

Research Letter

Diagnostic Code Ambiguity and Misclassification of Adults With Spinal Muscular Atrophy: Single-Center Chart Review

Gabriel Holly¹, MS; Bryce Bean², MS; Haidy Beshay¹, MD; Gabrielle Edwards², BS; Nicholas Streicher¹, MD, MPH

¹Department of Neurology, School of Medicine, Georgetown University, Washington, DC, United States

²MedStar Health Research Institute, Washington, DC, United States

Corresponding Author:

Nicholas Streicher, MD, MPH
Department of Neurology
School of Medicine, Georgetown University
3800 Reservoir Rd.
Washington, DC
United States
Phone: 1 202-877-1212
Email: streichn@umich.edu

Abstract

In a single US health system, fewer than one in three adults flagged by spinal muscular atrophy–associated diagnostic codes had molecularly confirmed spinal muscular atrophy (positive predictive value 27%), with the majority miscoded across clinically distinct categories.

JMIR Med Inform 2026;14:e103543; doi: [10.2196/103543](https://doi.org/10.2196/103543)

Keywords: spinal muscular atrophy; ICD-10; clinical coding; International Classification of Diseases; electronic health records; rare diseases; disease-modifying therapy; natural language processing; cohort identification; SMN1; secondary use of clinical data; diagnostic miscoding

Introduction

Spinal muscular atrophy (SMA) is an autosomal recessive disorder caused by biallelic loss of the survival motor neuron 1 (SMN1) gene [1]. Since 2016, three disease-modifying therapies have been approved, all now available to adults. The antisense oligonucleotide nusinersen and the oral splicing modifier risdiplam benefit adults as well as children [2]. The gene-replacement therapy onasemnogene abeparvovec, previously limited to infants, reached patients aged 2 and older through an intrathecal formulation (2025). Adults now make up roughly half the prevalent population, an estimated 8000-10,000 individuals [3]. Adults first diagnosed in childhood, when genetic confirmation was not routine and before treatment existed, may have had little reason to follow up, especially if mild or stable. Identifying them within a health system now relies on coded electronic health record (EHR) data.

The primary SMA diagnostic code is an unreliable marker. Diagnostic terminologies lack SMA-specific codes and group distinct disorders under shared rubrics, conflating true SMA

with similarly named conditions such as spinal and bulbar muscular atrophy (SBMA) [4]. In a claims-based cohort, SMA codes did not distinguish subtypes or confirm true cases [5], and in a genomic cohort, overlapping codes could not confirm SMN1-SMA without genetic testing [6].

We quantified the positive predictive value (PPV) of SMA-associated diagnostic codes against chart adjudication in one US health system and characterized the miscoded cases.

Methods

Ethical Considerations

We conducted a single-center retrospective chart review in an academic multihospital health system in the mid-Atlantic United States, approved by the MedStar Health Research Institute Institutional Review Board (Protocol STUDY00008773) and conducted in accordance with the Declaration of Helsinki under a Health Insurance Portability and Accountability Act waiver.

Study Design

On October 29, 2025, we identified all adults (≥ 18 y at the qualifying encounter) with any encounter in the preceding 10 years carrying an SMA-associated code in either *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM; G12.0, infantile SMA, type 1; G12.1, other inherited SMA; G12.8, other SMA and related syndromes; G12.9, SMA, unspecified) or SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms), since the EHR stores problem-list diagnoses in SNOMED CT, mapped to ICD-10-CM for billing. Each chart was attributed to the diagnostic code as entered by the provider, either ICD-10-CM or SNOMED CT. G12.0 was included because it can persist on adult charts through historical entries. Confirmed SMA required documented molecular SMN1 confirmation. Each chart, in which every patient had been evaluated by at least two neurologists, was reviewed for the documented diagnosis and responsible code, then classified as confirmed SMA, other appropriately coded motor neuron condition, SBMA, asymptomatic carrier, prenatal carrier screening, no relevant neuromuscular diagnosis, or unrelated. PPV was calculated with Wilson

score 95% CIs. The per-code distribution, demographic characteristics, and category definitions appear in [Multimedia Appendix 1](#).

Results

Of 60 charts, 22 (37%, 95% CI 26%-49%) were appropriately coded for SMA or a related motor neuron disorder. Sixteen (27%, 95% CI 17%-39%) had molecularly confirmed SMA, and the other 6 were appropriately coded motor neuron conditions (5 Hirayama disease, 1 distal SMA). The PPV of an SMA-associated code for molecularly confirmed SMA was therefore 27%, rising to 37% if related motor neuron disorders are included. Among the 16 confirmed patients, all seen after 2016, 10 (63%) had no documented disease-modifying therapy.

The remaining 38 charts (63%) were miscoded ([Table 1](#)), comprising conditions unrelated to SMA (n=10, 17%), asymptomatic SMA carriers (n=9, 15%), charts with no relevant neuromuscular diagnosis (n=9, 15%), SBMA (n=5, 8%), and prenatal carrier screening (n=5, 8%).

Table 1. Adjudicated diagnosis of 60 adult charts flagged by spinal muscular atrophy (SMA)–associated diagnostic codes.

Adjudicated category	n	%
Confirmed SMA (molecularly confirmed)	16	27
Other appropriately coded motor neuron condition	6	10
Spinal and bulbar muscular atrophy	5	8
Asymptomatic SMA carrier	9	15
Prenatal carrier screening	5	8
No relevant neuromuscular diagnosis	9	15
Unrelated diagnosis	10	17
Total	60	100

Discussion

Most adults whose chart carried an SMA code did not have SMA. Fewer than 1 in 3 had confirmed disease, well below the roughly 4 in 5 correct for the infantile SMA code in a national registry [7], the two-thirds for amyotrophic lateral sclerosis codes [8], and the one-half for muscular dystrophy [9]. No comparable estimate existed for the SMA-associated codes used in adults. Because the the ICD-10 codes provide no SMA subtype and bundle genetically distinct disorders, carriers, and screening encounters under shared rubrics, the code cannot by itself separate affected patients from unaffected ones.

Although SBMA and distal SMA can resemble SMA clinically, they are distinct genetic diseases. SBMA lacks a dedicated code and is often filed under SMA or adjacent codes, defensible administratively yet reflecting a classification ICD-10 no longer captures. Asymptomatic carriers and prenatal screening encounters represent unaffected individuals. Counting any as SMA inflates the apparent cohort; here, miscoded charts outnumbered confirmed cases more than 2 to 1, so a registry built on these codes would be wrong before its

first analysis. Because coding underlies disease queries, this undermines estimates of prevalence, treatment patterns, and natural history from single systems to national registries.

Correcting this requires two changes. First, the codes must be aligned with the clinical record so they are applied as documented and anchored on molecular SMN1 confirmation with type-specific subcodes so they are precise. Europe built Orphacodes, a dedicated rare-disease coding system, to capture diseases ICD-10 cannot distinguish [10], and a comparable US reform through the National Center for Health Statistics could require SMN1 confirmation for the main SMA code, add subcodes for types 0 to 4, and retire the catchall categories. Second, these codes must then be applied consistently to every note, old and new. In future work, disease-specific natural language processing could recode existing records and check new entries against the documented diagnosis.

As a single-center, retrospective estimate from a small sample, this rate may not transfer to other settings, though the pattern appears in claims data elsewhere. Adjudication drew on records in which every patient had been evaluated by at

least two neurologists, and confirmed cases were anchored on molecular testing rather than reviewer judgment. Because all confirmed patients were seen after 2016, undocumented therapy is not a pretreatment artifact of limited confirmation or stable disease; it more likely reflects care outside this system, which a single institution cannot distinguish from loss to follow-up. Privacy-preserving record linkage could close this gap.

Despite these limits, SMA-associated diagnostic codes are an unreliable marker for confirmed disease. Codes that are aligned, molecularly anchored, consistently applied, and audited would make identification reliable, strengthening prevalence estimates, treatment tracking, and natural-history research.

Acknowledgments

A generative AI assistant (Claude Opus 4.8; Anthropic) was used in preparing the manuscript and for copyediting; the authors reviewed all content and take full responsibility for it.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data supporting the findings of this study contain protected health information and are not publicly available. Deidentified data may be available from the corresponding author on reasonable request, subject to institutional and privacy approvals.

Author Contributions

NS: conceptualization, methodology, supervision, validation, writing – original draft and review and editing. GH: investigation, data curation, formal analysis, writing – review and editing. BB: investigation, data curation, writing – review and editing. HB: investigation, writing – review and editing. GE: investigation, writing – review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables.

[[DOCX File \(Microsoft Word File\)](#), 10 KB-Multimedia Appendix 1]

References

1. Mercuri E, Sumner CJ, Muntoni F, Darras BT, Finkel RS. Spinal muscular atrophy. *Nat Rev Dis Primers*. Aug 4, 2022;8(1):52. [doi: [10.1038/s41572-022-00380-8](https://doi.org/10.1038/s41572-022-00380-8)] [Medline: [35927425](https://pubmed.ncbi.nlm.nih.gov/35927425/)]
2. Hagenacker T, Wurster CD, Günther R, et al. Nusinersen in adults with 5q spinal muscular atrophy: a non-interventional, multicentre, observational cohort study. *Lancet Neurol*. Apr 2020;19(4):317-325. [doi: [10.1016/S1474-4422\(20\)30037-5](https://doi.org/10.1016/S1474-4422(20)30037-5)] [Medline: [32199097](https://pubmed.ncbi.nlm.nih.gov/32199097/)]
3. Cure SMA. State of SMA report (data year 2021). Cure SMA. 2022. URL: <https://www.curesma.org/wp-content/uploads/2023/03/2022-State-of-SMA-Report.pdf> [Accessed 2026-05-21]
4. Kennedy WR, Alter M, Sung JH. Progressive proximal spinal and bulbar muscular atrophy of late onset. A sex-linked recessive trait. *Neurology*. Jul 1968;18(7):671-680. [doi: [10.1212/wnl.18.7.671](https://doi.org/10.1212/wnl.18.7.671)] [Medline: [4233749](https://pubmed.ncbi.nlm.nih.gov/4233749/)]
5. Belter L, Cruz R, Kulas S, McGinnis E, Dabbous O, Jarecki J. Economic burden of spinal muscular atrophy: an analysis of claims data. *J Mark Access Health Policy*. Nov 8, 2020;8(1):1843277. [doi: [10.1080/20016689.2020.1843277](https://doi.org/10.1080/20016689.2020.1843277)] [Medline: [33224449](https://pubmed.ncbi.nlm.nih.gov/33224449/)]
6. Hall TS, Beaumont RN, Fasham J, et al. Characterisation of the SMN1/2 locus using a highly specific variant caller on whole-genome sequence data from 500,000 individuals. *Eur J Hum Genet*. Jul 2, 2026. [doi: [10.1038/s41431-026-02175-x](https://doi.org/10.1038/s41431-026-02175-x)] [Medline: [42393335](https://pubmed.ncbi.nlm.nih.gov/42393335/)]
7. National Disease Registration Service. Spinal muscular atrophy type 1: NCARDRS data briefing, appendix 1 (technical notes). NHS England. 2023. URL: <https://digital.nhs.uk/ndrs/data/data-outputs/spinal-muscular-atrophy-type-1-ncardrs-data-briefing/spinal-muscular-atrophy-type-1-appendix-1> [Accessed 2026-06-15]
8. Stickler DE, Royer JA, Hardin JW. Accuracy and usefulness of ICD-10 death certificate coding for the identification of patients with ALS: results from the South Carolina ALS Surveillance Pilot Project. *Amyotroph Lateral Scler*. Jan 2012;13(1):69-73. [doi: [10.3109/17482968.2011.614253](https://doi.org/10.3109/17482968.2011.614253)] [Medline: [21929354](https://pubmed.ncbi.nlm.nih.gov/21929354/)]
9. Smith MG, Royer J, Mann JR, McDermott S. Using administrative data to ascertain true cases of muscular dystrophy: rare disease surveillance. *JMIR Public Health Surveill*. Jan 12, 2017;3(1):e2. [doi: [10.2196/publichealth.6720](https://doi.org/10.2196/publichealth.6720)] [Medline: [28082256](https://pubmed.ncbi.nlm.nih.gov/28082256/)]

10. Angin C, Mazzucato M, Weber S, et al. Coding undiagnosed rare disease patients in health information systems: recommendations from the RD-CODE project. *Orphanet J Rare Dis*. Jan 27, 2024;19(1):28. [doi: [10.1186/s13023-024-03030-2](https://doi.org/10.1186/s13023-024-03030-2)] [Medline: [38280999](https://pubmed.ncbi.nlm.nih.gov/38280999/)]

Abbreviations

EHR: electronic health record

ICD-10-CM: *International Classification of Diseases, Tenth Revision, Clinical Modification*

PPV: positive predictive value

SBMA: spinal and bulbar muscular atrophy

SMA: spinal muscular atrophy

SMN1: survival motor neuron 1

SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms

Edited by Andrew Coristine; peer-reviewed by Kentaro Okamoto; submitted 04.Jun.2026; final revised version received 15.Jul.2026; accepted 17.Jul.2026; published 28.Jul.2026

Please cite as:

Holly G, Bean B, Beshay H, Edwards G, Streicher N

Diagnostic Code Ambiguity and Misclassification of Adults With Spinal Muscular Atrophy: Single-Center Chart Review
JMIR Med Inform 2026;14:e103543

URL: <https://medinform.jmir.org/2026/1/e103543>

doi: [10.2196/103543](https://doi.org/10.2196/103543)

© Gabriel Holly, Bryce Bean, Haidy Beshay, Gabrielle Edwards, Nicholas Streicher. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 28.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.