CHAPTER 6

A SYSTEMATIC REVIEW OF CASE-IDENTIFICATION ALGORITHMS BASED ON ITALIAN HEALTHCARE ADMINISTRATIVE DATABASES FOR TWO RELEVANT DISEASES OF THE RESPIRATORY SYSTEM: ASTHMA AND CHRONIC OBSTRUCTIVE PULMONARY DISEASE

Riccardo Di Domenicantonio,[1] Giovanna Cappai,[1] Mirko Di Martino,[1] Nera Agabiti,[1] Lorenzo Simonato,[2] Cristina Canova,[2] Claudio Barbiellini Amidei[2]
[1] Department of Epidemiology, Lazio Regional Health Service, ASL Roma 1, Rome (Italy)
[2] Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padua, Padua (Italy)
Corresponding author: Riccardo Di Domenicantonio; r.didomenicantonio@deplazio.it

ABSTRACT

OBJECTIVES: to identify and describe all asthma and Chronic Obstructive Pulmonary Disease (COPD) case-identification algorithms by means of Italian Healthcare Administrative Databases (HADs), through the review of papers published in the past 10 years.

METHODS: this study is part of a project that systematically reviewed case-identification algorithms for 18 acute and chronic conditions by means of HADs in Italy.

PubMed was searched for original articles, published between 2007 and 2017, in Italian or English. The search string consisted of a combination of free text and MeSH terms with a common part that focused on HADs and a disease-specific part.

All identified papers were screened by two independent reviewers; exclusion criteria were the following: no description of reported algorithms, algorithm developed outside of the Italian context, exclusive use of death certificates, pathology register, general practitioner or pediatrician data. Pertinent papers were classified according to the objective for which the algorithm had been used, and only articles that used algorithms for primary objectives (**I** disease occurrence; **II** population/cohort selection; **III** outcome identification) were considered for algorithm extraction. The HADs used (hospital discharge records, drug prescriptions, etc.), ICD-9 and ICD-10 codes, ATC classification of drugs, follow-back periods, and age ranges applied by the algorithms have been reported. Further information on specific objective(s), accuracy measures, sensitivity analyses and the contribution of each HAD, have also been recorded.

RESULTS: the search string led to the identification of 98 and 147 papers, respectively for asthma and COPD. By screening the references, 2 papers for asthma and 7 for COPD were added. At the end of the screening process, 14 pertinent papers were identified for asthma and 31 for COPD. Half of these used healthcare data covering a time period between 2008 and 2014. More than 75% considered the age range 6-17 for asthma and ≥45 for COPD. About one-third of the articles used algorithms to estimate the occurrence of these diseases.

Fourteen algorithms for asthma and 16 for COPD were extracted from the papers and characterized. The Drug Prescription Database (DPD) was used by almost all asthma case-iden-

WHAT IS ALREADY KNOWN

- **Determining occurrence and assessing quality of care for asthma and COPD are crucial elements to make accurate healthcare policy decisions.**
- **Several studies relying on case-identification algorithms, based on Italian healthcare administrative databases, have been published.**
- **No standard evaluation of the characteristics of each algorithm is available and comparison between results is therefore hampered.**

WHAT THIS PAPER ADDS

- **A complete overview of all the literature published in the past 10 years (2007-2017) on asthma and COPD case-identification algorithms, based on Italian healthcare administrative databases.**
- **The use of healthcare administrative databases to identify cases affected by asthma or COPD is relatively vast and algorithms are quite heterogeneous.**
- **For asthma, drug prescriptions (DPs) play a crucial role in almost all algorithms. Since asthma and COPD have a relevant therapeutic overlap, but also a certain age-specificity, cut-offs based on age have been extensively used.**
- **For COPD, the characteristics of the disease and patients' scarce adherence to therapy contribute to the low sensitivity of algorithms based on drug prescriptions. Hospital discharge records are a reliable source and can be used to select cohorts of patients.**
- **Algorithm validation processes for both these diseases are scarce and no standardized approach to estimate accuracy measures has yet been defined.**

tification algorithms, while only 7 COPD algorithms used this data source. The spectrum of active ingredients was strongly overlapping between the two diseases, with different combinations of drugs and administration routes, as well as specific number of prescriptions, follow-back years, and age ranges. Age class and chronic treatment were the main disease-specific traits that emerged from the algorithms.
Three external validation processes have been performed for asthma and three for COPD. High accuracy levels have been found for asthma. COPD sensitivity analyses were unsatisfactory, while a high specificity was found for algorithms based on hospital discharge records.
CONCLUSION: elements from the review on the use of healthcare administrative databases represent a useful tool to decide which algorithm to adopt, based on the algorithm's individual requirements, limits, and accuracy, taking into account the specific research objective.

Keywords: Algorithms, healthcare administrative databases, asthma, chronic obstructive pulmonary disease, COPD.

RIASSUNTO

OBIETTIVI: identificare e descrivere tutti i lavori pubblicati negli ultimi 10 anni che, utilizzando flussi amministrativi sanitari (FAS) italiani, hanno elaborato almeno un algoritmo originale per l'identificazione di soggetti affetti da asma e broncopneumopatia cronica ostruttiva (BPCO).
METODI: questo studio si inserisce all'interno di un progetto di 16 revisioni sistematiche per la valutazione dello stato dell'arte degli algoritmi per l'identificazione di 18 patologie acute e croniche. La revisione, effettuata da due revisori indipendenti, mira a identificare articoli originali pubblicati tra il 2007 e il 2017 in inglese o italiano, individuati su PubMed mediante una stringa di ricerca costituita sia da testo libero che da termini MeSH, con una parte comune a tutte le patologie e una parte patologia-specifica. I lavori pertinenti sono stati classificati a seconda dell'obiettivo per cui ciascun algoritmo è stato utilizzato e si sono estratti i dati solo dagli algoritmi con obiettivi primari (**I** occorrenza di malattia; **II** selezione di coorti/popolazioni; **III** identificazione di outcome). I criteri di esclusione erano i seguenti: assenza di una descrizione degli algoritmi riportati; sviluppo dell'algoritmo al di fuori del contesto italiano; uso esclusivo di: certificate di morte, registri di patologia, dati dei medici di medicina generali o dei pediatri di libera scelta.Le informazioni estratte per caratterizzare e confrontare gli algoritmi originali sono: i FAS utilizzati (schede di dimissione ospedaliera, prescrizioni farmaceutiche, etc.), i codici ICD-9 e ICD-10, la selezione dei farmaci secondo il sistema di classificazione ATC, i criteri di identificazione dei casi, il periodo di osservazione/follow-back, i criteri di selezione anagrafica applicati ed eventuali validazioni esterne con le relative misure di accuratezza (sensibilità, specificità, valori predittivi) riportate.
RISULTATI: la stringa di ricerca ha portato all'identificazione di 98 e 147 articoli, rispettivamente per asma e BPCO, con l'aggiunta dai riferimenti bibliografici di 2 articoli per l'asma e 7 per la BPCO. Alla fine del processo di selezione, sono stati identificati 14 lavori pertinenti per l'asma e 31 per la BPCO. La metà di questi hanno utilizzato dati sanitari relativi al periodo 2008-2014. Più del 75% ha considerato limiti d'età compresi tra 6-17 anni per l'asma e ≥45 anni per la BPCO. Quasi un terzo degli articoli hanno utilizzato gli algoritmi per stimare l'occorrenza di queste patologie.
Quattordici algoritmi per l'asma e 16 per la BPCO sono stati estratti dagli articoli e caratterizzati. I dati relativi alle prescrizioni farmaceutiche sono stati utilizzati da quasi tutti gli algoritmi per l'identificazione dell'asma, mentre sono stati considerati soltanto da 7 algoritmi per la BPCO. Lo spettro dei principi attivi impiegati nel trattamento di queste due patologie era in gran parte sovrapponibile, con differenze nella combinazione dei farmaci, nella via di somministrazione, così come nel numero di prescrizioni, anni di follow-back e gruppi di età. Proprio l'età e la terapia cronica costituiscono i principali elementi distintivi che sono emersi dalla revisione.
Sono state eseguite 3 validazioni esterne per l'asma e 3 per la BPCO. Si sono riscontrati elevati livelli di accuratezza per gli algoritmi relativi all'asma. La valutazione della sensibilità degli algoritmi per la BPCO ha invece mostrato risultati poco soddisfacenti, mentre si è evidenziata un'elevata specificità per gli algoritmi basati sulle schede di dimissione ospedaliera.
CONCLUSIONE: gli elementi emersi dalla revisione sull'uso dei FAS rappresentano uno strumento utile per decidere quali algoritmi scegliere, sulla base dei requisiti, dei limiti e dell'accuratezza di ogni singolo algoritmo, tenendo in considerazione gli specifici obiettivi di ricerca.

Parole chiave: algoritmi, dati amministrativi sanitari, asma, broncopneumopatia cronica ostruttiva, BPCO

INTRODUCTION

Asthma and chronic obstructive pulmonary disease (COPD) are two highly impactful chronic respiratory conditions in present society. Guidelines are being regularly updated and research working toward innovating specific drugs is constant.[1,2] These two chronic conditions require a long-term follow-up of patients. Asthma prevalently affects children and young adults and is characterized by frequent relapses, with a pharmaceutical treatment that aims at resolving acute attacks. COPD, on the other hand, mainly affects adults and elderly people and is a progressive condition that requires a continuous treatment, with clinical exacerbations that require acute care. Despite these different characteristics, a certain clinical and therapeutic overlap exists.
Determining prevalence and incidence of these diseases and enrolling cohorts of patients to assess quality of care would be of great importance to make accurate healthcare policy decisions and make it possible to set up a monitoring program on a population level. Healthcare administra-

tive data has been extensively used in Italy as source of information for these purposes.[3]

A number of studies in Italy reviewed algorithms to identify several chronic and acute conditions from healthcare administrative data;[4] some of these studies focused on asthma and/or COPD,[5-8] but these experiences only included validated strategies. Due to the importance and relatively vast use of algorithms in Italy, a systematic review of their characteristics and fields of utilization, performed through a comprehensive search strategy (not restricted only to validated work) is needed.

The objective of this systematic review is to describe the characteristics of algorithms that have actually been used in the past 10 years, in Italy, for asthma and COPD case identification.

METHODS

This review considered research articles published either in English or in Italian on PubMed/Medline, between January 1, 2007 and December 31st, 2017. All details on methods are available in a specific paper,[9] which reports the study protocol with complete information on the literature search (specific search string applied to retrieve administrative healthcare data papers, inclusion / exclusion criteria and data extraction), characterization of selected papers and algorithms (strategy to identify original algorithms, algorithm objective definition). The search string used to select PubMed records consisted of a part optimized to retrieve papers focused on Italian administrative healthcare data and a specific part for the condition under study, reported in box 1. We chose to use a single database (PubMed/Medline) for the literature search, as we believe that the types of papers to be included in the systematic review are published in journals indexed in this database. Moreover, all the bibliographic references in the identified articles are checked and relevant studies not identified by the search string are included.

Two independent researchers screened the articles and classified pertinent ones, according to the objective for which the papers' algorithms were used. Inclusion criteria for a detailed data extraction of the algorithm were instead that the article used an original case-identification algorithm for any of the following purposes ("rimary objectives): **I** estimating disease occurrence, **II** population/cohort selection, **III** outcome identification. Papers that used secondary objectives (**VI** to identify the disease as comorbidity for adjustments, **V** to identify the disease as exclusion criteria for other conditions, **VI** to calculate hospitalization rates or disease-specific drug prescription rates, **VII** other objectives) are expected to apply less elaborate algorithms, such as single-source algorithms (e.g., HDD to identify chronic conditions), so they were not considered for algorithm extraction.

In this paper, a change to what was reported in the protocol was made regarding age cut-offs used by each algorithm: algorithms were considered independent original algorithms when more than 5 years of difference in the age limits distinguished two otherwise identical algorithms.

To categorize the heterogeneity and facilitate the comparison between algorithms that reported different active ingredients, we defined several drug categories among drugs used for obstructive airway diseases and asthma (ATC: R03XXXX). Table S3 (see on-line supplementary materials) lists the ATC codes/drug groups, reported in detail precisely as in each algorithm, according to the ATC classification.

ASTHMA: "asthma"[MeSH Terms] OR "asthma"[title/Abstract]

COPD: (((COPD[Title/Abstract] OR COAD[Title/Abstract]) OR ("Pulmonary Disease, Chronic Obstructive"[Mesh]) OR (((((obstruct*[Title/Abstract]) AND (lung*[Title/Abstract] OR respirat*[Title/Abstract] OR pulmonar*[Title/Abstract])) AND disease*[Title/Abstract])) OR ((((pulmonar*[Title/Abstract] OR lung*[Title/Abstract] OR airway*[Title/Abstract] OR airflow*[Title/Abstract] OR bronch*[Title/Abstract] OR respirat*[Title/Abstract])))) AND ((chronic[Title/Abstract]) AND obstruct*[Title/Abstract])))))

Box 1. Search strings used to select records from PubMed.

RESULTS

ASTHMA AND COPD PAPERS IDENTIFIED WITH THE SEARCH STRATEGY

The search strategy led to the identification of 98 articles for asthma and 147 articles for COPD (table 1). Out of the selected articles, 74 and 109 papers, respectively for asthma and COPD, were excluded by title and abstract. This brought to full-text screening of 24 and 38 papers, resulting in 12 papers being considered pertinent for asthma and 24 for COPD. Most of the article exclusions were due to the following criteria: no disease specific algorithms reported, data exclusively collected from disease registers, algorithms were not defined (data not reported). References from the selected articles allowed the identification of respectively 2 and 7 more works, leading to a total of 14 pertinent papers for asthma and 31 for COPD. Out of the papers included from references, the 2 regarding asthma both presented algorithms with objective I and the 7 included for COPD prevalently had objective II (data not reported).

PERTINENT PAPERS FOR ASTHMA AND COPD

The chronological distribution of pertinent papers showed that approximately half of the articles were published in the last three years (2014-2017), respectively 6 out of 14 for asthma and 15 out of 31 for COPD (table 2). The

	ASTHMA	COPD
Papers identified by the string	98	147
Full-text readings	24	38
Pertinent papers	12	24
References added from bibliography	2	7
Total pertinent papers	14	31
Papers with objectives IV+*	2	7
Papers with objectives I-III*	12	24
Papers (with objective I-III) with at least one original algorithm	7	14
Papers (with objective I-III) with external validation	3	3
Original algorithms (with objective I-III)	**14**	**16**

* **I** to measure the occurrence of the disease; **II** to identify a population/cohort of subjects affected by the disease of interest; **III** to identify the disease as outcome; **IV** to identify the disease as comorbidity for statistical adjustments; **V** to identify the disease as exclusion criteria for other conditions; **VI** to calculate hospitalization rates or disease-specific drug prescription rates; **VII** other objectives

Table 1. Selection of papers published in PubMed between 2007 and 2017 and original algorithms included in the review according to the disease.

	Asthma	COPD
Year of Publication		
2007-2010	2	11
2011-2013	6	5
2014-2017	6	15
Journal		
Italian	2	9
International	12	22
Setting		
Sub-regional (LHU, cities,..)	3	9
Regional (entire region)	9	14
National multicenter	2	8
International multicenter	0	0
Data time frame for the identification of the disease		
1 year	5	3
> 1 year	9	28
Use of data (even partial) following 2007 (≥2008)	8	15
Objective*		
I	7[12,14-16,19-21]	11[10,11,13,14,17,21-26]
II	3[27-29]	13[30-42]
III	3[28,43,44]	0
IV	0	2[45,46]
V	1[13]	0
VI	1[20]	4[47-50]
VII	1[51]	1[52]

LHU: Local Health Unit
* **I** to measure the occurrence of the disease; **II** to identify a population/cohort of subjects affected by the disease of interest; **III** to identify the disease as outcome; **IV** to identify the disease as comorbidity for statistical adjustments; **V** to identify the disease as exclusion criteria for other conditions; **VI** to calculate hospitalization rates or disease-specific drug prescription rates; **VII** other objectives

Table 2. Characteristics of all pertinent papers published in PubMed between 2007 and 2017 included in the review according to the disease (References).

majority of the works, for both conditions, focused on a region-wide setting. Only 2 papers for asthma and 8 for COPD were based on a national multicenter context and none on an international multicenter one.
About half of the papers used administrative data (for case identification) that dated to 2008 or later. In 9 cases for asthma and in 28 cases for COPD, the data used for the analysis covered more than one year.
Children and young adults were included in all papers that focused on asthma, while for COPD for the most part individuals aged 40 years or above were considered.
Articles that used at least one algorithm for objectives **I**, **II**, or **III**, were respectively 7, 3, and 3 for asthma and 11, 13, and none for COPD. The complete list and several characteristics of the papers, using algorithms for objectives **I-III**, can be found in tables S1 and S2 (see on-line supplementary materials), along with other concomitant uses of the algorithms, for different objectives.
Prevalence for asthma across papers that estimated the occurrence of the disease (objective I), ranged from 2.7% to 11% (crude rates), with important differences even within the same article, according to the temporal and quantitative characteristics of the algorithms used. Among these works, only one estimated the incidence at 0.8% (table S1).
Prevalence estimates of COPD had crude rates that ranged from 3.1% to 9.4% and standardized rates from 1.5% to 6.7% (table S2).

Asthma Algorithms. Out of the 14 selected papers, 14 original algorithms focused on objectives I-III were identified (table 3A). Specific age ranges for case definition were present in 13 cases and considered only children or young adults (always ≤ 40 years old). Only 5 of these algorithms included subjects less than 5 years of age.
Thirteen algorithms used drug prescriptions for case identification. In 11 cases, drug prescriptions alone were sufficient to define a subject as asthmatic. Among all the algorithms using data derived from DPD, irrespective of the differences in the minimum number of prescribed drugs and the periods of follow-back that were used to define a case, five different patterns of drugs were highlighted. One (more sensitive) algorithm used all bronchodilators (algorithm 2 in table 3A), whereas a second pattern (algorithms 4-10 in table 3A), included only adrenergic inhaled with or without combined use of inhaled corticosteroids (ICS). The third pattern (algorithm 1 in table 3A) is characterized by inhaled adrenergics, and a selection of systemic selective Short-Acting Beta-Adrenoreceptor Antagonist (SABA), Long-Acting Beta-Adrenoreceptor Antagonist (LABA) and Leukotriene Receptor Antagonist (LTRA) . In the fourth pattern all active ingredients were specified with a 5th level ATC classification (algorithm 3 in table 3A) and included a tailored selection of inhaled drugs (adrenergic and non adrenergic), Xanthines, LTRA and steroids (not as exclusive therapy). In this case, the selection of inhaled adrenergics is restricted to those with a specific formulation (metered dose inhaler / dry powder inhaler, not nebulized formula), identified by means of the MinSan (Ministry of Health) code, which corresponds to a specific drug formulation. The last pattern, used in three algorithms (11-13 in table 3A) includes a selection of all drug categories included in the previously reported patterns, specified through the 5th level of ATC classification, without any distinction based on the formulation. The detailed list of every active ingredient used by the algorithms is reported in table S3 (see on-line supplementary materials).
The hospital discharge database (HDD), using ICD 9-CM codes 493.XX in the main or secondary diagnosis fields, and the co-payment exemption database (ECD) were included only in 3 algorithms. Three algorithms reported an incident case definition.
External validations were obtained by means of three different approaches: a population-based survey, a comparison with a group of general practitioner (GP) archives, and a group of pediatrician archives. One study used data from a questionnaire-based survey as reference and compared different algorithms according to several different follow-back periods and drug prescription cut-offs, reporting a sensitivity ranging from 32% to 70% and a specificity ranging from 83% to 98% (algorithms 5,7,9 in table 3A). Another study validated its algorithm by comparing algorithm-identified asthmatics ("potential asthmatics") with a pediatrician database and reported a sensitivity of 91% with a specificity of 98% (algorithm 3 in table 3A). The third study used 50 GP archives to validate the 3 proposed algorithms, estimating only sensitivity, which varied from 31% to 63% according to different quantitative drug prescription (DP) cut-offs (respectively ≥2 or ≥1) (algorithms 11-13 in table 3A).

COPD algorithms. We selected 31 papers using the search strategy, leading to 16 original algorithms focused on objectives I-III (table 3B); 14 algorithms applied specific age ranges for case definition, 6 included subjects aged 35 years or more, 4 algorithms selected a population of 40 years and above, 3 considered only subjects aged 44 years or more; 7 algorithms used DPs for case identification alone or in combination with other sources. Within these algorithms, considering only the combinations of active ingredients, regardless of the minimum number of DPs and the periods of follow-back used to define a case, 3 different patterns can be observed. The first, used by one algorithm (algorithm 1 in table 3B), includes the following drug groups: SABA, Short-Acting Muscarinic Antagonist (SAMA), Long-Acting Muscarinic Antagonist

Algorithm ID#	Author, year of publication of the original algorithm (following articles with the same algorithm)	Objective (I disease occurrence, II population-cohort selection, III outcome identification)	Identification of cases: incident (I)-prevalent (P)	SOURCES USED IN THE ALGORITHM			
				HDD ICD-9-CM code (Main diagnosis (M), Any diagnosis (A) Not reported (N))	ECD code	DPD code**	Other sources (code)
1	Tessari, 2008[12]	I	P	493 (A)	007.493	INHALED DRUGS: ADRENERGICS; SYSTEMIC DRUGS: [Selection of selective SABA / LTRA]	MRD (493)
2	Bianchi, 2009[15]	I	P	–	–	DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	–
3	Bianchi, 2011a[19] Bianchi, 2011b[20]	I	I, P	–	–	INHALED DRUGS in MDI/DPI formulations: Selection of [selective SABA / LABA / LABA+ICS / SAMA / ICS / Antiallergic agents (excl. corticosteroids)]; SYSTEMIC DRUGS in ORAL formulations: Selection of [Xanthines / (LTRA / Corticosteroids) not as exclusive therapy]	–
4	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
5	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
6	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
7	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
8	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
9	Bechtold, 2013[16]	I	P	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)]	–
10	Pitter, 2016a[43] Canova, 2015[28]	II	I	–	–	INHALED DRUGS: Selection of ADRENERGICS [selective SABA / LABA / Adrenergics in combination with corticosteroids or other drugs (excl. Anticholinergics)] / ICS	–
11	Biffi, 2017[14]	I	P	493 (N)	007.493	Age specific combinations of the following elements: INHALED DRUGS: Selection of [selective SABA / LABA / LABA+ICS / SAMA / ICS Antiallergic agents (excl. corticosteroids)]; SYSTEMIC DRUGS: Selection of [selective SABA / Xanthines / LTRA] / ANTIBACTERIALS	–
12	Biffi, 2017[14]	I	P	–	–	Combinations of the following elements: INHALED DRUGS: Selection of [selective SABA / LABA / LABA+ICS / SAMA / ICS / Antiallergic agents (excl. Corticosteroids)]; SYSTEMIC DRUGS: Selection of [selective SABA / Xanthines / LTRA] / Oral corticosteroids	–
13	Biffi, 2017[14]	I	P	–	–	Combinations of the following elements: INHALED DRUGS: Selection of [selective SABA / LABA / LABA+ICS / SAMA / ICS / Antiallergic agents (excl. Corticosteroids)]; SYSTEMIC DRUGS: Selection of [selective SABA / Xanthines / LTRA] / Oral corticosteroids	–
14	Talini, 2017[29]	II	P	–	007.493	–	–

HDD: hospital discharge record database; **ECD:** exemption from healthcare co-payment database; **DPD:** drug prescription database; **MRD:** mortality registry database; **Se:** sensitivity; **Sp:** specificity; **PPV:** positive predictive value; **NPV:** negative predictive value

* Negative values means that the criteria is assessed after the date of estimation.

** For detailed list of ATC codes see table S3 (on-line supplementary materials).

Table 3A. Characteristics of Asthma case-identification algorithms published in PubMed between 2007 and 2017.

(LAMA), LABA, LABA/ICS, and ICS. The specific list of ATC codes included in each drug group was not reported in the paper. The second pattern, used in three algorithms (algorithms 7, 9, 10 in table 3B), relies upon the combination of two main criteria: use of a particular drug dose for obstructive airway disease (2nd level ATC groups or "R03_XXXX") or use of different quantities of drugs, within different time windows, considering each chemical subgroup (4th level ATC groups or "R03AC_XX", "R03AK_XX", "R03BA_XX", etc.), separately, for any drug used to treat obstructive airway diseases. The last pattern, observed in three algorithms (algorithms 13-15 in table 3B), invokes a more complex and tailored combination of drug groups, with the inclusion of specific inhaled and systemic drugs combinations.

Regardless of the use of other sources, 13 algorithms based their case definitions on HDD, with approximately 3 different patterns of ICD-9-CM code combinations

CASE DEFINITION		Incidence:* criteria for the exclusion of prevalent cases (look-back time frame)	EVALUATION OF THE ALGORITHM			
Algorithm*	Age range (as definition criteria)		Algorithm derived from a previous published one (reference)	Source used for external validation	Accuracy measures of external validation (Se, Sp, PPV, NPV)	Sensitivity analysis (S) -contribution or coherence of the sources (C)
(HDD (1 year) OR DPD (1 year) OR ECD OR MRD (-1 year)	0-34	–	–	–	–	S C
Occasional users: 1 DPD (1 year) Low users: 2-3 DPD (1 year) High users: ≥ 4 DPD (1 year)	6-17 / ≥ 6	–	Clavenna 2003	–	–	S
DPD (1year)	6-17 / ≥ 6	≥ 1 (DPD (2yr) OR HDD (code 493) (3yr))	Bianchi 2009	138 Pediatricians	Se: 91% Sp: 98%	S
DPD (1 year)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	-	S
≥ 2 DPD (1 year)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	Se: 32% Sp: 98%	S
DPD (2 years)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	–	S
≥ 2 DPD (2 years)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	Se. 56% Sp: 94%	S
DPD (4 years)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	–	S
≥ 2 DPD (4 years)	6-7 / 13	–	–	Population survey, diagnosis based on standardized questionnaire	Se: 70% Sp: 83%	S
≥ 2 DPD (1 year)	1-18	First event in life	–	–	–	S
(HDD (1 year) OR DPD (1 year) OR ECD)	<40	–	–	50 GPs	Se: 39%	–
DPD (1 year)	<40	–	Bianchi 2011b	50 GPs	Se: 63%	–
≥ 2 DPD (1 year)	<40	–	Bianchi 2011b	50 GPs	Se: 31%	–
ECD	–	–	–	–	–	–

used. The first pattern (more sensitive), used in 8 algorithms (2, 3, 4, 7, 9, 10, 13, 16 in table 3B), includes a selection of ICD 9-CM codes (490-496) for Chronic Obstructive Pulmonary Disease Conditions, retrieved from both principal and secondary diagnoses. The second pattern, used by one algorithm (5 in table 3B), consists in the restriction of the criteria reported for the first pattern only to the main diagnosis. The last pattern, used by 4 algorithms (6, 8, 11, 12 in table 3B) uses a combination of two criteria: the presence of different selections of 490-496 codes as main diagnosis or as secondary diagnosis, if associated with the presence of a main diagnosis related to symptoms of the disease (see table 3B for details). Table 3B shows that algorithms characterized by the more sensitive pattern are more prone to be used for objective **I** (occurrence), whereas the second and third patterns, aimed at gaining specificity, were always used for objectives **II** and **III** (cohort selection or disease as outcome).

Algorithm ID#	Author, year of publication of the original algorithm (following articles with the same algorithm)	Objective (I disease occurrence, II population-cohort selection, III outcome identification)	Identification of cases: incident (I)-prevalent (P)	SOURCES USED IN THE ALGORITHM		
				HDD ICD-9-CM code (Main diagnosis (M), Secondary diagnosis (S), Any diagnosis (A) Not reported (N))	DPD code**	Other sources (code)
1	Anecchino, 2007[17]	I	P	–	Quantity / Frequency combination of Inhaled / Oral DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	–
2	Faustini, 2007a[22]	I	P	490-492, 494, 496 (A)	–	MRD (490-492, 494,496)
3	Faustini, 2007b[23]	I	P	490-492, 494,496 (A)	–	MRD (490-492, 494,496)
4	Faustini, 2008a[24]	I	P	490-492, 494,496 (A)	–	MRD (490-492, 494,496)
5	Faustini, 2008b[30]	II	P	490-492,494,496 (M)	–	–
6	P.Re.Val.E. 2008[31]	II	P	M (490-492, 494,496) OR (S (490-492, 494,496) AND M (518.81-518.84, 786.0/2/4))	–	–
7	ARS Toscana, 2009[10]	I	P	490-492, 494, 496 (A)	Quantity / Frequency combinations of DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	–
8	Agabiti, 2010[32]	II	P	M (490-492, 494,496) OR (S (490-492, 494,496) AND M (518.81-518.82, 786.0/2/4))	–	–
9	Faustini, 2012[13]	I	P	490-492, 494,496 (A)	Quantity / Frequency combinations of DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	MRD (490-492, 494,496)
10	Gini, 2013[11]	I	P	490-492, 494, 496 (A)	Quantity / Frequency combinations of DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	–
11	Blasi, 2014[34]	II	P	M (491.2) OR (M (480-486, 518.8, 512.0, 512.8, 428, 415) AND S (491.2))	–	–
12	Di Martino, 2014[35]	II	P	M (490-492, 494,496) OR (S (490-492, 494,496) AND M (518.81-518.84, 786.0/2/4))	–	–
13	Biffi, 2017[14]	I	P	491, 492, 496 (A)	Combinations of the following elements: INHALED DRUGS: Selection of [LABA / LABA+ICS / selective SABA / SAMA / LAMA / ICS] SYSTEMIC DRUGS: Selection of [selective SABA / Xantine]	–
14	Biffi, 2017[14]	I	P	–	INHALED DRUGS: Selection of [LABA /LABA+ICS / selective SABA / SAMA / LAMA / SABA+SAMA / ICS] SYSTEMIC DRUGS: Selection of selective SABA	–
15	Biffi, 2017[14]	I	P	–	INHALED DRUGS: Selection of [LABA /LABA+ICS / selective SABA / SAMA / LAMA / SABA+SAMA / ICS] SYSTEMIC DRUGS: Selection of selective SABA	–
16	Fedeli, 2017[41]	II	P	491, 492, 496 (A)	–	–

HDD: hospital discharge record database; **DPD:** drug prescription database; **MRD:** mortality registry database; **Se:** sensitivity; **Sp:** specificity; **PPV:** positive predictive value; **NPV:** negative predictive value

* Negative values means that the criteria is assessed after the date of estimation.

** For detailed list of ATC codes see table S3 (on-line supplementary materials).

Table 3B. Characteristics of COPD case-identification algorithms published in PubMed between 2007 and 2017.

Data from MRD were used in 4 algorithms, all with objective **I**. Only 3 algorithms among those with objective **I** used data covering more than one year; conversely, algorithms for objective **II-III** share longer follow-back periods (often more than one year). No algorithm based on DPD or MRD was used for objectives **II-III**. No algorithm reported any incident case definition. For one algorithm only (algorithm 4 in table 3B) the PPV is reported. This algorithm was based on HDD and estimated a PPV of 80.2% when using spirometry as gold standard. The sensitivity of three slightly different algorithms reported by Biffi, obtained by comparison with 50 GP databases, ranged from 51% to 72.9%, according to different cut-offs in DPs used. The algorithm's specificity was not reported. One medical chart review study on discharge data reported that among subjects discharged from hospital after COPD diagnosis, 90% of reviewed cases were confirmed.

CASE DEFINITION			EVALUATION OF THE ALGORITHM			
Algorithm*	Age range (as definition criteria)	Incidence:* criteria for the exclusion of prevalent cases (look-back time frame)	Algorithm derived from a previous published one (reference)	Source used for external validation	Accuracy measures of external validation (Se, Sp, PPV, NPV)	Sensitivity analysis (S) -contribution or coherence of the sources (C)
≥ 5 DPD (1 year)	45+	–	–	–	–	–
HDD (1 year) OR MRD (1 year)	35+	–	European Community Health Indicators	–	–	–
HDD (1 year) AND MRD (-1 year)	–	–	European Community Health Indicators	–	–	C
HDD (4 years) OR HDD (-1 year) OR MRD (-1 year))	35+	–	Faustini, 2007a	Clinical (hospital or outpatient) charts from referral center, diagnosis based on spirometry (Romanelli, 2016)	PPV of 80.2% for HDD contribution	C
HDD (3 years)	35+	–	European Community Health Indicators	–	–	–
HDD (2 years)	35+	–	–	Medical chart review from hospital (Fano, 2012)	Confirmation rate of 94% for COPD diagnosis	–
(HDD (1 year) OR DPD (1 year))	16+	–	–	–	–	–
HDD (5 years)	35+	–	–	–	–	–
HDD (4 years) OR HDD (-1 year) OR MRD (-1 year) OR DPD (-1 year)	35+	–	Faustini, 2008a Ars Toscana, 2009	–	–	S / C
HDD (6 years) OR DPD (6 years)	–	–	Faustini, 2007a; Anecchino, 2007	–	–	–
HDD (1 year)	40+	–	–	–	–	–
HDD (3 years)	45+	–	Fano, 2012	Medical chart review from hospital (Fano, 2012)	Confirmation rate of 94% for COPD diagnosis	–
HDD (1 year) OR DPD (1 year)	40+	–	Douglas, 2006	50 GPs	Se: 52.6%	–
≥2 DPD (1 year)	40+	–	Anecchino, 2007	50 GPs	Se: 51%	S
DPD (1 year)	40+	–	Anecchino, 2007	50 GPs	Se: 72.9%	–
HDD (3 years)	45+	–	–	–	–	–

DISCUSSION

This work reviewed and described the characteristics of case-identification algorithms applied in the last decade in Italy for two relevant clinical conditions.

Especially for COPD, we observed an important variability among the algorithms according to their objectives. For both diseases, a progressive increase in the complexity of algorithms, in terms of combination of data sources and codes used, is not negligible. For COPD, we observed, more frequently than for asthma, original algorithms designed for objectives other than prevalence estimation. This could be a direct consequence of the uncertain reliability of DPs in patients affected by COPD. This probably incentivized the development of more specific algorithms for objectives **II** and **III**.

For both diseases, as expected, we observed a low degree of variability in the ICD-9 diagnostic codes used by the algorithms. On the other hand, data on drug prescrip-

tions showed heterogeneous and complex inclusion criteria, leading to different drug combinations, variable cut-offs, and time intervals between prescriptions in order to define a case. This is the result of an effort to reduce the risk of misclassification associated with the use of drug prescriptions, as there are significant diagnostic and therapeutic overlaps between the two diseases. Since the onset of COPD is generally after the 3rd-4th decade of life, age is a constant criterion used to develop algorithms capable of discriminating between the two diseases (tables 3A and 3B). In the entire systematic review, among all COPD case-identification algorithms that used drug prescriptions, only two algorithms for COPD did not consider this age-specificity.[10,11] Nonetheless, the age cut-off, generally set at around 35-45 years, leaves room for misclassification,[12] especially among the older population, where subjects with asthma are certainly present.[13] Thus, since age cut-offs have a relevant role in the trade-off between sensitivity and specificity, their choice deserves particular attention that should be tailored on the research objective.

Since pharmacological treatments defined in guidelines for both diseases,[1,2] greatly overlapped, it is not surprising that the active ingredients included in algorithms for asthma and COPD case identification are quite similar: as reported in table S3, only the adrenergic/anticholinergic/steroid combination (SABA+SAMA/LABA+LAMA/LABA+LAMA+ICS) and tiotropium bromide proved specific for COPD algorithms, whereas LTRA, systemic fenoterol, and antiallergic drugs are only included in asthma case-identification algorithms.

The variability among algorithms may partly reflect updates in guidelines. The lack of several long-acting bronchodilators (including triple combination), recommended for severe COPD and available in Italy after 2010, highlights the need to promptly update algorithms, as new drugs are included in the guidelines for asthma and COPD. This aspect is more relevant when algorithms are defined by specific, well-described drugs (namely ATC classification 5th level: chemical substance) indexes.

QUANTITATIVE AND QUALITATIVE COMBINATIONS OF DRUGS IN ASTHMA AND COPD ALGORITHMS

Asthma. The most recently published paper in our review presented three algorithms for asthma, with different combinations of drugs and quantitative cut-offs in association with specific age ranges which arose from a combined effort that involved specialized pneumologists.[14] Nevertheless, the observed sensitivity was relatively poor (39%). Furthermore, the risk of requiring elaborate drug combination criteria for case identification might lead to identify subsets of patients with specific clinical characteristics. The paper also presented two other less complex algorithms and reported a marked reduction in sensitivity (from 63% to 31% for asthma) when passing from a permissive cut-off (≥1 DP/year) to a restrictive one (≥2 DP/year).

Only one approach, that stands out for its original contribution,[15] consisted in a restriction of the required drugs, depending on specific types of formulations – metered dose inhalers (MDI) and dry powder inhalers (DPI) – which require a specific code (MinSan) to be identified. This approach, as the one proposed by Biffi and a team of clinicians, is also designed to increase the algorithm's specificity.

One last article focused on developing algorithms with different follow-back periods.[16] A comparison was done between algorithms that required more than one disease-specific drug prescription (≥2 DP/year) over the past year, two years, or four years. This resulted in a progressive gain of sensitivity and a loss of specificity. This is explainable by considering the intermitting nature and the broad spectrum of symptoms that characterize asthma. Therefore, very mild cases, where subjects are given 1 or less than 1 DP per year, might not be captured by more specific algorithms. These results may favor an informed decision on the optimal balance between sensitivity and specificity, according to the specific research requirements of each study.

COPD. The same paper from Biffi also defined three algorithms for COPD, applying different combinations of drugs and quantitative cut-offs among subjects aged 40 years or more. As seen before, sensitivity was suboptimal (53%) here, too. Similarly to asthma, the paper presented other two, less elaborate algorithms, reporting a marked reduction in sensitivity (from 73% to 51% for COPD) when passing from a permissive cut-off (≥1 DP/year) to a restrictive one (≥2 DP/year).

All other previously published algorithms for COPD used selection criteria of DP cut-offs, based on the minimum length of treatment among a broad spectrum of drugs (at least 5 or more DPs of any bronchodilators in 12 months), or considered the utilization of any bronchodilator with subgroup separations (3 or more DPs in different minimum lengths of treatment periods). This approach focused more on the time span of the treatment with any respiratory drug than on the use of one or more specific active ingredients. This was an attempt to better reflect both the chronic course of COPD and the complexity of existing treatments in clinical practice. Since this approach is irrespective of the 5th level of ATC classification, the selection criteria are also less sensitive to changes as guidelines were updated, and might minimize the risk of selection bias. Moreover, considering only periods with a minimum duration of treatment may avoid the inclusion of short treatments, commonly associated with acute health events, as often experienced by asthmatic subjects.

Nonetheless, as noted for age cut-offs, these criteria arose from the general analysis of intervals between prescriptions through an empirical definition of "patients treated chronically".[17] This aspect must be taken into account when considering the application of such algorithms.

VALIDATION STUDIES

Results from this systematic review highlighted the paucity of experiences to evaluate algorithm accuracy and the heterogeneity of these approaches. The absence of a commonly accepted algorithm for asthma and COPD case identification by means of HADs is, in fact, mainly due to a difficulty in the validation process. For asthma, two papers reported both sensitivity and specificity of four algorithms (table 3A), whereas results from papers reporting only sensitivity, based on GP data, was discussed in previous paragraphs. The other two validation examples used different approaches (a population-based survey by standardized questionnaire and data from the pediatrician network) with very variable results as far as sensitivity is concerned. Although, at present, the most commonly accepted gold standard for asthma validation is questionnaire-based, the 2 validations that were reported show that there is no consensus in applying this procedure for an accurate validation. For completeness of information and the possibility of having a clinical overview of the patients, GP and pediatrician networks could be valuable resources, although they lack a population of clinically verified, unaffected subjects, to allow an estimate of specificity as well as sensitivity. Nevertheless, with such a vast variety of algorithms, the risk of introducing a selection bias is still present and an exhaustive and robust validation process is desirable.

For COPD, no algorithms were clearly validated in our review (table 3B). Partial elements to derive accuracy estimates rely on two experiences which reported high confirmation rates for hospital-based diagnosis of COPD and high positive predictive value for HDD in three algorithms. This observation led to the conclusion that it is not possible to identify a commonly accepted algorithm for COPD among those included in our analysis. Limitation in the availability of validation studies is due to a lack of population data regarding the disease and the difficulty in defining a reliable validation gold standard. However, since underdiagnosis and undertreatment of the disease, as well as low adherence to COPD-drug treatments is unneglectable, it is reasonable that algorithms included in this review may not be accurate in identifying subjects with COPD. This observation is supported by the sensitivity results we reported from Biffi (table 3B) and from a recent study which assesses the accuracy of one algorithm for COPD by means of a population-based survey, reporting a value of sensitivity and specificity of 38.5% and 91.7%, respectively.[18]

FUTURE DEVELOPMENTS

As of right now, none of the included articles in this systematic review correlated any pharmaceutical pattern, whether qualitatively and/or quantitatively defined, with the severity of the disease. This could be an interesting future development, as there is a correlation between the type and amount of drugs prescribed to an individual and the severity/therapeutic resistance of the disease as reported for asthma[1]. The same is true for COPD, where treatment with LABA in combination with LAMA, or LABA in combination with LAMA and ICS, is recommended in more severe cases.[2] A greater combination of qualitatively different drugs is most likely an indicator of poor control of asthma symptoms.[1] Similarly, a greater quantity of prescribed drugs suggests an increased need for therapeutic support of the individual with more frequent and/or more severe attacks. For COPD, due to the low adherence of patients to pharmacological treatments [3], efforts must be more focused on developing algorithms capable of excluding misclassification of cases. To achieve this objective, even in the light of the low accuracy of drug prescription data in identifying affected subjects,[18] a hospital-based diagnosis may be used conditioned on previous respiratory drug utilization to select a cohort of subjects with a homogeneous severity of illness. Definition of indicators of quality of care among subjects identified through application of these algorithms has been proposed to monitor quality of care of COPD patients in the context of a national evaluation program (data not published, personal communication to authors).

STRENGTHS AND LIMITATIONS

A major issue raised by this work, is whether it is possible to indicate the most suitable algorithm for a specific research question. To find an answer, further investigation, with ad hoc accuracy analysis, is required.

One of the limitations this systematic review could have encountered is related to the sensitivity of the string, which was designed to maximize the inclusion of articles that used algorithms for objectives **I**, **II**, or **III**. Articles that relied on single-source algorithms might have been missed; our results suggest that string sensitivity may be lower for COPD. As far as asthma is concerned, some papers[15] were initially missed because they were actually letters and not research articles, but they were captured, all the same, through the references. As the papers included in this review did not use data after 2012, it is plausible that several analyses have been performed at local level and have not been captured by our work.

CONCLUSION

Italian literature on the use of administrative healthcare data for asthma and COPD case identification is rela-

tively varied, with algorithms that are quite different one from the other. Evidence on the accuracy of the algorithms is still lacking and further efforts deserve to be made. This review highlights a progressive increase in the complexity of the algorithms. This is mainly due to the increasing amount of administrative healthcare data, especially regarding DPs. This growth has allowed to develop algorithms that can respond in an ever more specific manner to research objectives. Elements from the review of the use of healthcare administrative databases and their prompt updates represent a useful tool to inform decisions on what algorithm to adopt, what characteristics of each algorithm may be hampered, what the accuracy limits are, according to the specific objective of the research study.

Algorithms used at present for asthma case identification, especially by means of DPs, seem to be reliable, although a risk of misclassification in adults persists and evidence from validation studies is still limited. Variable follow-back periods and DP cut-offs are both elements that can be changed to modulate an algorithm's sensitivity and specificity.

As far as COPD is concerned, the characteristics of the disease, as well as patients' scarce adherence to therapy, contribute to a low accuracy of algorithms based on DPs, as shown by the relatively few validation experiences. Diagnoses derived from hospital discharge records are reliable and can be even more informative in the process of selection of cohorts of hospitalized subjects, when integrated with information on drug prescriptions, regarding the guideline-recommended active ingredients for specific stages of the disease. The use of ECD, recently implemented by certain Italian regions, will represent a possible alternative source of information.

Conflict of interest disclosure: none reported.

Funding disclosure: the "Algoritmi" Project was partially funded by the Department of Cardio-Thoraco-Vascular Sciences and Public Health (University of Padua) within the Projects BIRD (Integrated Budget for Research in Departments) in the year 2017.

REFERENCES

1. Paggiaro P (ed). Linee guida GINA italiane. Versione 2016-2017. Available from: http://ginasma.it/wp-content/uploads/materiali/2017/GINA_adattamento_ita_2017_doc.pdf (last accessed: October 8, 2018).
2. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2018 Report). Available from: https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf (last accessed: October 8, 2018).
3. Simonato L, Baldi I, Balzi D, et al. Objectives, tools and methods for an epidemiological use of electronic health archives in various areas of Italy. Epidemiol Prev 2008;32(3) Suppl:5-14.
4. Abraha I, Orso M, Grilli P, et al. The current state of validation of administrative healthcare databases in Italy: a systematic review. Int J Stat Med Res 2014;3(3):309-20.
5. Rimland JM, Abraha I, Luchetta ML, et al. Validation of chronic obstructive pulmonary disease (COPD) diagnoses in healthcare databases: a systematic review protocol. BMJ Open 2016;6(6):e011777.
6. Nissen F, Quint JK, Wilkinson S, Mullerova H, Smeeth L, Douglas IJ. Validation of asthma recording in electronic health records: protocol for a systematic review. BMJ Open 2017;7(5):e014694.
7. Nissen F, Quint JK, Wilkinson S, Mullerova H, Smeeth L, Douglas IJ. Validation of asthma recording in electronic health records: a systematic review. Clin Epidemiol 2017;9:643-56.
8. Souri S, Symonds NE, Rouhi A, et al. Identification of validated case definitions for chronic disease using electronic medical records: a systematic review protocol. Syst Rev 2017;6(1):38.
9. Canova C, Simonato L, Barbiellini Amidei C, et al. A systematic review of case-identification algorithms for 18 conditions based on Italian healthcare administrative databases: a study protocol. Epidemiol Prev 2019;43(4) Suppl 2:8-16.
10. User S. La banca dati MaCro delle malattie croniche in Toscana. ARS Agenzia Regionale di Sanità Toscana. Available from: https://www.ars.toscana.it/collana-documenti-ars/281-2009/768-la-banca-dati-macro-delle-malattie-croniche-in-toscana.html (last accessed: October 8, 2018).
11. Gini R, Francesconi P, Mazzaglia G, et al. Chronic disease prevalence from Italian administrative databases in the VALORE project: a validation through comparison of population estimates with general practice databases and national survey. BMC Public Health. 2013;13:15.
12. Tessari R, Migliore E, Balzi D, et al. [Asthma prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy]. Epidemiol Prev. 2008;32(3) Suppl:56-65.
13. Faustini A, Canova C, Cascini S, et al. The reliability of hospital and pharmaceutical data to assess prevalent cases of chronic obstructive pulmonary disease. COPD 2012;9(2):184-96.
14. Biffi A, Comoretto R, Arfe A, et al. Can healthcare utilization data reliably capture cases of chronic respiratory diseases? a cross-sectional investigation in Italy. BMC Pulm Med 2017;17(1):20.
15. Bianchi M, Clavenna A, Labate L, et al. Anti-asthmatic drug prescriptions to an Italian paediatric population. Pediatr Allergy Immunol 2009;20(6):585-91.
16. Bechtold P, Ranzi A, Gambini M, et al. Assessing paediatric asthma occurrence through dispensed prescription data and questionnaires. Eur J Public Health 2013;23(5):873-78.
17. Anecchino C, Rossi E, Fanizza C, et al. Prevalence of chronic obstructive pulmonary disease and pattern of comorbidities in a general population. Int J Chron Obstruct Pulmon Dis 2007;2(4):567-74.
18. Di Domenicantonio R, Cappai G, Cascini S, et al. Validation of algorithms for the identification of subjects with chronic disease using health information systems. Epidemiol Prev 2018;42(5-6):316-25.
19. Bianchi M, Clavenna A, Sequi M, Bonati M. Asthma diagnosis vs. analysis of anti-asthmatic prescriptions to identify asthma in children. Eur J Clin Pharmacol 2011;67(9):967-68.
20. Bianchi M, Clavenna A, Sequi M, et al. Anti-asthma medication prescribing to children in the Lombardy Region of Italy: chronic versus new users. BMC Pulm Med 2011;11:48.
21. Anello P, Cestari L, Canova C, et al. Investigation on the health effects of traffic-related air pollution from Mestre motorway (Veneto Region, Northern Italy). Epidemiol Prev 2015;39(4):251-60.
22. Faustini A, Marino C, D'Ovidio M, Perucci CA. The concurrent COPD mortality doubles the mortality estimate from COPD as underlying cause in Lazio, Italy. Respir Med 2007;101(9):1988-93.
23. Faustini A, Marino C, Dovidio M, Perucci CA. Mortality due to chronic obstructive pulmonary disease: indicators of impact, access to health services, and outcome of hospital treatment in an Italian region. Epidemiol Prev 2007;31(2-3):140-47.
24. Faustini A, Cascini S, Arca M, et al. Chronic obstructive pulmonary disease prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy. Epidemiol Prev 2008;32(3) Suppl:46-55.
25. Bruni A, Gianicolo EA, Vigotti MA, Faustini A. Chronic obstructive pulmonary disease (COPD) prevalence in Brindisi Province (Southern Italy) for the years 2005-2009. Epidemiol Prev 2014;38(2):108-15.
26. Romanelli AM, Raciti M, Protti MA, Prediletto R, Fornai E, Faustini A. How Reliable

Are Current Data for Assessing the Actual Prevalence of Chronic Obstructive Pulmonary Disease? PLoS One 2016;11(2):e0149302.
27. Bianchi M, Clavenna A, Sequi M, et al. Spirometry testing in a population of Italian children: age and gender differences. Respir Med 2012;106(10):1383-88.
28. Canova C, Pitter G, Ludvigsson JF, et al. Coeliac disease and asthma association in children: the role of antibiotic consumption. Eur Respir J 2015;46(1):115-22.
29. Talini D, Ciberti A, Bartoli D, et al. Work-related asthma in a sample of subjects with established asthma. Respir Med 2017;130:85-91.
30. Faustini A, Marino C, D'Ippoliti D, Forastiere F, Belleudi V, Perucci CA. The impact on risk-factor analysis of different mortality outcomes in COPD patients. Eur Respir J 2008;32(3):629-36.
31. P.Re.Val.E Programma Regionale Valutazione degli Esiti degli Interventi Sanitari 2016. Mortalità a 30 giorni dal primo accesso per broncopneumopatia cronica ostruttiva (BPCO) riacutizzata protocollo operativo. Available from: http://95.110.213.190/prevale2017/Main/protocolli/pro_21.pdf (last accessed: October 8, 2018).
32. Agabiti N, Belleudi V, Davoli M, et al. Profiling hospital performance to monitor the quality of care: the case of COPD. Eur Respir J 2010;35(5):1031-38.
33. Fano V, D'Ovidio M, del Zio K, et al. The role of the quality of hospital discharge records on the comparative evaluation of outcomes: the example of chronic obstructive pulmonary disease (COPD). Epidemiol Prev 2012;36(3-4):172-79.
34. Blasi F, Cesana G, Conti S, et al. The clinical and economic impact of exacerbations of chronic obstructive pulmonary disease: a cohort of hospitalized patients. PLoS One 2014;9(6):e101228.
35. Di Martino M, Agabiti N, Bauleo L, et al. Use patterns of long-acting bronchodilators in routine COPD care: the OUTPUL study. COPD 2014;11(4):414-23.
36. Kirchmayer U, Cascini S, Agabiti N, et al. One-year mortality associated with COPD treatment: a comparison of tiotropium and long-acting beta2-agonists in three Italian regions: results from the OUTPUL study. Pharmacoepidemiol Drug Saf 2016;25(5):578-89.
37. Di Martino M, Agabiti N, Cascini S, et al. The Effect on Total Mortality of Adding Inhaled Corticosteroids to Long-Acting Bronchodilators for COPD: A Real Practice Analysis in Italy. COPD 2016;13(3):293-302.
38. Di Martino M, Ventura M, Cappai G, et al. Adherence to Long-Acting Bronchodilators After Discharge for COPD: How Much of the Geographic Variation is Attributable to the Hospital of Discharge and How Much to the Primary Care Providers? COPD 2017;14(1):86-94.
39. Belleudi V, Di Martino M, Cascini S, et al. The impact of adherence to inhaled drugs on 5-year survival in COPD patients: a time dependent approach. Pharmacoepidemiol Drug Saf 2016;25(11):1295-304.
40. Ferroni E, Belleudi V, Cascini S, et al. Role of Tiotropium in Reducing Exacerbations of Chronic Obstructive Pulmonary Disease When Combined With Long-Acting beta2 -Agonists and Inhaled Corticosteroids: The OUTPUL Study. J Clin Pharmacol 2016;56(11):1423-32.
41. 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;12:443-50.
42. Cascini S, Kirchmayer U, Belleudi V, et al. Inhaled Corticosteroid Use in Chronic Obstructive Pulmonary Disease and Risk of Pneumonia: A Nested Case-Control Population-based Study in Lazio (Italy)-The OUTPUL Study. COPD 2017;14(3):311-17.
43. Pitter G, Ludvigsson JF, Romor P, et al. Antibiotic exposure in the first year of life and later treated asthma, a population-based birth cohort study of 143,000 children. Eur J Epidemiol 2016;31(1):85-94.
44. Pitter G, Canova C. Reply to Bianchi et al.: In Italy anti-asthmatic drug prescription is not always a reliable proxy of asthma. Eur J Epidemiol 2016;31(5):533.
45. Forastiere F, Stafoggia M, Tasco C, et al. Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or differential susceptibility. Am J Ind Med 2007;50(3):208-16.
46. Nuvolone D, Balzi D, Chini M, Scala D, Giovannini F, Barchielli A. Short-term association between ambient air pollution and risk of hospitalization for acute myocardial infarction: results of the cardiovascular risk and air pollution in Tuscany (RISCAT) study. Am J Epidemiol 2011;174(1):63-71.
47. Antonelli-Incalzi R, Ancona C, Forastiere F, Belleudi V, Corsonello A, Perucci CA. Socioeconomic status and hospitalization in the very old: a retrospective study. BMC Public Health 2007;7:227.
48. Trerotoli P, Bartolomeo N, Moretti AM, Serio G. Hospitalisation for COPD in Puglia: the role of hospital discharge database to estimate prevalence and incidence. Monaldi Arch Chest Dis 2008;69(3):94-106.
49. Arfe A, Nicotra F, Cerveri I, et al. Incidence, Predictors, and Clinical Implications of Discontinuing Therapy with Inhaled Long-Acting Bronchodilators among Patients with Chronic Obstructive Pulmonary Disease. COPD 2016;13(5):540-46.
50. Fabiani L, Mammarella L, Vittorini P, Zazzara F, Di Orio F. Preliminary analysis of the Respiratory Disease, based on Hospital Discharge Records, in the Province of L'Aquila, Abruzzo, Italy. Ann Ig 2016;28(6):392-403.
51. Bianchi M, Clavenna A, Sequi M, et al. Childhood asthma management pre- and post-incident asthma hospitalization. PloS One 2013;8(10):e76439.
52. Belleudi V, Agabiti N, Kirchmayer U, et al. Definition and validation of a predictive model to identify patients with chronic obstructive pulmonary disease (COPD) from administrative databases. Epidemiol Prev 2012;36(3-4):162-71.