










A Systematic Review of Codelists to Identify Chronic Bronchitis and Emphysema in Routine Electronic Healthcare Record Data and Derivation of a Standardized Codelist for Future Research

Harley Hiu Yung Kwok , Georgie May Massen , Alexander J Adamson , Constantinos Kallis , Adam Lahmami , Yaqoob Abbas Sheikh , Mohammed Hafzalla Elhag Elseddig , Nabaz Mutabchi , Jennifer Kathleen Quint 

School of Public Health, Imperial College London, London, UK

*These authors contributed equally to this work

Correspondence: Harley Hiu Yung Kwok, School of Public Health, Imperial College London, White City Campus, 90 Wood Lane, London, W12 0BZ, United Kingdom, Email h.kwok23@imperial.ac.uk

Abstract: Medical codes and codelists are essential to identify diseases, medicines, and related events when using electronic healthcare records (EHRs) for research. As multiple codes can be used to define the same medical condition, an inconsistency of applying codes by clinicians or researchers may complicate the identification of diseases. To minimize the inconsistency of codes and increase the comparability of research findings, we sought to integrate all medical codes and codelists generated from existing studies to produce a single list of codes for both chronic bronchitis and emphysema for future research. We systematically reviewed studies, which included codes used to define chronic bronchitis and/or emphysema regardless of code systems. We searched MEDLINE, Embase, Web of Science, Scopus, CINAHL, and Global Health from 1st January 2000 to May 2024 for relevant studies. Medical codes or codelists were identified, extracted, and compiled into single codelists for chronic bronchitis and emphysema separately based on their frequency of appearance across all included studies (PROSPERO: CRD42024529169). We identified 18 studies containing codelists, from a total of 1082 studies. Only International Classification of Disease (ICD) codes were found in studies, which defined chronic bronchitis and emphysema. Chronic bronchitis was defined as 490 and 491 in ICD-9 and J40, J41, J42, and J44 in ICD-10. Meanwhile, emphysema was defined as 492 in ICD-9 and J43 in ICD-10. Additionally, to fill the current gap, we created SNOMED CT codelists for both chronic bronchitis and emphysema, to standardize the definitions of these diseases in future studies.

Keywords: chronic obstructive pulmonary disease, bronchitis, emphysema, phenotype, code systems, disease definition

Introduction

Chronic bronchitis and emphysema are common phenotypes of chronic obstructive pulmonary disease (COPD). Given the significant burden of COPD worldwide, a significant number of research studies have been conducted in large-scale population-level electronic healthcare records (EHRs) to provide robust findings to solve public health problems.¹ In studies using EHRs, medical codes and codelists are used to identify diseases, symptoms, medicines, and procedures of interest.² Different coding systems are used in primary and secondary care. Common coding systems include systematized nomenclature of medicine – clinical terms (SNOMED CT), international classification of diseases (ICD), and Read codes. Clinical diagnoses or drugs can be coded in different ways, depending on the code systems used. Some code systems have greater granularity than others. Codes can contain information such as disease subtypes, stage, or disease



manifestations, accommodating varying levels of detail and specificity. For this reason, there may exist multiple codes for a disease within a coding system. This precision allows researchers to identify diseases with specific sub-conditions, potentially enhancing the generalizability of the study findings in specific populations.

Despite the advantages of having precise codes, the inconsistency of applying codes by clinicians in their work may complicate the identification of specific health conditions by researchers using EHRs. Regarding respiratory conditions, especially chronic bronchitis and emphysema, they are defined using various clinical codes,^{3,4} which may lead to considerable discrepancies in study findings and affect the comparability of findings.⁵ A systematic review on interstitial lung disease showed that studies using different case definitions defined by various codes produced a large difference in the estimated prevalence of disease even within the same country.⁶ The prevalence discrepancy inferred that the choice of code systems and medical codes for defining a medical condition was critical and a standardized set of codes should be explored to capture all possible cases of a condition of interest in the routine data, hence upholding the representativeness of the findings within the study population.

There are very few published codelists for emphysema or chronic bronchitis alone, and the majority of codes are contained within codelists for COPD. To identify chronic bronchitis and emphysema as independent diseases rather than subgroups of COPD, we systematically summarized the medical codes across coding systems for these two diseases separately. We aimed to integrate all codelists generated from various code systems into one document, from which we could produce a suggested list of codes for both chronic bronchitis and emphysema. We anticipate that the final list of codes could be used by researchers to identify chronic bronchitis and emphysema in EHRs in future studies.

Materials and Methods

We systematically searched the literature and code repositories to identify all codes relating to chronic bronchitis and emphysema. This study adopted a similar approach to MacRae et al's earlier work on deriving standardised codelists by searching existing literature and UK code repositories to identify and summarize disease codelists for asthma, COPD, and respiratory tract infections for respiratory research.⁷ The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were adhered to.⁸ The study protocol was registered with Prospective Register of Systematic Reviews (PROSPERO: CRD42024529169). Approval from an ethics committee was not required as we relied on published data only.

Data Sources and Searches in Literature

To identify relevant studies for the review, we searched MEDLINE (OVID) from 1st January 2000 to 9th May 2024, Embase (OVID) from 1st January 2000 to 9th May 2024, Web of Science from 1st January 2000 to 7th May 2024, Scopus from 2000 to 10th May 2024, CINAHL from 2006 to 10th May 2024, and Global Health (OVID) from 1st January 2000 to 10th May 2024. The detailed search strategy can be found in [Appendices 1–6](#). We also screened the reference lists of included studies to identify any additional papers. The search was limited to studies published in English. Searches were performed using key search terms for chronic bronchitis and emphysema separately. Our search strategy used similar terms in MacRae et al's study, such as medical records, medical record linkage, electronic health records, health informatics, and clinical coding, to identify observational secondary research studies using electronic health records and routine data. We included manuscripts containing the word “chronic obstructive pulmonary disease”, “chronic bronchitis” or “emphysema” in the title with at least one keyword relating to electronic health records or clinical coding. The search strategy was first developed in Medline and then adapted in searching other databases. We consulted medical librarians in developing search strategies throughout the process.

Study Selection

The search results from the different databases were exported to RIS file and imported to Covidence⁹ for further screening. We first removed duplicates from the database searches and subsequently performed title and abstract screening based on the inclusion and exclusion criteria. Two reviewers (any two of AL, MHEE, NM, and YAS) screened the titles and abstracts independently for inclusion of chronic bronchitis or emphysema separately. Any discrepancies for inclusion were resolved by discussion. A third reviewer (HHYK) was involved if a consensus was not reached. We

undertook full-text screening of the studies which potentially meet the eligibility but were not clearly discerned when screening the titles and abstracts alone. Supplementary material from the included studies was also reviewed to ensure we identified all codes. In the final stage, codelists published in the included studies were compiled into a Microsoft Excel spreadsheet. We documented the screening process using the PRISMA flowchart.⁸

Only peer-reviewed studies with full-text access were included. We included original observational population-based studies only, such as cohort, case control, and cross-sectional, if they reported codelists for chronic bronchitis and/or emphysema. We only included randomized controlled trials if they used EHRs for follow-up. We excluded all editorial, commentary, systematic review, government report, and pre-printed articles. Regarding the eligibility of study population, we only included literature with a study sample aged ≥ 18 years.

Data Extraction

Two reviewers (any two of AL, MHEE, NM, and YAS) independently extracted data from the included studies onto a data extraction form developed for this review. Information such as the codes used and how often they were used to define chronic bronchitis and emphysema in the literature was captured. The extraction form was piloted on four included studies prior to applying to all included studies. Disagreements were resolved by discussion between the two reviewers, if not reaching a consensus we involved a third reviewer (HHYK).

We collected the following items from the included studies: (1) general study information (authors, year of publication, study title, sample size, study sample characteristics, study period covered, and data source), (2) disease(s) identified (chronic bronchitis, emphysema, or both), (3) code systems used to identify disease(s) in routine data, and (4) medical codes of the disease. The specific codelists of interest included but were not limited to Read version 2 codes, SNOMED CT codes or Clinical Practice Research Datalink (CPRD) medcodeid codes, as well as ICD-9, ICD-10, and ICD-11 codes.

Searches in Code Repositories

In addition to the literature search, we searched existing code repositories including HDR UK Phenotype Library,¹⁰ the Cambridge University repository Apollo,¹¹ LSHTM Data Compass,¹² QResearch – QCode Group Library,¹³ ORCHID (Oxford-RCGP RSC),¹⁴ UK Biobank Data Showcase,¹⁵ and OpenSAFELY.¹⁶ We screened codelists and/or studies using the same inclusion criteria as used in the literature search. The included studies were added to the summary table, and all medical codes defining chronic bronchitis and/or emphysema were added to the list of codes identified from the literature search. Two reviewers (HHYK and GMM) independently extracted relevant information from the studies identified in code repositories search and cross-checked each other's work.

Data Synthesis and Analysis

We synthesized and presented a descriptive narrative summary for each included study by tabulating all data items extracted. We summarized data related to details of chronic bronchitis and emphysema identification. Data sources used in each publication and the country that the data sources originated from were identified. We extracted the codes used to define chronic bronchitis and emphysema from the literature and calculated the frequency of code used. We created codelists by including only the more frequently used codes for both chronic bronchitis and emphysema by code systems. We defined more frequently used as being used more than once across all included studies to ensure the codes were widely agreed beyond a single study under the same study context.

Quality assessment of included studies was not necessary for this review as we focused on summarizing the applied methodology instead of the findings of those studies. There is no existing critical appraisal tool or guidelines for papers aiming at deriving codelists either.

SNOMED CT Codelist Creation

Since researchers generally require SNOMED CT codes to extract cohorts specific to their study subjects from primary care EHR, we created two separate SNOMED CT codelists for chronic bronchitis and emphysema. We compiled these codelists using the National Health Service (NHS) SNOMED CT code browser, which is available in the Trusted

Reference Update Distribution (TRUD). We adopted similar methodologies suggested by Massen et al¹⁷ and followed the steps for codelist creation published on NHLI-SPH Respiratory EHR Group GitHub page (Link to page: NHLI-Respiratory-Epi/SNOMED-CT-codelists: How to: create SNOMED CT codelists (github.com)). We used an R package, “RDiagnosislist”, to identify SNOMED CT codes from the code browser and created codelists.¹⁸ This package contains SNOMED dictionary reserving the hierarchy of the medical codes.

We first identified search terms of codelists we wanted to create. Search terms of “chronic bronchitis” and “emphysema” were determined with the help of clinical experts. We then explored and extracted the relevant SNOMED CT concept ID from the SNOMED dictionary according to the search terms we identified using RDiagnosislist. Since there were no pre-existing codelists (using this terminology) for chronic bronchitis or emphysema, we did not include the step of comparing derived codelist with existing codelist. We directly exported the separate codelists of the two diseases and consulted medical experts for their opinions on which terms represented the true clinical conditions of patients with the diseases. We restricted the SNOMED CT codelists to include only those approved codes and finalized our codelists for both chronic bronchitis and emphysema.

Results

We identified 1,082 studies after the removal of duplicates from the six electronic literature databases, the reference lists of the included studies, and code repositories. We screened the full text of 782 articles, of which 18 studies were accepted for inclusion. A total of 764 studies were excluded due to: (1) not specifying the medical codes despite using EHRs in studies ($n = 278$), (2) focusing on chronic bronchitis and/or emphysema but not having specific medical codes for them ($n = 42$), (3) chronic bronchitis and/or emphysema not defined independently from COPD ($n = 315$), (4) not having full-text available ($n = 115$), (5) being written in non-English languages ($n = 3$), and (6) study designs outside of interest ($n = 11$) (Figure 1).

Study characteristics including settings, data sources, codelists and disease descriptions are presented in Table 1. We identified studies analyzing data from the United States (US), the United Kingdom (UK), Germany, Finland, Italy, Denmark, Sweden, Iceland, and China. Most studies were conducted in the US ($n = 8$). The UK and the Nordic countries including Finland, Denmark, Sweden, and Iceland used relatively large cohorts compared with those used in the rest of the countries. No codelist using the following terminologies were found: SNOMED CT, CPRD medcodeid, or Read version 2 code. All included studies used either ICD-9 ($n = 6$) or ICD-10 ($n = 8$), or both ($n = 4$) in identifying records of chronic bronchitis and/or emphysema in their analyses. Seven studies reported COPD in general but also specified the definition of chronic bronchitis and/or emphysema.^{19–25} While eight studies contained medical codes used to identify both chronic bronchitis and emphysema,^{26–33} three studies contained codes to define chronic bronchitis only.^{34–36}

Frequency of Medical Codes Used in Each Code System

A total of eight studies reported ICD-9 codes to identify chronic bronchitis.^{19,21,23,25,26,30,31,35} One study used only the ICD-9 code: 490²³ whilst four studies used ICD-9: 491 alone^{19,21,31,35} to define chronic bronchitis and three studies used both ICD-9: 490 and 491 to identify the disease.^{25,26,30} The described condition identified using these ICD-9 codes varied across studies. While most studies described the condition as chronic bronchitis, two studies using combined codes 490 and 491 additionally added “bronchitis, not specified as acute or chronic” for the description in their studies.^{25,26} Oelsner et al described the condition the most differently as “chronic obstructive bronchitis”.³⁰ Both ICD-9: 490 and ICD-9: 491 have been frequently used in studies for defining chronic bronchitis (Table 2).

Regarding the identification of emphysema using ICD-9 codes, a total of six studies used ICD-9: 492.^{19,21,23,26,30,31} There was no discrepancy in the descriptions from the studies that all studies described the condition as “emphysema” (Table 2). Westerling & Rosen combined the reporting of chronic bronchitis and emphysema²⁸ and Melzer et al described several respiratory diseases using five ICD-9 codes without specifying which code corresponded to which disease (Table 2).²⁷

Regarding studies using ICD-10 codes, nine and seven studies defined chronic bronchitis and emphysema, respectively. Six studies used ICD-10: J40, J41, and J42 to define chronic bronchitis.^{22,25,26,29,30,33} Zhong et al described the condition as “bronchitis”²⁹ and Oelsner et al described the condition as “chronic obstructive bronchitis”,³⁰ the remaining four studies^{22,25,26,33} described the condition according to the official ICD-10 terms: J40 (Bronchitis, not specified as

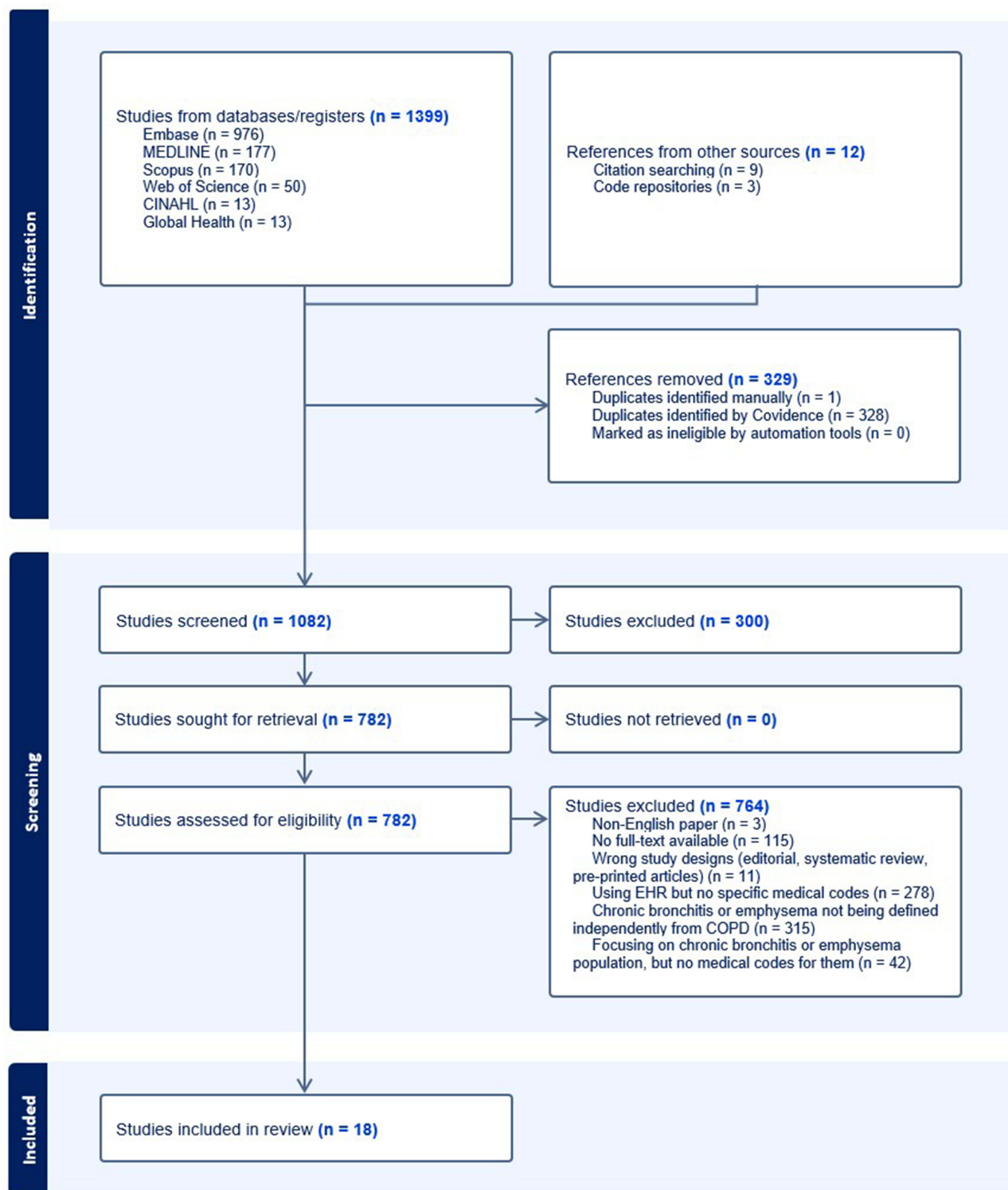


Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

acute or chronic), J41 (Simple and mucopurulent chronic bronchitis), and J42 (Unspecified chronic bronchitis).³⁷ Cook et al used only J41 and J42 to define chronic bronchitis.¹⁹ One study defined chronic bronchitis in a distinctive way, using the following codes, ICD-10: J43, J44, and J47 to describe chronic obstructive bronchitis.³⁴ ICD-10 codes of J40, J41, and J42 were frequently used to define chronic bronchitis in the included studies (Table 3).

Table 1 Background Characteristics of Included Studies (n = 18)

Study (Year of Publication)	Country Setting	Data Source(s)	Sample Size	Study Period Covered	Codelist(s) Reported	What Designated Disease Reported?
Jiang et al ³³ (2024)	United Kingdom	UK Biobank	427934	March 2022	ICD-10	Both Chronic bronchitis and Emphysema
Chan et al ²⁴ (2022)	China	China Kadoorie Biobank	512716	2004–2018	ICD-10	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Cook et al ¹⁹ (2022)	United States	Vanderbilt's Synthetic Derivative database, a deidentified version of Vanderbilt's electronic medical record	6065	1997–2017	ICD-9, ICD-10	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Loosen et al ²⁰ (2022)	Germany	Disease Analyzer database (IQVIA)	324492	2005–2019	ICD-10	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Nguyen et al ²⁶ (2022)	United States	Rochester Epidemiology Project records-linkage system	545	1988–2018	ICD-9, ICD-10	Both Chronic bronchitis and Emphysema
Carlsen et al ³⁶ (2021)	Iceland	The National Medicines Register; Primary care centers, and Landspítali (the National University Hospital)	215965	2010–2014	ICD-10	Chronic bronchitis only
Chu et al ²⁵ (2021)	United States	Mass General Brigham Biobank - MGB Research Patient Data Registry	282	1986 to data of analysis taken place	ICD-9, ICD-10	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Fromme et al ³² (2022)	United Kingdom	UK Biobank	482380	NA	ICD-10	Both Chronic bronchitis and Emphysema
Zhong et al ²⁹ (2021)	China	Xiamen Health and Medical Big Data Center	246568	2017–2019	ICD-10	Both Chronic bronchitis and Emphysema
Kivimaki et al ³⁴ (2020)	Finland	Two Finnish prospective cohort studies: 1) the Health and Social Support (HeSSup) study; 2) the Finnish Public Sector (FPS) study	109246	HeSSup study: 1998–2012 FPS study: 2000–2016	ICD-10	Chronic bronchitis only
Fedeli et al ²¹ (2017)	Italy	The archive of discharge records of public and private hospitals	27272	2008–2013	ICD-9	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Melzer et al ²⁷ (2017)	United States	Veterans Victory over Tobacco Study & Veterans Affairs (VA) National Patient Care Databases	3307	2009–2011	ICD-9	Both Chronic bronchitis and Emphysema
Hvidberg et al ²² (2016)	Denmark	Nationwide Danish public healthcare registers: 1) The Danish National Patient Register (NPR) 2) The Danish National Psychiatry Central Research Register (PCRR) 3) National Health Service Register (NHSR) 4) The Danish National Prescription Register (DNPR) 5) The Danish Civil Registration System (DCRS) and linkage to healthcare registers	4555439	1994–2013	ICD-10	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Oelsner et al ³⁰ (2016)	United States	1) Hispanic Community Health Study/Study of Latinos (HCHS/SOL) 2) Multi-Ethnic Study of Atherosclerosis (MESA)	HCHS/SOL: 16415 MESA: 6814	HCHS/SOL: 2008–2011 MESA: 2000–2013	ICD-9, ICD-10	Both Chronic bronchitis and Emphysema
Matthews et al ²³ (2014)	United States	DoD electronic medical records	371	2005–2009	ICD-9	COPD, whose codelist includes medical codes for Chronic bronchitis and/or Emphysema
Kuschner et al ³⁵ (2009)	United States	US Department of Veterans Palo Alto Health Care System	60	1999–2008	ICD-9	Chronic bronchitis only
Delea et al ³¹ (2008)	United States	PharMetrics Patient Centric Database	9217	1997–2005	ICD-9	Both Chronic bronchitis and Emphysema
Westerling & Rosen ²⁸ (2002)	Sweden	1) Swedish nationwide 1985 Population and Housing Census 2) Swedish Cause of Death Register	4812480	1986–1990	ICD-9	Both Chronic bronchitis and Emphysema

Abbreviations: ICD, International Classification of Diseases; COPD, chronic obstructive pulmonary disease; NA, not available.

Table 2 ICD-9 Codes Used to Define Chronic Bronchitis and Emphysema in Included Studies

Conditions	Study	ICD-9 Code	Description in Studies
Chronic bronchitis	Matthews et al ²³ (2014)	490	Chronic bronchitis
	Cook et al ¹⁹ (2022)	491	Chronic bronchitis
	Fedeli et al ²¹ (2017)	491	Chronic bronchitis
	Kuschner et al ³⁵ (2009)	491	Chronic bronchitis
	Delea et al ³¹ (2008)	491	Chronic bronchitis
	Nguyen et al ²⁶ (2022)	490, 491	Bronchitis, not specified as acute or chronic; Chronic bronchitis
	Chu et al ²⁵ (2021)	490, 491	Bronchitis, not specified as acute or chronic; Chronic bronchitis
Oelsner et al ³⁰ (2016)	490, 491	Chronic obstructive bronchitis	
Emphysema	Matthews et al ²³ (2014)	492	Emphysema
	Cook et al ¹⁹ (2022)	492	Emphysema
	Fedeli et al ²¹ (2017)	492	Emphysema
	Delea et al ³¹ (2008)	492	Emphysema
	Nguyen et al ²⁶ (2022)	492	Emphysema
	Oelsner et al ³⁰ (2016)	492	Emphysema
Combined reporting of disease	Melzer et al ²⁷ (2017)	491, 492, 493, 494, 496	Chronic bronchitis, chronic airway obstruction, emphysema, asthma, and bronchiectasis
	Westerling & Rosen ²⁸ (2002)	491, 492	Chronic bronchitis and emphysema

Abbreviation: ICD, International Classification of Diseases.

Table 3 ICD-10 Codes Used to Define Chronic Bronchitis and Emphysema in Included Studies

Condition	Study	ICD-10 Code	Description in Studies
Chronic bronchitis	Nguyen et al ²⁶ (2022)	J40	Bronchitis, not specified as acute or chronic
	Chu et al ²⁵ (2021)	J40	Bronchitis, not specified as acute or chronic
	Nguyen et al ²⁶ (2022)	J41	Simple and mucopurulent chronic bronchitis
	Chu et al ²⁵ (2021)	J41	Simple and mucopurulent chronic bronchitis
	Nguyen et al ²⁶ (2022)	J42	Unspecified chronic bronchitis
	Chu et al ²⁵ (2021)	J42	Unspecified chronic bronchitis
	Fromme et al ³² (2022)	J44	Chronic bronchitis
	Cook et al ¹⁹ (2022)	J41, J42	Chronic bronchitis
	Jiang et al ³³ (2024)	J40, J41, J42	Chronic bronchitis
	Zhong et al ²⁹ (2021)	J40, J41, J42	Bronchitis
	Hvidberg et al ²² (2016)	J40, J41, J42	Bronchitis, not specified as acute or chronic, simple and mucopurulent chronic bronchitis, and unspecified chronic bronchitis
	Oelsner et al ³⁰ (2016)	J40, J41, J42	Chronic obstructive bronchitis
	Kivimaki et al ³⁴ (2020)	J43, J44, J47	Chronic obstructive bronchitis
Emphysema	Jiang et al ³³ (2024)	J43	Emphysema
	Cook et al ¹⁹ (2022)	J43	Emphysema
	Nguyen et al ²⁶ (2022)	J43	Emphysema
	Fromme et al ³² (2022)	J43	Emphysema
	Zhong et al ²⁹ (2021)	J43	Emphysema
	Hvidberg et al ²² (2016)	J43	Emphysema
	Oelsner et al ³⁰ (2016)	J43	Emphysema
Combined reporting of disease	Loosen et al ²⁰ (2022)	J42, J43, J44	Chronic bronchitis and COPD
	Carlsen et al ³⁶ (2021)	J30-J39	Chronic bronchitis and sinusitis

Abbreviations: ICD, International Classification of Diseases; COPD, chronic obstructive pulmonary disease.

A total of seven studies used ICD-10: J43 to define emphysema and there was no discrepancy in the description of the condition in these studies (Table 3).^{19,22,26,29,30,32,33} While Loosen et al combined the reporting of chronic bronchitis with COPD using ICD-10: J42, J43, and J44,²⁰ Carlsen et al used J30-J39 to report chronic bronchitis and sinusitis simultaneously.³⁶

The finalized ICD codes for chronic bronchitis and emphysema are shown in Table 4. We recommend defining chronic bronchitis using 490 and 491 in ICD-9, as well as J40, J41, J42, and J44 in ICD-10. Meanwhile, we recommend that emphysema is defined using 492 in ICD-9 and J43 in ICD-10.

Codelist Development

We identified a total of 31 and 23 SNOMED CT codes to indicate events of chronic bronchitis and emphysema respectively after consulting clinical experts on the terms. The SNOMED CT codelists for both diseases are shown in Tables 5 and 6.

Table 4 Finalized ICD Codes for Chronic Bronchitis and Emphysema

	ICD-9	ICD-10
Chronic bronchitis	490, 491	J40, J41, J42, J44
Emphysema	492	J43

Abbreviation: ICD, International Classification of Diseases.

Table 5 SNOMED CT Codelist for Chronic Bronchitis

SNOMED CT Concept ID	Medical Term
8141004	Purulent chronic bronchitis (disorder)
61937009	Simple chronic bronchitis (disorder)
63480004	Chronic bronchitis (disorder)
71435009	Chronic obstructive bronchitis (disorder)
74417001	Mucopurulent chronic bronchitis (disorder)
84409004	Fetid chronic bronchitis (disorder)
89549007	Catarrhal bronchitis (disorder)
155566007	Chronic bronchitis (disorder)
155568008	Mucopurulent chronic bronchitis (disorder)
185086009	Chronic obstructive bronchitis (disorder)
195941006	Simple chronic bronchitis (disorder)
195942004	Simple chronic bronchitis (disorder)
195943009	Simple chronic bronchitis NOS (disorder)
195944003	Mucopurulent chronic bronchitis (disorder)
195945002	Purulent chronic bronchitis (disorder)
195946001	Purulent chronic bronchitis (disorder)
195947005	Mucopurulent chronic bronchitis NOS (disorder)

(Continued)

Table 5 (Continued).

SNOMED CT Concept ID	Medical Term
195952000	Obstructive chronic bronchitis NOS (disorder)
195953005	Mixed simple and mucopurulent chronic bronchitis (disorder)
195954004	Other chronic bronchitis (disorder)
195955003	Other chronic bronchitis NOS (disorder)
195956002	Chronic bronchitis NOS (disorder)
735465007	Protracted bronchitis caused by bacterium (disorder)
785736001	Chronic bronchitis co-occurrent with wheeze (disorder)
785737005	Bronchitis co-occurrent with chronic wheeze (disorder)
601581000000100	Obstructive chronic bronchitis NOS (disorder)
601591000000103	Other chronic bronchitis (disorder)
621061000000108	Other chronic bronchitis NOS (disorder)
621071000000101	Chronic bronchitis NOS (disorder)
649131000000101	Simple chronic bronchitis NOS (disorder)
649141000000105	Mucopurulent chronic bronchitis NOS (disorder)

Table 6 SNOMED CT Codelist for Emphysema

SNOMED CT Concept ID	Medical Term
4981000	Panacinar emphysema (disorder)
16003001	Giant bullous emphysema (disorder)
16846004	Obstructive emphysema (disorder)
31898008	Paraseptal emphysema (disorder)
45145000	Unilateral emphysema (situation)
49158009	Emphysema (morphologic abnormality)
68328006	Centriacinar emphysema (disorder)
87433001	Pulmonary emphysema (disorder)
125294002	Acute emphysema (morphologic abnormality)
125295001	Chronic emphysema (morphologic abnormality)
155573002	Emphysema (disorder)
195957006	Chronic bullous emphysema (disorder)
195958001	Segmental bullous emphysema (disorder)
195959009	Zonal bullous emphysema (disorder)
195961000	Chronic bullous emphysema NOS (disorder)
195964008	Emphysema: [acute interstitial] or [atrophic - senile] (disorder)

(Continued)

Table 6 (Continued).

SNOMED CT Concept ID	Medical Term
195966005	Emphysema NOS (morphologic abnormality)
233674008	Pulmonary emphysema in alpha-1 primary immunodeficiency deficiency (disorder)
266356006	Atrophic (senile) emphysema (disorder)
1010333003	Emphysema of left lung (disorder)
1010334009	Emphysema of right lung (disorder)
579661000000108	Chronic bullous emphysema NOS (disorder)
589221000000106	Emphysema NOS (morphologic abnormality)

Discussion

We conducted a systematic review of articles reporting the use of medical codes to identify chronic bronchitis and emphysema in EHRs. We suggested finalizing lists of codes for both chronic bronchitis and emphysema separately by summarizing the more frequently used medical codes. For ICD-9, 490 and 491 were found to be mostly used in the included articles to define chronic bronchitis, while 492 was consistently used to define emphysema. For ICD-10, J40-J42 and J44 were more frequently used to define chronic bronchitis. Similar to 492 in ICD-9, J43 alone was consistently used across all included articles to define emphysema. SNOMED CT codelists for defining chronic bronchitis and emphysema were also created.

This review is the first-of-a-kind study which systematically summarized the medical codes used to define chronic bronchitis and emphysema independently from defining COPD across studies using EHRs. Previous work has focused on updating and creating codelists for the identification of respiratory diseases with a broader definition, such as COPD and asthma.⁷ Under the umbrella of COPD, there are a few of conditions being largely overlooked as standalone conditions. Emphysema is a cause of COPD but not all people with emphysema develop COPD.³⁸ Similarly, around 6.2% of people with chronic bronchitis did not diagnosed with COPD.³⁹ The prevalence of chronic bronchitis also varied due to varying definitions of the disease,⁴⁰ notwithstanding that of emphysema could be more consistent as it has less variation in definition. Since chronic bronchitis and emphysema are conditions worth studying the impacts on patients' disease management and quality of life, specific codelists for these diseases should be created to fit studies tailored for these diseases using EHRs.

In this review, only ICD-9 and ICD-10 codes identifying chronic bronchitis and emphysema were found. Some common code systems used in primary care such as SNOMED CT and Read Codes were not recorded in the included studies which defined chronic bronchitis or emphysema. Nevertheless, ICD codelists are essential when using linked secondary care data, alongside primary care data, such as using CPRD databases alongside Hospital Episode Statistics. Secondary care databases such as Hospital Episode Statistics and Office for National Statistics death data use ICD-10 codes. Studies focusing on COPD populations used code systems beyond ICD to identify the disease. When searching code repositories, COPD was found to be defined by well-developed and validated codelists regardless of code systems. To strengthen the research quality on studying chronic bronchitis and emphysema, we created SNOMED CT codelists for these two diseases after gathering the information from ICD codes.

Different combinations of ICD-9 and ICD-10 codes were used to define the same disease across all included studies. For instance, some studies used either J40, J41, or J42 to identify chronic bronchitis, while some defined the disease using a combination of the codes. Although we identified codes, which were used more frequently than others, inconsistency of the codes used to define chronic bronchitis and emphysema in studies led to an extent of confusion, as manuscripts referred to the same medical condition whilst the codes used to define the condition varied. We only selected codes which were used in multiple studies to avoid instances where codes were used for a specific context in

some studies and were therefore less generalizable. We aimed to keep the codes included in the finalized codelists as generalized as possible.

We have not performed a quality assessment as currently there is no standardized appraisal tool to assess study quality in terms of methodology used to identify diseases using medical codes. A common practice for studies aiming to derive standardized codelists for diseases is to examine the frequency of code usage. We treated all codes equally regardless of whether the paper reported the methodology for collating clinical codes when we compiled the codes into a single codelist. Although this approach may compromise clinical precision as it does not account for potential inconsistencies in the code selection process by different researchers, the consistency of medical codes observed across the included studies for defining chronic bronchitis and emphysema was reassuring. The ICD codes included in our finalized codelists uphold their quality as much as possible.

Some limitations in our review should be noted. First, a variety of ICD codes were used to define chronic bronchitis across the included studies. To minimize the bias caused by study heterogeneity in terms of study aims or data sources, we only included medical codes appearing more than once into the finalized codelist. This may have impacted our results by not including all relevant terms into the derived codelist, yet it upheld the quality of codes being used and will be useful in future studies. Second, we screened the full-text of more than one hundred papers, which is uncommon in systematic reviews. Due to the nature and purpose of this review, however, we screened papers with titles relating to not only chronic bronchitis and emphysema, but also COPD in general and morbidities concurrent with COPD. We did so to retain all potential papers containing codes and/or codelists for the two diseases. We aimed to keep this review as robust as possible, despite the progress being laborious.

Conclusion

We summarized ICD codelists and developed SNOMED CT codelists for both chronic bronchitis and emphysema separately. Clinicians and researchers can use these standardized codelists in their studies to identify chronic bronchitis and/or emphysema irrespective of COPD diagnosis when using EHRs for research. The derived codelists, especially that of chronic bronchitis, can ensure the consistency of the disease definition, striving to produce high quality, comparable study findings.

Abbreviations

COPD, chronic obstructive pulmonary disease; CPRD, Clinical Practice Research Datalink; EHR, electronic health record; ICD, International Classification of Diseases; NHS, National Health Service; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SNOMED CT, systemized nomenclature of medicine – clinical terms; TRUD, Trusted Reference Update Distribution; UK, United Kingdom.

Data Sharing Statement

All data relevant to the study are included in the article or uploaded as [Supplementary Information](#). All works included in this analysis are referenced in the [Supplementary Information Document](#). No additional data not located within the manuscripts were used.

Ethics Approval and Informed Consent

Ethics approval or informed consent was not required for this work as we used published data only.

Consent for Publication

We consent that all contents including images in this work can be published.

Funding

This work is supported by Health Data Research UK Driver Program and the National Respiratory Audit Programme. The funders had no role in study design, data collection, analysis or interpretation, or manuscript writing. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Disclosure

HHYK, GMM, AJA, CK, AL, YAS, MHEE, and NM report no conflicts of interest in this work. JKQ reports grants from Medical Research Council, Health Data Research UK, GlaxoSmithKline, Bayer, Boehringer Ingelheim, asthma+lung UK, Chiesi and AstraZeneca and personal fees for advisory board participation or speaking fees from GSK, AZ, Chiesi and Insmad.

References

1. Massen GM, Blamires O, Grainger M, et al. UK electronic healthcare records for research: a scientometric analysis of respiratory, cardiovascular, and COVID-19 publications. *Pragmat Obs Res.* 2024;15:151–164. doi:10.2147/por.S469973
2. Bennett Institute for Applied Data Science. OpenSAFELY documentation - Introduction to codelists. University of Oxford. 2024. Available from: <https://docs.opensafely.org/codelist-intro/>. Accessed March 26, 2024.
3. Jennifer KQ, Hana M, Rachael LD, et al. Validation of chronic obstructive pulmonary disease recording in the clinical practice research datalink (CPRD-GOLD). *BMJ Open.* 2014;4(7):e005540. doi:10.1136/bmjopen-2014-005540
4. Simpson CR, Anandan C, Fischbacher C, Lefevre K, Sheikh A. Will systematized nomenclature of medicine-clinical terms improve our understanding of the disease burden posed by allergic disorders? *Clin Exp Allergy.* 2007;37(11):1586–1593. doi:10.1111/j.1365-2222.2007.02830.x
5. Tate AR, Dungey S, Glew S, Beloff N, Williams R, Williams T. Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect incidence estimates? Cross-sectional study using the CPRD database. *BMJ Open.* 2017;7(1):e012905. doi:10.1136/bmjopen-2016-012905
6. Gupta RS, Koteci A, Morgan A, George PM, Quint JK. Incidence and prevalence of interstitial lung diseases worldwide: a systematic literature review. *BMJ Open Respir Res.* 2023;10(1):e001291. doi:10.1136/bmjresp-2022-001291
7. MacRae C, Whittaker H, Mukherjee M, et al. Deriving a standardised recommended respiratory disease codelist repository for future research. *Pragmat Obs Res.* 2022;13:1–8. doi:10.2147/por.S353400
8. Moher DL, Alessandro T, Jennifer A, Douglas G. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Internal Med.* 2009;151(4):264–269. doi:10.7326/0003-4819-151-4-200908180-00135
9. *Covidence Systematic Review Software.* Melbourne, Australia: Veritas Health Innovation; 2024.
10. HDR UK Phenotype Library. HDR UK, 2024. Available from: <https://phenotypes.healthdatagateway.org/>. Accessed June 27, 2024.
11. University of Cambridge. Apollo - University of Cambridge Repository. University of Cambridge. Available from: <https://www.repository.cam.ac.uk/home>. Accessed August 14, 2024.
12. LSHTM data compass. London School of Hygiene and Tropical Medicine. Available from: <https://datacompass.lshtm.ac.uk/>. Accessed July 30, 2024.
13. QResearch. QCode group library. QResearch. Available from: <https://www.qresearch.org/data/qcode-group-library/>. Accessed August 14, 2024.
14. University of Oxford. ORCHID. University of Oxford. Available from: <https://orchid.phc.ox.ac.uk/>. Accessed August 14, 2024
15. UK Biobank. biobank - Search. University of Oxford. Available from: <https://biobank.ndph.ox.ac.uk/showcase/search.cgi>. Accessed August 14, 2024.
16. University of Oxford. OpenSAFELY - Secure analytics platform for NHS electronic health records. OpenSAFELY. Available from: <https://www.opensafely.org/>. Accessed August 14, 2024.
17. Massen GM, Stone PW, Kwok HHY, et al. Review of codelists used to define hypertension in electronic health records and development of a codelist for research. *Open Heart.* 2024;11(1):e002640. doi:10.1136/openhrt-2024-002640
18. Shah A. Manipulate SNOMED CT Diagnosis Lists. CRAN Repository. 2023. Available from: <https://cran.r-project.org/web/packages/Rdiagnosislist/Rdiagnosislist.pdf>. Accessed September 27, 2024.
19. Cook DP, Xu M, Martucci VL, et al. Clinical insights into pulmonary hypertension in chronic obstructive pulmonary disease. *Pulmonary Circulation.* 2022;12(1):e12006. doi:10.1002/pul2.12006
20. Loosen SH, Roderburg C, Curth O, et al. The spectrum of comorbidities at the initial diagnosis of heart failure a case control study. *Sci Rep.* 2022;12(1):2670. doi:10.1038/s41598-022-06618-5
21. Fedeli U, De Giorgi A, Gennaro N, et al. Lung and kidney: a dangerous liaison? A population-based cohort study in COPD patients in Italy. *Int J Chron Obstruct Pulmon Dis.* 2017;12:443–450. doi:10.2147/copd.S119390
22. Hvidberg MF, Johnsen SP, Glümer C, Petersen KD, Olesen AV, Ehlers L. Catalog of 199 register-based definitions of chronic conditions. *Scand J Public Health.* 2016;44(5):462–479. doi:10.1177/1403494816641553
23. Matthews T, Abraham J, Zacher LL, Morris MJ. The impact of deployment on COPD in active duty military personnel. *Mil Med.* 2014;179(11):1273–1278. doi:10.7205/milmed-d-14-00066
24. Chan KH, Wright N, Xiao D, et al. Tobacco smoking and risks of more than 470 diseases in China: a prospective cohort study. *Lancet Public Health.* 2022;7(12):e1014–e1026. doi:10.1016/S2468-2667(22)00227-4
25. Chu SH, Wan ES, Cho MH, et al. An independently validated, portable algorithm for the rapid identification of COPD patients using electronic health records. *Sci Rep.* 2021;11(1):19959. doi:10.1038/s41598-021-98719-w
26. Nguyen TT, Smith CY, Gazzuola Rocca L, Rocca WA, Vassallo R, Dulohery Scrodin MM. A population-based cohort study on the risk of obstructive lung disease after bilateral oophorectomy. *NPJ Prim Care Respir Med.* 2022;32(1):52. doi:10.1038/s41533-022-00317-4
27. Melzer AC, Clothier BA, Japuntich SJ, et al. Comparative effectiveness of proactive tobacco treatment among smokers with and without chronic lower respiratory disease. *Ann Am Thoracic Soc.* 2018;15(3):341–347. doi:10.1513/AnnalsATS.201707-582OC
28. Westerling R, Rosén M. 'Avoidable' mortality among immigrants in Sweden. *Eur J Public Health.* 2002;12(4):279–286. doi:10.1093/eurpub/12.4.279
29. Zhong TL, Zhang YW, Yao XY, et al. Association between atmospheric concentration of particulate matters and inpatient and outpatient visits for chronic respiratory diseases in Xiamen, China. *Chin Med J.* 134(13):1616–1618. doi:10.1097/cm9.0000000000001435

30. Oelsner EC, Loehr LR, Henderson AG, et al. Classifying chronic lower respiratory disease events in epidemiologic cohort studies. *Ann Am Thorac Soc*. 2016;13(7):1057–1066. doi:10.1513/AnnalsATS.201601-063OC
31. Delea TE, Hagiwara M, Dalal AA, Stanford RH, Blanchette CM. Healthcare use and costs in patients with chronic bronchitis initiating maintenance therapy with fluticasone/salmeterol vs other inhaled maintenance therapies. *Curr Med Res Opin*. 2009;25(1):1–13. doi:10.1185/03007990802534020
32. Fromme M, Schneider CV, Pereira V, et al. Hepatobiliary phenotypes of adults with alpha-1 antitrypsin deficiency. *Gut*. 2022;71(2):415–423. doi:10.1136/gutjnl-2020-323729
33. Jiang Y, Li H, Li Z, et al. A cross-trait study of lung cancer and its related respiratory diseases based on large-scale exome sequencing population. *Transl Lung Cancer Res*. 2024;13(3):512–525. doi:10.21037/tlcr-24-4
34. Kivimäki M, Batty GD, Pentti J, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. *Lancet Public Health*. 2020;5(3):e140–e149. doi:10.1016/S2468-2667(19)30248-8
35. Kuschner WG, Hegde S, Agrawal M. Occupational history quality in patients with newly documented, clinician-diagnosed chronic bronchitis. *Chest*. 2009;135(2):378–383. doi:10.1378/chest.08-1559
36. Carlsen HK, Ilyinskaya E, Baxter PJ, et al. Increased respiratory morbidity associated with exposure to a mature volcanic plume from a large Icelandic fissure eruption. *Nat Commun*. 2021;12(1):2161. doi:10.1038/s41467-021-22432-5
37. World Health Organization. International statistical classification of diseases and related health problems 10th Revision (ICD-10)-WHO Version for;2019-covid-expanded. World Health Organization. 2024. Available from: <https://icd.who.int/browse10/2019/en#/J40-J47>. Accessed August 27, 2024.
38. Cleveland Clinic. Emphysema. Cleveland Clinic. Available from: <https://my.clevelandclinic.org/health/diseases/9370-emphysema>. Accessed October 16, 2024.
39. de Oca MM, Halbert RJ, Lopez MV, et al. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. *Eur Respir J*. 2012;40(1):28–36. doi:10.1183/09031936.00141611
40. Kim V, Criner GJ. The chronic bronchitis phenotype in chronic obstructive pulmonary disease: features and implications. *Curr Opin Pulm Med*. 2015;21(2):133–141. doi:10.1097/mcp.000000000000145

International Journal of Chronic Obstructive Pulmonary Disease

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress
Taylor & Francis Group