American Journal of Medicine Open 10 (2023) 100038

Contents lists available at ScienceDirect

American Journal of Medicine Open

journal homepage: www.elsevier.com/locate/ajmo

Clinical Research Study

Prevalence and Clinical Burden of Idiopathic Dilated Cardiomyopathy in the United States ☆,☆☆

Yaa Ababio [a], Scott P. Kelly [a], Franca S Angeli [c], Joanne Berghout [b], Kui Huang [a], Kathy Liu [c], Sara Burns [d], Cynthia Senerchia [e], Rob Moccia [b,1], Gabriel C. Brooks [b,1,*]

[a] Pfizer, Inc, New York, N.Y.
[b] Pfizer, Inc, Boston, Mass
[c] Pfizer, Inc, Collegeville, PA
[d] Panalgo, Boston, Mass
[e] Optum Digital Research Network, Boston, Mass

ARTICLE INFO

Keywords:
Cardiomyopathy
Dilated cardiomyopathy
Epidemiology
Genetic testing
Heart failure
Heritable
Idiopathic
Prevalence

ABSTRACT

Background: Dilated cardiomyopathy (DCM) contributes significantly to heart failure prevalence, yet supporting epidemiologic data is sparse. This study sought to estimate the period prevalence of DCM and the proportion of idiopathic DCM in the United States using a large, diverse electronic health records (EHR) database.

Methods: This retrospective, observational study included 56,812,806 deidentified patients in Optum EHR with visits between 2017 and 2019. Suspected DCM cases were identified using ICD-10 coding. Deidentified clinical notes from 1000 randomly selected cases were manually reviewed to determine the diagnosis of DCM and estimate the proportion of idiopathic DCM. The period prevalence and clinical burden of DCM and idiopathic DCM were estimated.

Results: Manual clinical review demonstrated that our definition had a positive predictive value of 92.5% for DCM, with 46.3% estimated as the idiopathic DCM proportion. The estimated period prevalence of DCM between 2017 and 2019 was 118.33 per 100,000. Prevalence increased for adults ≥65 years of age, males, and African Americans. Extrapolation to the 2019 US population led to an overall estimated burden of roughly 388,350 patients. Adjusting for the proportion of cases with idiopathic DCM yielded an idiopathic DCM prevalence of 59.23 per 100,000 and a burden of 194,385 patients. Evidence of clinical genetic testing in this population was scarce, with less than 0.43% of DCM cases reporting a testing code.

Conclusions: This study establishes a conservative period prevalence for DCM and idiopathic DCM and demonstrates very low molecular genetic testing for DCM. These findings suggest that the clinical burden of genetic DCM may be underestimated.

Introduction

Dilated cardiomyopathy (DCM) is a cardiac muscle disorder characterized by left ventricular dilation and systolic dysfunction in the absence of abnormal loading conditions such as hypertension and valve disease or significant coronary artery disease.[1,2] DCM may be acquired (toxic exposure, peripartum cardiomyopathy, infectious/inflammatory) or idiopathic, with no discernable clinical or lifestyle factors sufficient to cause the pathology.[3] Idiopathic DCM often presents with familial patterns and in research cohorts with clinical sequencing, a causal genetic mutation can be identified in 26% of patients.[4] Over 250 different genes have been proposed as playing a role in DCM, with current evidence supporting a strong causal role for 19 genes.[5] Despite remarkable progress in understanding its genetic basis, the nomenclature of idiopathic DCM is still generally used when acquired causes of DCM have been ruled out and thus idiopathic DCM continues to be defined as a diagnosis of exclusion.[6,7] This term has been imperfectly deployed as a reference for, and estimate of, familial/genetic DCM.

☆ Disclaimer: This manuscript is not presently under consideration for publication elsewhere, nor have the contents of this manuscript been copyrighted or published previously.

☆☆ Funding/Support: This study was sponsored by Pfizer, and publication-related fees were met by Pfizer. Non-Pfizer authors received no funding.

* Corresponding author: Gabriel C Brooks, MD MAS FACC, Executive Director—Head of Rare Cardiology, Worldwide Research, Development and Medical, Rare Disease Research Unit, 1 Portland Street, 2nd floor, Cambridge, Mass.
E-mail address: Gabriel.Brooks@pfizer.com (G.C. Brooks).

[1] Dr. Rob Moccia and Dr. Gabriel Brooks contributed equally to the conception, execution and drafting of this research.

https://doi.org/10.1016/j.ajmo.2023.100038
Received 23 November 2022; Accepted 19 February 2023
Available online 25 February 2023

Accurate prevalence estimates of DCM worldwide and in the United States are limited, particularly for idiopathic DCM. Deriving more precise estimates of the true prevalence of idiopathic DCM is challenging because the diagnosis requires the completion of a full workup and subsequently requires that the diagnosis is accurately and consistently reported.

A prevalence of 1:2500 is commonly cited to estimate the prevalence of idiopathic DCM. This estimate relies on a population-based epidemiologic study conducted in Olmstead County, Minnesota, between 1975 and 1984 and is based on a mere 45 DCM cases in a limited demographic band.[8,9] Subsequent analyses, extrapolating from population studies for hypertrophic cardiomyopathy (HCM) using modern imaging methods and an assumed DCM to HCM ratio of 2:1, suggest a prevalence closer to 1:250.[6,10] Other estimates, leveraging the proportion of idiopathic DCM as a fraction of heart failure diagnoses, yield a prevalence estimate closer to 1:400.[6,8]

Improved estimates for DCM and the proportion of idiopathic cases are needed, similar to what has been conducted for HCM,[11] to reconcile these large differences in prior estimates, reflect current diagnostic practices, and ensure the inclusion of racial minorities, particularly African Americans who have increased risk for idiopathic DCM.[12] To estimate the current prevalence of DCM more precisely, a large, racially diverse real-world electronic health records (EHR) database including a vast array of diagnostic, laboratory, medication, and surgical procedure data was interrogated using International Classification of Diseases (ICD)–based algorithms to identify probable DCM cases. This study used 3 case definitions, ranging from a conservative base population using only patients with billing codes specific to DCM, to a broader definition that attempted to capture DCM patients coded under more general terms. While this study leveraged structured ICD codes, particularly I42.0 ("I42.0: Dilated Cardiomyopathy") to identify patients with DCM, the billing code itself does not differentiate between idiopathic and acquired etiologies for the observed pathology. To derive an estimate of the proportion of idiopathic DCM, manual review of deidentified notes and unstructured data was used for 1000 randomly selected patients satisfying our primary definition (definition 1) to evaluate evidence of acquired DCM and/or ischemic heart disease and eliminate these competing etiologies.

Methods

This study analyzed retrospectively collected data from the Optum deidentified Electronic Health Records database, a large, racially diverse population with clinical encounter data for over 101 million patients who belong to a provider network of 700 hospitals and 7000 clinics across all 50 states in the United States as of September 30, 2019. The data are sourced from multispecialty medical groups, integrated delivery networks, and hospital chains and then normalized, validated, and aggregated into a structured data format before use. The Optum EHR database integrates structured EHR data with prescribed medications and practice management data and captures a comprehensive collection of demographic, clinical, operational, and financial information from a patient's office visit and/or hospital stay (e.g. inpatient and outpatient data), including but not limited to the following: vital signs and other biometric measures, laboratory results, outpatient prescriptions written, inpatient medications administered, procedures performed, and inpatient/outpatient diagnoses. Data were deidentified in accordance with the Health Insurance Portability and Accountability Act (HIPAA). The study was conducted in accordance with the Code of Ethics of the World Medical Association (the Declaration of Helsinki).

Multiple case definitions were evaluated to identify patients with DCM during the study period, defined as January 1, 2017, to December 31, 2019. DCM case definition 1 (primary) was used to estimate a conservative base population of DCM in the United States. Cases were selected for inclusion if they had at least 1 inpatient DCM-specific ICD-10 code (I42.0: Dilated Cardiomyopathy) or 2 outpatient codes (I42.0) at least 30 days apart during the study period. Additional broader case definitions of DCM were performed to evaluate sensitivity of results to the choice of criteria. In DCM case definition 2, only 1 I42.0 ICD-10 code of any kind was required for inclusion. For DCM case definition 3, we included all patients identified by DCM case definition 1, plus those individuals lacking a DCM I42.0 ICD-10 code but who had codes present for other cardiomyopathies I42.8 and cardiomyopathies unspecified I42.9 as well as those with codes consistent with systolic heart failure/heart failure with reduced ejection fraction. Consistent with the prior case definitions, in this broader definition, patients were excluded if they had codes consistent with acquired cardiomyopathy including ischemic cardiomyopathy or coronary artery disease (full details in Supplemental Table S1).

Medical comorbidities were identified using ICD-10 coding during the baseline period (defined as the year prior to DCM diagnosis). Genetic testing utilization was determined based on Current Procedural Terminology (CPT) codes during the baseline period or during the follow-up period, defined as any time on or after the DCM diagnosis date until the end of available data (3/31/2020).

To validate DCM diagnoses in patients meeting the primary definition, a random sample of 1000 suspected DCM cases with at least one DCM-related clinical note available within the Optum database were selected for evaluation. The notes had sections of text containing DCM-related terms of interest extracted, deidentified, and loaded into an annotation tool by the Optum natural language processing team. Each term of interest was programmatically highlighted and labeled with the categorization predefined by the specification (e.g. genetic screening terms, family history, cardiomyopathy, arrhythmia, etc.). The annotation system was designed to allow a reviewer to read the notes and confirm that the term was relevant and correctly labeled, to extend the annotation to include more words to support the context of the term, or to remove the tag from the text if it related to a negation, for example, "no evidence of cardiomyopathy." Prior to annotation, 2 external clinically trained registered nurses (each with critical care experience) and 2 external physicians reviewed and agreed on the definitions of the terms. Each of the nurses independently reviewed notes from 500 patients and at the completion of their work the output was reviewed and curated by an Optum MD, who also read the notes and adjudicated the content annotated by the nurses. High-probability idiopathic DCM cases were determined by manual clinician and nursing review of natural language processing extracted unstructured data and clinical notes based on key search terms related to diagnosis of DCM (Supplemental Table S5). In order to discriminate between an acquired and idiopathic etiology of DCM, each of the 1000 patients was assigned one of the following categories: (a) idiopathic DCM; (b) not idiopathic DCM (ischemic DCM, acquired DCM, other DCM); (c) not DCM (e.g. HCM, muscular dystrophy, Takotsubo cardiomyopathy); or (d) undetermined, for mixed ischemic and nonischemic or not enough information in the notes to determine. Categorization of a case as idiopathic was made only when there was no evidence of a possible alternative etiology, yielding a conservative estimate.

Period prevalence of DCM and idiopathic DCM were defined as the number of suspected cases divided by the total study population at risk. The positive predictive value (PPV) of definition 1 for DCM was calculated as the number of clinical notes–validated DCM cases relative to the 1000 randomly selected diagnosis code–derived suspected DCM cases. The idiopathic DCM patient proportion was calculated as the total number of patients with validated idiopathic DCM divided by the 1000 randomly selected suspected DCM cases. Overall and stratum-specific prevalence across age groups (<12 years, 12-17, 18-29, 30-49, 50-64, 65+ years), gender (male, female), and race (Caucasian, African American, Asian, Other/Unknown) were calculated using the PPV of DCM and the idiopathic DCM proportion derived from the clinical notes validation process to yield overall and stratum-specific prevalence rates. Confidence intervals for a binomial proportion were calculated for each using the Clopper–Pearson exact method. To estimate the clinical bur-

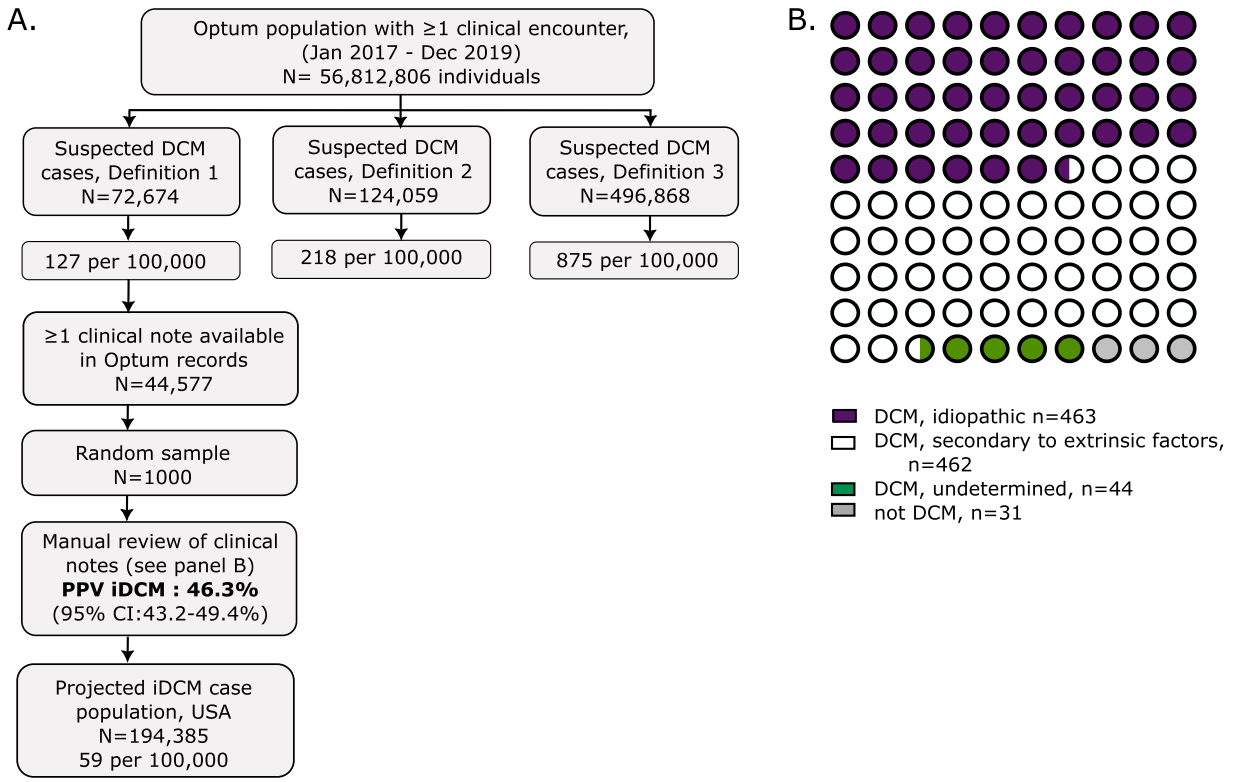

Figure 1. Estimation and validation of idiopathic DCM prevalence in the United States. (**A**) Flowchart diagram of study approach and results. Patients with a suspected dilated cardiomyopathy (DCM) were identified using three definitions: definition 1 required that the patient received at least 1 ICD-10 inpatient code for I42.0, or 2 I42.0 outpatient codes at least 30 days apart; definition 2 required only 1 instance of I42.0 in the clinical record; and definition 3 was constructed to capture potentially underdiagnosed patients with EHR features of nonischemic cardiomyopathy who may or may not have received an I42.0 code; see details in methods and Supplementary Table S1. Definition 1 was validated by review of 1000 randomly selected cases with clinical notes, and the positive predictive value (PPV) of this definition to identify idiopathic DCM was determined at 46.3%. Using this PPV and projecting the case prevalence of suspected DCM to the unweighted US Census population yielded an idiopathic DCM estimate of 194,385 people, or 59 cases per 100,000 people. (**B**) Manual review of 1000 randomly selected clinical notes from cases meeting DCM definition 1 found that $n = 463$ (46.3%) cases were idiopathic. Each circle represents 10 individuals, and the proportion of each color reflects the proportion of cases assigned to the indicated clinical category by Optum-contracted clinicians. DCM with suspected secondary or extrinsic factors was detected in the notes of 462 cases.

den within the US population, validated DCM and idiopathic DCM case counts from the Optum data were extrapolated to the entire US population using unweighted (crude) and weighted methodology that standardized counts based on key demographic group strata (age, sex, race). The crude extrapolations were calculated as US Census population counts as of 2019 divided by the Optum population counts among subjects at risk (denominator), multiplied by the Optum DCM case count. The methodology for the weighted extrapolations has been described in detail elsewhere,[13] but in brief, the Optum data were weighted by the demographic group stratum of interest (proportion of the US Census population counts[14] for a specific stratum divided by the equivalent proportion of the strata in the Optum data). All results are presented descriptively; no formal statistical comparisons are made. All statistical analyses were performed in SAS, version 9.4 (SAS Institute, Cary, NC, USA).

Results

Baseline Characteristics

During the study period, 56,812,806 patients had a clinical encounter. Based on the most conservative DCM case definition, DCM case definition 1 (primary), 72,674 patients were classified as DCM cases (Figure 1), of which 14,509 died during the study period (20.0%). The majority of DCM cases identified using the primary case definition were at least 50 years of age (84%), male (63%), and Caucasian (75%) (Table 1). The less strict DCM-specific case definition (DCM case definition 2) identified 124,059 DCM cases, while the broadest DCM case definition 3, which included additional possible DCM cases lacking the I42.0 billing code in their EHR, yielded 496,868 cases (Supplemental Table S4).

Clinical Validation of DCM

Among the 1000 patients with suspected DCM satisfying definition 1 that were randomly selected for manual clinical review and validation, 925 were validated as DCM (463 idiopathic, 462 not idiopathic), 31 patients did not have DCM, and 44 were undetermined. The PPV of case definition 1 for the presence of notes-validated DCM was 92.5% (95% confidence interval [95% CI]: 90.7%–94.1%). Idiopathic DCM was confirmed in 46.3% (95% CI: 43.2%-49.4%) of selected patients based on physician diagnostic notes in clinical charts (Figure 1).

Estimation of DCM Prevalence and Extrapolated Clinical Burden

Applying the PPV of definition 1 to the total patients satisfying the primary definition criteria, the estimated prevalence of clinical DCM during the study period was 118.33 per 100,000. Stratum-specific DCM prevalence rates were highest in individuals ≥ 65 years (337.4 per 100,000), males (167.93 per 100,000), and African Americans (213.54

Table 1
Period Prevalence Rates of Dilated Cardiomyopathy and Extrapolated Clinical Burden

Category	Period Prevalence Rates of DCM in Optum EHR Database from 2017 to 2019									
	Optum DCM Cases,* Prevalence, and Population at Risk					Extrapolated Clinical Burden and Population at Risk				
	DCM Cases	Optum Population at Risk	Rate per 100,000 (DCM)	Rate per 100,000 (Validated DCM)	Rate per 100,000 (Idiopathic DCM)	US Population at Risk	Weighted: US Census Projected Validated DCM	Weighted: US Census Projected Idiopathic DCM	Unweighted: US Census Projected Validated DCM	Unweighted: US Census Projected Idiopathic DCM
	Total									
All cases	72,674	56,812,806	127.92	118.33	59.23	328,239,523	388,388	194,404	388,350	194,385
	Age Group									
≤11	192	7,467,499	2.57	2.38	1.19	48,022,779	1142	572	1026	513
12-17	96	3,391,527	2.83	2.62	1.31	25,016,371	655	328	513	257
18-29	1253	8,719,802	14.37	13.29	6.65	53,728,222	7141	3575	6696	3352
30-49	9488	14,559,350	65.17	60.28	30.17	84,488,200	50,930	25,492	50,701	25,378
50-64	23,604	12,181,780	193.76	179.23	89.71	62,925,688	112,783	56,453	126,133	63,135
≥65	38,041	10,428,926	364.76	337.4	168.88	54,058,263	182,396	91,297	203,281	101,750
Missing	0	63,922								
	Gender									
Male	45,991	25,332,651	181.55	167.93	84.06	161,657,324	271,475	135,884	245,763	123,014
Female	26,634	31,362,183	84.92	78.95	39.32	166,582,199	130,858	65,500	142,325	71,239
Unknown	49	117,972								
	Race									
Caucasian	54,150	37,980,135	142.57	131.88	66.01	250,522,190	330,392	165,375	289,363	144,838
African American	13,635	5,906,321	230.85							
213.54	106.88	44,075,086	94,118	47,110	72,862	36,471				
Asian	611	1,262,259	48.41	44.78	22.41	19,504,862	8733	4371	3265	1634
Other/Unknown	4278	11,664,091	36.68	33.93	16.98	14,137,385	4796	2401	22,860	11,443

* DCM cases were obtained using DCM case definition 1, which was defined as at least 1 inpatient ICD-10 code, or 2 outpatient ICD-10 codes at least 30 days apart for I42.0.Abbreviations: EHR, Electronic Health Record; Rate, Period prevalence rate; PPV, Positive Predictive value; US, United States. Notes: Rates calculated as period prevalence (per 100,000) from 2017 to 2019. US Census counts are extrapolated to the 2019 US population.

per 100,000) (Table 1). DCM case definition 2 estimated a period prevalence of 201.98 per 100,000, while DCM case definition 3 estimated a prevalence of 808.98 per 100,000 (Supplemental Table S3). For both alternate case definitions, similar age groups, gender, and racial patterns were observed as DCM case definition 1 (primary).

When extrapolated to the 2019 US population, based on DCM case definition 1 (primary), the projected true DCM population in the United States was 388,350 patients based on the unweighted methodology, while the weighted methodology yielded a similar projection of 388,388 patients. Patients ages 65 and older (52.3%) and males (63.3%) comprised the largest proportion of their respective age and sex categories (Table 1).

Estimation of Idiopathic DCM Prevalence and Extrapolated Clinical Burden

The proportion of patients with idiopathic DCM (46.3%, Figure 1B) was multiplied by the estimated prevalence of DCM to yield an estimate of the idiopathic DCM prevalence. This estimated prevalence for idiopathic DCM across strata was 59.23 per 100,000 (Table 1), which varied by demographics. By age group, idiopathic DCM prevalence was highest in patients ages ≥65 years (168.88 per 100,000), compared to patients ages 50-64 (89.71 per 100,000), and substantially lower in younger age groups. When examining prevalence by gender and race, rates were higher in males (84.06 per 100,000) compared to females (39.32 per 100,000), and African Americans (106.88 per 100,000) compared to Caucasians (66.01 per 100,000), with Asian groups having the lowest population prevalence (22.41 per 100,000). After extrapolating the estimated idiopathic DCM prevalence rate to 2019 US Census data, 194,385 patients (unweighted methodology) were projected to have idiopathic DCM in the United States (Figure 1 and Table 1). Demographic proportions of the projected idiopathic DCM population were extrapolated in an identical manner to those of the projected DCM population.

Comorbidities and Genetic Testing Utilization

The most prevalent comorbidities among cases using DCM case definition 1 (primary) included other forms of heart disease (72.31%), hypertension (55.05%), and diabetes mellitus (27.39%) (Table 2). Similar comorbidities were observed for the other DCM case definition populations; however, the proportion of comorbidities was reduced among the DCM case definition 3 population. When genetic testing utilization was examined, only 0.16% of DCM case definition 1 (primary) patients had a relevant CPT code during the baseline period, and only 0.27% had one during follow-up (Supplemental Tables S2 and S3) for a total of less than 0.43% of patients overall. Similar testing rates were exhibited for DCM case definitions 2 and 3.

Sensitivity Analyses

In sensitivity analyses of both DCM and idiopathic DCM extrapolations to the US Census, results were robust and there were no material differences in age and sex after weighting and standardization; however, extrapolated results for the clinical burden did highlight differences by race.

Discussion

This retrospective, observational study identified the patient proportion and period prevalence of DCM and idiopathic DCM using Optum Electronic Health Records. Multiple case definitions of DCM were evaluated. The occurrence of diagnosed validated DCM was estimated to be 118.33 per 100,000. When this prevalence rate was extrapolated to the 2019 US Census population, approximately 388,000 patients were identified as having DCM in the United States. The proportion of patients confirmed to have idiopathic DCM after clinical validation of diagnosis

Table 2
Comorbidities Among Dilated Cardiomyopathy Cases in the Optum EHR Database, 2017-2019

Comorbidities Among DCM Cases in Optum EHR Database from 2017 to 2019			
	DCM Case Definition 1 (primary)*	DCM Case Definition 2†	DCM Case Definition 3††
DCM cases, *N* (%)	72,674	124,059	496,868
Diabetes mellitus, *n* (%)	19,903 (27.39%)	33,888 (27.32%)	65,051 (13.09%)
Hypertension, *n* (%)	40,005 (55.05%)	69,798 (56.26%)	136,462 (27.46%)
Obesity, *n* (%)	6911 (9.51%)	12,460 (10.04%)	20,295 (4.08%)
Other forms of heart disease, *n* (%)	52,547 (72.31%)	95,490 (76.97%)	165,125 (33.23%)
Cardiac conduction disease/arrhythmias, *n* (%)	35,893 (49.39%)	64,864 (52.28%)	104,852 (21.10%)
Other heart disorders in diseases classified elsewhere, *n* (%)	14 (0.02%)	24 (0.02%)	41 (0.01%)
Heart failure, *n* (%)	187 (0.26%)	328 (0.26%)	261 (0.05%)
Heart transplant, *n* (%)	292 (0.40%)	581 (0.47%)	659 (0.13%)
Ischemic heart diseases, *n* (%)	27,346 (37.63%)	50,075 (40.36%)	76,991 (15.50%)
Myocardial infarction, *n* (%)	5221 (7.18%)	10,961 (8.84%)	8685 (1.75%)
Cerebrovascular diseases, *n* (%)	6376 (8.77%)	11,439 (9.22%)	20,485 (4.12%)
Stroke, *n* (%)	2700 (3.72%)	5120 (4.13%)	8397 (1.69%)
Renal impairment, *n* (%)	10,415 (14.33%)	17,490 (14.10%)	26,990 (5.43%)

Note: The baseline time period to determine comorbidities was defined as the period 1 year prior to DCM diagnosis.

* DCM Case Definition 1 (primary): At least 1 inpatient ICD-10 code, or 2 outpatient ICD-10 codes at least 30 days apart for I42.0.

† DCM Case Definition 2: Only 1 ICD-10 code for I42.0 of any kind is required.

†† DCM Case Definition 3: Individuals meeting DCM Case Definition 1 supplemented by individuals without a DCM I42.0 ICD code but exhibiting codes for other cardiomyopathies I42.8 and/or cardiomyopathies unspecified I42.9, as well as those with codes consistent with systolic heart failure/heart failure with reduced ejection fraction. Patients were excluded if they had codes consistent with ischemic cardiomyopathy and/or coronary artery disease (at least 1 inpatient ICD-10 code or 2 outpatient ICD-10 codes at least 30 days apart, or only 1 NDC were required).

and ruling out of other causes was 46.3%. Using these estimates, the prevalence of idiopathic DCM in the United States was estimated to be 59.23 per 100,000, which, when extrapolated to the US Census population, projected 194,385 patients with idiopathic DCM during this period. Importantly, there was little evidence of genetic/molecular testing either prior to or after a patient's diagnosis of DCM.

Idiopathic DCM including genetic DCM is a differential diagnosis that requires consideration of patient medical history and careful clinical assessment. Traditionally, DCM cases with no known external cause are classified as idiopathic before genetic causes are considered.[6] Challenges to deriving accurate prevalence estimates from diagnostic coding data for idiopathic DCM include imprecise sensitivity due to inaccurate provider coding and incomplete diagnostic workups, as well as inclusion of nonischemic, acquired, and other cause DCM diagnoses that may still be billed under I42.0. The use of chart reviews to estimate a PPV for idiopathic DCM from all diagnosed patients allows for a conservative (lower bound) estimate of the prevalence of idiopathic, as this method only adjusts for imperfect specificity in ICD-10 coding and not the imperfect sensitivity.

Despite adopting a conservative approach to estimating the prevalence of idiopathic DCM, our estimates were higher than those observed by previous population-based studies. The population-based study in Olmstead County, Minnesota, previously reported an idiopathic DCM prevalence rate of 36.5 per 100,000.[8] However, the Olmstead County study may not be generalizable to the current US population due to its limited sample size (45 patients), homogenous population (predominately Caucasian), and changes in disease epidemiology and clinical care since its publication. Similar limitations arise when comparing our results with a survey-based study from Japan reporting a prevalence estimate of 14 per 100,000.[15] A recent population-based cohort study leveraging the UK Clinical Practice Research Datalink research database estimated the 2018 prevalence of idiopathic DCM in patients ≥36 years old to be 2.7 cases per 10,000 in females and 5.9 cases per 10,000 in males,[16] which compare favorably to our estimates of 7.9 per 10,000 and 16.8 per 10,000, respectively. Differences in prevalence estimates from the UK research database may stem from differences in definitions from code lists, data extraction methodology from clinical records, and population included. Importantly, a common theme is the acknowledgment that cardiomyopathies are underdiagnosed and that diagnoses are often delayed.[6,16] In our analysis, the extrapolation of DCM rates in a care-seeking population to the broader US population who may or may not seek care likely does not overestimate disease burden, due to well-established disparities in health care access within the country.

The distribution of comorbid conditions in this study differed from previous estimates. Rates of diabetes and hypertension (17.9% and 51.7%, respectively) were lower in the UK Clinical Practice Research Datalink cohort,[16] relative to the DCM population in the US Optum EHR (27.39 and 55.05%, respectively). These differences are likely due to variations in the geographic distribution of these comorbidities, as the United States has one of the highest prevalences of diabetes and metabolic syndrome.[17] Rates of stroke were slightly higher in DCM patients enrolled in the EURObservational Research Programme, a European prospective multinational registry, when compared to this Optum cohort (4.5% vs 3.72%, respectively), which may also reflect geographic and lifestyle variation in primary risk factors for atherosclerotic vascular disease.[18]

Ours is the first study of this scale to explore current EHR-derived practices in genetic testing within these populations and strongly suggests that the burden of genetic DCM is underappreciated as genetic testing is underutilized when viewed at the national scale. Genetic testing rates for hereditary cardiomyopathy ranged from 0.38% to 0.43% across the three DCM case definitions under evaluation (Supplemental Table S2). Current clinical guidelines suggest genetic testing in cases of familial cardiomyopathy for the most affected family member and cascade testing for family members at risk for pathogenic variants.[9] Our estimates suggest a discrepancy in DCM-related genetic testing uptake that warrants further investigation.

Strengths and Limitations

Strengths of this study include access to a large, clinically rich, heterogenous population that enabled the ability to assess disease prevalence by age, sex, and race and allowed for a more robust, representative source population generalizable to the national scale rather than relying on previously conducted single-center DCM studies. Furthermore, clinical notes review of unstructured data among a large subsample of the DCM cases enabled estimates of idiopathic DCM, as ICD-10 coding

for DCM lacks granularity and there is no information regarding the reported cause of disease.

This study possesses some limitations that are important to note. First, there are inherent limitations due to the nature of the study design and use of retrospective observational data. Utilizing Optum EHR as the sole database for analysis excludes any patients who belong to hospitals and clinics outside of this provider network. Subclinical DCM cases may not be adequately represented in this study population, except for DCM case definition 3, as those who seek clinical care are more likely to have advanced disease. Patients reflected in this database may also be more likely to have insurance coverage when compared to the general US population. Furthermore, for the denominator, only patients with any clinical encounter between 2017 and 2019 were included, and the population at risk could be higher or lower depending on alternate denominator criteria, such as if all patients in the Optum database were included during all years or only those with yearly clinical encounters in each and every year of the study period. Validation of DCM diagnosis was limited to a subset of patients who had clinical notes associated with their records. The inclusion and quality of clinical notes were dependent on the healthcare organization that the patient belonged to, disease severity, and whether a procedure was performed. Thus, there may be biases present (i.e. location, level of care received) that drive differences in which DCM patient subpopulations possessed clinical notes. Consequently, the proportion of idiopathic DCM generated from this sample may not adequately reflect the true proportion of idiopathic DCM in the full DCM patient population. As this study used Optum EHR data, individuals with and without health insurance are included; however, this cohort selects for individuals with DCM who are more often in need of medical care, thereby limiting the generalizability to individuals with DCM who do not seek care which may result in some degree of underestimation of the DCM prevalence rate. Lastly, genetic testing rates are complex and difficult to ascertain due to an evolving understanding of the genetic basis of DCM and rapidly updating guidelines in the field as new practice recommendations are made. The unique subset of CPT codes included in this study may not adequately reflect the full breadth of DCM-related genetic testing codes, and the imprecision of these in current billing practice also means that we cannot rule out that genetic testing (when present) was performed for other comorbid pathologies. Additionally, genetic testing performed under industry-sponsored programs for patients who meet specific eligibility criteria may not be captured by CPT coding.

Conclusion

In summary, we believe that the conservative population-based prevalence estimate generated from this study serves as a floor and plausibly underestimates the true prevalence of idiopathic DCM. Additionally, genetic testing rates among DCM cases are extremely low, suggesting that the true prevalence of genetic DCM in the United States is not well captured. Future studies should consider clinically validating multiple case definitions of DCM and utilizing genetic testing to better understand the full clinical burden of potentially genetic DCM. Genetic biobank databases and natural history studies may also help expand the knowledge base surrounding the mutations that underlie genetic DCM and ultimately aid in identifying additional DCM-related genetic testing codes in claims data.

Authors' Contributions

Study concept and design: All authors. Acquisition: Burns, Liu, Senerchia. Analysis and interpretation of data: All authors. Drafting of the manuscript: All authors. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Burns, Liu. Administrative, technical, or material support: All authors. Study supervision: Angeli, Brooks, Kelly, Moccia.

Declaration of Competing Interest

Pfizer is a biopharma company engaged in developing novel drugs to treat diseases with high unmet medical need. Pfizer is engaged in the development of cardiovascular drugs including a marketed a drug for the treatment of wild-type or hereditary transthyretin amyloidosis in adult patients with cardiomyopathy. Pfizer may target other cardiomyopathies including heritable dilated cardiomyopathies for therapeutic drug development.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ajmo.2023.100038.

References

1. Mazzarotto F, Tayal U, Buchan RJ, et al. Reevaluating the genetic contribution of monogenic dilated cardiomyopathy. *Circulation*. 2020;141(5):387–398.
2. Jordan E, Peterson L, Ai T, et al. Evidence-based assessment of genes in dilated cardiomyopathy. *Circulation*. 2021;144(1):7–19.
3. Schultheiss H-P, Fairweather D, Caforio ALP, et al. Dilated cardiomyopathy. *Nat Rev Dis Primers [Internet]*. 2019;5(1). Available at. http://www.nature.com/articles/s41572-019-0084-1. Accessed November 14, 2021.
4. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. Prevalence of hypertrophic cardiomyopathy in a general population of young adults: echocardiographic analysis of 4111 subjects in the CARDIA study. *Circulation*. 1995;92(4):785–789.
5. US Census BureauNational population by characteristics: 2010. *Census.gov*. 2019. Available at. https://www.census.gov/data/datasets/time-series/demo/popest/2010s-national-detail.html . Accessed January 20, 2022.
6. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*. 2013;10(9):531–547.
7. Miura K. Epidemiology of idiopathic cardiomyopathy in Japan: results from a nationwide survey. *Heart*. 2002;87(2):126–130.
8. Brownrigg JR, Leo V, Rose J, et al. Epidemiology of cardiomyopathies and incident heart failure in a population-based cohort study. *Heart*. 2021 doi:10.1136/heartjnl-2021-320181.
9. Codd MB, Sugrue DD, Gersh BJ, Melton LJ. Epidemiology of idiopathic dilated and hypertrophic cardiomyopathy. A population-based study in Olmsted County, Minnesota, 1975-1984. *Circulation*. 1989;80(3):564–572.
10. Lin X, Xu Y, Pan X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci Rep*. 2020;10:14790.
11. Maron MS, Hellawell JL, Lucove JC, Farzaneh-Far R, Olivotto I. Occurrence of Clinically Diagnosed Hypertrophic Cardiomyopathy in the United States. *Am J Cardiol*. 2016;117(10):1651–1654. doi:10.1016/j.amjcard.2016.02.044.
12. Charron P, Elliott PM, Gimeno JR, et al. The Cardiomyopathy Registry of the EURObservational Research Programme of the European Society of Cardiology: baseline data and contemporary management of adult patients with cardiomyopathies. *Eur Heart J.*. 2018;39(20):1784–1793.
13. Hershberger RE, Givertz MM, Ho CY, et al. Genetic evaluation of cardiomyopathy: a clinical practice resource of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2018;20(9):899–909.
14. Coughlin SS, Gottdiener JS, Baughman KL, et al. Black-white differences in mortality in idiopathic dilated cardiomyopathy: the Washington, DC, dilated cardiomyopathy study. *J Natl Med Assoc*. 1994;86(8):583–591.
15. Wasser T, Wu B, Yčas J, Tunceli O. Applying weighting methodologies to a commercial database to Project US Census Demographic Data. *Am J Accountable Care*. 2015;3(3). Available at:. https://www.ajmc.com/view/applying-weighting-methodologies-to-a-commercial-database-to-project-us-census-demographic-data. Accessed March 15, 2022.
16. Elliott P, Andersson B, Arbustini E, et al. Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J*. 2007;29(2):270–276.
17. Pinto YM, Elliott PM, Arbustini E, et al. Proposal for a revised definition of dilated cardiomyopathy, hypokinetic non-dilated cardiomyopathy, and its implications for clinical practice: a position statement of the ESC Working Group on Myocardial and Pericardial Diseases. *Eur Heart J*. 2016;37(23):1850–1858.
18. Sinagra G, Elliott PM, Merlo M. Dilated cardiomyopathy: so many cardiomyopathies!. *Eur Heart J.*. 2020;41(39):3784–3786.