Testing and Extending Strategies for Identifying Genetic Disease-Related Encounters in Pediatric Patients

Lisa P. Spees, PhD[1,2], **Karen Hicklin, PhD**[3], **Michael C. Adams, MD**[4], **Laura Farnan, PhD**[2], **Jeannette T. Bensen, PhD**[2,5], **Donna B. Gilleskie, PhD**[6], **Jonathan S. Berg, MD, PhD**[7], **Bradford C. Powell, MD, PhD**[7], **Kristen Hassmiller Lich, PhD**[1]

[1]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill (UNC-CH);

[2]Lineberger Comprehensive Cancer Center, UNC-CH;

[3]Department of Industrial and Systems Engineering, Herbert Wertheim College of Engineering, University of Florida;

[4]Division of Pediatric Genetics and Metabolism, UNC-CH;

[5]Department of Epidemiology, Gillings School of Global Public Health, UNC-CH;

[6]Department of Economics, UNC-CH;

[7]Department of Genetics, UNC-CH

Abstract

Purpose: To better understand healthcare utilization and develop decision support tools, methods for identifying patients with suspected genetic disease (GD) are needed. Gonzaludo (2019) identified inpatient-relevant ICD codes that were *possibly, probably,* or *definitely* indicative of GDs. We assessed whether these codes identified GD-related inpatient, outpatient, and emergency department (ED) encounters among pediatric patients with suspected GDs from a prior study (NCGENES).

Methods: Using the electronic medical records of 140 pediatric patients from the NCGENES study, we characterized the presence of ICD codes representing *possible*, *probable*, or *definite* GD-related codes across encounter types. Additionally, we examined codes from encounters where initially no GD-related codes had been found and determined if these codes were indicative of a GD.

Corresponding Author: Lisa P. Spees, 1102-G McGavran-Greenberg, CB#7411, Gillings School of Global Public Health, Chapel Hill, NC 27599, lspees21@live.unc.edu, (919)-966-6609.

Results: Among NCGENES patients with visits between 2014–2017, 92% of inpatient, 75% of ED, and 63% of outpatient encounters included ≥1 GD-related code. Encounters with highly specific (i.e., *definite*) GD codes had fewer low-specificity GD codes than encounters with only low-specificity GD codes. We identified an additional 32 ICD9 and 56 ICD10 codes *possibly* indicative of a GD.

Conclusion: Code-based strategies can be refined to assess pediatric patients' healthcare utilization and may contribute to a systematic approach to identify patients with suspected GDs.

INTRODUCTION

Children with genetic diseases (GDs) comprise only about 1–4% of the pediatric population (1–3). However, up to one-third of total pediatric healthcare spending is attributed to these patients (3). Not surprisingly, inpatient GD-related visits among pediatric patients cost on average $12,000 to $77,000 more than non-GD inpatient visits, and in total, aggregate costs for pediatric patients with GD include about $57 billion dollars per year in hospitalizations alone (4). Moreover, pediatric patients with GDs have high utilization of other healthcare resources; they are more likely to visit the emergency room or an ambulatory care setting compared to the general pediatric population (5,6). Even compared to pediatric patients with chronic conditions, pediatric patients with GDs need almost twice as much specialty care every year (5,7).

In the United States, assessment of the overall healthcare utilization of pediatric patients with suspected or diagnosed GDs has been limited to the inpatient setting (4,8,9). In order to understand the burden of GD and to identify opportunities to improve care for patients with suspected or diagnosed GDs, a broader characterization of their healthcare utilization patterns is needed. Furthermore, enhancing our ability to identify patients with suspected GDs could shorten diagnostic odysseys, improve decision-making for genetic testing and care management, increase patient quality of life and outcomes, and ultimately lower healthcare costs.

Analyses focused on examining the burden of complex rare disorders have used International Classification of Disease (ICD) codes in administrative data to examine related diagnoses, including GD-related diagnoses (3,4,6,8). As a notable example, in Gonzaludo et al., a medical geneticist manually curated a list of ICD codes into three categories: *definite* GD (i.e., 100% of discharges will yield a genetic diagnosis with appropriate testing), *probable* GD (i.e., >50% of discharges will yield a genetic diagnosis with appropriate testing), or *possible* GD (i.e., >10% of discharges will yield a genetic diagnosis with appropriate testing). These lists of GD-related codes were then applied to pediatric inpatient discharges to assess the burden of GD in a nationally representative sample of children in the US (4). Improved diagnosis and management of pediatric patients with rare GDs may be enabled by their identification in clinical or claims data; diagnostic code-based identification strategies provide a viable starting point. With this goal in mind, the present analysis

had two objectives. First, among a cohort of pediatric patients enriched for suspected or diagnosed GD, we sought to assess the extent to which the Gonzaludo et al. approach identified GD-related hospitalizations and to identify additional GD codes from unflagged hospitalizations that should be added to the identification strategy. Second, we replicated this process in two new encounter settings to assess its generalizability among these patients' outpatient and emergency department visits.

METHODS

Study Population & Data Sources

The study sample included a subset of participants enrolled between 2012–2015 in the North Carolina Clinical Genomic Evaluation by Next-Generation Exome Sequencing (NCGENES1), an NIH-funded feasibility study seeking to explore optimal applications of exome sequencing (ES) (10). NCGENES included children and adults with suspected Mendelian GDs that have a variety of indications including cancer, cardiogenetic diseases, neurodevelopmental disorders, and retinal diseases. For the present analysis, we included 140 pediatric patients who displayed the most commonly found phenotypes among this population, which included neuromuscular disorders, syndromic and nonsyndromic intellectual disability/autism, dysmorphology, and epilepsy. Electronic medical records (EMR) data, available between April 4, 2014 and December 29, 2017, was retrieved from the Carolina Data Warehouse for Health (CDW-H), a central data repository containing clinical, research, and administrative data sourced from the UNC Health Care System (11). These data included patient demographics, laboratory test results, medications, and encounter types (i.e., inpatient, ED, and outpatient). We categorized any outpatient or ED encounters that immediately preceded or overlapped with an inpatient encounter as an inpatient encounter only; consequently, 40% of observed ED encounters were ultimately categorized as inpatient encounters.

To identify GD-related codes, we used mutually exclusive lists of ICD codes that were categorized as *possible, probable*, or *definite* GD (4). Following Gonzaludo's approach, these lists were stratified based on patient age at the time of encounter and classified as newborn (birth to 28 days old) or pediatric (29 days old to 18 years old). Neonate and pediatric ICD9 GD-related codes were included as supplemental files in Gonzaludo's article while neonate and pediatric ICD10 GD-related codes were obtained directly from the authors. We added an additional category labeled *other* GD, which included ICD codes listed as a *possible*, *probable*, or *definite* GD diagnosis for one age group but excluded from another age group; for example, if an ICD code was listed as a *possible* GD for newborns but was not listed as a *possible, probable,* or *definite* GD for pediatric patients, the ICD code would be included as an *other* GD code for pediatric patients. Appendix Tables 1–3 present the lists of ICD codes identifying *possible, probable, definite,* or *other* GDs by age group.

Data Analysis

In order to evaluate a code-based strategy for identifying GD-related encounters, we first described the demographic characteristics of the NCGENES patient cohort overall and by encounter type. We then characterized the frequency of ICD codes representing

possible, *probable*, and *definite* GD-related diagnoses among patients' inpatient, ED, and outpatient encounters. Specifically, for each encounter type, we described the (1) total number of encounters, (2) proportion of encounters with at least one GD-related code, and (3) proportion of encounters with the highest specificity of GD code (i.e., proportion of encounters with one or more *definite* codes; of those without a *definite* code, the proportion of encounters with one or more *probable* codes; and, of those without a *probable* code, the proportion of encounters with one or more *possible* codes). Among the stratified groups, based on highest specificity of GD-related code, we calculated the average number of *definite/probable/possible* GD-related codes included per encounter. Chi-square tests and t-tests were employed to identify differences between encounter types.

Additionally, we extended the lists of GD-related codes. Specifically, a single study clinician (MA) examined a list of the ICD codes from encounters where initially no GD-related codes were identified and manually assessed whether these diagnoses were potentially indicative of a GD using the same definitions used by the clinical geneticist who manually curated the list of GD-related ICD codes in Gonzaludo's article. For example, to identify a *possible* GD, we assessed if "10% of all children with this code applied to them have a genetic disease." A 5% random sample of these diagnoses and their assigned GD-classifications was then reviewed by two study clinical geneticists (JSB, BCP) to ensure concordance in diagnostic classification. Discordant classifications were resolved by consensus among the three clinical experts. We then re-analyzed the NCGENES cohort using the updated lists of GD-related codes. While these groupings of ICD codes are only expert opinion-based estimates, they do provide sets that would be expected to have varying "specificity" for identifying patients who might benefit from genetic testing, which is of interest for developing decision support for providers.

RESULTS

Table 1 provides descriptive characteristics of the NCGENES cohort at the encounter-level. In the NCGENES cohort, participants accumulated 166 inpatient encounters, 48 ED visits, and 2,378 outpatient encounters from 2014–2017. Across all encounter types, the modal age category was between ages 1 and 4 (61% among inpatient encounters, 69% among ED encounters, and 44% among outpatient encounters). The female share of encounters (61% for inpatient encounters, 60% among ED encounters, and 58% among outpatient encounters) is larger than their 53% representation among the 140 patients. White patients comprise a similar proportion of the encounters (64% for inpatient encounters, 60% among ED encounters, and 66% among outpatient encounters) as the patient sample (69%). Overall, these encounter-level frequencies reflected patient-level frequencies (Appendix Table 4).

In Table 2, we present the proportion of inpatient, ED, and outpatient encounters with 1) at least one GD-related code and 2) with the highest specificity of a GD code. The majority of inpatient, ED, and outpatient encounters included at least one GD-related code; specifically, among inpatient, ED, and outpatient encounters, respectively, 92%, 75%, and 63% included at least 1 *definite*, *probable*, or *possible* GD-related code. However, the frequency of these GD-related codes varied by code specificity and by encounter type; for example, a *definite* GD-related diagnosis code was found in 23% of inpatient encounters but only in 8%

of ED encounters and 9% of outpatient encounters (*P*<0.0001). For 38% of inpatient hospitalizations and 31% of ED encounters, the highest specificity of GD-related diagnosis code was in the *probable* category; half as many outpatient encounters (16%) had the highest specificity of GD-related diagnosis code in the *probable* category (*P*<0.0001). Lastly, similar proportions of *possible* GDs (as the highest specificity of GD-related diagnosis code) were found across encounter types (31% of inpatient encounters, 35% of ED encounters, and 38% of outpatient encounters, *P*=0.16). *Other* GD codes were found only among a small number of inpatient (n=7, 4%) and ED (n=1, 2%) encounters. All of these *other* GD code encounters were among pediatric patients who were assigned '*possible* GD codes for newborns' and included: failure to thrive in childhood (n=4), encephalopathy (n=1), unspecified lack of expected normal physiological development in childhood (n=1), and delay in development (n=1).

Additionally, in Table 2, per encounter, we present the average number of *definite/probable/ possible* GD-related codes. Compared to encounters with a *definite* GD-related code, encounters where the highest specificity of a GD-related code was *probable* had a higher average number of *probable* GD-related codes. For example, among outpatient encounters with a *definite* GD-related code, the average number of *probable* GD-related codes was 0.2 whereas outpatient encounters with only *probable* GD-related codes had an average of 1.2 *probable* GD-related codes (*P*<0.0001, t-test). In the same way, compared to encounters where a *definite* GD-related code was identified, encounters where the highest specificity of GD-related code was in the *possible* category included a higher average number of *possible* GD-related codes; among outpatient encounters with a *definite* GD-related code, the average number of *possible* GD-related codes was 0.8 whereas outpatient encounters with only *possible* GD-related codes had an average of 1.7 *possible* GD-related codes (*P*<0.0001, t-test). Similarly, among inpatient encounters with *definite* GD-related codes, there was an average of 2.5 additional *possible* GD-related codes in those encounters; in contrast, when the highest specificity of GD-related code was *probable* or *possible*, there was an average of 4.4 (*P*=0.01, t-test) and 3.9 (*P*=0.03, t-test) additional *possible* GD-related codes included in that inpatient encounter, respectively.

Finally, Table 3 lists the additional ICD codes potentially indicative of a GD. In encounters where no GD-related codes had initially been found, we identified an additional 32 ICD9 and 56 ICD10 codes to classify as *possibly* indicative of a GD for pediatric patients. Among ICD9 code clusters (based on the sections of the ICD manual), the majority of these codes were categorized as "Symptoms, Signs, and Illness" (N=11) and "Factors Influencing Health Status" (N=7). The largest ICD10 code cluster was "Symptoms, Signs, and Abnormal Clinical and Lab Findings" (N=18), followed by "Diseases of the Nervous System" (N=15). Inclusion of these additional *possible* GD codes increased the total number of encounters with a *definite, probable,* or *possible* GD from 92% to 99% for inpatient hospitalizations, 75% to 90% for ED encounters, and 63% to 80% for outpatient encounters (Appendix Table 5). No *other* GD codes were identified using the updated GD code list; ICD codes previously labeled as *other* GD diagnoses were reclassified as *possible* GDs for pediatric patients.

DISCUSSION

Computational medicine approaches can develop powerful tools to leverage data generated during the course of clinical care. These data can be used to understand (with caveats) the patterns of care, healthcare utilization, and costs for patients with particular conditions, and subsequently enable measures to be implemented that will reduce the length of time to diagnosis and optimize management. Computational approaches may assist in the identification of patients at risk of having a rare GD (often referred to as being on a "diagnostic odyssey") and who might benefit from genomic sequencing. When deployed across an entire healthcare system, electronic decision support would allow primary care providers to better identify patients for referral to appropriate specialists early in their diagnostic odyssey, thereby improving consistency across the population, enabling identification of potential disparities in access to specialty care, and facilitating quality improvement efforts.

The present work demonstrates that the categories of ICD codes previously characterized as being *definite, probable,* or *possible* GD-related codes are frequently utilized in both inpatient and outpatient clinical encounters among patients with suspected GDs. However, there are differences in the frequency of these GD-related codes by encounter type and by the code's specificity. Recognizing these patterns will be useful for constructing algorithms to identify patients earlier in their diagnostic odyssey who would benefit most from genetic testing. Additionally, we identified additional ICD codes to include in the *possible* category, thus expanding the repertoire of codes used to analyze the impact and burden of genetic disease in the pediatric population.

We noted intriguing differences in the frequency of GD-related codes by encounter type. These differences may reflect the distinct medical activities and procedures associated with specific encounter types. For example, we hypothesize that since outpatient encounters cover a wide variety of services including preventative care, annual physicals, and medication management, the application of ICD codes may be more selectively focused on the purpose of the visit. In other words, outpatient encounters are less likely to include a broader range of GD-related codes describing chronic symptoms not addressed during the visit. In contrast, ED and inpatient encounters may be initiated by acute or serious medical issues or hard-to-explain symptoms and, subsequently, may be characterized by a more detailed and more comprehensive listing of symptom-related or phenotypic codes. Additionally, ICD codes frequently used among outpatient encounters associated with GD-related diagnoses may differ from those cited in ED or outpatient encounters. For example, in our analysis, ICD codes that broadly suggested a child was experiencing delays in their physical or mental development (e.g., ICD9 code 315.X and ICD10 codes R6251, R6252) were more often found in the outpatient setting. Specifically, among all instances of each of these codes, 81%, 83%, and 94% of, respectively, were associated with outpatient encounters. These generic codes are more likely to be used by primary care providers in outpatient settings (potentially early in the diagnostic odyssey), who would then refer patients to specialists for a more specific diagnosis as well as for potential treatment. In our analysis, we substantially increased the number of GD-related codes found among outpatient encounters by 17%, from

63% to 80%, through the inclusion of these additional ICD codes to the list of *possible* GD codes.

Interestingly, when a highly specific (i.e., *definite*) GD diagnosis was documented during an encounter, our findings suggest that clinicians used fewer GD-related codes to describe the encounter. This difference may indicate that establishing a more specific diagnosis obviated the need to document a number of ancillary symptoms or phenotypic features via ICD codes, at least for the purpose of billing an encounter. Conversely, the higher average number of *possible* or *probable* GD-related codes (i.e., less specific GD-related codes) found within an encounter without a *definite* GD diagnosis may indicate that clinicians are more likely to comprehensively document several present features or symptoms in order to justify the subsequent tests conducted or procedures performed. In other words, when a *definite* GD diagnosis is documented during an encounter, clinicians can more concisely explain the diagnosis and are less likely to need to meticulously document all of a patient's symptoms as billable codes.

While this analysis is an essential initial step in the development of algorithms to identify pediatric patients on diagnostic odysseys, there are a few limitations that should be addressed in future research. First, while our analysis includes only 140 pediatric patients over a three-year period, a broader analysis may increase the generalizability of our findings. Second, while we have a rich dataset of administrative and clinical data, we could not capture encounters that may have occurred outside the UNC Healthcare system. Furthermore, because EMRs are not primarily intended for research purposes, poor standardization in the type and quality of data recorded could lead to incomplete encounter information. While previous research has found significant overlap in EMRs and administrative claims data for outpatient encounters (12), future research should continue to examine the comprehensiveness and accuracy of clinical data as a measure of both utilization and healthcare information by comparing it to the information included in health insurance claims data across encounter types. Third, while this work focuses on the occurrence of GD-related codes among encounters involving patients suspected of having an underlying genetic diagnosis, it does not consider the frequency with which these codes are applied in encounters among other types of patients, which would better quantify specificity of the codes. Lastly, assignment of ICD codes as *possible, probable,* or *definite* GD codes was based on singular classification in a single encounter. However, it is likely that groups of ICD codes, particularly groups of *possible* or *probable* GD codes, within an encounter as well as across encounters, may further increase our ability to differentiate and identify pediatric patients needing genetic testing.

The goal of the present analysis was to assess the frequency with which *possible*, *probable*, and *definite* GD codes occurred in encounters of pediatric patients with suspected GDs. While our analysis increased the sensitivity of identifying patients, the next step is to increase this method's specificity. Consequently, in ongoing work, we are examining the trajectories of patients' healthcare utilization and resource use in order to identify patients earlier on in their diagnostic odysseys. In particular, we are examining three different types of trajectories: 1) diagnoses and symptoms using ICD codes, 2) medical procedures using Current Procedural Terminology (CPT) codes, and 3) timing, frequency, and type

of encounter. This additional level of detail available in EMR data may allow further refinement of predictors that a patient is in early phases of a diagnostic odyssey, the development of longitudinal "patterns of care," and the documentation of the overall cost effectiveness of genomic sequencing in patients undergoing a diagnostic odyssey. Ultimately, our goal is to develop an algorithm that would be employed in a healthcare system's EMRs to identify patients at the beginning of their diagnostic odyssey who would benefit from genetic testing.

Current approaches to determine which patients should receive a genetic test are inconsistent and inequitable (13–15). Even among patients identified as eligible for genetic testing, evidence suggests that the majority do not receive a genetic test (16–22). A systematic approach to identifying patients suspected of having a GD would improve decision-making for genetic testing. In our analysis, we show that the extended GD categorizations may be useful for identifying patients with suspected GDs. We conclude that these refined lists of GD-related codes contribute to examining pediatric patients' healthcare utilization and resource trajectories and, ultimately, to improving the healthcare management of pediatric patients with GDs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Data Availability

The individual EMR dataset (even de-identified) used and/or analyzed during the current study is not publicly available due to Carolina Data Warehouse (CDW-H) policies. Collaboration requests and data use agreements with CDH-W (https://tracs.unc.edu/index.php/services/informatics-and-data-science/cdw-h) are necessary to obtain access to the de-identified EMR data.

REFERENCES

1. Blackburn C, Read J, Spencer N. Children with neurodevelopmental disabilities [Internet]. 2012 [cited 2021 Mar 22]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/252659/33571_2901304_CMO_Chapter_9.pdf

2. Deprivation Emerson E., ethnicity and the prevalence of intellectual and developmental disabilities. J Epidemiol Community Health. 2012 Mar;66(3):218–24.

3. Cohen E, Berry JG, Camacho X, Anderson G, Wodchis W, Guttmann A. Patterns and costs of health care use of children with medical complexity. Pediatrics. 2012;130(6).

4. Gonzaludo N, Belmont JW, Gainullin VG, Taft RJ. Estimating the burden and economic impact of pediatric genetic disease. Genet Med. 2018;21(8).

5. Marshall DA, Benchimol EI, MacKenzie A, Duque DR, MacDonald KV., Hartley T, et al. Direct health-care costs for children diagnosed with genetic diseases are significantly higher than for children with other chronic diseases. Genet Med. 2019 May 1;21(5):1049–57. [PubMed: 30245512]

6. Chang JC, Mandell DS, Knight AM. High Health Care Utilization Preceding Diagnosis of Systemic Lupus Erythematosus in Youth. Arthritis Care Res. 2018 Sep 1;70(9):1303–11.

7. Dragojlovic N, van Karnebeek CDM, Ghani A, Genereaux D, Kim E, Birch P, et al. The cost trajectory of the diagnostic care pathway for children with suspected genetic disorders. Genet Med. 2020 Feb 1;22(2):292–300. [PubMed: 31462755]

8. Berry JG, Poduri A, Bonkowsky JL, Zhou J, Graham DA, Welch C, et al. Trends in resource utilization by children with neurological impairment in the United States inpatient health care system: A repeat cross-sectional study. PLoS Med. 2012 Jan;9(1).

9. Gjorgioski S, Halliday J, Riley M, Amor DJ, Delatycki MB, Bankier A. Genetics and pediatric hospital admissions, 1985 to 2017. Genet Med 2020 2211 [Internet]. 2020 Jun 19 [cited 2021 Jul 12];22(11):1777–85. Available from: https://www.nature.com/articles/s41436-020-0871-9 [PubMed: 32555541]

10. Foreman AKM, Lee K, Evans JP. The NCGENES project: exploring the new world of genome sequencing. N C Med J. 2013 Nov 1;74(6):500–4. [PubMed: 24316776]

11. Carolina Data Warehouse for Health [Internet]. [cited 2021 Apr 8]. Available from: https://tracs.unc.edu/index.php/services/informatics-and-data-science/cdw-h

12. Kharrazi H, Chi W, Chang HY, Richards TM, Gallagher JM, Knudson SM, et al. Comparing Population-based Risk-stratification Model Performance Using Demographic, Diagnosis and Medication Data Extracted from Outpatient Electronic Health Records Versus Administrative Claims. Med Care. 2017;55(8):789–96. [PubMed: 28598890]

13. Suther S, Kiros GE. Barriers to the use of genetic testing: A study of racial and ethnic disparities. Genet Med. 2009 Sep;11(9):655–62. [PubMed: 19752639]

14. Underhill ML, Jones T, Habin K. Disparities in cancer genetic risk assessment and testing. Oncol Nurs Forum. 2016 Jul 1;43(4):519–23. [PubMed: 27314195]

15. Carroll NM, Blum-Barnett E, Madrid SD, Jonas C, Janes K, Alvarado M, et al. Demographic differences in the utilization of clinical and direct-to-consumer genetic testing. In: Journal of Genetic Counseling. John Wiley and Sons Inc.; 2020. p. 634–43.

16. Noll AJ Parekh P, Zhou M, Weber TK, Ahnen D, Wu XC, et al. Barriers to Lynch Syndrome Testing and Preoperative Result Availability in Early-onset Colorectal Cancer: A National Physician Survey Study. Clin Transl Gastroenterol. 2018 Sep 1;9(9).

17. Moreno-De-Luca D, Kavanaugh BC, Best CR, Sheinkopf SJ, Phornphutkul C, Morrow EM. Clinical genetic testing in autism spectrum disorder in a large community-based population sample. Vol. 77, JAMA Psychiatry. American Medical Association; 2020. p. 979–81. [PubMed: 32401282]

18. Bleiker EMA, Esplen MJ, Meiser B, Petersen HV, Patenaude AF. 100 years lynch syndrome: What have we learned about psychosocial issues? Vol. 12, Familial Cancer. Fam Cancer; 2013. p. 325–39. [PubMed: 23670341]

19. Ropka ME, Wenzel J, Phillips EK, Siadaty M, Philbrick JT. Uptake rates for breast cancer genetic testing: A systematic review. Vol. 15, Cancer Epidemiology Biomarkers and Prevention. Cancer Epidemiol Biomarkers Prev; 2006. p. 840–55.

20. Finlay E, Stopfer JE, Burlingame E, Evans KG, Nathanson KL, Weber BL, et al. Factors determining dissemination of results and uptake of genetic testing in families with known BRCA1/2 mutations. Genet Test. 2008 Mar 1;12(1):81–91. [PubMed: 18373407]

21. McGuinness JE, Trivedi MS, Silverman T, Marte A, Mata J, Kukafka R, et al. Uptake of genetic testing for germline BRCA1/2 pathogenic variants in a predominantly Hispanic population. Cancer Genet. 2019 Jun 1;235–236:72–6.

22. Allen CG, Roberts M, Guan Y. Exploring predictors of genetic counseling and testing for hereditary breast and ovarian cancer: Findings from the 2015 U.S. national health interview survey. J Pers Med. 2019 Jun 1;9(2).

Table 1.

Descriptive characteristics of NCGENES cohort (encounter-level)

	Inpatient hospitalizations (N=166)		ED encounters (N=48)		Outpatient encounters (N=2378)	
	N	%	N	%	N	%
Age at visit						
<1 year	19	(11)	0	(0)	118	(5)
1–4 years	101	(61)	33	(69)	1048	(44)
5–9 years	25	(15)	10	(21)	676	(28)
10–14 years	17	(10)	5	(10)	412	(17)
15–<18 years	4	(2)	0	(0)	124	(5)
Sex						
Male	65	(39)	19	(40)	1009	(42)
Female	101	(61)	29	(60)	1369	(58)
Race						
White	106	(64)	29	(60)	1573	(66)
Black	3	(2)	7	(15)	113	(5)
Asian	24	(14)	7	(15)	203	(8)
Other	29	(17)	5	(10)	399	(17)
Unknown	4	(2)	0	(0)	95	(4)
Ethnicity						
Hispanic	25	(15)	5	(10)	263	(11)
Non-Hispanic	136	(82)	42	(88)	2023	(85)
Unknown	5	(3)	1	(2)	92	(4)

Table 2.

Frequency of genetic disease (GD) codes among encounters of the NCGENES cohort

	Inpatient	ED encounters	Outpatient encounters	*P*
Total number of encounters	166	48	2378	
Encounters with 1 *definite*, *probable*, or *possible* GD code	153 (92%)	36 (75%)	1492 (63%)	<0.0001
Encounters with highest specificity of GD code being *definite*	39 (23%)	4 (8%)	203 (9%)	<0.0001
Average number of *definite* GD codes per encounter	1.3	1.3	1.1	
Average number of *probable* GD codes per encounter	0.8	1.0	0.2	
Average number of *possible* GD codes per encounter	2.5	0.8	0.8	
Encounters with highest specificity of GD code being *probable*	63 (38%)	15 (31%)	385 (16%)	<0.0001
Average number of *probable* GD codes per encounter	1.2	1.1	1.2	
Average number of *possible* GD codes per encounter	4.4	1.3	1.4	
Encounters with highest specificity of GD code being *possible*	51 (31%)	17 (35%)	904 (38%)	0.16
Average number of *possible* GD codes per encounter	3.9	1.3	1.7	
Encounters with *other* GD codes	7 (4%)	1 (2%)	0 (0%)	<0.0001
Encounters with *no* GD codes	6 (4%)	11 (23%)	888 (37%)	<0.0001

Note. All percentages were calculated based on total number of inpatient/ED/outpatient encounters. *P*-values are based on chi-square tests.

Table 3.

Key additional p*ossible* genetic disease (GD) diagnoses for pediatric patients (ages day 29 to >18 years) among NCGENES cohort, by ICD9/10

	ICD9 clusters		
Diseases of the Circulatory System		**Diseases of the Digestive System**	
42789	Other Specified Cardiac Dysrhythmias	5363	Gastroparesis
Diseases of the Nervous System and Sense Organs		**Mental Disorders**	
32751	Periodic Limb Movement Disorder	3154	Developmental Coordination Disorder
3589	Myoneural Disorders, Unspecified	3158	Other Specified Delay in Development
3899	Unspecified Hearing Loss	3159	Unspecified Delay in Development
36021	Progressive High (Degenerative) Myopia	31531	Expressive Language Disorder
		31539	Other Developmental Speech or Language Disorder
Factors Influencing Health Status			
V440	Tracheostomy Status		
V441	Gastrostomy Status	**Symptoms, Signs, and Illness**	
V444	Status of Other Artificial Opening of Gastrointestinal Tract	78039	Other Convulsions
V4611	Dependence on Respirator, Status	7812	Abnormality of Gait
V4613	Failure to Wean from Mechanical Ventilation	7813	Lack of Coordination
V8279	Other Genetic Screening	7833	Feeding Difficulties and Mismanagement
V8551	Body Mass Index, Pediatric, Less Than 5th Percentile for Age	78321	Loss of Weight
		78322	Underweight
		78340	Lack of Normal Physiological Development, Unspecified
Diseases of the Respiratory System		78341	Failure to Thrive in Childhood
51883	Chronic Respiratory Failure	78342	Delayed Milestones
51884	Acute and Chronic Respiratory Failure	78343	Short Stature
		78459	Other Speech Disturbance

	ICD10 clusters		
Diseases of the Circulatory System		**Diseases of the Digestive System**	
I459	Conduction Disorder, Unspecified	K3184	Gastroparesis
I7389	Other Specified Peripheral Vascular Diseases		
		Diseases of the Eye and Adnexa	
Diseases of the Ear and Mastoid Process		H47619	Cortical Blindness, Unspecified Side of Brain
H9190	Unspecified Hearing Loss, Unspecified Ear	H5589	Other Irregular Eye Movements
H9193	Hearing Loss, Bilateral		
		Diseases of the Musculoskeletal System and Connective Tissue	
Diseases of the Genitourinary System		M359	Systemic Involvement of Connective Tissue, Unspecified
N319	Neuromuscular Dysfunction of Bladder, Unspecified	M6281	Muscle Weakness
		M6289	Other Specified Disorders of Muscle
Diseases of the Respiratory System			
J9610	Chronic Respiratory Failure, Unspecified Whether with Hypoxia or Hypercapnia	**Symptoms, Signs, and Abnormal Clinical and Lab Findings**	
		R251	Tremor, Unspecified

Diseases of the Nervous System

G248	Other Dystonia
G249	Dystonia, Unspecified
G253	Myoclonus
G40812	Lennox-Gastaut Syndrome, Not Intractable, Without Status Epilepticus
G40813	Lennox-Gastaut Syndrome, Intractable, With Status Epilepticus
G40814	Lennox-Gastaut Syndrome, Intractable, Without Status Epilepticus
G40909	Epilepsy, Unspecified, Not Intractable, Without Status Epilepticus
G6289	Other Specified Polyneuropathy
G629	Polyneuropathy, Unspecified
G63	Polyneuropathy Associated with Underlying Disease
G709	Myoneural Disorder, Unspecified
G909	Disorder of The Autonomic Nervous System, Unspecified
G9340	Encephalopathy, Unspecified
G9349	Other Encephalopathy
G808	Other Cerebral Palsy

Endocrine, Nutritional, and Metabolic Diseases

E162	Hypoglycemia, Unspecified
E230	Hypopituitarism
E43	Unspecified Severe Protein-Calorie Malnutrition

Mental, Behavioral, and Neurodevelopmental Disorders

F800	Phonological Disorder
F802	Mixed Receptive-Expressive Language Disorder
F82	Specific Developmental Disorder of Motor Function

R258	Other Abnormal Involuntary Movements
R259	Unspecified Abnormal Involuntary Movements
R262	Difficulty in Walking, Not Elsewhere Classified
R2689	Other Abnormalities of Gait and Mobility
R269	Unspecified Abnormalities of Gait and Mobility
R270	Ataxia, Unspecified
R278	Other Lack of Coordination
R279	Unspecified Lack of Coordination
R293	Abnormal Posture
R531	Weakness
R569	Unspecified Convulsions
R6250	Unspecified Lack of Expected Normal Physiological Development in Childhood
R6251	Failure to Thrive (Child)
R6252	Short Stature (Child)
R29898	Other Symptoms and Signs Involving the Musculoskeletal System
R203	Hyperesthesia
R208	Other Disturbances of Skin Sensation

Factors Influencing Health Status

Z6851	Body Mass Index (BMI) Pediatric, <5th Percentile for Age
Z930	Tracheostomy Status
Z931	Gastrostomy Status
Z934	Other Artificial Openings of Gastrointestinal Tract Status
Z9911	Dependence on Respirator (Ventilator) Status