47.9	Sleep disorder, unspecified
R06.81	Apnea, not elsewhere classified

