Prevalence of Multiple Sclerosis in Tuscany (Central Italy): A Study Based on Validated Administrative Data

Daiana Bezzini\textsuperscript{a,b} Laura Policardo\textsuperscript{c} Giuseppe Meucci\textsuperscript{d} Monica Ulivelli\textsuperscript{e} Sabina Bartalini\textsuperscript{e} Francesco Profili\textsuperscript{c} Mario Alberto Battaglia\textsuperscript{a,b} Paolo Francesconi\textsuperscript{c}

\textsuperscript{a}Department of Life Science, University of Siena, Siena, \textsuperscript{b}Fondazione Italiana Sclerosi Multipla, Genova, \textsuperscript{c}Agenzia Regionale di Sanità della Toscana, Firenze, \textsuperscript{d}Unit of Neurology, USL6, Livorno, and \textsuperscript{e}Department of Medicine, Surgery and Neuroscience, UOC Neurology and Clinical Neurophysiology, Siena, Italy

\section*{Conclusions:} We found a prevalence higher than the rates present in literature. Our algorithm, based on administrative data, can accurately identify MS patients; moreover, the resulting cohort is suitable to monitor disease care pathways.

\section*{Introduction}

Multiple sclerosis (MS) is the most common non-traumatic disease of the central nervous system that causes permanent disability in young adults \cite{WHO_MSIF}. According to WHO/MSIF MS Atlas, it affects 2.3 million people worldwide, with a global median prevalence of 35 cases and a median incidence of 4 cases per 100,000 \cite{WHO_MSIF}. In a recent literature review, the authors reported a median estimated prevalence of 112.0 (with a range of 5.2–335.0) and a median estimated incidence of 5.2 (range 0.5–20.6) per 100,000 \cite{Leber}. In Europe, there are more than 600,000 patients with a median prevalence of 100 cases and a median incidence of 5.5 cases per 100,000 \cite{WHO_MSIF}.

Italy is classified as a high-risk area for MS, with highest rates in the island of Sardinia, and with no evidence of the latitude gradient \cite{Nicolini}. The Italian MS patient society...
(AISM) estimates that in Italy there are 72,000 cases with an incidence of 2,000 cases per year [5].

The latest rates published in Italy show a prevalence of around 140–170 cases per 100,000 (in 2009, 2007 and 2005) [6–8] with the exception of Sardinia, where prevalence raised up to about 224 cases per 100,000 in 2009 [9]. The incidence was of about 5.5 cases per 100,000 in Continental Italy in 2009 [6] and 9.7 cases per 100,000 in Sar- dinia in 2011 [10]. In Tuscany, the latest estimates show a prevalence of 56 per 100,000 (1991) [11] and a mortality of 0.4 for males and 0.6 for females (2002–2006) [12]. Nowadays, in Italy, prevalence is absolutely higher than the above-mentioned estimates because incidence exceeds mortality [13] and this difference will probably increase, according to the notion that MS incidence is increasing over time [6, 13–16]. Some authors even hypothesize an increase in the incidence due not only to the improvement of diagnostic technique and earlier diagnosis but probably also to some still unknown environmental etiological factors [15].

The aim of our study is to update the prevalence rate of MS in Tuscany (Central Italy) as at 2011.

For the surveillance of MS and other chronic diseases, there are disease-specific clinical registries that can assist in the routine clinical management and they can also facilitate the long-term monitoring of patients, such as the ones treated by disease-modifying drugs (DMDs) in MS.

The most utilized database in Italy is the Multiple Sclerosis Database Network based on iMed, an electronic clinical database [17, 18].

There are also 2 regional population registries of MS, in Tuscany and in Liguria, both promoted by the Italian Multiple Sclerosis Foundation (FISM).

Population-based studies are preferable to monitor disease epidemiology, comorbidities and care pathways, but also to describe the burden of this disease and to plan its prevention, treatment, as well as the management strategies and resource allocation necessary to cope with it [19, 20]. Administrative data offer a unique opportunity to study the population-based prevalence of chronic diseases such as MS [19, 21–29], or Parkinson’s disease [30].

In fact, administrative data are cost-effective, continuous and standardized, easy to be obtained and processed, and cover all the resident population enrolled in publicly funded health systems, such as the Italian system [25], with the possibility to link to individual consumption of resources. However, administrative data have some limitations: they lack comprehensive clinical data, they do not include some population subgroups, nor those services uncovered by the National Health Service (NHS); moreover, the quality of data is not always perfect. For these reasons, it will be desirable to merge clinical registries and administrative data to obtain a complete view of MS, collecting both clinical and health care information [21].

In Italy, NHS is tax-based and organized in 3 levels (national, regional and local). Administrative data, such as inpatient and outpatient care, and drug dispensations, are routinely collected by Italian local health authorities (LHA). Data are sent in an anonymous form to the national level, without the unique personal identification code; so, record-linkage can be performed only in the local and regional levels. These homogeneous data, collected at the local level, form the Italian administrative databases [29].

However, because administrative data are collected not for epidemiological surveillance, but for the health system management and for reimbursement by the NHS, it is necessary to assess the capacity of administrative data sources to create a model able to detect prevalent cases and to monitor care pathways.

Therefore, in our study, we verified the validity of case-finding algorithm, based on administrative data, using a group of well-documented MS cases as the gold standard, to determine the prevalence of MS in Tuscany.

**Methods**

**Data Sources**

We obtained data from the Tuscan health administrative databases in observance of the privacy law. The Tuscany region (population at December 31, 2011: 3,667,780 with 1,759,289 males and 1,908,491 females) [31] is divided into 12 LHA that provide health care for their residents.

Our data sources were hospital discharge records, drug-dispensing records, disease-specific payment exemptions from co-payment to health care, home and residential long-term care and inhabitant registry (table 1).

We made a record-linkage using the unique personal identification code.

**Case Ascertainment**

We created a case-finding algorithm in which we linked hospital discharge records, drug-dispensing records, disease-specific exemptions and long-term care up to the end of 2011.

The cohort was composed by all inhabitants alive at the prevalence day who met at least one of the following criteria: (a) at least one hospital discharge record (SDO) with MS diagnosis, (b) one active payment exemption (SEA) at December 31, 2011 for MS, (c) at least 2 drug prescriptions (FED + SPF) with different dates for at least one of the 5 drugs that are specific for MS (table 2), (d) MS diagnosis in home and residential long-term care data (AD/ RSA).
We analyzed data from the beginning of the recording of administrative data to the index date (December 31, 2011), except for disease-specific payment exemptions regarding only last year. Then, we calculated crude and standardized (for Tuscan population as at December 31, 2011) prevalence at the index date.

**Gold Standard Cohort**

To test sensitivity, we used a true-positive reference cohort of individuals with MS, using data extracted by the Tuscan MS register. Data were transmitted from the registry to the Regional Health Agency (Agenzia Regionale di Sanità della Toscana) in observance of privacy law, decoding the patients’ fiscal codes into the unique personal identification numbers. This cohort of patients was formed by all individuals with MS alive as at December 31, 2011, that had been inserted in the registry by specialists of 2 MS centers, the one in Livorno and the one in Siena. In total, the cohort was formed by 302 individuals with a definite MS diagnosis. Then, we linked the true-positive reference cohort with the one resulting from our algorithm to obtain a Venn diagram with the percentage of individuals of the reference cohort who were included in our cohort.

To test specificity, we used another cohort of individuals who were presumably not affected by MS. For this purpose, we created a presumably true-negative reference cohort composed by individuals who were resident in Tuscany and alive as at the prevalence day, had never undergone either cranial or spinal cord CT scan or MRI and had never received a neurological outpatient visit within the NHS. We obtained a cohort of 2,644,094 subjects presumably not affected by MS.

<table>
<thead>
<tr>
<th>Table 1. List and details of sources of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital discharge records (SDO)</td>
</tr>
<tr>
<td>Drug dispensing records (FED and SPF)</td>
</tr>
<tr>
<td>Disease-specific exemptions from copayment to health care (SEA)</td>
</tr>
<tr>
<td>Home and residential long-term care (AD/RSA)</td>
</tr>
<tr>
<td>Inhabitant registry</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. List of drugs specific for MS used for our algorithm and their ATC code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacophore</td>
</tr>
<tr>
<td>Glatiramer acetate</td>
</tr>
<tr>
<td>Interferon beta 1A</td>
</tr>
<tr>
<td>Interferon beta 1B</td>
</tr>
<tr>
<td>Fingolimod</td>
</tr>
<tr>
<td>Natalizumab</td>
</tr>
</tbody>
</table>

Results

As at December 31, 2011, the algorithm identified 6,890 prevalent cases, with 4,738 females and 2,152 males. Most identified individuals with MS were 16–64 years old (table 3).

The total crude prevalence obtained was 187.9 per 100,000, 248.3 per 100,000 in females and 122.3 per 100,000 in males (table 4).

Analyzing prevalence rates by each of the 12 LHA of residence, we observed the highest rate in LHA 1 and the lowest ones in LHA 12 and 2, also considering standardized rates (table 4). A total of 5,839 cases (85%) had at least one MS-related hospitalization, 4,115 cases (60%) had an MS-specific payment exemption, 2,844 cases (41%) had at least 2 prescriptions for an MS-specific drug, and only 28 cases were in long-term care for MS (fig. 1). These 4 groups of patients overlap as shown in figure 1.

In addition to Venn diagram of the whole region, we analyzed the diagrams of each LHA (data not shown). We observed that data-capture patterns were similar for all LHAs: in fact, the most representative patterns were SDO alone (with percentage that ranged from 38 to 18 excepting LHA 1 with a higher value of 69), SDO + drugs + SEA (from 21 to 38% excepting LHA 1 and 2 with 15 and 8% respectively) and SDO + SEA (from 29 to 15% excepting LHA 1 and 2 with 8 and 57% respectively). For LHA 1 and 2, we observed that only 20 and 14% of patients were captured by drug-dispensing records, respectively, versus a regional mean of 41%.

When sensitivity was tested, we linked the true-positive reference cohort with the one obtained through our algorithm. We calculated that 296 individuals of the true-positive reference cohort were included in our cohort,
and only 6 individuals with a certain MS diagnosis were not included. So, the sensitivity was 98%. We did not calculate the positive predictive value because we did not have a cohort representative of all subjects with MS for all LHAs of Tuscany. Analyzing the 6 false negative cases, we found that one had a drug exemption being indeed a missing datum, whereas the others could not be captured because they had neither a drug exemption nor a recent hospitalization, and they were not using specific drugs.

Regarding specificity, we first linked the true-negative reference cohort with our cohort and we obtained a specificity of 99% with only 353 patients (in our cohort of 6,890) being false positive.

### Discussion

We estimated the prevalence of subjects with MS in Tuscany by creating a cohort of patients using administrative data. Our result shows a higher value than the latest available prevalence rate of 56 cases per 100,000 reported in 1991 by Meucci et al. [11]. One of the possible reasons for this difference is the different methodologies used to capture patients. In fact, the study of 1991 was based only on referrals from general practitioners, while, in our study, we used a case-finding algorithm based on various data sources, thereby allowing the possibility of capturing a larger cohort of patients. However, addition-
al factors that occurred after 1991 have to be taken into consideration:

1. Revision of diagnostic criteria [32] and the wide use of MRI have facilitated and anticipated the diagnosis [33, 34];
2. Life duration is increased in the general population; in addition, in MS patients, therapeutic approaches and management of symptoms and comorbidities ameliorated life expectancy [35].

Our results show a prevalence higher than the latest available data (2009) in Italy (188 vs. 140 cases per 100,000) and this reported prevalence increase might not be accounted solely for the 2 previous points. Rather, it could reflect a possible actual increase of incidence, as previously hypothesized [6, 13–16] but not yet demonstrated.

Although our validity study demonstrated a high level of sensibility, we could have missed some patients, especially individuals with a severe form of MS, who did not access the healthcare system and who did not use the DMDs included in our algorithm. We decided to exclude other DMDs because they are used not only for MS but also for other many diseases and their inclusion would reduce the specificity of our algorithm. Moreover, MS-specific prescriptions had a low sensitivity, due to the incompleteness of recording of these drugs, and it represents a limitation of our study. In fact, in 2 LHAs (1 and 2), we observed a very low percentage of patients captured by drug-dispensing records and this may be due to a systematic error during recording of drugs directly dispensed by hospital pharmacies, because these drugs need no reimbursement by the NHS. Regarding the other differences observed between LHAs, such as the highest prevalence rate in LHA 1, we are actually performing a new study to observe the distribution of MS in Tuscany, and to determine if there are some clusters of disease.

As expected from the extremely specific inclusion criteria, we found a high specificity and the 353 patients false-positive could be true-positive because either they might have received a neurological visit or a CT scan or MRI in private structures, or their neurological visits or tests could be not registered. In fact, it is possible that some neurological visits were not registered with the code of neurology, or some CT scan or MRI tests were not registered in the hospital discharge record.

Conclusions

With our study, we confirmed that Tuscany is a high-risk area for MS and that the prevalence is increasing over time since 1991. Despite some limitations of our findings and of our method, we have also demonstrated that our algorithm, based on administrative data, can accurately identify patient cohorts. Our future aims are to investigate the geographical distribution of MS in Tuscany and to create an integrated dataset with administrative and clinical data to monitor not only prevalence but also disease-care pathways and proxy to outcomes [36–38].

Disclosure Statement

Conflict of interest: None.

References
