Coding of Multiple Sclerosis Subtypes in German Administrative Claims Data: A Validation Study Using the MSDS-AOK PLUS Linked Database

Zhuleku E[1], Ziemssen T[2], Dillenseger A[2], Maywald U[3], Wilke T[4], Ghiani M[4]

1. Cytel, Berlin, Germany, 2. Zentrum für Klinische Neurowissenschaften, Dresden, Germany,

3. AOK PLUS, Dresden, Germany, 4. Institut für Pharmakoökonomie und Arzneimittellogistik e.V. (IPAM), Wismar, Germany

RWD58

Overview

Research Objective

To investigate multiple sclerosis (MS) subtype coding practices in administrative claims data and assess the predictive performance of claims-based definitions for MS subtypes.

Method

Using the MSDS-AOK PLUS linked database, we validated MS subtype coding in claims data against true recordings in patient medical records.

Highlights

- Using a definition requiring at least 1 inpatient and/or 2 outpatient MS subtype diagnoses in a one-year period surrounding an MS subtype recording in patient medical charts, approx. 13% of subtype recordings could not be assigned any MS diagnosis in claims data and 40% were only assigned an unspecified MS subtype.
- MS subtype recording in claims data as validated against MS patient medical records presented with low sensitivity/recall across all subtypes, highlighting the need for linked real-world databases that combine detailed clinical information and administrative claim data for real-world studies in MS.

Background

Multiple sclerosis (MS) is a chronic autoimmune disease that targets the central nervous system[1]. MS may present in various forms based on the disease course, commonly subclassified as clinically isolated syndrome (CIS), relapsing-remitting (RR) MS, primary progressive (PP) MS, or secondary progressive (SP) MS.

RRMS is the most frequent course, represented by 70-80% of patients at initial onset, whereas PPMS manifests in 15-20% of patients with gradual deterioration from onset of MS without relapse[1]. In contrast, SPMS is characterized by gradual deterioration after the initial RR course.

Overall, MS subtypes are clinically relevant to inform treatment and health economic decisions. As such, defining the MS subtype is important for real-world evidence studies. While various definitions have been used to identify MS subtypes in administrative claims data, coding practices have not been assessed.

Methods

We used the MSDS-AOK PLUS database, which links claims data from the German AOK PLUS sickness fund and patient charts from the Multiple Sclerosis Management System 3D (MSDS^{3D})[2-4] to assess congruency of claims-based MS subtype coding (ICD-10-GM G35.-) and MSDS^{3D} patient records.

We selected a cohort of patients with at least one recorded MS subtype and date of recording (index) in MSDS^{3D} between 01/10/2015 and 30/06/2019. Claims data six months before and after index was used to define the closest subtype diagnosis using a definition requiring 1 inpatient and/or 2 outpatient diagnoses in different quarters from any specialty or GP. Patients and/or observations (recordings) without continuous insurance within this period from each index were excluded from the sample.

The predictive performance of claims-based diagnoses was assessed against MSDS^{3D} through a multi-class confusion matrix of the first MS subtype observation per patient and all observations across all patients. Subsequently, predictive performance measures were assessed including sensitivity, specificity, the positive and negative prediction values (PPV, NPV), and F1-score for each subtype.

Results

PATIENT CHARACTERISTICS

- Overall, 179 continuous insured patients were included (75.42% female, mean 48.13 years at index, **Table 1**).
- From 179 patients, there were 1,163 total MS subtype recordings in MSDS^{3D} (mean 6.50 recordings across all patients).
- Among patients with more than one MS subtype recording in the inclusion period, 18 patients (8.94%) presented with more than one different subtype.

Table 1: Patient characteristics at first recording

Characteristics	N = 179
Demographics	
Female, n (%)	135 (75.42)
Age, mean years (SD)	48.13 (11.79)
Clinical characteristics	
Time since diagnosis*, mean years (SD)	11.02 (8.35)
MS subtype at index (MSDS^{3D}), n (%)	
ND/CIS	5 (2.79)
RRMS	147 (82.12)
PPMS	6 (3.35)
SPMS	20 (11.17)
Unspecified	1 (0.56)
EDSS at index, n (%)	
1-1.5	36 (20.11)
2-2.5	48 (26.82)
3-3.5	35 (19.55)
4-4.5	25 (14.00)
5-5.5	5 (2.79)
6-6.5	23 (12.85)
7-7.5	6 (3.35)
8-8.5	1 (0.56)

*Missing for 10 patients; ND, newly diagnosed

MULTICLASS CONFUSION MATRIX

- For 148 observations (12.73%), an MS diagnosis could not be assigned in claims when requiring 1 inpatient or 2 outpatient diagnoses of an MS subtype in the one-year period surrounding the index.
- For 465 cases (39.98%), only a G35.9 unspecified diagnosis was assigned in claims data vs. 0.69% true unspecified cases in MSDS^{3D} (**Table 2**).
- True positives were most commonly observed for RRMS (G35.1); false positives were most commonly predicted as unknown (G35.9) (**Table 3**).

Table 2: Multi-class confusion matrix

		Predicted: *Claims Data*				
		G35.0	**G35.1**	**G35.2**	**G35.3**	**G35.9**
True: *MSDS^{3D}*	**G35.0**	2	4	0	1	1
	G35.1	51	342	59	26	440
	G35.2	2	4	21	1	6
	G35.3	0	9	14	10	15
	G35.9	4	0	0	0	3

Table 3: Positive/negative predictions by MS subtype

MS Subtype	True Positive	True Negative	False Positive	False Negative
G35.0: *ND/CIS*	2	1097	57	7
G35.1: *RRMS*	342	87	17	717
G35.2: *PPMS*	21	1051	73	18
G35.3: *SPMS*	10	1087	28	38
G35.9: *Unknown*	3	693	462	5

ND, newly diagnosed

PREDICTIVE PERFORMANCE

- Congruency across all observations in the two datasets was 32.50% (378/1,163).

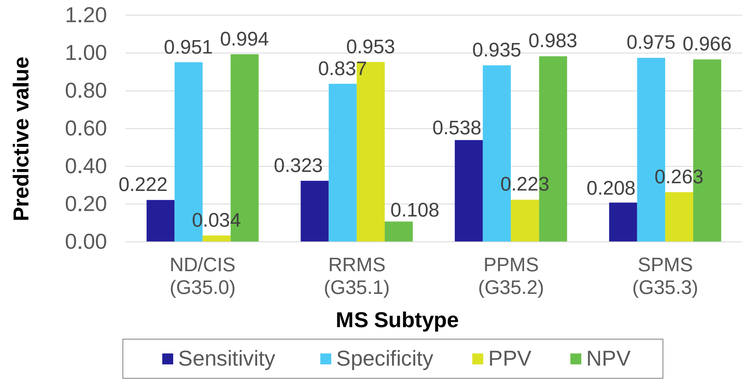

Figure 1: Predictive performance of subtype coding

- PPV was highest for RRMS (G35.1, 0.953) and lowest for newly diagnosed/CIS (G35.0, 0.034), whereas NPV was highest for PPMS (G35.2, 0.983) and lowest for RRMS (0.108; **Figure 1**).
- Sensitivity (recall) was generally low, with the highest value achieved for PPMS (G35.2, 0.538); In contrast, specificity was high across all subtypes (>0.837; **Figure 1**).
- The mean F1-score across all subtypes including unknown (G35.0 - G35.9) was 0.220, indicating poor overall predictive performance.
- Results remained consistent when analyzing only the first recorded MS subtype per patient (1 index observation per patient).

Limitations

- The analysis assessed one definition (1 inpatient and/or at least 2 outpatient diagnoses) in a one-year period. It is possible that alternative diagnoses requirements or longer periods within claims data may increase sensitivity.
- The analysis additionally looked at both neurology specialty and GP specialty; whereby it is more likely that the MS subtype is unknown. Despite this, many cases were assigned only an unknown subtype.

Conclusions

- While MS diagnoses are well recorded in administrative claims data sources, the specific subtype is poorly captured, with approx. 40% of cases assigned an unspecified MS type (G35.9).
- The overall low predictive performance of specific MS subtypes in claims data highlights the importance of linking to patient medical records to supplement detailed clinical information for real-world evidence studies in MS.

References

1. Tafti D, Ehsan M, Xixis KL. Multiple Sclerosis. 2022 Apr 9. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan–. PMID: 29763024.
2. Ziemssen T, Kempcke R, Eulitz M, Großmann L, Suhrbier A, Thomas K, et al. Multiple sclerosis documentation system (MSDS): moving from documentation to management of MS patients. J Neural Transm (Vienna). 2013;120 Suppl 1:S61-6.
3. Ziemssen, T., Kern, R., Voigt, I. & Haase, R. Data Collection in Multiple Sclerosis: The MSDS Approach. Front Neurol 11, 445 (2020).
4. Kern, R., Haase, R., Eisele, J. C., Thomas, K. & Ziemssen, T. Designing an Electronic Patient Management System for Multiple Sclerosis: Building a Next Generation Multiple Sclerosis Documentation System. Interact J Medical Res 5, e2 (2016).

Disclosures

M.G. and T.W are employees of IPAM and E.Z. is an employee of Cytel. A.D. has received compensation and travel grants from Biogen, Celgene, Janssen, Roche and Sanofi for speaker activity. T.Z. has received consulting fees, grants, and research support from various pharmaceutical companies e.g. Bayer, Biogen, Merck, Roche, Sanofi. T.W. has received honoraria from several pharmaceutical/consultancy firms e.g. Roche, Abbvie, Merck, BMS, Bayer.

Contact information

evi.zhuleku@cytel.com

www.cytel.com

Cytel

IPAM
INSTITUT FÜR PHARMAKOÖKONOMIE UND ARZNEIMITTELLOGISTIK
AN-INSTITUT DER HOCHSCHULE WISMAR