A systematic assessment of medical diagnoses preceding the first diagnosis of multiple sclerosis

Christiane Gasperi, MD[1*], Alexander Hapfelmeier, PhD[2,5*], Tanja Daltrozzo[3], MD, Antonius Schneider[5], MD, Ewan Donnachie, Msc[4], and Bernhard Hemmer, MD[1,6]

Author Affiliations:

1 Department of Neurology, Klinikum rechts der Isar, TUM School of Medicine, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany.

2 Institute of Medical Informatics, Statistics and Epidemiology, TUM School of Medicine, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany.

3 Department of Psychosomatic Medicine and Psychotherapy, Klinikum rechts der Isar, TUM School of Medicine, Technical University of Munich, Langerstr. 3, 81675 Munich, Germany.

4 Bavarian Association of Statutory Health Insurance Physicians, Elsenheimer Straße 39, 80687 Munich, Germany.

5 Institute of General Practice and Health Services Research, TUM School of Medicine, Technical University Munich, Germany

6 Munich Cluster for Systems Neurology (SyNergy), Germany.

[*]These authors contributed equally to the manuscript

Corresponding Author:

Bernhard Hemmer

Department of Neurology, Klinikum rechts der Isar, Technical University of Munich

Ismaninger Str. 22

81675 Munich

Germany

Phone +498941404601

Fax +498941407681

Email: hemmer@tum.de

Supplemental data:

Data available from Dryad (Tables 5-10, Figure 4)

The statistical analysis was performed by Christiane Gasperi and Alexander Hapfelmeier.

Word count paper: 4488

Word count abstract: 250

Character count title: 96

Study funding:

Bernhard Hemmer, Christiane Gasperi and Alexander Hapfelmeier are associated with DIFUTURE (Data Integration for Future Medicine, BMBF 01ZZ1804[A-I]). Bernhard Hemmer received funding for the study by the European Union's Horizon 2020 Research and Innovation Program (grant MultipleMS, EU RIA 733161) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy – ID 390857198).

Search Terms: [41] Multiple Sclerosis, [53] Case control studies, [109] All Health Services Research, [132] Autoimmune diseases

Disclosures

Dr. C. Gasperi received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), the Hertie Foundation and the Hans and Klementia Langmatz Stiftung.

Dr. A. Hapfelmeier received a speaker honorarium from Biogen for the Biogen Symposium on Statistical Methods in Real World Evidence 2017 and funding from the Bavarian Ministry of Economic Affairs, Regional Development and Energy, not related to this study.

Dr. T. Daltrozzo reports no disclosures.

Dr. A. Schneider reports no disclosures.

Ewan Donnachie reports no disclosures.

Dr. B. Hemmer has served on scientific advisory boards for Novartis; he has served as DMSC member for AllergyCare, Polpharma and TG therapeutics; he or his institution have received speaker honoraria from Desitin; His institution received research grants from Regeneron for MS research. He holds part of two patents; one for the detection of antibodies against KIR4.1 in a subpopulation of patients with MS and one for genetic determinants of neutralizing antibodies to interferon. All conflicts are not relevant to the topic of the study.

Abstract

Objective: To explore the occurrence of diseases and symptoms in the five years prior to diagnosis in patients with multiple sclerosis (MS) in a case-control study.

Methods: Using ambulatory claims data we systematically assessed differences in the occurrence of diseases and symptoms in the five years prior to first diagnosis in patients with MS (n=10,262) as compared to patients with two other autoimmune diseases – Crohn's disease (n=15,502) and psoriasis (n=98,432) - and individuals without these diseases (n=73,430).

Results: Forty-three ICD-10 codes were recorded more frequently for patients with MS before diagnosis as compared to controls without autoimmune disease. Many of these findings were confirmed in a comparison to the other control groups. A high proportion of these ICD-10 codes represent symptoms suggestive of demyelinating events or other neurological diagnoses. In a sensitivity analysis excluding patients with such recordings prior to first diagnosis, no association remained significant. Seven ICD-10 codes were associated with lower odds ratios of MS, four of which represented upper respiratory tract infections. Here, the relations with MS were even more pronounced in the sensitivity analysis.

Conclusions: Our analyses suggest that patients with MS are frequently not diagnosed at their first demyelinating event but often years later. Symptoms and physician encounters before MS diagnosis seem to be related to already ongoing disease rather than a prodrome. The observed association of upper respiratory tract infections with lower ORs of MS diagnosis suggests a link between protection from infection and MS that however needs to be validated and further investigated.

Introduction

In patients with multiple sclerosis (MS), early diagnosis enabling early treatment initiation is crucial to prevent damage in the CNS caused by inflammation and neurodegeneration and leads to better long-term outcomes with less disability accumulation[1, 2].

Previous studies reported that patients with MS showed an altered behavior with increased rates of physician and hospital encounters related to neurological, musculoskeletal and genitourinary as well as psychiatric symptoms up to 10 years before first diagnosis[3, 4]. Additionally, increased levels of disability pension and sick leaves have been reported up to 15 years before first diagnosis[5]. These observations prompted discussions about a possible prodromal phase of MS. Alternatively, missed first demyelinating events in patients later diagnosed with MS would also explain many of these findings.

The aims of the present study were to investigate whether higher recording rates for different diagnoses and symptoms in the years before first diagnosis can be observed in patients with MS from southern Germany, and to gain insight into whether these might represent missed demyelinating events or rather symptoms of a possible MS prodrome. Using two additional control cohorts of patients with other autoimmune diseases allowed us to further assess whether the observed associations are specific for MS or shared between different autoimmune diseases. A better characterization of the health care use and the occurrence of symptoms in patients with MS in the years before first diagnosis could facilitate earlier diagnosis and treatment.

Methods

Data

According to the official membership statistics of the German Federal Ministry of Health there were about 11 million members of the statutory health insurance in the German federal state of Bavaria in 2017. This equals a coverage of about 85% of the general population of Bavaria[6]. Anonymous ambulatory claims data, coded according to the German Modification of the International Statistical Classification of Diseases and Related Health Problems (ICD-10-GM), was made available for analysis by the Bavarian Association of Statutory Health Insurance Physicians (BASHIP) for each quarterly billing period between 2005 and 2017.

We used the data to define a cohort of cases with new-onset MS and three control cohorts with either new-onset Crohn's disease, psoriasis or none of these autoimmune diseases. Except in the latter, patients were required to have at least two secured ICD-10 diagnoses of the respective disease, that is G35 (MS), K50 (Crohn's disease) or L40 (psoriasis), in separate quarterly periods between 2010 and 2017 and not before. Patients with more than one of these three autoimmune diseases were removed from the analysis. All patients with MS should have been diagnosed using the 2010 revision of the McDonald criteria[7]. Participants in the MS cohort further had to have a recorded visit with a neurologist and no diagnosis of a clinically isolated syndrome (CIS) at any time. The ICD-10 code G04 ("Encephalitis, myelitis and encephalomyelitis") - although rater unspecific - is typically used to record CIS in Germany. Additionally, we excluded all patients with a recording of secondary progressive MS (SPMS) at first diagnosis. We randomly selected controls without record of the aforementioned autoimmune diseases from the BASHIP data without replacement and matched them to the MS cohort by year of birth, sex, and district of residence in a 5:1 ratio. For these patients we chose a synthetic quarter of first diagnosis equal to the one of their matching partner. We excluded patients with a recorded demyelinating event in the 5-year period prior to the first diagnosis from each cohort. For patients in the MS cohorts, recording of such events could indicate an earlier onset of disease. For the control cohorts, the rationale was to remove all patients with possible demyelinating diseases even if they have not been

diagnosed with MS. To implement this restriction, we used ICD-10 codes related to encephalitis, myelitis and encephalomyelitis as well as inflammatory diseases of the central nervous system (G04, G05, G09), demyelinating diseases of the central nervous system (G36, G37), retrobulbar neuritis and other disorders of the optic nerve and the visual pathways (H46, H47, H48), and abnormal findings on diagnostic imaging of central nervous system (R90). In an even more conservative sensitivity analysis, we further removed patients with any record of neurological or cerebrovascular ICD-10 codes as well as any ICD-10 codes that were associated with MS in the primary analysis and could be interpreted as symptoms of a MS relapse within the 5-years period prior to first diagnosis from each cohort. For this restriction, we used the ICD-10 codes for "Diseases of the nervous system" (G00-G99), "Cerebrovascular diseases" (I60-I69), "Visual disturbances" (H53), "Disorders of vestibular function" (H81), "Neuromuscular dysfunction of bladder, not elsewhere classified" (N31), "Disturbances of skin sensation" (R20), "Abnormalities of gait and mobility" (R26), "Unspecified urinary incontinence" (R32), "Other and unspecified symptoms and signs involving the urinary system" (R39), "Dizziness and giddiness" (R42), and "Speech disturbances, not elsewhere classified" (R47). In any case, we restricted age to 21 to 70 years due to small sample sizes in the age groups < 21 years and > 70 years.

We used the official documentation of the ICD-10-GM, to track changes of ICD-10 codes between 2005 and 2017 and updated all codes that have been subject to change to the coding valid in 2017. We removed codes that violated age and sex restrictions and considered only secured ICD-10 codes for analysis. Finally, we excluded all ICD-10 codes that are sex-specific as well as all ICD-10 codes from the chapter XXI, which covers factors influencing health status and contact with health services.

Analysis

In the main analyses, we investigated the five years prior to first diagnosis for each patient (figure 4, available from Dryad) to compare the MS cohort to the control cohort without autoimmune disease. For any ICD-10 code we used a binary predictor variable to indicate whether the specific ICD-10 code was recorded at least once (yes/no). We investigated the relation of these predictor variables to the outcome MS (yes/no) by means of unconditional logistic regression models and quantified the relation

by respective odds ratios (ORs) as effect measure[8]. Firth's bias-reduced logistic regression was used in case of complete or quasi-complete separation[9, 10].

We built separate models to estimate main effects as well as the interaction effects of sex and the age categories (21-30 yrs., 31-40 yrs, …, 61-70 yrs.) to obtain adjusted effect estimates. Additionally, for ICD-10 codes associated with MS in the main analysis, we performed single year analyses for each of the five years separately. The last quarter before first diagnosis was excluded from any analysis to conservatively rule out any effects that might already result from the onset of the disease rather than precede it.

Despite the exploratory nature of these analyses we addressed the potential multiple testing problem by use of Sidak's correction to strictly control the familywise error rate (FWER) at a 5% significance level. Comparing the MS cohort to the control cohort without autoimmune diseases the number of recorded ICD-10 codes and respective tests was 1438. In addition, we subsequently excluded ICD-10 codes with a relative frequency of less than 0.5% in both cohorts from further investigations to rule out the exploration of data artefacts.

We used all significant findings for further exploratory where we compared the MS cohort to the cohorts with Crohn's disease and psoriasis in similarity to the main analysis. Further descriptive statistics are given for the frequency of encounters with and diagnoses made by different medical specialists.

We computed all analyses with R 3.6.1 (The R Foundation for Statistical Computing, Vienna, Austria).

Data availability statement

Data protection regulations prohibit the open distribution of the underlying data. Interested researchers may contact the BASHIP or the corresponding author to request access.

Standard protocol approvals, registrations, and patient consents

In this retrospective case-control study, we analyzed anonymous claims data held by the BASHIP. Approval by an ethical standards committee on human experimentation (institutional or regional) for any experiments using human participants was not needed according to the Guidelines and Recommendations for Good Practice of Secondary Data Analysis[11]. Approval was obtained from the responsible data protection officer of the BASHIP. There was no need for written informed consent from participants. No photographs, videos, or other information of recognizable persons are used in this article. Authorization for disclosure was therefore not necessary.

Results

Descriptive statistics of the study cohorts

According to the aforementioned definitions of the cohorts a database query resulted in 10,262 patients newly diagnosed with MS and a total of 187,364 controls, that is 98,432 patients newly diagnosed with psoriasis, 15,502 patients newly diagnosed with Crohn's disease and 73,430 individuals with none of these autoimmune diseases (table 1). The cohorts with psoriasis or Crohn's disease were older and included less women than the other cohorts. We consequently took these factors into account in the statistical analysis.

Excluding patients with previous symptoms suggestive of a demyelinating event or neurological or cerebrovascular disorders resulted in samples sizes of 4,029 patients with MS, 45,632 patients with psoriasis, 7,500 patients with Crohn's disease and 35,512 patients with no autoimmune disease. The numbers of patients with recordings of the ICD-10 codes used for this exclusion are depicted in table 5 (available from Dryad).

Mainly neurological and psychiatric ICD-10 codes are recorded more frequently in patients with MS before first diagnosis

A total of 43 ICD-10 codes were significantly more frequently recorded for patients with MS than for patients without any of the considered autoimmune diseases in the five years prior to first diagnosis (table 2). Of these, 14 were neurological or cerebrovascular disorders or symptoms linked to neurological diseases. Additionally, ICD-10 codes related to visual disturbance, disc disorders, dizziness or giddiness, motor impairment, disorders of the urinary system and abnormal skin sensation were more frequently recorded in patients with MS. We further observed an association of six psychiatric ICD-10 codes with higher ORs of MS diagnosis including depressive and persistent mood disorders, dissociative disorder and nicotine dependence. Interestingly, age seemed to be related to the spectrum of ICD-10 codes recorded. ORs for *Abnormalities of gait and mobility (R26)* and *Dizziness and giddiness*

(R42) were more pronounced in older and *Disturbances of skin sensation (R20)* or *Visual disturbances (H53)* in younger people (figure 1).

To investigate when the differences between patients with MS and the control cohorts first become apparent, we performed analyses for each of the five years before first diagnoses separately. We could observe that the ORs of MS were > 1.0 for all 43 ICD-10 codes and all five years before diagnosis with an exception for a single year and ICD-10 code. Most of these relations were more pronounced for time intervals closer to first diagnosis (figure 2).

To further investigate whether these effects were MS specific we used patients newly diagnosed with psoriasis or Crohn's disease as controls, resulting in 35 and 19 of the 43 ICD-10 codes still being significantly associated with higher ORs of MS, respectively (data available from Dryad, table 6). This was true mostly for neurological ICD-10 codes. It has to be noted however that the Crohn's disease cohort is significantly smaller than the other two control cohorts (table 1).

Most of the diagnoses listed in table 2 were recorded by general practitioners (37.0% of the recordings for patients with MS, 37.5-39.0% for the other cohorts) and only a small percentage of the recordings were made by neurologists (4.8% and 2.6-3.4%). Considering only neurological ICD-10 codes (ICD-10 chapter 6), we found that these were also recoded most often by general practitioners (47.3% for patients with MS, 50.4-52.4% for the control cohorts). The respective relative frequencies of diagnoses recorded by neurologists are 28.3% and 22.7%-23.5% (data available from Dryad, table 7).

Besides neurological diagnoses, the ICD-10 codes recorded more frequently for patients with MS included urinary, vestibular symptoms and visual disturbances that could be interpreted as symptoms of a MS relapse. In the sensitivity analysis excluding patients with such or other neurological symptoms, the resulting ORs of MS were below 1.0 for 11 of the 20 ICD-10 codes that could still be analyzed. For 8 ICD-10 codes the ORs of MS were closer to 1.0 as compared to the primary analysis. None of the 20 ICD-10 codes analyzed in this analysis were significantly and positively associated with MS diagnosis after correction for multiple testing (data available from Dryad, table 8).

Infections, gastritis and duodenitis, and abdominal and pelvic pain are recorded less frequently in patients with MS before first diagnosis

We could identify seven ICD-10 codes that were less frequently recorded in patients with MS than in patients without any of the considered autoimmune diseases in the five years before first diagnosis (table 3). These included four ICD-10 codes for infections of the upper respiratory tract, *Conjunctivitis (H10), Gastritis and duodenitis (K29),* and *Abdominal and pelvic pain (R10)*.

In the sensitivity analysis all ICD-10 codes identified in the primary analysis were still significantly associated with lower ORs of MS compared to controls without any of the three autoimmune diseases (data available from Dryad, table 9). Here, the ORs of MS were lower as compared to the first analysis, suggesting an even stronger relation of these seven ICD-10 codes with lower ORs of MS.

When analyzing the five years before first diagnosis separately, the ORs were continuously below 1.0 for each year and all seven ICD-10 codes (figure 3).

Referring to patients with Crohn's disease as controls, six of the seven ICD-10 codes associated with lower ORs of MS in the primary analysis were still significantly associated with MS (data available from Dryad, table 10). Unsurprisingly, *Abdominal and pelvic pain (R10)* and *Gastritis and duodenitis (K29)* showed the strongest association with lower ORs of MS. In comparison to patients with psoriasis two ICD-10 codes were still associated with lower ORs of MS after correction for multiple testing - *Acute tonsillitis (J03),* and *Conjunctivitis (R10).*

Patients with MS have more medical encounters with different medical specialists before first diagnosis

Before first diagnosis patients with MS had more frequent encounters with medical doctors of different specialties compared to matched controls without any of the three autoimmune diseases (table 4). This was most pronounced for neurologists and neurosurgeons. Other medical specialists that were more frequently seen by patients with MS included urologists, specialists for physical and rehabilitative

medicine, radiologists, nuclear medicine doctors, ophthalmologists, orthopedists, surgeons, psychiatrists and psychotherapists, as well as ENT (ear-nose-throat) specialists (table 4).

In the sensitivity analysis, we observed lower rates of doctor's encounters of almost all specialties in patients with MS (table 4).

Discussion

Previous studies have reported more frequent doctor encounters for a number of different medical specialists[12] and for different conditions including disorders and symptoms of the nervous, sensory, musculoskeletal and genito-urinary systems as well as for psychiatric disorders and various types of pain[3, 4, 13]in patients with MS up to 10 years before first diagnosis or a first recorded demyelinating event. These results prompted a discussion about a possible prodromal phase of MS. The concept of a pre-clinical or sub-clinical phase of MS is widely recognized[2, 14] and a recent study reported increased levels of neurofilament light chains in the serum of patients with MS years before clinical onset[15]. Whether a prodromal phase of MS with non-specific, clinical symptoms occurring before and not typical for a distinct demyelinating event exists, however, is not clear. Another possible explanation for many of the findings from the aforementioned studies would be that some patients with MS might experience demyelinating events even years before first diagnosis that are not recognized as such and therefore not recorded accordingly.

In the present study, we used ambulatory claims data of 197,626 individuals from Bavaria to investigate differences in the recording rates of diseases and symptoms in the five years prior to first diagnosis between patients with MS and three different control cohorts. We identified a total of 43 ICD-10 codes that were more frequently recorded in patients with MS in the years before first diagnosis. However, many of these ICD-10 codes either represent neurological or cerebrovascular disorders or symptoms that could be caused by a demyelinating event. We therefore performed a sensitivity analysis for which we excluded all patients with neurological diagnosis or recorded symptoms suggestive of a demyelinating event in the five years before first diagnosis. We could not identify any ICD-10 code still significantly and positively associated with MS in this sensitivity analysis, suggesting that many of the symptoms recorded more frequently in patients with MS in the years before first diagnosis could represent demyelinating events that have not been recognized as such. The fact that most of these ICD-10 codes were recorded by general practitioners further supports this hypothesis.

Due to the design of our sensitivity analysis we cannot exclude that neurological diagnosis codes recorded more frequently in patients with MS before first diagnosis actually represent the occurrence of prodromal symptoms not typical of and likely not caused by demyelination as these ICD-10 codes were removed from the sensitivity analysis. However, we did not identify any other non-neurological ICD-10 codes to be positively associated with MS diagnosis in this sensitivity analysis.

Wijnands et al. reported more physician encounters across 16 of 17 ICD-10 chapters in the five years before the first recorded demyelinating event in patients with MS as compared to matched controls[3]. In a second, significantly smaller cohort with available clinical data, the clinical index date was set to the time point of the earliest known symptoms suggestive of MS as recalled by the patients. In this cohort, MS cases had significantly more physician encounters or hospital admissions only for 7 ICD-10 chapters and the observed effects were considerably smaller. The most pronounced results were observed for diseases of the nervous system, diseases of the sense organs, and diseases of the musculoskeletal system and connective tissues. These were the ICD-10 chapters that also showed the strongest association to MS diagnosis in our data and an evaluation of the single ICD-10 codes driving these associations in our analysis suggest that many of these diagnoses or symptoms actually are suggestive of demyelinating events. Additionally, in the clinical cohort studied by Wijnands et al., MS cases had more hospital admissions for diseases of the skin and subcutaneous tissues. These encounters are unlikely to be related to demyelinating events, however, we did not observe any association of MS diagnosis with recording rates for ICD-10 codes from this ICD-10 chapter and to our knowledge no other study investigating a possible prodromal phase of MS reported a similar association.

Another previous study specifically studied the occurrence of fatigue in patients with MS before first diagnosis and reported that approximately 30% of patients with MS were labeled with “chronic fatigue” or “malaise or fatigue” up to three years prior to the diagnosis[16]. We did not observe higher recording rates for the ICD-10 code R53 (“Malaise and Fatigue”) in this study.

In our study, the mean age at first diagnosis was 40 years in the primary analyses and 37.6 in the sensitivity analyses. This is considerably lower than the reported mean age at first recorded diagnosis of 43 years in the study by Wijnands et al.[3] and 47 years in the study by Disanto et al.[4] In the smaller clinical cohort studied by Wijnands et al., the mean age at MS symptom onset 36.5 years. This suggests, that first demyelinating events might not have been recorded also in these studies.

Psychiatric disorders have been identified as frequent comorbidities in patients with MS in a number of previous studies[17-21]. In the present study we also recorded more psychiatric ICD-10 codes in patients who were later diagnosed with MS in our initial analysis. However, in our sensitivity analysis we did not observe significantly higher recording rates for any psychiatric ICD-10 codes. These results suggest that psychiatric disorders are not preceding the onset of MS but are rather related to an already ongoing and symptomatic but not yet diagnosed disease.

Interestingly, in the present study we observed a number of ICD-10 codes that were associated with lower ORs of MS. We observed lower rates of upper respiratory tract infections in the five years before first diagnosis of MS as well as lower recording rates for conjunctivitis, gastritis and duodenitis, and for abdominal pain. These associations were even more pronounced in the sensitivity analyses. The association of upper respiratory tract infections with lower ORs of MS diagnosis is unexpected as several previous studies found associations of infections (e.g. influenza) with the development of relapses in patients with MS[22, 23]. In our study, when comparing patients with MS with the control groups of patients with Crohn's disease or psoriasis a few of these infections were still associated with MS, but the associations where not as pronounced. These results indicate that the association of infections with lower ORs of MS could be an effect at least to some degree shared by different autoimmune diseases and therefore not specific for MS. While our results suggest a possible link between a protection from respiratory infections and the occurrence of MS, this needs to be further investigated.

We further analyzed the recorded visits with different medical specialists in the years before first diagnosis and observed higher encounter rates for patients with MS as compared to controls for a number of medical specialties. These results were most pronounced for neurologists, neurosurgeons, urologists, specialists for physical and rehabilitative medicine, and radiologists. This is in accordance with the hypothesis that many patients with MS show symptoms even before diagnosis and probably see doctors of different specialties receiving a number of different diagnoses. For the cohorts from our sensitivity analysis, we still observed slightly higher rates of doctor encounters for neurologists and neurosurgeons, while for most of the other medical specialties the encounter rates for patients with MS were similar or lower than the encounter rates for controls. This is in accordance with the hypothesis that most symptoms that patients with MS develop before diagnosis are in fact caused by undiagnosed demyelinating events.

The study is subject to several limitations. The underlying diagnoses are not audited and reflect to some extend the coding and clinical practices of the physicians. Furthermore, they do not cover symptoms that haven't been diagnosed for patients who did not consult a physician despite experiencing symptoms. Hospital claims data is also not covered. Such coding habits and possibly missing information may result in both decreased and increased frequencies of diagnoses in each of the cohorts studied, possibly to varying degrees. It is therefore difficult to assess a possible bias in the comparison of cohorts. A similar limitation concerns the definition of cases. The utilized claims data are not subject to systematic audit and their validity is therefore subject to the clinical judgement of the physicians consulted. However, distinctions between suspected, secured and excluded diagnoses are made in the data. Our cohort should not therefore contain patients who were merely under evaluation for a possible MS diagnosis. Furthermore, the requirement to have two diagnoses in separate quarterly periods ensures that only patients with an established medical history are included. This is a standard procedure used, for example, in the German health insurance risk adjustment scheme (morbiRSA). However, the approach differs from approaches used for inpatient and outpatient health data from the United States and Canada, where it has been recommended that three or more

MS-related claims within a year are required for a case definition[24]. We investigated the robustness of the results through sensitivity analyses in the present study, excluding patients with signs of a demyelinating event before disease onset. Further and even stricter definitions of cohorts could have been chosen. In this regard patients with preceding signs of the disease defining the respective control cohort could have been excluded from the sensitivity analysis. We determined that the processing of data in this way might have involved a certain degree of arbitrariness and that a fair and unbiased comparison of the cohorts might also have suffered from treating them differently. We therefore decided to apply the same exclusion rules to each cohort. The database provided by the BASHIP covers about 85% of the general population, resulting in a high degree of generalizability. The 15% not covered by the study data are persons with private health insurance, that is mostly civil servants, the self-employed and those earning above a set income threshold. The impact of these factors on the presented results is considered to be minor. Further limitations are the potential of confounding which might be induced by the non-experimental study design and by using ambulatory claims data only, the missing clinical confirmation of demyelinating events, the coding of neurological diseases by general practitioners.

Conclusion

The results of this study suggest that many patients with MS are not diagnosed at the time point of the first demyelinating event but often several years later. These patients appear to have more medical encounters with doctors from different specialties and to be diagnosed with a number of different diseases or symptoms – including many neurological ICD-10 codes – that might represent not recognized first demyelinating events. Previous studies reported higher occurrences of diagnoses and symptoms in the years before a first recorded demyelinating event suggesting non-specific prodromal symptoms. In this study, we could evaluate the single ICD-10 codes associated with MS diagnosis and after removal of patients who had recorded neurological diagnoses or symptoms suggestive of demyelinating events in the years before first MS diagnosis we could not observe any ICD-10 codes more frequently recorded for MS patients. While these results argue against prodromal symptoms not typical of demyelinating events in patients with MS, this study is not designed to directly investigate possible prodromal symptoms of MS and more studies with larger numbers of MS patients with detailed clinical data are needed. Many of the ICD-10 codes recorded more frequently for patients with MS and possibly related to not recognized demyelinating events were not recorded by neurologists, but by general practitioners of other medical specialists; suggesting that timely consultation of neurologists might facilitate earlier diagnosis at least for some of the patients. Interestingly, we could observe a number of ICD-10 codes that were less frequently recorded for patients with MS as compared to the different control groups including four ICD-10 codes of infections of the upper respiratory tract. While these results are surprising considering that MS relapses have been reported to be associated with acute infections, we believe that they could suggest a possible link between the occurrence of autoimmunity in MS and a protection from infections. However, this observed association needs to be further studied and validated. Future research will address the investigation of longer time frames to be able to more precisely specify when the differences in the use of the health system first occur. The analysis of cohorts with available detailed clinical data could help to pinpoint the symptoms that are likely not to be recognized as caused by demyelinating events. Furthermore, it

will be interesting to perform similar analyses on patients diagnosed with the 2017 revision of the McDonald criteria[25] of MS as these criteria allow for an earlier diagnosis.

Appendix 1

Name	Location	Contribution
Christiane Gasperi, MD	Technical University Munich, Munich, Germany	Designed and conceptualized the study; performed statistical analyses; drafted the manuscript.
Alexander Hapfelmeier, PhD	Technical University Munich, Munich, Germany	Designed and conceptualized the study; performed statistical analyses; drafted the manuscript.
Tanja Daltrozzo, MD	Technical University Munich, Munich, Germany	Interpreted the data; revised the manuscript for intellectual content.
Antonius Schneider, MD	Technical University Munich, Munich, Germany	Interpreted the data; revised the manuscript for intellectual content.
Ewan Donnachie, Msc	Bavarian Association of Statutory Health Insurance Physicians, Munich, Germany	Acquired the data; revised the manuscript for intellectual content.
Bernhard Hemmer, MD	Technical University Munich, Munich, Germany	Designed and conceptualized the study; supervised the research; revised the manuscript for intellectual content

References

1. Brown JWL, Coles A, Horakova D, et al. Association of Initial Disease-Modifying Therapy With Later Conversion to Secondary Progressive Multiple Sclerosis. JAMA 2019;321:175-187.

2. Compston A, Coles A. Multiple sclerosis. Lancet 2008;372:1502-1517.

3. Wijnands JM, Zhu F, Kingwell E, et al. Five years before multiple sclerosis onset: Phenotyping the prodrome. Mult Scler 2019;25:1092-1101.

4. Disanto G, Zecca C, MacLachlan S, et al. Prodromal symptoms of multiple sclerosis in primary care. Ann Neurol 2018;83:1162-1173.

5. Landfeldt E, Castelo-Branco A, Svedbom A, Lofroth E, Kavaliunas A, Hillert J. Sick leave and disability pension before and after diagnosis of multiple sclerosis. Mult Scler 2016;22:1859-1866.

6. Daltrozzo T, Hapfelmeier A, Donnachie E, Schneider A, Hemmer B. A Systematic Assessment of Prevalence, Incidence and Regional Distribution of Multiple Sclerosis in Bavaria From 2006 to 2015. Front Neurol 2018;9:871.

7. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011;69:292-302.

8. Kuo CL, Duan Y, Grady J. Unconditional or Conditional Logistic Regression Model for Age-Matched Case-Control Data? Front Public Health 2018;6:57.

9. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80:27-38.

10. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med 2002;21:2409-2419.

11. Swart E, Gothe H, Geyer S, et al. Gute Praxis Sekundärdatenanalyse (GPS): Leitlinien und Empfehlungen. Gesundheitswesen 2015;77:120-126.

12. Marrie RA, Yu N, Wei Y, Elliott L, Blanchard J. High rates of physician services utilization at least five years before multiple sclerosis diagnosis. Mult Scler 2013;19:1113-1119.

13. Yusuf FLA, Ng BC, Wijnands JMA, Kingwell E, Marrie RA, Tremlett H. A systematic review of morbidities suggestive of the multiple sclerosis prodrome. Expert Rev Neurother 2020;20:799-819.

14. De Stefano N, Giorgio A, Tintore M, et al. Radiologically isolated syndrome or subclinical multiple sclerosis: MAGNIMS consensus recommendations. Mult Scler 2018;24:214-221.

15. Bjornevik K, Munger KL, Cortese M, et al. Serum Neurofilament Light Chain Levels in Patients With Presymptomatic Multiple Sclerosis. JAMA Neurol 2020;77:58-64.

16. Berger JR, Pocoski J, Preblick R, Boklage S. Fatigue heralding multiple sclerosis. Mult Scler 2013;19:1526-1532.

17. Marrie RA. Comorbidity in Multiple Sclerosis: Some Answers, More Questions. Int J MS Care 2016;18:271-272.

18. Marrie RA, Cohen J, Stuve O, et al. A systematic review of the incidence and prevalence of comorbidity in multiple sclerosis: overview. Mult Scler 2015;21:263-281.

19. Panda SP, Das RC, Srivastava K, Ratnam A, Sharma N. Psychiatric comorbidity in multiple sclerosis. Neurol Neurochir Pol 2018;52:704-709.

20. Patten SB, Svenson LW, Metz LM. Psychotic disorders in MS: population-based evidence of an association. Neurology 2005;65:1123-1125.

21. Johansson V, Lundholm C, Hillert J, et al. Multiple sclerosis and psychiatric disorders: comorbidity and sibling risk in a nationwide Swedish cohort. Mult Scler 2014;20:1881-1891.

22. De Keyser J, Zwanikken C, Boon M. Effects of influenza vaccination and influenza illness on exacerbations in multiple sclerosis. J Neurol Sci 1998;159:51-53.

23. Oikonen M, Laaksonen M, Aalto V, et al. Temporal relationship between environmental influenza A and Epstein-Barr viral infections and high multiple sclerosis relapse occurrence. Mult Scler 2011;17:672-680.

24. Culpepper WJ, Marrie RA, Langer-Gould A, et al. Validation of an algorithm for identifying MS cases in administrative health claims datasets. Neurology 2019;92:e1016-e1028.

25. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018;17:162-173.

Figure legends

Figure 1: Age dependent associations of ICD-10 codes with MS diagnosis

Some ICD-10 codes showed age dependent associations with MS in the analysis using individuals without any of the three autoimmune diseases. The associations of *Disturbances of skin sensation (R20)* (A) and *Visual disturbances (H53)* (B) were more pronounced in younger individuals whereas the associations of *Abnormalities of gait and mobility (R26)* (C) and *Dizziness and giddiness (R42)* (D) increased with increasing age.

Figure 2: Single-year analysis on ICD-10 codes associated with higher odds ratios of MS

Odds ratios (ORs) of MS are above 1.0 for the ICD-10 codes associated with higher ORs of MS suggestive of an MS relapse for each of the five years before first diagnosis (A). We observed a trend with higher ORs in the years closer to first diagnosis. In a joint analysis on all neurological and neurovascular ICD-10 codes associated with MS for each five years before first diagnosis separately (B) ORs for MS are also increasing for years closer to first diagnosis. *not elsewhere classified.

Figure 3: Single-year analysis on ICD-10 codes associated with lower odds ratios of MS

Odds ratios (ORs) of MS are below 1.0 for each of the five years before first diagnosis for all six ICD-10 codes associated with lower ORs of MS in the primary analyses (A) as well as in the analyses on patients without previous recordings of neurological diseases of symptoms suggestive of a relapse (B). *of multiple and unspecified sites.

Table 1: Descriptive statistics of the cohorts

	Cohort	ICD-10	Size (n)	Distinct ICD-10 codes within 5 years before first diagnosis, median (IQR)	Age at first diagnosis, years (mean ± SD)	Women (%)
Primary analysis	MS	G35	10,262	26 (13-40)	40.0 ± 11.8	7,119 (69.4)
	Psoriasis	L40	98,432	27 (14-41)	47.5 ± 13.3	50,643 (51.4)
	Crohn's disease	K50	15,502	28 (15-42)	40.8 ± 13.6	8,693 (56.1)
	no autoimmune	None of G35, L40, K50	73,430	26 (16-39)	40.4 ± 11.9	50,793 (69.2)
Sensitivity analysis	MS	G35	4,029	13 (1-24)	37.6 ± 10.4	2,665 (66.1)
	Psoriasis	L40	45,632	16 (6-26)	44.2 ± 13.1	20,673 (45.3)
	Crohn's disease	K50	7,500	16 (5-27)	38.1 ± 12.6	3,732 (49.8)
	no autoimmune	None of G35, L40, K50	35,512	19 (11-28)	38.5 ± 11.2	23,358 (65.8)

Abbreviations: IQR – interquartile range, SD – standard deviation

Table 2: ICD-10 codes associated with higher odds ratios of MS diagnosis

ICD - Description	N MS	N Controls	OR	adjusted p
A69 - Other spirochetal infections	190	988	1.40 (1.20-1.64)	3.44×10^{-02}
E53 - Deficiency of other B group vitamins	88	371	1.72 (1.36-2.18)	7.09×10^{-03}
F06 - Other mental disorders due to known physiological condition	110	489	1.66 (1.34-2.04)	2.99×10^{-03}
F17 - Nicotine dependence	1184	7240	1.20 (1.13-1.28)	4.97×10^{-05}
F32 - Major depressive disorder, single episode	2619	17204	1.14 (1.08-1.19)	4.12×10^{-04}
F33 - Major depressive disorder, recurrent	842	4837	1.29 (1.19-1.39)	1.90×10^{-07}
F34 - Persistent mood [affective] disorders	522	2824	1.35 (1.23-1.49)	8.52×10^{-07}
F44 - Dissociative and conversion disorders	105	404	1.87 (1.51-2.33)	1.65×10^{-05}
G25 - Other extrapyramidal and movement disorders	219	963	1.69 (1.45-1.96)	8.69×10^{-09}
G45 - Transient cerebral ischemic attacks and related syndromes	102	378	2.02 (1.62-2.53)	5.34×10^{-07}
G50 - Disorders of trigeminal nerve	164	633	1.89 (1.59-2.24)	9.62×10^{-10}
G54 - Nerve root and plexus disorders	388	1538	1.88 (1.68-2.10)	2.99×10^{-24}
G55 - Nerve root and plexus compressions in diseases classified elsewhere	295	1409	1.56 (1.37-1.77)	1.92×10^{-08}
G56 - Mononeuropathies of upper limb	898	4592	1.47 (1.37-1.59)	1.21×10^{-20}
G57 - Mononeuropathies of lower limb	268	1017	1.96 (1.71-2.24)	9.07×10^{-19}
G62 - Other and unspecified polyneuropathies	336	982	2.66 (2.34-3.02)	9.58×10^{-48}
G81 - Hemiplegia and hemiparesis	165	306	4.10 (3.38-4.96)	5.59×10^{-44}
G83 - Other paralytic syndromes	191	311	4.58 (3.82-5.49)	3.04×10^{-57}
G93 - Other disorders of brain	303	1004	2.21 (1.94-2.52)	1.13×10^{-29}
H43 - Disorders of vitreous body	360	1863	1.44 (1.28-1.62)	8.63×10^{-07}
H53 - Visual disturbances	1573	8783	1.35 (1.27-1.43)	1.74×10^{-20}
H81 - Disorders of vestibular function	667	3209	1.54 (1.42-1.68)	9.40×10^{-20}

I63 - Cerebral infarction	123	320	2.97 (2.40-3.67)	7.85×10^{-21}
I64 - Stroke, not specified as hemorrhage or infarction	122	327	2.90 (2.35-3.59)	7.94×10^{-20}
I67 - Other cerebrovascular diseases	272	1025	2.09 (1.82-2.41)	2.88×10^{-22}
I70 - Atherosclerosis	226	1280	1.37 (1.18-1.59)	3.59×10^{-02}
M42 - Spinal osteochondrosis	1088	6482	1.27 (1.18-1.36)	2.63×10^{-08}
M47 - Spondylosis	1906	12317	1.17 (1.11-1.24)	2.29×10^{-05}
M48 - Other spondylopathies	396	1818	1.67 (1.49-1.87)	8.85×10^{-16}
M50 - Cervical disc disorders	586	3002	1.46 (1.33-1.60)	1.04×10^{-12}
M51 - Thoracic, thoracolumbar, and lumbosacral intervertebral disc disorders	1822	9810	1.46 (1.38-1.55)	3.65×10^{-36}
M79 - Other and unspecified soft tissue disorders, not elsewhere classified	2179	13997	1.15 (1.10-1.22)	4.79×10^{-05}
M81 - Osteoporosis without current pathological fracture	246	1356	1.38 (1.19-1.59)	1.58×10^{-02}
N31 - Neuromuscular dysfunction of bladder, not elsewhere classified	159	501	2.36 (1.97-2.83)	1.40×10^{-17}
N32 - Other disorders of bladder	241	1270	1.39 (1.21-1.60)	5.35×10^{-03}
R20 - Disturbances of skin sensation	1302	3321	3.10 (2.90-3.32)	1.60×10^{-230}
R25 - Abnormal involuntary movements	318	1645	1.41 (1.25-1.59)	5.85×10^{-05}
R26 - Abnormalities of gait and mobility	368	548	5.20 (4.54-5.95)	3.35×10^{-123}
R32 - Unspecified urinary incontinence	249	1265	1.46 (1.27-1.68)	1.58×10^{-04}
R39 - Other and unspecified symptoms and signs involving the genitourinary system	284	1348	1.57 (1.38-1.79)	1.40×10^{-08}
R42 - Dizziness and giddiness	1708	9011	1.44 (1.36-1.52)	6.78×10^{-33}
R47 - Speech disturbances, not elsewhere classified	98	323	2.23 (1.78-2.80)	5.97×10^{-09}
R52 - Pain, unspecified	1249	7901	1.16 (1.09-1.24)	5.25×10^{-03}

Associations of ICD-10 codes with higher odds ratios of MS which reach statistical significance in the comparison to controls without autoimmune disease and with a relative frequency of more than 0.5%. Abbreviations: N - count of patients with at least one diagnosis, OR - Odds Ratio, adjusted p - p-value adjusted for multiple testing.

Table 3: ICD-10 codes associated with lower odds ratios of MS diagnosis

ICD - Description	N MS	N Controls	OR	adjusted p
H10 - Conjunctivitis	1484	11854	0.88 (0.83-0.93)	1.55×10^{-02}
J02 - Acute pharyngitis	1910	14968	0.88 (0.84-0.93)	5.99×10^{-03}
J03 - Acute tonsillitis	1514	12755	0.80 (0.76-0.85)	4.56×10^{-10}
J06 - Acute upper respiratory infections of multiple and unspecified sites	4507	35085	0.84 (0.81-0.88)	2.46×10^{-12}
J20 - Acute bronchitis	2374	18592	0.88 (0.84-0.93)	1.30×10^{-03}
K29 - Gastritis and duodenitis	1804	14223	0.89 (0.84-0.94)	3.29×10^{-02}
R10 - Abdominal and pelvic pain	3691	28923	0.85 (0.82-0.89)	2.57×10^{-09}

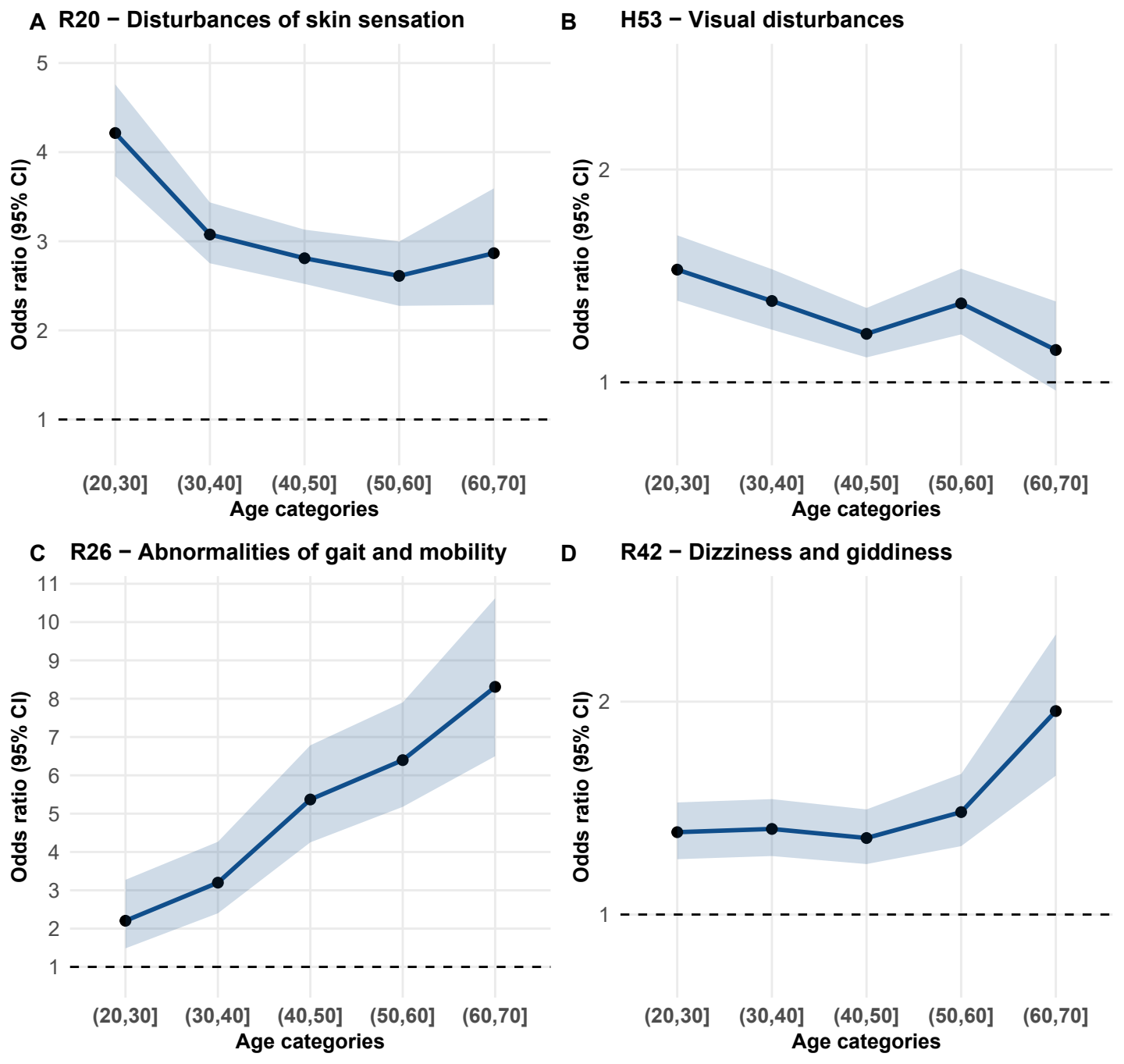
Associations of ICD-10 codes with lower odds ratios of MS which reach statistical significance in the comparison to controls without autoimmune disease and with a relative frequency of more than 0.5%. Abbreviations: N - count of patients with at least one diagnosis, OR - Odds Ratio, adjusted p - p-value adjusted for multiple testing.

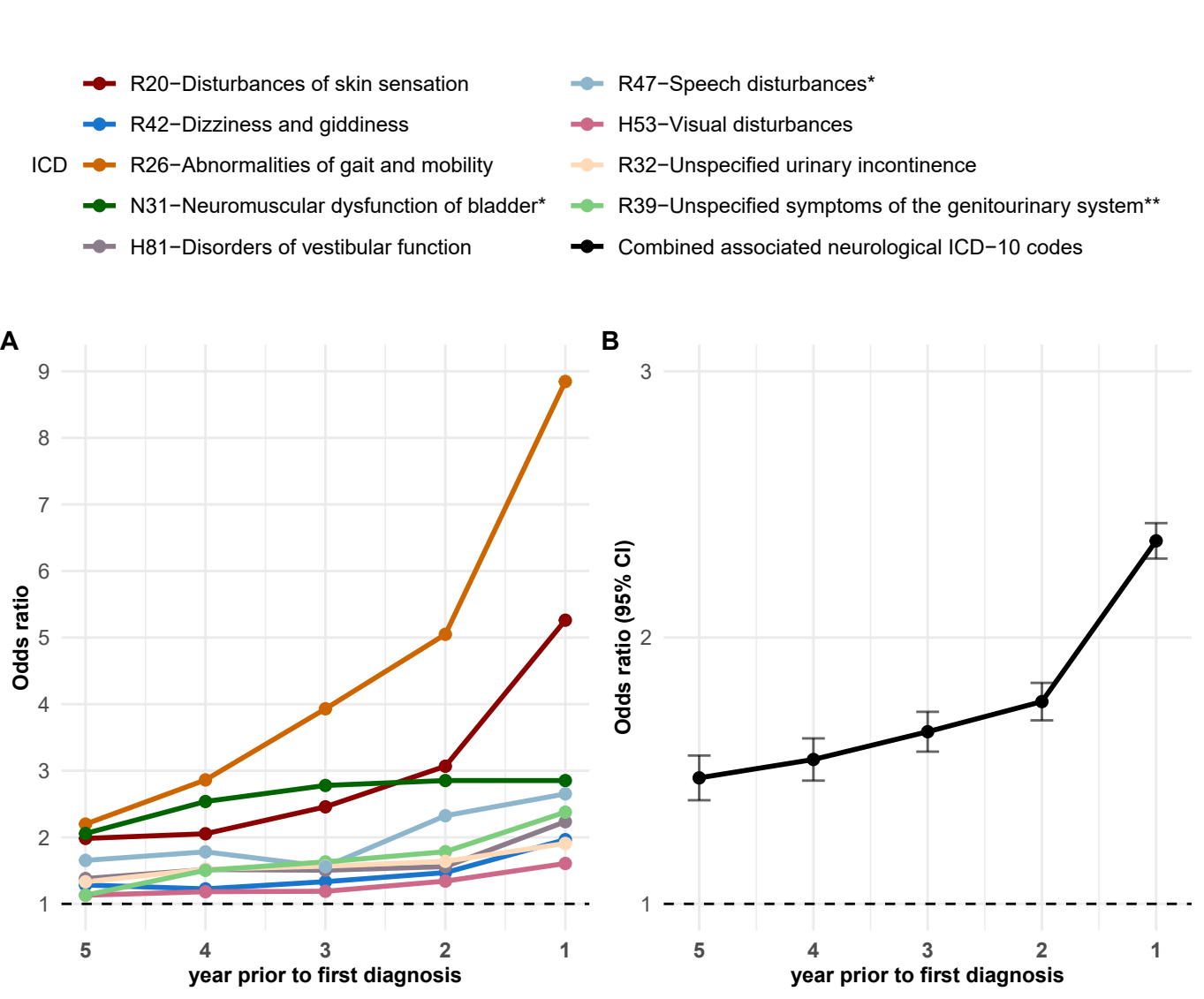
Table 4: Medical encounters with different medical specialists in the five years before first diagnosis.

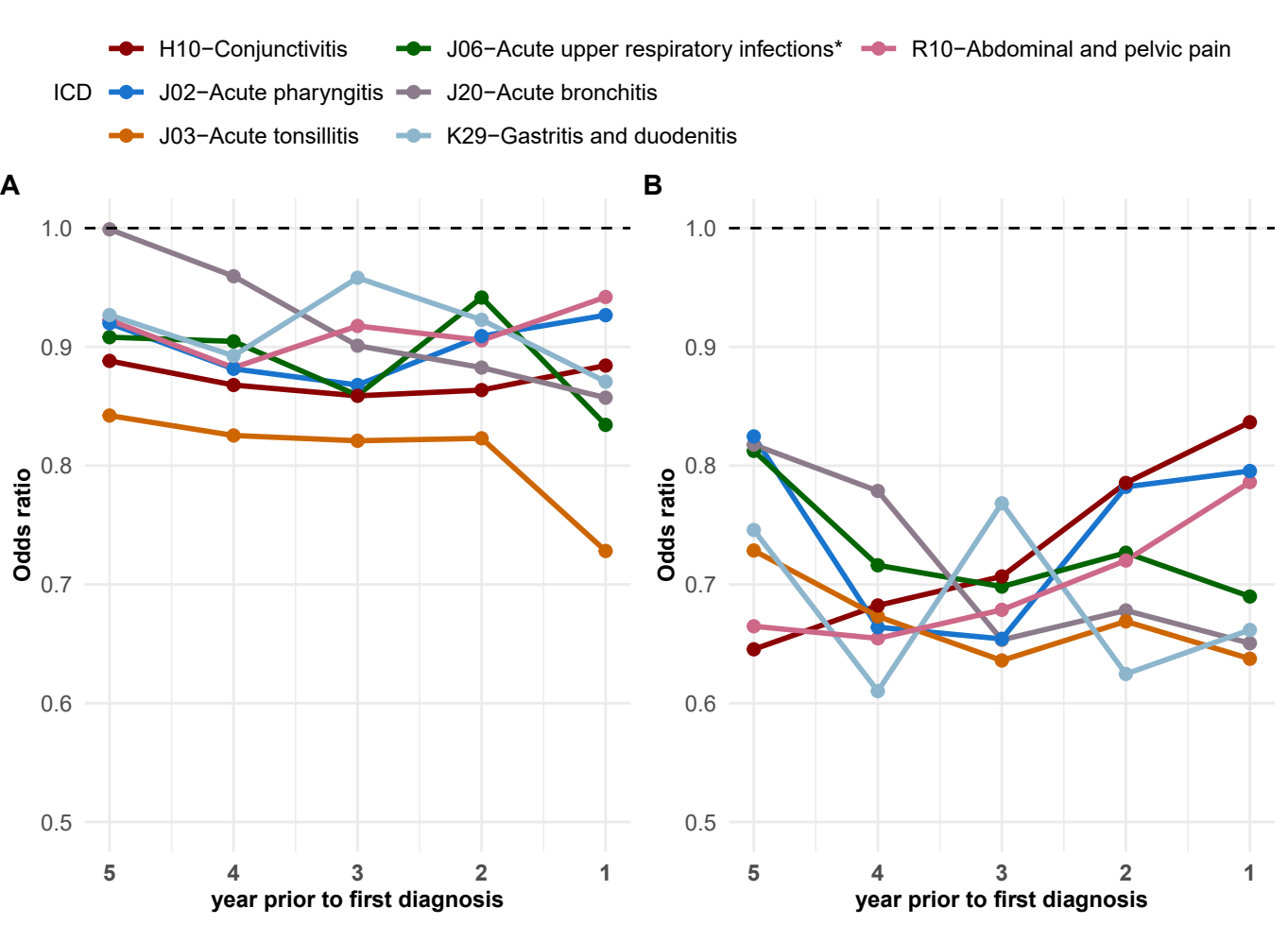
Medical specialty	Number of visits per person during the investigated time frame (mean, standard deviation)			
	Primary analysis cohorts		Sensitivity analysis cohorts	
	MS	Controls	MS	Controls
All	32.046, 25.899	29.368, 20.178	19.434, 15.187	21.194, 14.784
Anesthesiology	0.233, 0.866	0.216, 0.74	0.123, 0.435	0.141, 0.506
Dermatology	1.414, 2.452	1.358, 2.388	1.074, 2.107	1.148, 2.181
Emergency medicine	0.032, 0.238	0.028, 0.207	0.016, 0.133	0.016, 0.142
ENT (ear-nose-throat)	1.227, 2.305	1.13, 2.186	0.747, 1.68	0.781, 1.764
Family medicine	12.554, 9.214	12.314, 7.076	8.738, 6.29	9.702, 6.202
Internal medicine	1.545, 3.313	1.443, 3.105	0.74, 2.049	0.919, 2.362
Laboratory medicine	4.965, 6.193	4.679, 6.088	3.044, 4.407	3.427, 4.813
Maxillofacial surgery	0.015, 0.182	0.015, 0.166	0.013, 0.176	0.012, 0.15
Neurology	1.533, 3.446	0.81, 2.616	0.314, 1.568	0.232, 1.472
Neurosurgery	0.121, 0.672	0.076, 0.549	0.017, 0.179	0.015, 0.206
Nuclear medicine	0.239, 0.77	0.205, 0.733	0.122, 0.534	0.135, 0.623
Opthalmology	1.293, 2.355	1.123, 2.055	0.675, 1.326	0.697, 1.506
Orthopedics	2.049, 3.233	1.785, 2.909	1.077, 2.042	1.125, 2.079
Other	0.046, 0.286	0.044, 0.316	0.026, 0.211	0.033, 0.254
Pathology	0.92, 1.593	0.913, 1.592	0.709, 1.473	0.753, 1.461
Pediatrics	0.081, 0.642	0.084, 0.646	0.078, 0.595	0.074, 0.56
Physical and rehabilitative medicine	0.171, 0.896	0.135, 0.791	0.087, 0.589	0.072, 0.491
Psychiatry and psychotherapy	0.795, 2.843	0.706, 2.758	0.398, 1.945	0.419, 2.102
Radiology	1.433, 2.207	1.134, 1.869	0.673, 1.341	0.689, 1.367
Surgery	0.907, 1.893	0.804, 1.677	0.548, 1.278	0.573, 1.343
Urology	0.473, 1.694	0.364, 1.374	0.214, 0.953	0.231, 0.989

The number of visits with different medical specialists per person in the five years before first diagnosis have been calculated for the MS cohort and the cohort of patients without any of the considered autoimmune

diseases. This was done for the cohorts from the primary analysis as well as for the cohorts from the sensitivity analyses.


A R20 − Disturbances of skin sensation
B H53 − Visual disturbances
C R26 − Abnormalities of gait and mobility
D R42 − Dizziness and giddiness
Odds ratio (95% CI)
Age categories
(20,30]
(30,40]
(40,50]
(50,60]
(60,70]


ICD
R20−Disturbances of skin sensation
R42−Dizziness and giddiness
R26−Abnormalities of gait and mobility
N31−Neuromuscular dysfunction of bladder*
H81−Disorders of vestibular function
R47−Speech disturbances*
H53−Visual disturbances
R32−Unspecified urinary incontinence
R39−Unspecified symptoms of the genitourinary system**
Combined associated neurological ICD−10 codes
A
Odds ratio
year prior to first diagnosis
B
Odds ratio (95% CI)
year prior to first diagnosis


ICD
H10−Conjunctivitis
J02−Acute pharyngitis
J03−Acute tonsillitis
J06−Acute upper respiratory infections*
J20−Acute bronchitis
K29−Gastritis and duodenitis
R10−Abdominal and pelvic pain
A
B
Odds ratio
1.0
0.9
0.8
0.7
0.6
0.5
5
4
3
2
1
year prior to first diagnosis