CHAPTER 5

A SYSTEMATIC REVIEW OF CASE-IDENTIFICATION ALGORITHMS BASED ON ITALIAN HEALTHCARE ADMINISTRATIVE DATABASES FOR THREE RELEVANT DISEASES OF THE NERVOUS SYSTEM: PARKINSON'S DISEASE, MULTIPLE SCLEROSIS, AND EPILEPSY

Cristina Canova,[1] Silvia Danieli,[1] Claudio Barbiellini Amidei,[1] Lorenzo Simonato,[1] Riccardo Di Domenicantonio,[2] Giovanna Cappai,[2] Anna Maria Bargagli[2]
[1] Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padua, Padua (Italy)
[2] Department of Epidemiology, Lazio Regional Health Service, ASL Roma 1, Rome (Italy)
Corresponding author: Cristina Canova; cristina.canova@unipd.it

ABSTRACT

BACKGROUND: Parkinson's Disease (PD), Multiple Sclerosis (MS), and Epilepsy are three highly impactful health conditions affecting the nervous system. PD, MS, and epilepsy cases can be identified by means of Healthcare Administrative Databases (HADs) to estimate the occurrence of these diseases, to better monitor the adherence to treatments, and to evaluate patients' outcomes. Nevertheless, the absence of a validated and standardized approach makes it hard to quantify case misclassification.
OBJECTIVES: to identify and describe all PD, MS, and epilepsy case-identification algorithms by means of Italian 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. 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 strategy led to the identification of 70 papers for PD, 154 for MS, and 100 for epilepsy, of which 3 papers for PD, 6 for MS, and 5 for epilepsy were considered pertinent. Most articles were published in the last three years (2014-2017) and focused on a region-wide setting. Out of all pertinent articles, 3 original algorithms for PD, 4 for MS, and 4 for epilepsy were identified. The Drug Prescription Database (DPD) and Hospital Discharge record Database (HDD) were used by almost all PD, MS, and epilepsy case-identification algorithms. The Exemption from healthcare Co-payment Database (ECD) was used by all PD and MS case-identification algorithms, while only 1 epilepsy case-identification algorithm used this source. All epilepsy case-identification algorithms were based on at least a combination of electroencephalogram (EEG) and drug prescriptions.
An external validation had been performed by 2 papers for MS, 2 for epilepsy, and only 1 for PD.
CONCLUSION: the results of our review highlighted the scarce use of HADs for the identification of cases affected by neurological diseases in Italy. While PD and MS algorithms are not so heterogeneous, epilepsy case-identification algorithms have increased in complexity over time. Further validations are needed to better understand the specific characteristics of these algorithms.

WHAT IS ALREADY KNOWN

- **The use of HADs is an accessible tool to monitor frequency and trends of several diseases.**
- **Limited studies validated algorithms based on claims data to capture subjects affected from PD, MS, and epilepsy.**

WHAT THIS PAPER ADDS

- **Relatively few Italian published papers focused on identifying PD, MS, and epilepsy by means of HADs.**
- **DPD is the most relevant source among all HADs to estimate PD prevalence.**
  The most important HADs to identify MS are HDD, DPD, and ECD.
- **A combination of EEG and any antiepileptic drug is the most efficient way to capture epileptic cases.**

Keywords: algorithms, healthcare administrative databases, Parkinson's disease, multiple sclerosis, epilepsy

RIASSUNTO

INTRODUZIONE: il morbo di Parkinson (MP), la sclerosi multipla (SM) e l'epilessia sono tre patologie neurologiche con un forte impatto sanitario. L'uso integrato dei flussi amministrativi sanitari (FAS) consente l'identificazione dei casi affetti da MP, SM ed epilessia e permette di ottenere una stima dell'occorrenza di tali malattie, per monitorare l'aderenza ai trattamenti e per valutare gli outcome dei pazienti. Tuttavia, l'assenza di un approccio validato e standardizzato rende difficile quantificare la misclassificazione dei casi.

OBIETTIVI: identificare e descrivere tutti i lavori pubblicati negli ultimi 10 anni che, utilizzando FAS italiani, hanno elaborato almeno un algoritmo originale per l'identificazione di pazienti affetti da MP, SM ed epilessia.

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 condizioni 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 che combina testo libero con termini MeSH, in parte comune a tutte le patologie e in parte patologia-specifica. Gli articoli pertinenti sono stati classificati secondo l'obiettivo di utilizzo degli algoritmi di identificazione delle patologie considerate e solo gli articoli con obiettivo primario (**I** stima di occorrenza; **II** identificazione di popolazioni/coorti; **III** identificazione della patologia come esito) sono stati inclusi per estrarne informazioni. Per ogni algoritmo sono state estratte informazioni sui FAS utilizzati (schede di dimissione ospedaliera – SDO, esenzioni ticket – ET, prescrizioni farmaceutiche – PF), i codici ICD versione 9 e 10, il sistema di classificazione ATC per i farmaci, i criteri di identificazione dei casi, il periodo di osservazione/follow-back e le fasce d'età considerate. Sono state inoltre estratte ulteriori informazioni su eventuali validazioni esterne con le opportune misure di accuratezza riportate, analisi di sensibilità ed il contributo di ciascun FAS.

RISULTATI: la strategia di ricerca ha portato all'identificazione di 70 articoli per il MP, 154 per la SM e 100 per l'epilessia. Di questi, sono stati considerati pertinenti 3 documenti per il MP, 6 per la SM e 5 per epilessia. La maggior parte degli articoli è stata pubblicata negli ultimi tre anni (2014-2017) e si concentra su un setting regionale. Tra gli articoli pertinenti, sono stati identificati 3 algoritmi originali per il MP, 4 per la SM e 4 per epilessia. Le PF e le SDO sono state utilizzate da quasi tutti gli algoritmi di identificazione dei casi di MP, SM ed epilessia. Le ET sono state utilizzate da tutti gli algoritmi di identificazione dei casi il MP e la SM, mentre solo 1 algoritmo di identificazione di casi di epilessia ha utilizzato questa fonte. Tutti gli algoritmi di identificazione di casi di epilessia sono basati su almeno una combinazione di elettroencefalogramma e PF.

Le validazioni esterne sono state eseguite da 1 articolo per il MP, 2 per la SM e 2 per l'epilessia.

CONCLUSIONI: i risultati della revisione hanno messo in evidenza lo scarso utilizzo di FAS per l'identificazione di soggetti affetti da malattie neurologiche in Italia. Gli algoritmi di identificazione dei casi di MP e di SM sono abbastanza similari tra loro, mentre gli algoritmi di identificazione dei casi di epilessia sono aumentati di complessità nel tempo. Sono necessarie ulteriori studi di validazione esterne per comprendere meglio le caratteristiche specifiche di questi algoritmi.

Parole chiave: algoritmi, database amministrativi sanitari, morbo di Parkinson, sclerosi multipla, epilessia

INTRODUCTION

Parkinson's Disease (PD), Multiple Sclerosis (MS) and epilepsy are three common neurological diseases that have a significant impact on patients and on the healthcare system in terms of frequency, disabling outcomes, and costs for the National Healthcare Service (NHS). Impact of PD and epilepsy is destined to increase in Western Countries, especially with an ageing population.[1,2]

PARKINSON'S DISEASE

Parkinson's Disease (PD) is a chronic, progressive neurodegenerative disorder without known etiology[3] and a leading cause of neurological disability in adults (50+).[4] PD affects 1-2‰ of the entire population.[5] Its prevalence increases with age and affects up to 1% of the population above 60 years of age[1] and can reach a peak of 9% among people aged 80-84.[6] Its prevalence is expected to increase dramatically, almost doubling in Western Countries by 2050.[7]

PD diagnosis is essentially based on clinical criteria. Medical history and neurologic examination highlight the core motor features of PD, such as bradykinesia, stiffness, resting tremor, rigidity, and postural instability;[8] however, non-motor symptoms such as depression, cognitive deficits, and disorders of autonomic functions may also be paramount.[9,10]

PD treatment is mainly based on the use of two broad categories of pharmaceuticals: symptomatic drugs and neuroprotective drugs. All Anti-Parkinson Drugs (APDs), are recorded in the Drug Prescription Database (DPD) and are identified by a specific ATC (Anatomical Therapeutic Chemical Classification System). All APDs fall into the category N04 (N = Nervous system, 04 = anti-Parkinson drugs), require a medical prescription, and their delivery is electronically registered in DPD.[9] The NHS in Italy provides co-payment exemption benefits and a specific code for PD (code 038), which allows a better identification of patients affected by this disease. For patients with advanced PD, palliative care, instead of specific pharmacological therapy, is recommended.[10]

MULTIPLE SCLEROSIS

Multiple Sclerosis (MS) is the most common demyelinating disease of the central nervous system and the first non-traumatic cause of disability in young adults.[11] This disease usually occurs between 20 and 40 years of age. In 1% of cases it occurs before the age of 10, while it is very rare before the age of 6.[12] MS is much more common among women (2-3 women/men ratio). The etiology of MS is unknown, although it is probably multifactorial and likely develops in genetically susceptible individuals, in the presence of environmental exposures. In 2013, an estimated 2.3 million people worldwide were affected by MS, of whom 600,000 were in Europe.[13] Italy is considered a high-prevalence Country, with about 118,000 people affected (about 195 per 100,000), with significant differences among regions.[14,15] The distribution of the disease is not uniform: MS is more common among people who live in temperate areas.[11]

MS has a variable clinical course and affects each person differently. However, two common patterns of the disease can be described: relapsing and progressive. The most common types of MS are: Relapsing-Remitting MS (RRMS), Secondary-Progressive MS (SPMS), Primary-Progressive MS (PPMS). Approximately 80%-85% of MS patients are initially diagnosed with RRMS. Patients can experience a worsening of symptoms, known as relapses, exacerbations, or attacks. Symptom-free periods can last months or even years, although over a 10-year period, about 50% of cases have a progression of the disease. A less frequent form of MS is Progressive-Remitting MS (PRMS), characterized by a steady worsening of symptoms and relapses, with or without remissions. The extreme heterogeneity in the clinical presentation of MS corresponds to a substantial variability in the response to Disease Modifying Drugs (DMD), that slow the progression of the disease. MS patients may benefit from healthcare co-payment exemptions after having certified their condition with a clinical diagnosis.

EPILEPSY

Epilepsy is a central nervous system disorder, in which an abnormal brain activity causes seizures or periods of unusual behavior, sensorial perceptions and sometimes loss of consciousness. Worldwide, the prevalence of this disease is approximately 4-8‰[16,17] and the annual incidence in developed Countries is nearly 0.5‰.[18] Epilepsy affects all ages, races, and socioeconomic statuses. Several epileptogenic conditions appear to be age-related,[19] higher among children (25% of all epileptic subjects are below the age of 15).[18] In the last few years, in Italy and in other developed Countries, prevalence and incidence among the elderly has increased, while it has decreased in pediatric subjects.[20,21]

The diagnosis of epilepsy is generally based on the referral of epileptic seizures. For a precise diagnosis, specific tests such as the electroencephalogram (EEG), are required. Other diagnostic tests (e.g., magnetic resonance imaging or cerebral computed tomography, laboratory tests) are indicated to ascertain or exclude specific causes of epileptic seizures. Antiepileptic Drugs (AEDs) (ATC category N03: N = Nervous system, 03 = antiepileptic drugs) have been the principal therapeutic option for epileptic disorders.[18] International guidelines recommend treatment should begin with a monotherapy, when possible.[22]

In Italy, people affected by epilepsy can benefit from an ad hoc disease-specific co-payment exemption (code 017).

In the last few years there has been an increasing interest in the employment and analysis of HADs to assess healthcare resource use, health outcomes, and costs of these three diseases.[23-25]

This systematic review aimed at identifying and describing all PD, MS, and epilepsy case-identification algorithms based on Italian HADs used in the past 10 years.

METHODS

All details regarding the methods used in these systematic reviews are available in a specific paper.[26] This article reports the study protocol and all details on the literature search (specific search strings to retrieve papers on PubMed that used Italian administrative healthcare data, inclusion/exclusion criteria and data extraction) and on the characterization of the selected papers and algorithms (strategy to identify original algorithms, algorithm objective definition). For all aspects not reported in the present section, we refer the reader to the protocol paper [26].The search string used to select PubMed records consisted of a combination of free text and MeSH terms, with a common part that focused on HADs (see Canova et al.[26] 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.

PARKINSON'S DISEASE: "Parkinson Disease"[Mesh] OR "parkinson"[title/Abstract]

MULTIPLE SCLEROSIS: Multiple Sclerosis [MeSH Terms] OR "Multiple Sclerosis"[Title/Abstract] OR "Multiple Sclerosis"[All Fields] OR "multiple sclerosis"[Title/Abstract] OR "multiple sclerosis"[All Fields]

EPILEPSY: "Epilepsy"[Mesh] OR "Epilepsy"[title/Abstract]

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

Three independent reviewers screened the articles published between 2007 and 2017 and classified pertinent ones, according to the objective for which the papers' algorithms were used. Only articles that used algorithms for "primary objectives" (**I** disease occurrence, **II** population/cohort selection, **III** outcome identification) were considered for algorithm extraction, since those with "secondary objectives" (**IV** 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). No changes were made to the methods agreed for all revisions and reported in the protocol.[26]

RESULTS

PARKINSON'S DISEASE, MULTIPLE SCLEROSIS, AND EPILEPSY: PAPERS IDENTIFIED WITH THE SEARCH STRATEGY

The search strategy led to the identification of 70 papers for Parkinson's disease, 154 for multiple sclerosis, and 100 papers for epilepsy (table 1).

Out of the selected articles, 64, 142, and 89 papers, respectively for PD, MS, and epilepsy, were excluded by title and abstract. This brought to respectively 6, 12, and 11 full-text screenings, out of which 3 papers for PD, 6 for MS, and 5 for epilepsy were considered pertinent.

All reasons for exclusions can be grouped into the following: "no specific disease algorithm reported" (all PD papers excluded, 66% of MS papers excluded) or "data collected exclusively from specific disease registries" (all epilepsy papers excluded, 33% of MS papers excluded).

References from the selected articles did not identify any additional relevant works, thus confirming the number of pertinent papers obtained from full-text screening.

Finally, out of all pertinent papers, 3 original algorithms for PD, 4 for MS, and 4 for epilepsy were identified.

PERTINENT PAPERS FOR PARKINSON'S DISEASE, MULTIPLE SCLEROSIS, AND EPILEPSY

The chronological distribution of pertinent papers showed that most of the articles were published in the years 2014-2017: all 3 articles for PD, 5 out of 6 for MS, and 4 out of 5 for epilepsy (table 2).

For all conditions, most of the works focused on a regional setting. In fact, only 1 paper for PD, 1 for MS, and 2 for epilepsy were based on a sub-regional context and 1 paper for PD was based on a national multicenter context (8 Local Health Units from 8 separate Italian regions).[27] No article was based on a multicenter, international context.

All papers (excluding 1 for PD[27] and 1 for MS[28]) used data from HADs, that was at least partly collected after 2008. In 2 out of 3 cases for PD, 5 out of 6 cases for MS, and in all 5 cases for epilepsy, the analyzed period was longer than one year.

Most works did not apply age restrictions in the inclusion criteria. Only 1 paper for PD[27] and 1 for epilepsy[29] focused on an elderly population (65+), while only 1 for epilepsy[18] focused on a young population (0-16).

Among all pertinent articles, one for PD [27] and one for epilepsy[18] have objective **IV+**, while the remaining 2 papers for PD, 6 papers for MS, and 4 papers for epilepsy have objectives **I-III**.

A complete list of the main characteristics of each pertinent article (regardless of the algorithm's objective), including the estimate of prevalence and/or incidence (only for articles with objective **I**) is shown in tables S1, S2, and S3 (see on-line supplementary materials) respectively for PD, MS, and epilepsy.

In particular, prevalence for PD across papers that estimated the occurrence of the disease (objective **I**), ranged from 0.28% to 0.49% (crude rates), with important differences even within the same article, according to the time frames and quantitative cut-offs that were considered (table S1) [9].

Among papers that studied MS using algorithms with objective **I**, prevalence estimates ranged from 0.89‰ to 1.95‰ (crude rates). One paper[30] estimated standardized prevalence rates for MS at 1.19‰ (European Standard Population for reference), while the crude prevalence calculated on the same data is 1.30‰ (table S2).

Finally, crude prevalence estimates of epilepsy ranged from 0.46% to 0.96% and the only crude incidence rate estimated was 0.53% (table S3).

PARKINSON'S DISEASE ALGORITHMS

Out of the 2 selected papers (focused on objectives **I-III**),[9,31] 3 original algorithms, all with objective I, were identified, and they all aimed to estimate the disease's prevalence: algorithm 1,[31] algorithm 2,[9] algorithm 3[9] (table 3A). None of these algorithms applied age restrictions.

For case identification, all algorithms used the Hospital Discharge records Database (HDD) (ICD-9-CM 332.* for Parkinson's Disease), Exemption from healthcare Co-payment Database (ECD code ICD-9-CM for Parkinson's Disease: 038), and APDs (ATC: N04*) in DPD.

Algorithm 1, which considered all APDs (codes N04*), applied several conditions to increase specificity: it excluded all cases with less than 3 different prescription dates. In addition, all cases with monotherapy (both considered as exclusive use of a single drug or prescription of different drugs of the same class) with non-specific APD drugs (anticholinergics, dopamine agonists, monoamine

	PARKINSON'S DISEASE	MULTIPLE SCLEROSIS	EPILEPSY
Papers identified by the string	70	154	100
Full-text readings	6	12	11
Pertinent papers	3	6	5
References added from bibliography	0	0	0
Total pertinent papers	3	6	5
Papers with objectives IV+*	1	0	1
Papers with objectives I-III*	2	6	4
Papers (with objective I-III)* with at least one original algorithm	2	4	4
Papers (with objective I-III)* with external validation	1	2	2
Original algorithms (with objective I-III)	**3**	**4**	**4**

* **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.

	PARKINSON'S DISEASE	MULTIPLE SCLEROSIS	EPILEPSY
Year of Publication			
2007-2010	0	1	0
2011-2013	0	0	1
2014-2017	3	5	4
Journal			
Italian	1	0	0
International	2	6	5
Setting			
Sub-regional (LHU, cities,..)	1	1	2
Regional (entire region)	1	5	3
National multicenter	1	0	0
International multicenter	0	0	0
Data time frame for the identification of the disease			
1 year	1	1	0
> 1 year	2	5	5
Use of data (even partial) following 2007 (≥2008)	2	5	5
Objective*			
I	2[9,31]	4[15,30,32,33]	3[19,34,43]
II	0	2[28,48]	0
III	0	0	1[29]
IV	0	0	0
V	0	0	0
VI	1[27]	0	1[18]
VII	0	0	0

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.

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			CASE DEFINITION		Incidence: criteria for the exclusion of prevalent cases (look-back time frame)	EVALUATION OF THE ALGORITHM			
				HDD ICD-9-CM code (Main diagnosis (M), Secondary diagnosis (S) Any diagnosis (A) Not reported (N))	ECD code	DPD code	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)
1	Tominz, 2015[31]	I	P	332,332.0 (A)	Parkinson: 038	N04*	[>=3 DPD (1year) NOT (monotherapy N04 non-specific drugs (anticholinergics, dopamine agonists, monoaminooxidases and beta inhibitors in absence of HDD in the previous 5 years)] OR ECD OR HDD (5 years)	All ages	–	–	–	–	S - C
2	Baldacci, 2015[9]	I	P	332.* (A)	Parkinson exemption	(Levodopa + benserazide) (N04BA02) OR (levodopa + carbidopa) (N04BA02), (Levodopa + carbidopa + entacapone) (N04BA03), (Melevodopa + carbidopa) (N04BA05), Ropinirole (N04BC04) (tablets >=2 mg or prolonged release >=4 mg), Pramipexole (N04BC05)(tablets >=0.7 mg or prolonged release >=0.52 mg), Apomorphine (N04BC07), Rotigotine (N04BC09), Selegiline (N04BD01), Rasagiline (N04BD02), Tolcapone (N04BX01), Entacapone (N04BX02), Triesifenidile (N04AA01), Biperiden (N04AA02)	HDD OR ECD OR >=2 DPD (1 year at least 6 months apart)	All ages	–	–	sample patients from the databases of the movement disorders clinics of 5 institutions located in different areas of the Tuscan Public Health System	Se: 91.20%	C
3	Baldacci, 2015[9]	I	P	332.* (A)	Parkinson exemption	(Levodopa+ carbidopa + entacapone) (N04BA03), Ropinirole (N04BC04) (tablets >=2 mg or prolonged release >=4 mg), Pramipexole (N04BC05) (tablets >=0.7 mg or prolonged release >=0.52 mg), Apomorphine (N04BC07), Rotigotine (N04BC09), Selegiline (N04BD01), Rasagiline (N04BD02), Tolcapone (N04BX01), Entacapone (N04BX02)	HDD OR ECD OR >=2 DPD (1 year at least 6 months apart)	All ages	–	–	sample patients from the databases of the movement disorders clinics of 5 institutions located in different areas of the Tuscan Public Health System	Se: 75.86%	C

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

Table 3A. Characteristics of Parkinson's Disease case-identification algorithms published in PubMed between 2007 and 2017.

Epidemiol Prev 2019; 43 (4) Suppl 2:62-74. doi: 10.19191/EP19.4.S2.P062.093

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				CASE DEFINITION		Incidence: criteria for the exclusion of prevalent cases (look-back time frame)	EVALUATION OF THE ALGORITHM			
				HDD ICD-9-CM code (Main diagnosis (M) Any diagnosis (A) Not reported (N))	ECD code	DPD code	Other sources (code)	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)
1	Pedotti, 2009[28]	II	I	–	340	–	–	1ECD(1year)	All ages	new activeted exempion in the study period	–	–	–	–
2	Bergamaschi, 2013[32]	I	P	340; 341.8; 341.9 (A)	046	"File F"	–	HDD (6 years) OR DPD (6 years) OR ECD (6 years) (2004-2009)	All ages	–	–	–	–	–
3	Bezzini, 2016[33] Battaglia, 2017[15]	I	P	340 (A)	046	>1 (Glatiramer acetate L03AX13, Interferon beta 1A L03AB07, Interferon beta 1B L03AB08, Fingolimod L04AA27, Natalizumab L04AA23)	RC 340 HC 340	HDD (13 years) OR 2 DPD (9 years) OR ECD (1 years) OR 1 RC/HC (2 years)	All ages	–	–	–	Tuscan MS register+ list of presumably unaffected subjects extracted from SIS	Se: 98.00%; Sp: 99.00%	C
4	Bargagli, 2016[30] Colais, 2017[48]	I; II	P	340 (A)	046	[Interferon beta-1a (Avonex, Rebif), Interferon beta-1b (Betaferon, Extavia), Copaxone (Glatiramer Acetate) and Tysabri (Natalizumab)]	–	HDD (6 years) OR DPD (6 years) OR ECD (6 years)	All ages	–	–	–	5 clinical centers enrolled in the same period (2006-2011)	Se: 85.00%	C

HDD: hospital discharge record database; **ECD:** exemption from healthcare co-payment database; **DPD:** drug prescription database; **RC:** residential long-term care; **HC:** home care; **Se:** sensitivity; **Sp:** specificity; **PPV:** positive predictive value; **NPV:** negative predictive value

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

oxidase type B inhibitors), without a hospitalization with ICD-9 code = 332 (Parkinson's disease) or 332.0 (agitating paralysis) in the period of study, were excluded.
Algorithms 2 and 3 only differed for quantitative cut-off differences in APD prescriptions (belonging to the N04 ATC family) and used as inclusion criteria the presence of a minimum of two prescriptions for at least one APD, dispensed at least 6 months apart one from the other in the study year.
Only algorithm 1 reported data on the internal coherence of the sources, while the other 2 algorithms presented the absolute and exclusive contribution of the sources: absolute contribution showed that drugs alone were sufficient to capture almost 95% of the estimated PD population, while exclusive contribution of the sources revealed that DPD was the only source to capture 72.70% of PD cases, HDD exclusively captured only 2.98% of cases, and ECD captured 2.20%.
Validation by means of an external independent source was present for algorithms 2 and 3 and was obtained with a sample of patients taken from the databases of 5 clinics for movement disorders, located in different areas in Tuscany. The study reported a sensitivity of 91.20% for algorithm 2 and 75.86% for algorithm 3.[9]

MULTIPLE SCLEROSIS ALGORITHMS

Out of the 6 selected papers (4 with objective I and 2 with objective II), 4 original algorithms were identified: 3 aimed to estimate the disease's prevalence (Bergamaschi et al.:[32] algorithm 2; Bezzini et al.:[33] algorithm 3; Bargagli et al.:[30] algorithm 4) and 1 to identify cases, to evaluate the association with other diseases (Pedotti et al.:[28] algorithm 1) (table 3B). No age restrictions were applied to any of the algorithms. To identify cases, 3 out of 4 original algorithms used HDD: algorithms 3 and 4 used ICD-9 code 340.* for MS while algorithm 2 also included codes 341.8 ("other demyelinating diseases of the central nervous system") and 341.9 ("other demyelinating diseases of the central nervous system, unspecified"). All algorithms used the ECD code 046 specific for MS. Algorithm 1 used this database as exclusive source of data for case identification. Home and residential long-term care databases were considered by algorithm 3 as an additional source of information (ICD-9 code 340.*), besides HDD, ECD, and DPD. All the algorithms aiming to identify cases to estimate the disease's prevalence used DPD and considered the ATC codes of different disease modifying drugs (DMD) specific for MS. In particular, the type of drug considered in the algorithms depended on the year of marketing authorization and the time span considered in the study. Algorithm 3 considered at least two DMD prescriptions to identify an MS case, while in the other algorithms the presence of only one prescription was sufficient. The papers that used algorithms 3 and 4 provided information on the absolute and exclusive contribution of the sources. In algorithm 3, HDD identified about 85% of all MS patients, while ECD included most individuals affected by MS (67%) captured by the algorithm. In general, the exclusive contribution of HDD was higher than that of other databases, although some differences between algorithms were present (30% in algorithm 3 vs. 16% in algorithm 4). Validation by means of an external independent source was present for both algorithm 3 and 4. In the study by Bezzini et al. (2016), the validation analysis included patients extracted from the regional MS Registry and a cohort of individuals who had never received healthcare services related to MS within the Regional Healthcare Service.[33] The study estimated a sensitivity of 98.00% and a specificity of 99.90%.[33] Bargagli et al. (2016) validated their algorithm by comparing MS cases, identified through the algorithm, with the population of patients that entered any of the five specialized high-volume centers and estimated a sensitivity of 85.00%.[30]

EPILEPSY ALGORITHMS

Out of the 5 selected papers, 4 original algorithms with objectives **I** and **III** were identified (table 3C). Three algorithms considered all ages, while 1 algorithm included exclusively subjects over 65 years of age.
Three algorithms with objective **I** were used to estimate the prevalence of epilepsy, while one algorithm with objective **III** was used to estimate incident seizures.
All algorithms are based on at least a combination of EEG (code 89.14*, 89.191 in Ambulatory Care-service Databases (ACDs)) and an antiepileptic drug (AED) in DPD. All AEDs included in the algorithms fall into a specific ATC category (N03), except for one active ingredient that belongs to the subgroup of psycholeptic drugs (ATC N05): Clobazam (N05BA09), which is included in all original algorithms (table S3.1; see on-line supplementary materials).
Three algorithms use HDD (ICD-9 codes: 333.2 Myoclonus, 345.* Epilepsy and recurrent seizures, 779.0 Convulsions in newborn, 780.3 Convulsions, 781.0 Abnormal involuntary movements), while only 1 algorithm used ECD (code 017).[34]
The less recent paper in our review presented 15 algorithms for epilepsy with different combinations of sources: ECD, HDD, ACDs, and DPD. Overall, the one that maximized the Area Under the Curve (AUC), was "EEG AND DPD" and was considered to be the "best algorithm (successively implemented by Giussani et al.:[19] algorithm 2), with an AUC value of 0.9478 (IC 0.9121-0.9836).[34]
To validate this, a total of 11 General Practitioners (GPs)

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				CASE DEFINITION		Incidence: criteria for the exclusion of prevalent cases (look-back time frame)	EVALUATION OF THE ALGORITHM			
				HDD ICD-9-CM code (Main diagnosis (M) Any diagnosis (A) Not reported (N))	ECD code	DPD code	Other sources (code)	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)
1	Franchi, 2013[34]	I	P	345.x, 333.2, 780.3, 779.0, 781.0 (N)	epilepsy: 017	(1) At least two of Carbamazepine (N03AF01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Primidone (N03AA03), Barbexaclone (N03AA04), Clonazepam (N03AE01), Ethosuximide (N03AD01), Valproate (N03AG01), Valpromide (N03AG02), Clobazam (N05BA09), Vigabatrin (N03AG04), Felbamate (N03AX10), Tiagabine (N03AG06), Pregabalin (N03AX16), Oxcarbazepine ((N03AF02), Gabapentin (N03AX12), Topiramate (N03AX11), Levetiracetam (N03AX14), Zonisamide (N03AX15); (2) One between Carbamazepine (N03AF01), Valproate (N03AG01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Primidone (N03AA03), Barbexaclone (N03AA04), Clonazepam (N03AE01), Ethosuximide (N03AD01), Clobazam (N05BA09), Vigabatrin (N03AG04), Felbamate (N03AX10), Tiagabine (N03AG06), Oxcarbazepine ((N03AF02), Levetiracetam (N03AX14), Zonisamide (N03AX15)	ACD: EEG (code not specified)	15 different combinations of diagnostic codes (in 9 year)	–	–	–	11 GP	Best algorithm: EEG AND DPD: Se 85.90%, Sp 99.80%, PPV 64.20% NPV 99.90%	S
2	Giussani, 2014[19]	I	I-P	–	–	(1) At least two of Carbamazepine (N03AF01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Primidone (N03AA03), Barbexaclone (N03AA04), Clonazepam (N03AE01), Ethosuximide (N03AD01), Valproate (N03AG01), Valpromide (N03AG02), Clobazam (N05BA09), Vigabatrin (N03AG04), Felbamate (N03AX10), Tiagabine (N03AG06), Pregabalin (N03AX16), Oxcarbazepine ((N03AF02), Gabapentin (N03AX12), Topiramate (N03AX11), Levetiracetam (N03AX14), Zonisamide (N03AX15); (2) One between Carbamazepine (N03AF01), Valproate (N03AG01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Primidone (N03AA03), Barbexaclone (N03AA04), Clonazepam (N03AE01), Ethosuximide (N03AD01), Clobazam (N05BA09), Vigabatrin (N03AG04), Felbamate (N03AX10), Tiagabine (N03AG06), Oxcarbazepine ((N03AF02), Levetiracetam (N03AX14), Zonisamide (N03AX15)	ACD: EEG (code not specified)	EEG (in 9 years) AND DPD (in 1 year)	–	4 years	Best algorithm (from validation study) of Franchi, 2013	–	–	–
3	Baviera, 2017[29]	III	I	345.x, 780.3, 781.0 (A)	–	Carbamazepine (N03AF01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Barbexaclone (N03AA04), Clobazam (N05BA09), Clonazepam (N03AE01), Etosuccimide (N03AD01), Felbamate (N03AX10), Gabapentin (N03AX12), Lamotrigine (N03AX09), Levitiracetam (N03AX14), Oxcarbazepine (N03AF02), Pregabalin (N03AX16), Primidone (N03AA03), Tiagabine (N03AG06), Topiramate (N03AX11), Valproate (N03AG01), Valpromide (N03AG02), Vigabatrin (N03AG04), Zonisamide (N03AX15)	ACD: EEG (8914, 89141, 89142, 89143, 89144, 89145, 89191)	HDD OR (EEG AND DPD (6 months before or after the SVD))	–	2 years	Franchi, 2013	–	–	–

4	Bellini, 2017[43]	I	P	345.x (A)	–	Carbamazepine (N03AF01), Phenytoin (N03AB02), Phenobarbital (N03AA02), Oxcarbazepine (N03AF02), Lamotrigine (N03AX09), Topiramate (N03AX11), Pregabalin (N03AX16), Gabapentin (N03AX12), Clobazam (N05BA09), Clonazepam (N03AE01) (not specific for epilepsy) Levitiracetam (N03AX14), Lacosamide (N03AX18), Perampanel (N03AX22), Eslicarbazepine (N03AF04), Zonisamide (N03AX15), Rufinamide (N03AF03), Stiripentol (N03AX17), Tiagabine (N03AG06), Felbamate (N03AX10), Vigabatrin (N03AG04) (specific for epilepsy)	ACD: EEG (89.14*)	[(EEG AND ≥2 DPD at a distance of 12 months) OR (≥2 DPD "specific for epilepsy" at a distance of 12 months) OR (HDD)] AND (DPD OR HDD in 1 year: 2015, EEG not specified, available from 2002)	–	–	–	Two neurological centres for the treatment of epilepsy in Tuscany	Se: 87.30%, Sp: 99.90%	C

HDD: hospital discharge record database; **ECD:** exemption from healthcare co-payment database; **DPD:** drug prescription database; **ACD:** ambulatory care-services database; **EEG:** electroencephalogram; **Se:** sensitivity; **Sp:** specificity; **PPV:** positive predictive value; **NPV:** negative predictive value

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

working in the study area were involved and represented the gold standard. Each GP kept electronic records of his/her patients and these records were examined in search of patients with epilepsy.[34] The validation of this algorithm estimated a sensitivity of 85.90%, a specificity of 99.80%, a positive predictive value of 64.20%, and a negative predictive value: 99.90%.

Algorithm 3 (table 3C) identified an epileptic subject according to the criteria proposed by algorithm 1, with the addition of cases identified by an HDD diagnosis of seizure or epilepsy.

Algorithm 4 (table 3C) presented a more complex structure, that considered HDD, EEG, and DPD as sources of case identification. In addition to the already used "EEG AND DPD (excluding monotherapy)" or "HDD specific for epileptic seizures", the algorithm added at least 2 drugs "specific for epilepsy" prescribed 12 months apart. This algorithm was validated against two neurological centers for the treatment of epilepsy in Tuscany, estimating a sensitivity of 87.30% and specificity of 99.90%. This paper also quantified the absolute and exclusive contribution of the sources: combinations of EEG "AND" any AED, alone, were able to capture about 74% of cases affected by epilepsy while 20% of cases affected by epilepsy were only identified with specific AEDs and 4% of cases were captured exclusively by HDD.

DISCUSSION

This work reviewed and described all PD, MS, and epilepsy case-identification algorithms that have been published in the last decade in Italy. We chose to focus on a national setting, since experiences developed in other Countries often rely on data sources that are not present in Italy (mainly information retrieved from outpatient claims of ambulatory care, GP, specialists and pediatricians). Moreover, we chose to perform a comprehensive review, regardless of the presence of an algorithm's validation, to give a complete overview of the characteristics and fields of utilization of all algorithm-based approaches used and published in the literature, within the Italian context.

Despite the fact that use of HADs has recently become an accessible tool to monitor frequency and trends of several diseases,[35] the number of Italian published papers that focused on identifying PD, MS, and epilepsy by means of HADs, found by this review, is relatively small. Evidence deriving from this work suggests the following:

Parkinson's disease. The identified algorithms for PD were quite similar to one another, with an exclusive use of the three sources DPD, HDD, and ECD. DPD is the most relevant source of PD case-identification: this source alone is sufficient to capture about 95% of the estimated PD population;[9]

Multiple Sclerosis. The most important administrative databases to identify MS patients are HDD, DPD and ECD, while the contribution of home and residential long-term care databases is minimal.

Epilepsy. The most used algorithm among pertinent papers identified by our research consists in a combination of EEG "AND" any AED prescription, which was sufficient to capture around 74% of epileptic cases.[34] The inclusion of the EEG in algorithm 1 (table 3C) significantly improved the positive predictive value and according to the study of Holden et al. (2005)[36] it significantly helped to exclude non-epileptic subjects.

ALGORITHM OBJECTIVES

Among algorithms for case identification of PD and epilepsy, we did not notice any significant differences based on the algorithms' objectives, perhaps because the papers used the algorithms for similar purposes. All the algorithms, excluding one for epilepsy, share the same objective.

Some differences are present among the MS algorithms, in relation to the study objectives. When estimating prevalence or spatial distribution of MS, researchers used many (largely overlapping) data sources and applied very similar criteria for case identification. Pedotti et al. (2009) used the population of MS patients identified on the basis of a single data source, i.e., ECD, to examine the association between MS and allergic disorders (objective **II**)[28]. In this case, the algorithm identified a group of MS patients to detect and validate the outcome under study. Bergamaschi et al. (2014) used different HADs to select a population affected by MS.[32] In this case, the Authors did not specify how the algorithm was developed and what specific drugs were considered, limiting the reproducibility of the algorithm. Moreover, this study included patients discharged with an ICD-9 code for other demyelinating disease of central nervous system or unspecified demyelinating disease, diagnoses that were not considered in other studies, aiming to estimate the prevalence of the disease (objective **I**).

THE ROLE OF DRUG PRESCRIPTIONS IN PD, MS, AND EPILEPSY ALGORITHMS

Parkinson's disease. In algorithm 1 (table 3A) DPD contributed with 81.57% of the total cases and provided an exclusive contribution of 41.70%. In algorithm 2 (table 3A) subjects identified with pharmaceuticals were 94.57% of the entire population and 72.70% were captured by DPD exclusively. Hence, a good reproducibility of the results found in these papers is very likely, at least in all Countries that keep electronic records of drug prescriptions.

The presence of a prescription for any anti-Parkinson medication has been reported to have a positive predictive value of 85.50% and a negative predictive value of 48.60%, with the highest sensitivity (74.10%) but the lowest specificity (62.40%) for levodopa (L-DOPA).[37] For these reasons, Baldacci et al. (2015)[9] and Tominz et al. (2015),[31] in accordance with the literature,[38,39] excluded all subjects that presented only one pharmaceutical prescription or subjects in monotherapy with anticholinergics, dopamine agonists, and monoamine oxidase type B inhibitors (if absent in the HDD archive) (9, 30). While algorithm 1 considered the entire N04 family of APD (following the appropriate rules for the exclusion of non-univocal cases), algorithms 2 and 3 considered specific combinations of drugs, chosen a priori, as shown in detail in table 3A.

Multiple sclerosis. Most of the algorithms proposed in the last years to identify MS cases from HADs rely on drug prescriptions [40,41]. It should be noted that through DPD the only patients that can be identified are those who are in a stage of the disease in which pharmacological treatment with disease-modifying drugs is recommended. In algorithm 3 (table 3B), DPD contributed to identify 41% of the total cases and provided an exclusive contribution of 5%. In algorithm 4 (table 3B) patients identified through drug prescriptions were 61% of the entire population and 10% were present only in DPD. The sole difference between the two algorithms was the number of drug claims required to define cases (2 in algorithm 3 and only 1 in algorithm 4), as shown in detail in table 3B.

Epilepsy. Active ingredients and their combinations, selected in algorithm 1 (table 3C), are consistent with the prescribing patterns for drug-resistant epilepsy used nowadays in all reference centers in Italy [42]. Patients receiving monotherapy with drugs also used for clinical conditions other than epilepsy were excluded. The exceptions were patients receiving only Carbamazepine or Valproic acid [34]. This algorithm has become the reference point for all papers published later.

The authors that proposed algorithm 4 (table 3A) observed that the combination of EEG and AEDs resulted in a percentage of about 15% false negatives and 30% false positives. Among false positives, about 50% took AEDs for seizure prophylaxis and the remaining had a single unprovoked seizure or another clinical condition [43]. Therefore, the authors added the condition of having a concomitant EEG in the algorithm only when the dispensed AEDs were not specific for epilepsy (for details, see table S3, on-line supplementary materials). By doing this, the authors expected to find a lower percentage of false negatives, although the percentage of false positives was not supposed to vary.

VALIDATION STUDIES

Results from this systematic review highlighted the paucity of experiences to evaluate algorithm accuracy and the heterogeneous approaches to case identification. The absence of standardized HAD-based case-identification algorithms for these neurological conditions is mainly due to a difficulty in the validation process.

One out of two papers on PD presented a validation study. The sensitivity of their prevalence estimations was calculated by using a sample of clinically diagnosed PD patients in 5 clinics for movement disorders.

Among the studies on MS included in this review, 2 performed a validation of the algorithm based on administrative databases. Notably, algorithm 3 estimated very high values of sensitivity and specificity, respectively 98% and 99%,[33] while algorithm 4 reported a sensitivity of 85%.[30] These results suggest that the use of HDD, DPD, or ECD allows to identify MS cases with a good level of accuracy.

Two papers for epilepsy presented validation studies. The first used a group of 11 GPs (including 3 pediatricians), with 15,728 affiliates as gold standard.[34] The second paper tested sensitivity through specialized clinical centers in the study area.[43] The two validations showed encouraging sensitivity and specificity of algorithms employing DPD "AND" EEG both alone and in combination with other sources. The paucity of experiences of external validation is not just an Italian limit: worldwide, there have been few studies that validated HADs for capturing cases affected by epilepsy.[44]

STRENGTHS AND LIMITATIONS

Difficulties in case identification of Parkinson's disease. As mentioned by Baldacci et al. (2015), identifying a population of patients affected by PD, based on HADs, represents a complicated issue for several, non-mutually exclusive reasons.[9] It is important to mention the absence of specific diagnostic tests or procedures to ascertain the presence of PD,[45] the lack of a precise diagnosis on drug prescriptions, or even the possible absence of PD specific exemptions, as patients may already have exemptions for other causes.

In addition, in the majority of studies based on HADs, the diagnosis was not clinically verified.[37,45] Algorithms 2 and 3 (table 3A) showed an estimated sensitivity on a population of PD patients with "the best possible diagnosis" made by experts in movement disorders, based on clinical data. It is worth noting that the accuracy of PD diagnosis strongly depends on the expertise of the clinician.[46]

Difficulties in case identification of multiple sclerosis. As reported by the majority of studies dealing with the identification of MS cases from HADs, algorithms tend to miss patients who do not access healthcare services. It appears that patients with a high level of disability or those in an early stage of the disease are more difficult to capture through administrative databases. Furthermore, algorithms based on HADs do not allow to disentangle different clinical forms of neurological disorders that entail both a different clinical management and use of healthcare resources.

Difficulties in epilepsy case identification. One of the most difficult HAD-based differential diagnoses is between epilepsy and migraine, which is itself an epileptic symptom.[47] Diagnostic and pharmaceutical overlap are present among these conditions.[47] This makes it theoretically impossible to exclude subjects who suffer from migraine from those incorrectly identified as epileptic (false positive). In any case, the analysis of false positives in the study conducted by Franchi et al. (2013) showed that almost two-thirds were patients who took AEDs for seizure prophylaxis or had a single unprovoked seizure, while one-third of these suffered from another clinical condition requiring treatment with AEDs.[34] None of them suffered from migraine.

CONCLUSION

This review did not find particularly relevant differences in complexity among algorithms, especially for PD.

PD, MS, and epilepsy share a scarcity of external validation studies. An extensive validation process is however essential to better understand the specific characteristics of algorithms, to assess their performance, and to choose the most appropriate one according to the study's aim. It would be advisable to introduce stricter requirements for HAD-based case-ascertainment algorithms, to enforce the assessment of their validity.

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. de Lau LML, Breteler MMB. Epidemiology of Parkinson's disease. Lancet Neurol 2006;5(6):525-35.
2. Dorsey ER, Constantinescu R, Thompson JP, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology 2007;68(5):384-86.
3. Guttman M, Slaughter PM, Theriault M-E, DeBoer DP, Naylor CD. Burden of parkinsonism: a population-based study. Mov Disord Off J Mov Disord Soc 2003;18(3):313-19.
4. Damiano AM, Snyder C, Strausser B, Willian MK. A review of health-related quality-of-life concepts and measures for Parkinson's disease. Qual Life Res 1999:8(3):235-43.
5. Tysnes OB, Storstein A. Epidemiology of Parkinson's disease. J Neural Transm (Vienna) 2017;124(8):901-05.
6. Nerius M, Fink A, Doblhammer G. Parkinson's disease in Germany: prevalence and incidence based on health claims data. Acta Neurol Scand 2017;136(5):386-92.
7. Bach J-P, Ziegler U, Deuschl G, Dodel R, Doblhammer-Reiter G. Projected numbers of people with movement disorders in the years 2030 and 2050. Mov Disord Off J Mov Disord Soc 2011;26(12):2286-90.
8. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;51(6):745-52.
9. Baldacci F, Policardo L, Rossi S, et al. Reliability of administrative data for the identification of Parkinson's disease cohorts. Neurol Sci 2015;36(5):783-86.
10. Cartabellotta A, Berti F, Patti AL, Quintana S, Eleopra R. Linee guida per la diagnosi e il trattamento della malattia di Parkinson. Evidence 2018;10(4):e1000181.
11. Reich DS. Imag(in)ing multiple sclerosis: Time to take better pictures. J Neuroimmunol 2017;304:72-80.
12. Pena JA, Lotze TE. Pediatric Multiple Sclerosis: Current Concepts and Consensus Definitions. Autoimmune Dis 2013;2013:673947.
13. Multiple Sclerosis International Federation. Available from: https://www.msif.org/about-us/who-we-are-and-what-we-do/advocacy/atlas/
14. Associazione Italiana Sclerosi Multipla (AISM). Il Barometro della SM 2018. Available from: https://www.aism.it/il_barometro_della_sm_2018
15. Battaglia MA, Bezzini D. Estimated prevalence of multiple sclerosis in Italy in 2015. Neurol Sci 2017;38(3):473-79.
16. Forsgren L, Beghi E, Oun A, Sillanpää M. The epidemiology of epilepsy in Europe - a systematic review. Eur J Neurol 2005;12(4):245-53.
17. World Health Organization. Neurological Disorders: Public Health Challenges. Geneva: World Health Organization; 2006.
18. Putignano D, Clavenna A, Campi R, et al. Antiepileptic drug use in Italian children over a decade. Eur J Clin Pharmacol 2017;73(2):241-48.
19. Giussani G, Franchi C, Messina P, Nobili A, Beghi E, the EPIRES Group. Prevalence and incidence of epilepsy in a well-defined population of Northern Italy. Epilepsia 2014;55(10):1526-33.
20. Casetta I, Pugliatti M, Faggioli R, et al. Incidence of childhood and adolescence epilepsy: a community-based prospective study in the province of Ferrara and in Copparo, Italy, 1996-2005: Incidence of childhood epilepsy in Italy. Eur J Neurol 2012;19(2):312-16.
21. Hauser WA. Recent developments in the epidemiology of epilepsy. Acta Neurol Scand 1995;92:17-21.
22. NICE - National Institute for Health and Care Excellence. Epilepsies: diagnosis and management. January 2012. Available from: https://www.nice.org.uk/guidance/cg137
23. Kurth T, Lewis BE, Walker AM. Health care resource utilization in patients with active epilepsy: Health care Resource Utilization. Epilepsia 2010;51(5):874-82.
24. St. Germaine-Smith C, Liu M, Quan H, Wiebe S, Jette N. Development of an epilepsy-specific risk adjustment comorbidity index: Epilepsy-Specific Risk Adjustment. Epilepsia 2011;52(12):2161-67.
25. Thurman DJ, Beghi E, Begley CE, et al. Standards for epidemiologic studies and surveillance of epilepsy. Epilepsia 2011;52 Suppl 7:2-26.
26. 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.
27. Rossini M, Viapiana O, Adami S, et al. Medication Use Before and After Hip Fracture: A Population-Based Cohort and Case-Control Study. Drugs Aging 2014;31(7):547-53.
28. Pedotti R, Farinotti M, Falcone C, et al. Allergy and multiple sclerosis: a population-based case-control study. Mult Scler J 2009;15(8):899-906.
29. Baviera M, Roncaglioni MC, Tettamanti M, et al. Diabetes mellitus: a risk factor for seizures in the elderly – a population-based study. Acta Diabetol 2017;54(9):863-70.
30. Bargagli AM, Colais P, Agabiti N, et al. Prevalence of multiple sclerosis in the Lazio region, Italy: use of an algorithm based on health information systems. J Neurol 2016;263(4):751-59.
31. Tominz R, Marin L, Mezzarobba S. Stima della prevalenza della malattia di Parkinson con l'utilizzo dl archivi sanitari elettronici correnti. Recenti Prog Med 2015;106:97-102.
32. Bergamaschi R, Agnello M, Colombo E, et al. Detection of clinical relapses in multiple sclerosis cohorts: construction and validation of a model based on administrative data. Neurol Sci 2014;35(2):265-69.
33. Bezzini D, Policardo L, Meucci G, et al. Prevalence of Multiple Sclerosis in Tuscany (Central Italy): A Study Based on Validated Administrative Data. Neuroepidemiology 2016;46(1):37-42.
34. Franchi C, Giussani G, Messina P, et al. Validation of healthcare administrative data for the diagnosis of epilepsy. J Epidemiol Community Health 2013;67(12):1019-24.
35. Langan, Benchimol, Sørensen H, et al. Setting the RECORD straight: developing a guideline for the REporting of studies Conducted using Observational Routinely collected Data. Clin Epidemiol 2013:5:29-31.
36. Holden EW, Grossman E, Nguyen HT, et al. Developing a computer algorithm to identify epilepsy cases in managed care organizations. Dis Manag DM 2005;8(1):1-14.
37. Szumski NR, Cheng EM. Optimizing algorithms to identify Parkinson's disease cases within an administrative database. Mov Disord 2009;24(1):51-56.
38. Chini F, Pezzotti P, Orzella L, Borgia P, Guasticchi G. Can we use the pharmacy data to estimate the prevalence of chronic conditions? a comparison of multiple data sources. BMC Public Health 2011;11:688.
39. Maio V, Yuen E, Rabinowitz C, et al. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. J Health Serv Res Policy 2005;10(4):232-38.
40 Blozik E, Rapold R, Eichler K, Reich O. Epidemiology and costs of multiple sclerosis in Switzerland: an analysis of health-care claims data, 2011-2015. Neuropsychiatr Dis Treat 2017;13:2737-45.
41. Palmer AJ, Hitchens PL, Simpson S, O'Leary B, Colman S, Taylor BV. A novel method for calculating prevalence of multiple sclerosis in Australia. Mult Scler J 2013;19(13):1704-11.
42. Malerba A, Ciampa C, De Fazio S, et al. Patterns of prescription of antiepileptic drugs in patients with refractory epilepsy at tertiary referral centres in Italy. Epilepsy Res 2010;91(2-3):273-82.
43. Bellini I, Policardo L, Zaccara G, et al. Identification of prevalent patients with epilepsy using administrative data: the Tuscany experience. Neurol Sci 2017;38(4):571-77.
44. Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. BMJ Open 2018;8(6):e020824.
45. Swarztrauber K, Anau J, Peters D. Identifying and distinguishing cases of parkinsonism and Parkinson's disease using ICD-9 CM codes and pharmacy data. Mov Disord Off J Mov Disord Soc 2005;20(8):964-70.
46. Hughes AJ, Daniel SE, Lees AJ. Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. Neurology 2001;57(8):1497-99.
47. Bigal ME, Rapoport AM, Lipton RB, Tepper SJ, Sheftell FD. Assessment of migraine disability using the migraine disability assessment (MIDAS) questionnaire: a comparison of chronic migraine with episodic migraine. Headache 2003;43(4):336-42.
48. Colais P, Agabiti N, Davoli M, et al. Identifying Relapses in Multiple Sclerosis Patients through Administrative Data: A Validation Study in the Lazio Region, Italy. Neuroepidemiology 2017;48(3-4):171-78.