Examining the Prevalence of Previously Recorded Phenotypically Related Diagnoses Among Fee-for-Service Medicare Enrollees Newly Diagnosed with Mendelian Conditions

J Gen Intern Med
DOI: 10.1007/s11606-020-06469-8

Using Vanderbilt's patient database, Batarache et al. found that constellations of billing codes could be used to identify patients with previously unidentified Mendelian (gene-borne) diseases.[1,2] Artificial intelligence-informed, billing-record-based[3] physician decision support at the point of care might enable earlier diagnosis and treatment. Among the fee-for-service Medicare population, we sought to examine the prevalence with which cases of newly diagnosed Mendelian conditions had phenotypically related diagnoses in previous years

METHODS

We used Medicare inpatient, outpatient, and part B files to identify individuals who were fully enrolled in fee-for-service Medicare between 2016 and 2019 and had a 2019 ICD-10 diagnosis of any of 12 Mendelian genetic conditions (each having at least 150 newly diagnosed cases between 2016 and 2019) listed in Table 1 that had not been recorded in 2016, 2017, or 2018.

For beneficiaries with any of these diagnoses, we examined 2016–2018 billing records to identify ICD-10 codes phenotypically associated with each specific condition, as described by Wu et al.[4] and provided through the Phenome Wide Association Studies Resources website.[5] For each condition, we enumerated condition-specific-related ICD-10 diagnostic codes (for example, macrocephaly for achondroplasia) and calculated the proportion of cases for which at least 5 and at least 10 phenotypically related codes were listed in billing records in 2018 and between 2016 and 2018. We also examined distributions of cases across white, black, and other race, limiting to those categories because black and white race constituted the majority of cases.

We had IRB and Centers for Medicare and Medicaid Services (CMS) approval to conduct this work through CareJourney's Virtual Research Data Center.

RESULTS

In 2019, of 39,917,598 beneficiaries fully enrolled in fee-for-service Medicare between 2016 and 2019, 28,377 had a newly coded diagnosis of at least 1 of the Mendelian diseases that we studied (Table 1). Polycythemia vera was the most common, representing 44% of all conditions examined; achondroplasia was the least common. With the exceptions of hereditary hemochromatosis and polycythemia vera, older white women accounted for most cases.

Depending on the disorder, between 60.5 and 87.8% of patients with a newly diagnosed Mendelian disorder in 2019 had at least 5 phenotypically related diagnoses in the previous year; between 73.6 and 97.3% had at least 5 phenotypically related diagnoses coded across the previous 3 years (Table 2). In 2018, between 29.7 and 58.7% of patients newly diagnosed in 2019 had at least 10 phenotypically related diagnoses; aggregating data from 2016 to 2018, those proportions grew to between 50.4 and 84.8%.

DISCUSSION

We used 4 years of Medicare fee-for-service data to identify beneficiaries with any of 12 newly coded Mendelian diseases and examined the prevalence of phenotypically related diagnoses in the 3 preceding years' billing records. For a given condition, up to 87.8% of identified patients had 5 or more related diagnoses in the year immediately preceding a new Mendelian condition diagnosis; up to 97.3% had 5 or more across the preceding 3 years.

Our findings suggest that—even in the older, Medicare-insured population that we studied—artificial intelligence-informed decision support might help providers identify patients with Mendelian disorders by aggregating constellations of diagnoses recorded in the recent past that suggest an overarching one.[3] Surprisingly, a not insubstantial

Received October 1, 2020
Accepted December 13, 2020

Published online: 21 January 2021

Table 1 Number and Distribution Across Race and Age Categories of New Diagnoses of 12 Mendelian Genetic Conditions in 2019 that Had Not Been Diagnosed in 2016, 2017, or 2018. Blank Cells Represent < 11 Patients; CMS Does Not Allow Reporting Exact Numbers for Cell Sizes less than 11. Data for Ages 18–30 Are Not Reported as the Vast Majority of Cells Had < 11 Patients

Disease	Total	White race	Black race	Other race	White race						Black race					
					Female			Male			Female			Male		
					31–44	45–64	65–99	31–44	45–64	65–99	31–44	45–64	65–99	31–44	45–64	65–99
Achondroplasia	158	131	12	15		15	39		29	25						
Alpha-1-antitrypsin deficiency	2924	2733	109	82	37	226	1293	36	189	940		23	42		12	28
Classical phenylketonuria	1087	906	79	102		54	462		48	327			37			21
DiGeorge's syndrome	255	242			11		102			94						
Duchenne or Becker muscular dystrophy	703	603	53	47	21	42	83	123	184	87					14	
Fragile X chromosome	375	321	27	27	20	23	51	64	77	64						
Hereditary hemorrhagic telangiectasia	3734	3522	98	114	19	113	2076		73	1229		19	45			26
Hereditary hemochromatosis	5568	5245	163	160	41	246	2224	59	301	2367	13	23	54		22	48
Neurofibromatosis, type 1	676	586	60	30	24	56	229	25	45	193			17		11	
Neurofibromatosis, type 2	175	159	12			18	66		18	43						
Polycythemia vera	12,655	11,447	611	597	62	545	4579	161	930	5136		57	199	34	119	185
Tuberous sclerosis	233	184	36	13	12	31	59	12	22	43						

Table 2 Proportion of Patients with Newly Coded Mendelian Disease in 2019 Who Had 5 or More (Left) or 10 or More (Right) Phenotypically Related Diagnoses in 2018 and in the 3-Year Period 2016–2018

Disease	Total new in 2019	% with 5 or more related diagnoses		% with 10 or more related diagnoses	
		In 2018	In 2016–2018	In 2018	In 2016–2018
Achondroplasia	158	66.5%	90.5%	29.7%	62.7%
Alpha-1-antitrypsin deficiency	2924	87.8%	96.9%	58.7%	84.8%
Classical phenylketonuria	1087	87.2%	97.1%	55.8%	84.4%
DiGeorge's syndrome	255	85.9%	97.3%	52.5%	82.4%
Duchenne or Becker muscular dystrophy	703	75.5%	92.9%	43.2%	71.6%
Fragile X chromosome	375	60.5%	73.6%	30.9%	50.4%
Hereditary hemorrhagic telangiectasia	3734	86.5%	96.9%	52.9%	83.7%
Hereditary hemochromatosis	5568	82.2%	95.2%	47.5%	78.4%
Neurofibromatosis, type 1	676	71.3%	87.9%	37.3%	67.8%
Neurofibromatosis, type 2	175	72.0%	89.1%	45.7%	72.0%
Polycythemia vera	12,655	81.2%	94.0%	47.3%	77.8%
Tuberous sclerosis	233	74.2%	88.8%	44.6%	73.0%

number of Medicare beneficiaries might be identified with the disorders we studied.

Our study has several limitations. First, it is possible that Mendelian disorder diagnostic codes were simply not recorded for 3 years before reappearing in 2019. While we recently found a fairly dramatic year-to-year drop-off in diagnostic coding of chronic conditions,[6] Mendelian conditions tend to be life-long, disabling, and, frequently, visually apparent; it should be somewhat surprising for them not to be recorded. Second, we were not able to confirm the diagnoses we studied with genetic testing, as Batarache et al. were able to do.[2] Third, our study was limited by its reliance on relatively recent administrative datasets wherein final reconciliation delays might trivially impact dataset completeness. Finally, future research should explore whether artificial intelligence-based decision support using recent phenotypically related diagnoses is appropriate for the Medicare-insured population.

Nonetheless, our analysis demonstrates that there are relatively large numbers of individuals in the fee-for-service Medicare beneficiary population that might be identified as having a Mendelian genetic disorder by screening phenotypically related diagnostic billing codes. Among those for whom the diagnosis is indeed novel, earlier genetic testing and diagnosis of these Mendelian disorders might lead to better treatment and outcomes.

William B. Weeks, MD, PhD, MBA[1]
Grace Huynh, MD, PhD[1]
Stacey Y. Cao, MHS[2]
Jeremy Smith, MPH[3]
Chaitanya Bangur, PhD, MBA, MS[1]
James N. Weinstein, DO, MS[1,4,5,6]

[1]Microsoft Corporation, Microsoft Research, Redmond, WA, USA
[2]CareJourney, Arlington, VA, USA
[3]White River Junction VA Outcomes Group, WRJ, Hartford, VT, USA
[4]The Dartmouth Institute, Lebanon, NH, USA
[5]Kellogg School of Business, Evanston, IL, USA
[6]Amos Tuck School of Business, Hanover, NH, USA

Corresponding Author: *William B. Weeks, MD, PhD, MBA; Microsoft Corporation, Microsoft Research, Redmond, WA, USA (e-mail: william.weeks@microsoft.com).*

Compliance with Ethical Standards:

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*
CareJourney has IRB approval ("Developing Care Model, Provider and Network Evaluation Tools using Medicare Claims and Assessment" Solutions IRB Study #2019, CMS DUA 52882) to analyze CMS data through a Virtual Research Data Center.

REFERENCES

1. **Bastarache L, Hughey JJ, Goldstein JA, et al.** Improving the phenotype risk score as a scalable approach to identifying patients with Mendelian disease. J Am Med Inform Assoc. 2019;26:1437-47.
2. **Bastarache L, Hughey JJ, Hebbring S, et al.** Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. Science. 2018;359:1233-9.
3. **Dias R, Torkamani A**. Artificial intelligence in clinical and genomic diagnostics. Genome Med. 2019;11:70.
4. **Wu P, Gifford A, Meng X, et al.** Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow Development and Initial Evaluation. JMIR Med Inform. 2019;7:e14325.
5. Phenome Wide Association Studies Resources. Human Phyotype Ontology to ICD10 maps. https://phewascatalog.org/. Accessed July 9, 2020.
6. **Weeks WB**, **Huynh G, Cao SY, Smith J, Weinstein JN**. Assessment of year-to-year patient specific comorbid conditions reported in the Medicare Chronic Conditions Data Warehouse. JAMA Netw Open. 2020;3(10):e2018176.