Medicines & Healthcare products Regulatory Agency

CPRD

1

General information

Protocol reference Id

22_001887

Study title

Chronology of healthcare resource use and comorbidities in people with obstructive sleep apnoea and narcolepsy before and after diagnosis: a descriptive study

Research Area

Disease Epidemiology

Economics

Health Services Delivery

Does this protocol describe an observational study using purely CPRD data?

Yes

Does this protocol involve requesting any additional information from GPs, or contact with patients?

Yes

CPRD query reference number: 00097017

2 Research team

Role	Chief Investigator
Title	Assistant professor
Full name	Helen Strongman
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	helen.strongman@lshtm.ac.uk
Will this person be analysing the data?	Yes
Status	Confirmed

Role	Corresponding Applicant
Title	Assistant professor
Full name	Helen Strongman
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	helen.strongman@lshtm.ac.uk
Will this person be analysing the data?	Yes
Status	Confirmed

Role	Collaborator
Title	Associate professor of biostatistics
Full name	Aurelien Belot
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	Aurelien.Belot@lshtm.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Lecturer in Statistical Epidemiology & National Institute for Health Research Postdoctoral Fellow
Full name	Krishnan Bhaskaran
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	Krishnan.Bhaskaran@lshtm.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Consultant neurologist and Honorary Associate Professor
Full name	Sofia Eriksson
Affiliation/organisation	University College London (UCL)
Email	sofia.eriksson@ucl.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Associate Professor (Reader)
Full name	Michelle Miller
Affiliation/organisation	University of Warwick
Email	Michelle.Miller@warwick.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Associate Professor in Clinical Trials and Health Economics
Full name	Hema Mistry
Affiliation/organisation	University of Warwick
Email	Hema.Mistry@warwick.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Professor of Health Services and Systems Research
Full name	Ellen Nolte
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	Ellen.Nolte@lshtm.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

Role	Collaborator
Title	Associate Professor
Full name	Charlotte Warren-Gash
Affiliation/organisation	London School of Hygiene & Tropical Medicine (LSHTM)
Email	Charlotte.Warren-Gash1@lshtm.ac.uk
Will this person be analysing the data?	No
Status	Confirmed

3

Access to data

Sponsor

London School of Hygiene & Tropical Medicine (LSHTM)

Funding source for the study

Is the funding source for the study the same as Chief Investigator's affiliation?

Yes

Funding source for the study

London School of Hygiene & Tropical Medicine (LSHTM)

Institution conducting the research

Is the institution conducting the research the same as Chief Investigator's affiliation?

Yes

Institution conducting the research

London School of Hygiene & Tropical Medicine (LSHTM)

Method to access the data

Indicate the method that will be used to access the data

Institutional multi-study licence

Is the institution the same as Chief Investigator's affiliation?

Yes

Institution name

London School of Hygiene & Tropical Medicine (LSHTM)

Extraction by CPRD

Will the dataset be extracted by CPRD

No

Multiple data delivery

This study requires multiple data extractions over its lifespan

No

Data processors

Data processor is	Same as the chief investigator's affiliation
Processing	Yes
Accessing	Yes
Storing	Yes
Processing area	UK

4

Information on data

Primary care data

CPRD GOLD

CPRD Aurum

Do you require data linkages

Yes

Patient level data

HES Accident and Emergency

HES Admitted Patient Care

HES Outpatient

ONS Death Registration Data

NCRAS data

Covid 19 linkages

Area level data

Do you require area level data?

Yes

Practice level (UK)

Practice Level Carstairs Index (Excluding Northern Ireland)

Patient level (England only)

Rural-Urban Classification

Withheld concepts

Are withheld concepts required?
No

Linkage to a dataset not listed

Are you requesting a linkage to a dataset not listed?
No

Patient data privacy

Does any person named in this application already have access to any of these data in a patient identifiable form, or associated with an identifiable patient index?
No

5

Protocol information

Lay Summary

Obstructive Sleep Apnoea (OSA) and narcolepsy are sleep disorders associated with multiple other symptoms caused by problems with:
- breathing at night (OSA),
- the inability of the brain to regulate sleep and wake patterns (narcolepsy).

People with OSA and narcolepsy experience delayed diagnosis and poor access to hospital care and treatments.

We will use anonymous information collected from patient records by the Clinical Practice Research Datalink (CPRD). Data describing diagnoses and symptoms, including OSA and narcolepsy, are limited to codes that can be linked to a medical dictionary. We will confirm that this coded information is sufficient to study OSA and narcolepsy by asking GPs to complete a questionnaire using more detailed information, including hospital letters, that are available to them.

We will describe (1) how healthcare services are used by people with these sleep disorders and (2) other medical conditions they experience, both before and after diagnosis. In doing so, we might uncover misdiagnoses, common referral routes, preventable and coexisting conditions, and inequalities in care. We will provide data comparing healthcare costs for people with and without these conditions before and after diagnosis.

Findings will be used by people with OSA and narcolepsy, and the organisations and doctors who support them, to influence future pathways that reduce time to diagnosis and improve access to healthcare, thus providing better joined up treatment of sleep disorders and related conditions. The research and Patient Public Involvement work will be expanded in future to study other uncommon or under-recognised conditions.

Technical Summary

We aim to measure healthcare resource use and medical conditions associated with Obstructive Sleep Apnoea (OSA) and narcolepsy over time before and after diagnosis. Additionally, we will validate coded records of these conditions in primary care and linked data. Our findings will inform policies regarding the diagnosis and treatment of sleep disorders and methods for investigating these and other under-researched conditions.

The primary cohort will include OSA, narcolepsy and comparison groups selected from CPRD primary care data. The sleep disorder groups will have a first record of the condition between 01/01/1990 and the end of data collection. Comparison groups will be matched by sex, year of birth, practice, and registration time. A secondary cohort will be restricted to people eligible for linkage to linked data sources.

Main outcomes and estimates will be:
- Positive Predictive Value of coded records for OSA and narcolepsy compared to hospital confirmed diagnoses measured using a GP questionnaire.
- Annual point prevalence and incidence of OSA and narcolepsy.
- Rates, rate ratios (RRs), rate differences (RDs), mean costs and differences in mean cost of healthcare resource use measured using primary care data and linked Hospital Episode Statistics Admitted Patient Care, Outpatient and Accident & Emergency data.
- Rates, RRs and RDs of common or associated medical conditions measured in primary care and linked data.

We will estimate RRs using generalised linear Poisson models with a classic log link (comorbidities) or a negative binomial link (resource use) including an interaction term between exposure status and time relative to diagnosis (categorical and restricted cubic spline) and adjusting for matching variables and time updating BMI (OSA only). We will predict adjusted rates and calculate RDs. Differences in mean costs will be estimated using generalised linear models. Analyses will be stratified by calendar time, demographics and BMI (OSA only).

Outcomes to be measured

VALIDATION
Coded and hospital confirmed diagnosis and diagnosis dates for OSA and narcolepsy

HEALTHCARE RESOURCE USE
Consultations/visits and costs
- Primary care
- Outpatient (overall and by specialty)
- Accident & Emergency
- Hospital admissions (overall and by ICD-10 chapter)

Primary care prescriptions and costs
- Overall and by BNF chapter
- Treatments for excessive daytime sleepiness
- Treatments for cataplexy (narcolepsy only)

Specific procedures and tests related to the diagnosis and treatment of OSA and narcolepsy

COMORBIDITIES
Conditions that are common in people with OSA and narcolepsy before and after diagnosis and less common conditions thought to be associated with OSA or narcolepsy.

Objectives, specific aims & rationale

OVERALL AIM

To answer the overarching question: “Before and after diagnosis with OSA and narcolepsy, how do people use healthcare resources, what comorbidities do they experience, and how could pathways to diagnosis and care be improved”?

By doing this, we will provide evidence to support policy promoting:
- Earlier recognition of the symptoms of sleep disorders leading to reduced misdiagnoses, unnecessary referrals, and development of avoidable comorbidities.
- Improved access to healthcare through expansion of sleep services and training and better joined up care of sleep disorders and related comorbidities.

In addition to this CPRD study, the programme of work includes sleep clinic-based questionnaires to assess the sensitivity of recording of OSA and narcolepsy in primary care data and qualitative research to understand barriers to diagnosis. The programme will increase confidence in the use of routinely collected healthcare data for research, provide reproducible research methods, and ultimately empower people, patient groups and clinicians to improve access to healthcare for other under-researched medical conditions.

Study background

The UK Strategy for Rare Diseases(1) describes the need for accessible and impactful research empowering patient groups to improve access to care for rare and neglected medical conditions. This need is confirmed by Narcolepsy UK’s Charter, informed by a survey of narcolepsy sufferers and supporters,(2) and underlined by their lived experiences(3), with other works showing that these problems also affect more common but low profile conditions such as OSA(4).

MISSED POTENTIAL FOR USE OF ROUTINELY COLLECTED HEALTH DATA

Over 2000 peer reviewed publications have been generated from electronic health records (EHRs) collected in UK general practices and linked to additional health and demographic datasets(5).

To date, use of UK EHR data has focused on the epidemiology of high-profile conditions and pharmaco-epidemiology (5). Data have been underused to study uncommon and low-profile conditions. While EHR data are increasingly used to inform health policy e.g. through NICE technology appraisals(6), for rare conditions they have had more limited influence. This may be due to a lack of: confidence in the validity of these data; understanding about how these data can be used to study disease epidemiology and healthcare resource outcomes; and qualitative insights into reasons for observed trends.

PREVALENCE

Estimates of the UK prevalence of OSA and narcolepsy (1.5 million and 30,000 people, respectively) and their symptoms are based, mainly, on self-reported interviews of affected people in the UK and European populations(7–9). Other estimates are based on predicted prevalence based on the distribution of risk factors, showing substantial regional variations and inequalities in access to healthcare resources for OSA (10).

DIAGNOSTIC PATHWAYS AND BARRIERS TO SEEKING CARE

There is very little recent evidence on pathways to diagnosis and health service use in the UK for people with sleep disorders. Morrish et al. looked at factors associated with delays in the diagnosis of narcolepsy, based on survey data, but this work dates to 2004(11).

HEALTHCARE RESOURCE USE

Case control and cohort studies in the USA, Denmark, Canada and Israel describe elevated health costs for undiagnosed OSA patients compared to individuals without OSA(12,13) and increased healthcare use in the years following diagnosis(14–17). Most of these studies have not adjusted for Body Mass Index raising questions about the extent to which increased healthcare resource use is caused by comorbid obesity. Guest et al. found no published UK resources describing health care use by people with OSA for inclusion in Markov models estimating the

cost-effectiveness of Continuous positive airway pressure (CPAP) (18). Instead, estimates were based on interviews with 19 sleep service clinicians.

Increased use of healthcare resources has been observed in people with narcolepsy compared to matched comparison groups in Danish EHR(19) and US claims data(20,21). In Denmark, healthcare use increased more than a decade before diagnosis(19). There are no related UK investigations.

COMORBIDITIES

The associations between OSA and comorbidities are not yet fully understood. Associations between OSA and multiple comorbidities including cardiovascular disease(22,23), hypertension(24), diabetes (25), respiratory(26,27) and anxiety and depression(28) have been reported in meta-analyses and systematic reviews. Most studies considered in reviews were based on prospective population and community-based cohorts, and prospective or retrospective clinic-based cohorts. Non-prospective cohort studies or those without confirmation of OSA by polysomnography were excluded(23,25,28). While this approach recognises the benefits of prospective, cross-sectional, and clinical cohorts in accurately and extensively phenotyping specific diseases, it does not recognise the complementary benefits of studies based on routinely collected data, leading to inconclusive summaries. Evidence of associations between OSA and other uncommon conditions such as specific cancers and renal disease is limited(29).
Likewise, there is limited evidence on the association between narcolepsy and comorbidities such as obesity, diabetes and mental health disorders. Existing work consists, largely, of small clinical studies investigating specific comorbidities(30), and a small number of studies studying multiple-comorbidities in the USA(31–34), Denmark(35) and Taiwan(36,37). These vary in design and are often limited by small sample sizes. As with OSA, the relative timing of diagnosis of narcolepsy, compared to related comorbidities, and the implications of this for diagnostic and treatment pathways, has not been thoroughly explored.

OPPORTUNITIES OFFERED BY UK HEALTHCARE DATA

Routinely collected UK healthcare data offer unprecedented opportunities to complement existing evidence from cross-sectional and prospective cohort studies by allowing us to (1) study large numbers of people from a population-based sample (2) follow subjects for long periods of time before and after diagnosis (3) compare people with and without sleep disorders from the same population, and (4) use detailed information captured prospectively including medical diagnoses, primary care prescriptions and primary and secondary care healthcare resource use. We can therefore study multiple exposure outcome associations and identify subgroups with different risk profiles. Through validating definitions of study variables, and demonstrating how CPRD data can be used to improve access to diagnosis and care, we will open the use of routinely collected data

in the UK to improve healthcare for people with sleep disorders and other under-researched conditions. This study will also generate hypotheses for future research exploring causal pathways.

Study type

We will first evaluate how accurately OSA and narcolepsy can be identified in CPRD data (Objective 1). We will then complete a descriptive hypothesis generating study describing patterns of disease incidence and prevalence over calendar time and healthcare resource use and comorbidities relative to time since diagnosis (Objectives 2 and 3). We aim to describe associations rather than to answer questions of causality and will interpret our findings with reference to existing literature and clinical and patient experts.

Study design

Objective 1: Validation questionnaire study.

Objectives 2/3: We will use a hybrid cohort/case control study design including sleep disorder groups matched to a general population comparison group. Rates, rate differences, rate ratios, and differences in mean costs will be estimated stratified by time relative to diagnosis by looking both forwards (cohort design) and backwards (case control design) since diagnosis. This approach is adapted from Collin et al's CPRD study of healthcare resource before and after a diagnosis of chronic fatigue syndrome(38). Prevalence and incidence of OSA and narcolepsy will be estimated using the denominator files.

Feasibility counts

The following feasibility counts are based on the March 2022 CPRD GOLD and Aurum builds and set 21 of the linkage data. Current patients are defined as those under follow-up beyond 1st December 2021.

OSA

We counted 164 413 adults with OSA diagnosed at least 12 months after registration in a CPRD practice between 01/01/1990 and 31/12/2019. Median (IQR) follow up before and after diagnosis was 10.7 years (5.1, 17.9) and 4.3 years (1.8, 8.4) respectively. When restricting the sample to linkage eligible patients and the HES APC, ONS mortality, HES outpatient and HES A&E coverage periods, respectively 131 720, 131 067, 121 972 and 107 759 people would be included. Approximately 91 202 people in the primary care cohort are currently registered with the practice.

NARCOLEPSY

We counted 2 883 people with narcolepsy diagnosed at least 12 months after registration in a CPRD GOLD practice between 01/01/1990 and 31/12/2019. Median (IQR) follow up before and after diagnosis was 7.3 years (3.4, 13.0) and 5.8 years (2.5, 11.2) respectively. When restricting the sample to linkage eligible patients and the HES APC, ONS mortality, HES outpatient and HES A&E coverage periods, 2 112, 2 075, 1 701 and 1 296 people are included. Approximately 1 170 people in the primary care cohort are currently registered with the practice.

Sample size considerations

VALIDATION STUDY (OBJECTIVE 1) A systematic review estimated a mean PPV of 90% for all medical conditions(39). With a sample size of 100 for each sleep disorder, the 95% confidence intervals (based on a normal distribution) would be 84.1% to 95.9%. Subset analyses would also have reasonable precision (e.g. expected 95% CI of 81.7 to 98.3 for a subset including 50 people). HEALTHCARE RESOURCE USE AND COMORBIDITIES (OBJECTIVES 2 AND 3) With a sample of 164 413 in the OSA group, in years relative to diagnosis where the proportion of events in the OSA group is 1% (n=1644) and 5% (n=8221), we will have 90% power to detect a rate ratio of 1.09 or 1.04 or higher, respectively, and a risk difference of 0.09% or 0.19% or higher, respectively, with alpha at 5%. We will be able to make precise detailed comparisons over time even for rare events. With a sample of 2 883 in the narcolepsy group, in years relative to diagnosis where the proportion of events in the narcolepsy group is 1% (n=29) or 5% (n=144), we will have 90% power to detect a rate ratio of 1.94 or 1.34 or higher, respectively, and a risk difference of 0.54% or 1.33% or higher, respectively, with alpha at 5%. This will allow more detailed comparisons than previously published studies of the associations of narcolepsy with healthcare use and prevalent outcomes over time since diagnosis(19,35). Statistical power will be reduced for incident comorbidities within each year relative to diagnosis. In assessing the strength of the evidence, we will consider how patterns change between years and formal statistical tests for categorical year groups.

Planned use of linked data and benefit to patients in England and Wales

The study will include primary care only and linked cohorts. The primary care only cohort will be used for the primary analysis, when secondary care data is not essential, to maximise sample size. The linked cohort will be used for primary analysis where the outcome is measured using linked data only or validation studies have demonstrated low sensitivity with the use of primary care data only, and in sensitivity analyses for other outcomes. This will be restricted to English practices due to linkage eligibility.

Should the number of practices contributing to CPRD Aurum increase substantially before data extraction, the study will be restricted to CPRD Aurum and the linked cohort will be used for the primary analysis as the proportional gain in sample size from using CPRD GOLD and primary care data only will be much reduced.

The following linked datasets will be used in the linked cohort:
- HES patient file (to measure ethnicity as a covariate)
- HES APC data (to measure narcolepsy and OSA, hospital admissions outcomes, and comorbidities for the sensitivity analysis)
- HES Outpatient data (to measure outpatient admissions outcomes overall and by specialty)
- HES A&E data (to measure A&E visit outcomes)
- ONS mortality data (to measure date of death and potentially fatal comorbidities that may not be recorded in primary care or HES APC data)

We will also use the practice area based Carstairs index of deprivation and practice level urban rural data as stratifying covariates.

This will benefit patients in England and Wales by informing national and local healthcare policies that improve access to diagnosis and care for people with narcolepsy and OSA through an understanding of healthcare resource use requirements and potential improvements.

Definition of the study population

The study will include a primary care and linked study population for each sleep disorder. The primary care study population will include people registered in CPRD GOLD or Aurum practices for a minimum baseline period, during the study period, and with research quality data (CPRD acceptable flag). The linked study population will include people registered in CPRD GOLD or Aurum practices whose records are eligible for linkage to HES APC, ONS mortality and area-based data. We will exclude practices contributing to both CPRD GOLD and CPRD Aurum from the CPRD GOLD cohort as the CPRD Aurum cohort is likely to include more recent data. The minimum baseline period will be defined following visual inspection of a plot of the incidence of narcolepsy and OSA over time since registration; it is likely to be less than the standard 12 months used in CPRD studies which will increase the size of the exposure groups(6).
The study period start and end dates will be 1st January 1990 and the end of CPRD data collection for the primary care cohort and 2nd January 1998 (start of data collection for ONS mortality data) and the end of HES APC/ONS mortality coverage for the linked cohort. Follow-up for each person in the study population will start at the latest of the study period start date and the minimum baseline period after the (current) registration date. Follow-up will end at the earliest of the study end date, practice last collection date, transfer out date, and CPRD derived death date.

The primary care and linked study cohorts will include sleep disorders groups (OSA and narcolepsy) and matched comparison groups (see selection of comparison group(s) or controls). People may be included in both OSA and narcolepsy groups if they have a diagnosis for both conditions.

The sleep disorders group will include the subset of the study population who have a medical code for the sleep disorder in the primary care record (primary care cohort) or in either the primary care record or any field in the HES APC data (linked cohort). The diagnosis date will be the first ever record of OSA or narcolepsy.

We have attached code lists for OSA and narcolepsy in Appendix A. We will use a broad code list for OSA including OSA, sleep apnoea unspecified and OSA/sleep apnoea syndrome. We will exclude people with a record for central or primary sleep apnoea prior to the diagnosis date from the OSA cohort and censor follow-up if these codes are recorded after diagnosis. We will include descriptive analyses in the validation study (Objective 1) to assess the impact of non-specific coding and run sensitivity analyses for Objectives 2 and 3 using narrower code lists as appropriate based on our findings.

With the exception of estimating sleep disorder prevalence, analyses will be restricted to people diagnosed on or after the start of follow-up (i.e. incident diagnoses). With the exception of estimating OSA prevalence and incidence, OSA analyses will be restricted to people aged 18 or over at diagnosis.

For the validation study (Objective 1), we will select a random subset of 143 people for each sleep disorder leading to 200 (100 x 2) responses, assuming a 70% response rate.

Selection of comparison groups/controls

The comparison group pool for each individual in the sleep disorder group will include the study population with the end of follow-up truncated at diagnosis of the sleep disorder (or other specified sleep apnoea for OSA) to represent the full population at risk (incidence density sampling). People diagnosed with the sleep disorder (or another specified sleep apnoea for OSA) prior to the start of follow-up will therefore be excluded from the pool.

Each person in the sleep disorder group will be matched to up to 5 people on year of birth (plus/minus 3 years), sex, practice, and registration time whereby the matched person will be currently registered on the diagnosis date of the matched person with a sleep disorder and have at least the minimum period of prior registration required for the sleep disorder group.

Exposures, outcomes and covariates

For this hybrid cohort/case control study, we have defined the OSA, narcolepsy and comparison groups as exposures and healthcare resource use and comorbidities as outcomes.

EXPOSURE
OSA or narcolepsy and matched comparison group (see definition of the study population and selection of comparison group(s) or controls).

VALIDATION STUDY OUTCOMES (OBJECTIVE 1)
Hospital confirmed diagnosis and date of diagnosis indicated by GP questionnaire. Hospital confirmed diagnoses will be ascertained through direct questions to GPs about whether the patient has been diagnosed or treated for narcolepsy/OSA by a hospital specialist at any time and when the patient was diagnosed by the hospital. The outcome and specific covariates (see below) will be measured using a GP questionnaire.

HEALTHCARE RESOURCE OUTCOMES (OBJECTIVE 2)
Objective 2 (healthcare resource use)
2a) Prevalence and incidence of OSA and narcolepsy (rates and numbers)
2b) Primary care consultations and costs
- Overall (in-person and remote consultations with clinicians identified using the consultation and staff files. Methods will be developed and tested for Aurum in the absence of existing published algorithms)
2c) Hospital visits and costs
- Outpatient attendance overall and by specialty (HES outpatient data from 01/04/2003 to end of data collection)
- Accident & Emergency visits (HES A&E data from 01/04/2007 to 31/03/2020). We will explore the potential to stratify by reason for admission and other key variables but are unable to plan this as we have found no published data describing completeness of the relevant fields.
- Hospital admissions overall and by ICD-10 Chapter (full linked study period)
2d) Primary care prescriptions and costs
- Overall and by BNF chapter
- Treatments for excessive daytime sleepiness (modafinil, dexamphetamine, methylphenidate as mono- and combination therapy)
- Treatments for cataplexy (fluoxetine, venlafaxine, clomipramine, other antidepressants; narcolepsy only)
Code lists will be developed using text searches of drug substance and brand names and searching relevant BNF chapters.
2e) Specific procedures and tests related to the diagnosis and treatment of OSA and narcolepsy (Polysomnography/ multiple Sleep Latency Test (MSLT), electroencephalogram (EEG) telemetry, oxygen saturation and blood gas values, oximetry, lumbar puncture (narcolepsy only)). These will be identified in HES APC data using OPCS codes and primary care data using Read, Snomed and EMIS code lists.
Costs will be based on NHS reference costs, Personal Social Services Research Unit (PSSRU) unit and electronic drug tariff costs for the latest year that these data are available to ensure comparability.

COMORBIDITY OUTCOMES (OBJECTIVE 3)
We will include conditions that are hypothesised to be associated with sleep disorders for any reason (causal association, differential diagnosis, common risk factors) and common conditions that are treated in primary and secondary care. Conditions that are hypothesised to be associated with sleep disorders have been identified with reference to existing literature, and in consultation with sleep clinicians and the Patient Public Involvement group for this study (see below). For common conditions, we will use Steps 1 to 3 of procedures developed for High Dimensional Propensity Score generation(40) to identify highly prevalent conditions in the 3 years before and after a patient's diagnosis in each sleep disorder cohort. Dimensions will use 4-digit ICD-10 codes mapped from Snomed codes in primary care(41). When we have identified dimensions with the highest prevalence, we will refine algorithms and code lists for related comorbidities based on clinical knowledge.

Final code lists will be based on validated or clinically informed code lists and algorithms or developed in collaboration with clinicians. Unless otherwise stated, in the primary analysis, definitions will include Read, Snomed or EMIS coded records of the condition in primary care; in analyses using linked data, we will add ICD-10 codes from HES APC data (all codes) and ONS mortality causes of death (all codes) for potentially fatal outcomes that might not be recorded in primary care or HES.

We will measure incident events (i.e. first ever record with follow-up time censored at the incident event unless otherwise specified) and prevalent diagnoses (from first ever record to end of follow-up for chronic conditions; until end of defined episode for acute/resolving conditions).

Conditions with hypothesised association with sleep disorders
Benign and primary malignant cancers (29) (prevalent until 5 years following diagnosis, first malignant cancer diagnosis only due to difficulty with differentiating between primary and secondary malignant cancers; use linked cohort including ONS mortality data for primary analysis)
- 20 most common cancers
Cardiovascular disease (29,30) (use linked cohort including ONS mortality data for primary analysis)
- All combined
- Coronary Artery Disease
- Cerebrovascular disease
- Heart failure / cardiomyopathy
- Venous thromboembolism (incident only)
- Arrhythmia
- Hypertension (time updating - incident at hypertension diagnosis, flag on practice hypertension register, or 2 subsequent blood pressure measures of 140/90 mm Hg within a 3-month period; prevalent until low blood pressure reading and no hypertension treatment 6 months prior)
- Treatment resistant hypertension (prioritised for treatment under NICE guidelines, hypertension treatment and 2 subsequent blood pressure measures of 140/90 mm Hg within a 3-month period)
Digestive diseases (42)
- Autoimmune (Autoimmune liver disease, Coeliac disease, Crohn’s disease, Irritable bowel syndrome (30))
- Gastro-oesophageal reflux disease (GORD)
- Nonalcoholic fatty liver disease (NAFLD)
- Peptic ulcers (episode duration 1 year)
- Chronic liver disease (use linked cohort including ONS mortality data for primary analysis)
Ear
- Tinnitus(43) (replace prevalent with seeking care for tinnitus – see mental health outcomes)
Endocrine
- Diabetes (combined, type 1, type 2) (primary care only) (29,30)
- Complications of diabetes (some evidence that these may be more severe in people with sleep apnoea)
- Underweight and obesity (classes I, II and III) (BMI records from primary care; time updated at each recorded measurement) (30)
- Thyroid disease (hypothyroidism is a risk factor for OSA in NICE guidelines)
- Acromegaly (risk factor for OSA in NICE guidelines)
- Gout (common risk factors with OSA)
Eye (44)
- Non-arteritic anterior ischaemic optic neuropathy (risk factor for OSA in NICE guidelines)
- Glaucoma
- Floppy eye syndrome
- Macular degeneration
Genitourinary
- Acute Kidney Injury and change in Glomerular Filtration Rate (GFR) (29,30) (time updating, using serum creatinine records from primary care)
- Chronic Kidney Disease (29,30) (serum creatinine records or codes)

- Polycystic ovary syndrome (risk factor for OSA in NICE guidelines)
- Urinary incontinence (association with sleep disturbance)

Haematological / Immunological

- Anaemia (iron deficiency anaemia is associated with sleep disorders, events defined using lab tests, iron prescriptions and diagnostic codes, replace prevalent with seeking care for – see mental health outcomes)

Infections (prevalent for 28 days)

- Upper respiratory tract infections (narcolepsy causal pathway(45))
- Lower respiratory tract
- Ear infections

Musculoskeletal

- Carpal tunnel syndrome (associated with poor sleep, episode duration 1 year)
- Back pain (associated with poor sleep)
- Rheumatoid arthritis (possible autoimmune link to narcolepsy)

Neurological

- Chronic fatigue syndrome
- Restless Legs Syndrome
- Epilepsy (association with sleep disorders; differential diagnosis for narcolepsy)
- Migraine (using codes and prescriptions for prophylactic treatment, episode duration 1 year) (30)
- Fibromyalgia (30)
- Neurodegenerative disease including dementia
- Down's Syndrome (risk factor for OSA in NICE guidelines)
- Narcolepsy (OSA)

Psychiatric / mental health (*prevalent analysis will be replaced with "actively seeking treatment" for these conditions i.e. code in the last 6 months)

- Major psychiatric disorder (schizophrenia (30), bipolar affective disorder (30), personality disorder)
- Panic disorder (30)
- Stress (30)*
- Substance misuse*
- Neuro-atypical (Autism, ADHD (30))
- Alcohol misuse*
- Anxiety* and anxiety disorders (30)
- Depression* (30) (codes for symptoms and diagnoses)
- Fatal and non-fatal* self harm
- Insomnia diagnosis* (medical codes and treatments) (30)

Respiratory (29,30)

- Allergic/chronic rhinitis (common and a consideration for management of OSA in NICE guidelines)
- Asthma (published definition by Jenni Quint et al, resolve if no code or prescription for 3 years)
- COPD
- OSA (narcolepsy)

Skin (46)

- Acne (incident only)
- Atopic dermatitis
- Psoriasis

Symptoms (recurring for incidence analysis; counted once if recorded in the same week)

- Fatigue / excessive daytime sleepiness (tiredness, sleepiness, epworth sleepiness scale)
- Tests for infections associated with fatigue (Epstein Barr, HIV/AIDs, Hepatitis B or C, Tuberculosis, Lyme disease)
- Disrupted night-time sleep (nightmares, hypnogogic hallucinations, hypnopompic hallucinations, sleep paralysis, sleep fragmentation, insomnia symptoms, dream enactment (30))
- Apnoea (choking during sleep, gasping, snoring)
- Headaches (30)
- Cognitive dysfunction or memory impairment (30) (including automatic behaviours(30))
- muscle weakness

COVARIATES

All objectives
- Calendar time (categorical)
- Patient demographics (age group, sex, ethnicity))
- Practice characteristics (region, area-based deprivation, urban rural, practice size)
- WHO BMI continuous for model adjustment, categorical for stratification (below 18.5 Underweight; 18.5–24.9 Normal weight; 25.0–29.9 Pre-obesity; 30.0–34.9 Obesity class I; 35.0–39.9 Obesity class II; Above 40 Obesity class III) measured using the last available record; OSA only.
- OSA coding (sleep apnoea vs sleep apnoea syndrome; OSA versus unspecified apnoea)
Objective 1 only
- Methods used to diagnose the patient in hospital recorded in primary or secondary care?
- Suspected narcolepsy/OSA, reasons why the diagnosis has not been confirmed (e.g. not

referred to specialist, on hospital waiting list, narcolepsy/OSA excluded)
- Narcolepsy/diagnosis excluded after initial diagnosis
- Additional detail i.e. cataplexy for narcolepsy and severity for OSA

Data/statistical analysis

VALIDATION QUESTIONNAIRE (OBJECTIVE 1)
We will estimate the Positive Predictive Value (PPV) of a coded diagnosis for OSA and narcolepsy using hospital confirmed diagnosis as a gold standard and describe differences between the CPRD derived diagnosis date and GP confirmed diagnosis date. We will describe the proportion of the people with narcolepsy in the validation study who have type 1, type 2 or unspecified narcolepsy and the proportion of people with OSA who have mild, moderate or severe OSA. We will estimate the sensitivity, PPV, specificity and Negative Predictive Value (NPV) of a coded diagnosis of type 1 and type 2 narcolepsy. Descriptive analysis of diagnostic methods will be used to ensure that the majority of hospital recorded diagnoses are based on guidelines. Descriptive analyses of recording and follow-up of suspected OSA/narcolepsy, exclusion of OSA/narcolepsy after initial diagnosis, and severity of OSA will be used to explore the origin of false positive diagnoses. Analyses for OSA will be stratified by OSA code category (OSA, sleep apnoea NOS, sleep apnoea syndrome) to understand the impact of using broad code lists.

HEALTHCARE RESOURCE USE AND COMORBIDITIES (OBJECTIVES 2 AND 3)
Incidence and prevalence
We will estimate the point prevalence of diagnosed OSA and narcolepsy for each calendar year in the study period by dividing the number of people in the sleep disorder group under follow-up at this time with the total number of people in the study population under follow-up at this time. We will calculate annual incidence rates by dividing the number of people diagnosed within the calendar year by the total person-time contribution in the comparison group pool. 95% confidence intervals (CIs) will be calculated using the Poisson distribution. To estimate and compare annual incidence and prevalence rates, taking into account changes in population structure, we will directly standardise all rates to the 2020 ONS population by age and sex. To estimate and compare the number of people with prevalent and incident diagnoses of narcolepsy and OSA each year, we will directly standardise age and sex specific rates to ONS population figures for each year.

Healthcare resource use (frequency of service use) and comorbidity outcomes
We will calculate rate ratios and 95% CIs for healthcare resource use count and comorbidity outcomes in people with sleep disorders compared to the comparison group by fitting generalised linear Poisson models including an interaction between exposure (i.e. OSA or narcolepsy) status and categorical year relative to diagnosis. A classic log link will be used for comorbidity outcomes and a negative binomial link will be used for healthcare resource outcomes which are commonly over dispersed. All analyses will be adjusted for the matching variables (age modelled with restricted cubic spline with 3 knots and sex). Due to strong causal associations between BMI and development of OSA, and BMI and healthcare resource use/comorbidities, primary analyses for OSA will be adjusted for time-updated BMI (3 knotted restricted cubic spline); people with missing BMI will be excluded from these analyses. To estimate adjusted rates and rate differences, the above generalised linear models will be repeated with restricted cubic splines replacing categorical year relative to diagnosis. We will use these models to predict rates in the sleep disorder and comparison groups and calculate rate differences with bootstrapped. In sensitivity analyses, we will also use generalised additive models with penalised regression splines for modelling age and year of diagnosis as an alternative; this should produce smoother curves but may have convergence issues for some outcomes preventing their use in the primary analysis.

Healthcare resource use (costs)
We will estimate mean (standard deviation) costs of healthcare visits and prescriptions/tests for each year relative to diagnosis using methods described by Lin et al to account for censoring (48) and bootstrapping to estimate 95% confidence intervals. To compare mean costs in the sleep disorder and comparison groups, we will use generalised linear models adjusted for age, sex, and BMI (OSA only) with our choice of generalised linear model family informed by the modified Park

test and Box-Cox test, respectively(49).

SECONDARY ANALYSES
We will stratify analyses by calendar time, the listed demographic covariates and BMI (OSA only).

In the validation analysis (objective 1), we will compare CPRD GOLD and CPRD Aurum, and people in CPRD GOLD diagnosed before and after the practice up-to-standard-date in CPRD GOLD. For patients eligible for linkage, we will compare the accuracy of diagnosis dates estimated using primary care data only with diagnosis dates estimated using linked data.
In objectives 2 and 3, we will estimate rate ratios and cost differences without adjustment for BMI to describe the extent to which BMI explains the association between OSA and healthcare resource use/comorbidities.

SENSITIVITY ANALYSES
To maximise sample size, primary analyses will be based on the primary care cohort for outcomes that do not require secondary care data. We will run a sensitivity analysis in the linked data cohort.

Due to changes in healthcare services during the COVID-19 pandemic, the primary analysis will end on 31/12/2019. We will include data from 01/01/2020 to the end of data collection in a sensitivity analysis to understand how including this time period might impact future studies.
Should the validation study indicate that the type of OSA code is associated with OSA severity, we will run sensitivity analyses for objectives 2 and 3 using narrower code lists.
We will compare rate ratios estimated using Poisson models adjusted for age and sex to models with no further adjustment beyond matching. This will inform our understanding of the impact of adjusting for matching variables in studies using our design (50,51). We will also compare rates and rate differences estimated using basic methods [i.e. summing the number of events during each year before and after the cohort entry date (diagnosis date for matched person with sleep disorder) and dividing by the total time contribution in that year] with those predicted using generalised linear models and estimated using generalised additive models.

Additional sensitivity analyses to explore the impact of missing data are described in the Missing Data section.

Plan for addressing confounding

This is a description study aimed at informing care pathways and future research rather than establishing causality. Adjustment for confounding will be restricted to matching by year of birth, sex and practice and adjusting for age, sex and BMI (OSA only).

Plans for addressing missing data

For the majority of variables in this study, people with no coded record of the exposure or outcome will be assumed not to have the exposure or outcome. Exceptions are BMI and ethnicity.

BMI will be missing until the first ever record of BMI in the patient record. Time periods prior to this record will be excluded from all OSA analyses and from narcolepsy analyses stratified by BMI. We anticipate this to have minimal impact on the OSA group as there are strong clinical reasons for GPs to measure and record BMI. The control group will be weighted towards people with a BMI measurement which may remove healthy controls or those who do not visit their GP practice. We will describe numbers and proportions of people with missing data in the exposed and comparison group and complete the following sensitivity analyses to inform our understanding of the impact of this:
(1) Estimating rate ratios using models adjusted for age and sex prior to excluding time periods with missing BMI data.
(2) An alternative definition of BMI that excludes time periods where BMI has not been measured for 5 years. This will be a more accurate measure of BMI but may exclude larger time periods.

In stratified analyses by ethnicity, we will include a missing category.

Patient or user group involvement

This research is part of an NIHR Advanced Fellowship held by Helen Strongman. Helen is a person with narcolepsy and Trustee & Director for Narcolepsy UK.

A Patient Public Involvement group will work with Helen and the team throughout the fellowship. They have reviewed the lay summary for this protocol and contributed to conversations about comorbidities, stratifying covariates and adjusting for BMI in primary analyses for OSA. They will be involved in reviewing manuscripts and designing and presenting materials to disseminate findings.

The fellowship steering group includes Amanda Molloy who is a GP and Trustee & Director of Narcolepsy UK and Chris Rogers, Secretary of the Sleep Apnoea Trust.

Plans for disseminating & communicating

Findings will be shared through research publications; clinical, research and patient conference presentations and workshops, social media communications, infographics and policy briefings to maximise impact on care. Outputs will be informed by a stakeholder workshop towards the end of the fellowship and co-produced with the PPI group.

This work on sleep disorders will provide an exemplar for the use of routinely collected healthcare data for other under-researched medical conditions. Code lists and coding for data management and analysis will therefore be published on Github and linked to LSHTM's data management system.

Conflict of interest statement

Helen Strongman is a volunteer Trustee & Director of Narcolepsy UK.

Sofia Eriksson has received honoraria for educational activities from Eisai, Fidia, Lincoln and UCB pharma.

Michelle Miller is an unpaid Trustee and Executive Committee Member of the British Sleep Society.

All other authors have no conflicts of interest to declare.

Limitations of study design

In common with all CPRD studies, this research relies on the generalisability of CPRD and linked data and the validity of definitions used to measure exposures, outcomes and covariates. Limitations specific to this study include:

/*Selection bias*/
- Using a GP questionnaire for the validation study will not provide a perfect gold standard as it relies on hospital letters being transferred to GP practices and GP practice staff being able to interpret these.
- Given the high proportion of people with OSA who are undiagnosed (10), the matched control population could include a large proportion of people who have undiagnosed OSA. This will lead to an underestimation of associations between OSA and healthcare resource use/comorbidities.
- We will not be able to differentiate by severity of OSA as this is not recorded in the primary care record and use of terms such as OSA syndrome to indicate mild disease has varied over time and is unlikely to be consistent in primary care data. We will we able to estimate the proportion of our sample that has mild, moderate or severe OSA through the validation study (Objective 1).
- We may include a small proportion of people with central or primary sleep apnoea in the OSA group. We will describe the proportion of people in our validation sample with other types of sleep apnoea.

/*Measurement bias*/
- We have not found published algorithms estimating numbers of GP consultations in CPRD Aurum. We will develop these algorithms based on the completeness of key variables and describe limitations in the manuscript informed by a sensitivity analysis comparing CPRD GOLD and CPRD Aurum.
- We will only capture treatments for narcolepsy that are prescribed in primary care; this will exclude specialist only third or later line therapies including sodium oxybate, pitolisant and solriamfetol.
- For the majority of chronic conditions, we will consider diagnoses to be prevalent from first diagnosis until the end of follow-up. Misdiagnoses will therefore continue to be counted.
- Recording of symptoms is less reliable and consistent than recording of diagnoses. They are only like to be coded if they are a key presenting symptom for a consultation and the diagnosis is unknown. We will interpret our analyses of sleep disorder symptoms with this in mind. For comorbidities where variability in coding practices make it difficult to distinguish between acute symptoms and chronic medical conditions or for which there is variability in frequency of visits and time to resolution (e.g. insomnia, depression), we will replace analysis of disease prevalence with analysis of people actively seeking care.
- We will not be able to differentiate between narcolepsy subtypes due to infrequent use of subtype specific codes. We will not therefore be able to stratify analyses by subtype. Including people with type 2 narcolepsy may reduce diagnostic security.
- We are planning to include data from the early days of electronic health record systems to increase the sample size, especially for narcolepsy. Data recording is likely to be less complete at the beginning of this study period. Should CPRD recruit considerably more CPRD Aurum practices before data extraction increasing the sample size, we will therefore use the linked cohort and coverage period as the primary analysis.

/*Analysis methods*/
- Our study includes multiple tests which increases the potential for false positive findings. We will therefore interpret our findings cautiously within the context of the current literature and by describing the strength of the evidence based on patterns of associations with each outcome over time rather than basing conclusions on an arbitrary p-value for a single estimate.

References

1. Department of Health. The UK Strategy for Rare Diseases [Internet]. 2013 [cited 2020 Aug 27]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/2
Narcolepsy Charter | Narcolepsy UK [Internet]. [cited 2020 Jun 11]. Available from:

https://www.narcolepsy.org.uk/resources/narcolepsy-charter
3. #NarcolepsyStories | Narcolepsy UK [Internet]. [cited 2020 Oct 20]. Available from: https://www.narcolepsy.org.uk/resources/narcolepsystories
4. British Lung Foundation. Toolkit for commissioning and planning local NHS services in the UK Obstructive Sleep Apnoea (OSA) [Internet]. 2015 [cited 2020 Oct 20]. Available from: www.blf.org.uk/OSA
5. Vezyridis P, Timmons S. Evolution of primary care databases in UK: a scientometric analysis of research output. BMJ Open [Internet]. 2016 [cited 2019 Apr 25];6(10):e012785. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27729352
6. Leahy TP, Ramagopalan S, Sammon C. The use of UK primary care databases in health technology assessments carried out by the National Institute for health and care excellence (NICE). BMC Health Serv Res [Internet]. 2020 Jul 22 [cited 2020 Nov 4];20(1). Available from: https://pubmed.ncbi.nlm.nih.gov/32698805/
7. Ohayon MM, Guilleminault C, Priest RG, Caulet M. Snoring and breathing pauses during sleep: Telephone interview survey of a United Kingdom population sample. Br Med J [Internet]. 1997 Mar 22 [cited 2020 Oct 26];314(7084):860–3. Available from: https://www.bmj.com/content/314/7084/860
8. Lechner M, Breeze CE, Ohayon MM, Kotecha B. Snoring and breathing pauses during sleep: interview survey of a United Kingdom population sample reveals a significant increase in the rates of sleep apnoea and obesity over the last 20 years - data from the UK sleep survey. Sleep Med [Internet]. 2019 Feb 1 [cited 2020 Oct 26];54:250–6. Available from: https://pubmed.ncbi.nlm.nih.gov/30597439/
9. Ohayon MM, Priest RG, Zulley J, Smirne S, Paiva T. Prevalence of narcolepsy symptomatology and diagnosis in the European general population. Neurology [Internet]. 2002 Jun 25 [cited 2020 Oct 27];58(12):1826–33. Available from: https://pubmed.ncbi.nlm.nih.gov/12084885/
10. Steier J, Martin A, Harris J, Jarrold I, Pugh D, Williams A. Predicted relative prevalence estimates for obstructive sleep apnoea and the associated healthcare provision across the UK. Thorax [Internet]. 2014 Apr 1 [cited 2020 Oct 23];69(4):390–2. Available from: http://www.amednews.com/
11. Morrish E, King MA, Smith IE, Shneerson JM. Factors associated with a delay in the diagnosis of narcolepsy. Sleep Med [Internet]. 2004 Jan [cited 2020 Oct 26];5(1):37–41. Available from: https://pubmed.ncbi.nlm.nih.gov/14725825/
12. Tarasiuk A, Reuveni H. The economic impact of obstructive sleep apnea. Curr Opin Pulm Med [Internet]. 2013 Nov [cited 2020 Jun 30];19(6):639–44. Available from: http://journals.lww.com/00063198-201311000-00009
13. Wickwire EM, Tom SE, Vadlamani A, Diaz-Abad M, Cooper LM, Johnson AM, et al. Older adult US Medicare beneficiaries with untreated obstructive sleep apnea are heavier users of health care than matched control patients. J Clin sleep Med JCSM. 2020 Jan;16(1):81–9.
14. Vo TN, Kats AM, Langsetmo L, Taylor BC, Schousboe JT, Redline S, et al. Association of sleep-disordered breathing with total healthcare costs and utilization in older men: the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) study. Sleep. 2020 Jan;43(1):zsz209.
15. Kao L-T, Lee H-C, Lin H-C, Tsai M-C, Chung S-D. Healthcare Service Utilization by Patients with Obstructive Sleep Apnea: A Population-Based Study. PLoS One. 2015;10(9):e0137459.
16. Jennum P, Kjellberg J. Health, social and economical consequences of sleep-disordered breathing: A controlled national study. Thorax [Internet]. 2011 [cited 2020 May 27];66(7):560–6. Available from: https://thorax.bmj.com/content/66/7/560
17. Diaz K, Faverio P, Hospenthal A, Restrepo MI, Amuan ME, Pugh MJ V. Obstructive sleep apnea is associated with higher healthcare utilization in elderly patients. Ann Thorac Med [Internet]. 2014 Apr;9(2):92–8. Available from: https://pubmed.ncbi.nlm.nih.gov/24791172/
18. Guest JF, Helter MT, Morga A, Stradling JR. Cost-effectiveness of using continuous positive airway pressure in the treatment of severe obstructive sleep apnoea/hypopnoea syndrome in the UK. Thorax. 2008 Oct;63(10):860–5.
19. Jennum P, Ibsen R, Petersen ER, Knudsen S, Kjellberg J. Health, social, and economic consequences of narcolepsy: A controlled national study evaluating the societal effect on patients and their partners. Sleep Med [Internet]. 2012 Sep [cited 2020 Jul 1];13(8):1086–93. Available from: https://pubmed.ncbi.nlm.nih.gov/22841027/
20. Black J, Reaven NL, Funk SE, McGaughey K, Ohayon M, Guilleminault C, et al. The Burden

of Narcolepsy Disease (BOND) study: Health-care utilization and cost findings. Sleep Med. 2014;15(5):522–9.
21. Villa KF, Reaven NL, Funk SE, McGaughey K, Black J. Changes in Medical Services and Drug Utilization and Associated Costs After Narcolepsy Diagnosis in the United States. Am Heal Drug Benefits [Internet]. 2018 [cited 2020 Jul 1];11(3):137. Available from: www.AHDBonline.com
22. Fu Y, Xia Y, Yi H, Xu H, Guan J, Yin S. Meta-analysis of all-cause and cardiovascular mortality in obstructive sleep apnea with or without continuous positive airway pressure treatment. Sleep Breath [Internet]. 2017 Mar 1 [cited 2020 Jul 6];21(1):181–9. Available from: https://link.springer.com/article/10.1007/s11325-016-1393-1
23. Li M, Hou WS, Zhang XW, Tang ZY. Obstructive sleep apnea and risk of stroke: A meta-analysis of prospective studies. Int J Cardiol [Internet]. 2014 Mar 15 [cited 2020 Jul 6];172(2):466–9. Available from: https://pubmed.ncbi.nlm.nih.gov/24452224/
24. Xia W, Huang Y, Peng B, Zhang X, Wu Q, Sang Y, et al. Relationship between obstructive sleep apnoea syndrome and essential hypertension: a dose–response meta-analysis. Sleep Med [Internet]. 2018 Jul 1 [cited 2020 Jul 8];47:11–8. Available from: https://pubmed.ncbi.nlm.nih.gov/29880142/
25. Reutrakul S, Mokhlesi B. Obstructive Sleep Apnea and Diabetes: A State of the Art Review [Internet]. Vol. 152, Chest. Elsevier Inc; 2017 [cited 2020 Jul 6]. p. 1070–86. Available from: /pmc/articles/PMC5812754/?report=abstract
26. Shawon MSR, Perret JL, Senaratna C V., Lodge C, Hamilton GS, Dharmage SC. Current evidence on prevalence and clinical outcomes of co-morbid obstructive sleep apnea and chronic obstructive pulmonary disease: A systematic review [Internet]. Vol. 32, Sleep Medicine Reviews. W.B. Saunders Ltd; 2017 [cited 2020 Jul 6]. p. 58–68. Available from: https://pubmed.ncbi.nlm.nih.gov/28169105/
27. Prasad B, Nyenhuis SM, Imayama I, Siddiqi A, Teodorescu M. Asthma and Obstructive Sleep Apnea Overlap: What Has the Evidence Taught Us? Am J Respir Crit Care Med [Internet]. 2020 Jun 1 [cited 2020 Jul 8];201(11):1345–57. Available from: https://www.atsjournals.org/doi/10.1164/rccm.201810-1838TR
28. Garbarino S, Bardwell WA, Guglielmi O, Chiorri C, Bonanni E, Magnavita N. Association of Anxiety and Depression in Obstructive Sleep Apnea Patients: A Systematic Review and Meta-Analysis [Internet]. Vol. 18, Behavioral Sleep Medicine. Routledge; 2020 [cited 2020 Jul 7]. p. 35–57. Available from: https://www.tandfonline.com/doi/abs/10.1080/15402002.2018.1545649
29. Bonsignore MR, Baiamonte P, Mazzuca E, Castrogiovanni A, Marrone O. Obstructive sleep apnea and comorbidities: A dangerous liaison [Internet]. Vol. 14, Multidisciplinary Respiratory Medicine. BioMed Central Ltd.; 2019 [cited 2020 Jul 8]. Available from: /pmc/articles/PMC6374907/?report=abstract
30. Panossian LA, Avidan AY. Narcolepsy and Other Comorbid Medical Illnesses. In: Goswami M, Thorpy MJ, Pandi-Perumal SR, editors. Narcolepsy: A Clinical Guide [Internet]. Second. Springer, Cham; 2015 [cited 2020 Jul 13]. p. 147–59. Available from: https://www.springer.com/gp/book/9783319237381
31. Cohen A, Mandrekar J, St. Louis EK, Silber MH, Kotagal S. Comorbidities in a community sample of narcolepsy. Sleep Med [Internet]. 2018 Mar 1 [cited 2020 Jul 8];43:14–8. Available from: https://pubmed.ncbi.nlm.nih.gov/29482805/
32. Ohayon MM. Narcolepsy is complicated by high medical and psychiatric comorbidities: a comparison with the general population. Sleep Med [Internet]. 2013 [cited 2020 Jul 10];14(6):488–92. Available from: https://pubmed.ncbi.nlm.nih.gov/23643648/
33. Black J, Reaven NL, Funk SE, McGaughey K, Ohayon MM, Guilleminault C, et al. Medical comorbidity in narcolepsy: findings from the Burden of Narcolepsy Disease (BOND) study. Sleep Med [Internet]. 2017 May 1 [cited 2020 Jul 13];33:13–8. Available from: https://pubmed.ncbi.nlm.nih.gov/28449892/
34. Flores NM, Villa KF, Black J, Chervin RD, Witt EA. The Humanistic and Economic Burden of Narcolepsy. J Clin Sleep Med [Internet]. 2016 Mar 15 [cited 2020 Jul 1];12(03):401–7. Available from: http://jcsm.aasm.org/doi/10.5664/jcsm.5594
35. Jennum P, Ibsen R, Knudsen S, Kjellberg J. Comorbidity and Mortality of Narcolepsy: A Controlled Retro- and Prospective National Study. Sleep [Internet]. 2013 Jun 1 [cited 2020 Jul 10];36(6):835–40. Available from: /pmc/articles/PMC3649826/?report=abstract
36. Lee MJ, Lee SY, Yuan SS, Yang CJ, Yang KC, Lee TL, et al. Comorbidity of narcolepsy and depressive disorders: a nationwide population-based study in Taiwan. Sleep Med [Internet]. 2017

Nov 1 [cited 2020 Jul 8];39:95–100. Available from: https://pubmed.ncbi.nlm.nih.gov/29157595/
37. Yeh JY, Shyu YC, Lee SY, Yuan SS, Yang CJ, Yang KC, et al. Comorbidity of Narcolepsy and Psychotic Disorders: A Nationwide Population-Based Study in Taiwan. Front Psychiatry [Internet]. 2020 Mar 25 [cited 2020 Jul 8];11. Available from: https://pubmed.ncbi.nlm.nih.gov/32269533/
38. Collin SM, Bakken IJ, Nazareth I, Crawley E, White PD. Health care resource use by patients before and after a diagnosis of chronic fatigue syndrome (CFS/ME): A clinical practice research datalink study. BMC Fam Pract [Internet]. 2017 May 5 [cited 2022 Jan 17];18(1):1–13. Available from: https://bmcprimcare.biomedcentral.com/articles/10.1186/s12875-017-0635-z
39. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. Br J Clin Pharmacol [Internet]. 2010 Jan [cited 2015 Jan 7];69(1):4–14. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2805870&tool=pmcentrez&rendertype=al
Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology [Internet]. 2009 [cited 2022 May 5];20(4):512. Available from: /pmc/articles/PMC3077219/
41. Tazare J, Smeeth L, Evans SJW, Williamson E, Douglas IJ. Implementing high-dimensional propensity score principles to improve confounder adjustment in UK electronic health records. Pharmacoepidemiol Drug Saf [Internet]. 2020 Nov 1 [cited 2022 May 5];29(11):1373–81. Available from: https://pubmed.ncbi.nlm.nih.gov/32926504/
42. Vernia F, Di Ruscio M, Ciccone A, Viscido A, Frieri G, Stefanelli G, et al. Sleep disorders related to nutrition and digestive diseases: a neglected clinical condition. Int J Med Sci [Internet]. 2021 [cited 2022 Apr 6];18(3):593–603. Available from: https://pubmed.ncbi.nlm.nih.gov/33437194/
43. Koo M, Hwang JH. Risk of tinnitus in patients with sleep apnea: A nationwide, population-based, case-control study. Laryngoscope [Internet]. 2017 Sep 1 [cited 2022 Apr 6];127(9):2171–5. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/lary.26323
44. West SD, Turnbull C. Obstructive sleep apnoea. Eye [Internet]. 2018 May 1 [cited 2022 Apr 6];32(5):889. Available from: /pmc/articles/PMC5944625/
45. Bassetti CLA, Adamantidis A, Burdakov D, Han F, Gay S, Kallweit U, et al. Narcolepsy — clinical spectrum, aetiopathophysiology, diagnosis and treatment [Internet]. Vol. 15, Nature Reviews Neurology. Nature Publishing Group; 2019 [cited 2020 Nov 4]. p. 519–39. Available from: https://pubmed.ncbi.nlm.nih.gov/31324898/
46. Soundararajan V, Lor J, Fishbein AB. Sleep Apnea and Skin. Curr sleep Med reports [Internet]. 2020 Sep 1 [cited 2022 Apr 8];6(3):94. Available from: /pmc/articles/PMC7986992/
47. Kirkwood BR, Sterne JAC. Comparing Two Rates. In: Essential Medical Statistics. 2nd ed. Oxford: Blackwell Science Ltd; 2003. p. 240–1.
48. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating Medical Costs from Incomplete Follow-Up Data. Biometrics. 1997 Jun;53(2):434.
49. Deb P, Norton EC. Modeling Health Care Expenditures and Use. Annu Rev Public Health [Internet]. 2018 Apr 1 [cited 2022 Apr 13];39:489–505. Available from: https://pubmed.ncbi.nlm.nih.gov/29328879/
50. Sjölander A, Greenland S. Ignoring the matching variables in cohort studies - when is it valid and why? Stat Med [Internet]. 2013 Nov 30 [cited 2022 Feb 25];32(27):4696–708. Available from:

https://pubmed.ncbi.nlm.nih.gov/23761197/
51. Pearce N. Analysis of matched case-control studies. BMJ [Internet]. 2016 Feb 25 [cited 2022 Feb 25];352. Available from: https://www.bmj.com/content/352/bmj.i969

Appendices

appendix-a.pdf

Grant ID
NIHR301730