ORIGINAL RESEARCH—CLINICAL

Clinical Phenotypes May be Able to Identify Populations With Nonalcoholic Fatty Liver-Spectrum Disease

Bridget M. Geyer,[1] Feier Chang,[1] Peng Wu,[1] Benjamin A. Goldstein,[1] Kara Wegermann,[1] Sunny L. Chung,[2] Matthew Phelan,[3] Joseph Wawrzynski,[1] Jacqueline B. Henson,[1] Howard Lee,[4] Phil Ambery,[5] Cynthia A. Moylan,[1] and Neha Pagidipati[1]

[1]Duke University, Durham, North Carolina; [2]Yale School of Medicine Section of Digestive Diseases, New Haven, Connecticut; [3]Target RWE, Durham, North Carolina; [4]Section of Gastroenterology and Hepatology, Department of Medicine, Baylor College of Medicine, Houston, Texas; and [5]Cardiovascular Renal and Metabolic Late-Stage Clinical Research, AstraZeneca, Gothenburg, Sweden

BACKGROUND AND AIMS: Despite causing significant morbidity and mortality, nonalcoholic fatty liver disease (NAFLD) is underdiagnosed. Clinical indices developed to identify hepatic steatosis are often used by providers but their potential for use at the population level remains unexplored. We assessed clinical phenotypes for their ability to identify potential patients with NAFLD and nonalcoholic steatohepatitis (NASH) in the electronic health record. **METHODS:** We conducted a single-center retrospective cohort study of adult patients from January 1, 2016, to December 31, 2022. We developed 4 phenotypes: clinical NAFLD (C-NAFLD), clinical NASH (C-NASH), NAFLD with diagnosis (D-NAFLD) and NASH with diagnosis (D-NASH) and compared characteristics across them to identify differences between patients with and without International Classification of Diseases diagnoses. **RESULTS:** Each of the (C) phenotypes identified a cohort of patients who had clinical evidence suggestive of disease without a documented diagnosis. Black patients were overrepresented in the (C) relative to (D) groups (C-NAFLD 24.3% vs D-NAFLD 21.2%; C-NASH 28.5% vs D-NASH 14.0%). Patients with D-NASH were more likely to be prescribed medications that may be effective in treating NAFLD-NASH spectrum disease, ie, glucagon-like peptide 1 receptor agonists (C-NASH 5.0% vs D-NASH 16.7%, $P < .001$). Fewer patients with D-NASH had cardiovascular (C-NASH 58.0% vs D-NASH 46.3%, $P < .001$) and heart failure (C-NASH 33.9% vs D-NASH 24.8%, $P < .001$) hospitalizations than those with C-NASH. **CONCLUSION:** Noninvasive clinical indices may improve identification of patients with or at risk for NAFLD-NASH at the population level. Systematic differences between populations with and without International Classification of Diseases diagnoses of NAFLD-spectrum disease suggest disparities in the application of screening and diagnostic procedures.

Keywords: Fatty Liver Disease; MASLD/MASH/NAFLD/NASH; Electronic Health Record; Population Health; Disparities

Introduction

Nonalcoholic fatty liver disease (NAFLD), now known as metabolic dysfunction–associated steatotic liver disease, comprises a spectrum of pathophysiologic states ranging from macrovesicular steatosis without hepatocellular injury to metabolic dysfunction–associated steatohepatitis (formerly nonalcoholic steatohepatitis or NASH) and eventual cirrhosis. NAFLD and NASH terminology will be used in this paper as the study, which relies on phenotypes based on these terms, was conducted under the former nomenclature. Associated with increased hepatic, cardiovascular (CV), and overall mortality, NAFLD is a leading cause of morbidity among US adults.[1,2] Timely diagnosis is important, as NAFLD-NASH confers an increased risk for both non–liver-related CV events as well as liver-related events such as the development of cirrhosis, hepatocellular carcinoma, and liver transplantation for end-stage liver disease.[2] Further, specific drug regimens may be beneficial for certain NAFLD patients with biopsy-proven NASH, obesity, or type 2 diabetes mellitus[3] and several targeted therapies for NASH are in late-stage clinical trials and will likely be available in the next few years.[4,5]

Despite this, NAFLD and NASH remain significantly underdiagnosed,[5,6] possibly due to largely asymptomatic disease in the early stages, lack of consensus on the need for screening, poor correlation between levels of liver aminotransferase elevations and liver damage from NAFLD, and/or an invasive gold standard for the diagnosis of NASH in

Abbreviations used in this paper: **ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; C-NAFLD, clinical nonalcoholic fatty liver disease; C-NASH, clinical nonalcoholic steatohepatitis; CV, cardiovascular; D-NAFLD, nonalcoholic fatty liver disease with diagnosis; D-NASH, nonalcoholic steatohepatitis with diagnosis; DUHS, Duke University Health System; ED, emergency department; EHR, electronic health record; FIB-4, Fibrosis-4; FSI, framingham steatosis index; GLP-1, glucagon-like peptide 1; HDL, high-density lipoprotein; HF, heart failure; ICD, International Classification of Diseases; IQR, interquartile range; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; PCP, primary care provider.**

Most current article

2772-5723
https://doi.org/10.1016/j.gastha.2024.100611

liver biopsy. As a result, there is a need to examine evidence related to barriers to diagnosis and to develop noninvasive, accurate, and easy-to-use tools to identify NAFLD and NASH at the population level.

Certain clinical scoring indices, such as the Framingham Steatosis Index (FSI) and the Fibrosis-4 score (FIB-4), have high validity when used to diagnose NAFLD and advanced hepatic fibrosis, respectively, for individuals in the clinical setting.[7,8] Whether the populations of patients identified by these tools (FSI and FIB-4) differ from those with International Classification of Diseases (ICD) codes for NAFLD or NASH remains unclear.

The present study is a retrospective analysis of electronic health record (EHR) data in which we sought to address these questions by using liver enzyme elevations, FSI, FIB-4, and ICD diagnosis codes to develop clinical phenotypes for NAFLD and NASH, which we employed to understand the characteristics of populations with likely or diagnosed NAFLD-NASH. As a secondary aim, we compared populations with NAFLD or NASH based on ICD codes to populations with markers of disease but without a diagnosis in the EHR, as we hypothesized that there are patients that remain underdiagnosed with these diseases who would likely benefit from being connected to appropriate care. This study expands the application of validated noninvasive clinical markers of NAFLD-NASH from the individual level to the population level. Such an approach may be able to broadly identify populations that are underdiagnosed, are at risk for NAFLD and/or NASH, may need improved lifestyle and medical management, and/or are eligible for clinical trials.

Methods

Study Population

We conducted a single center retrospective cohort study of adult patients using the Duke University Health System (DUHS), a midsize health system in North Carolina. EHR data was abstracted from our research ready datamart which includes information regarding patient encounters, demographics, diagnoses, prescriptions, and laboratory tests.[9]

The study was centered around an index date of January 1, 2021 and includes a prospective window from January 1, 2021, to December 31, 2022 (outcomes data) and a retrospective window from January 1, 2016, to December 31, 2020 (all other data). To target a study population receiving primary care at DUHS, patients were eligible for inclusion if they had a visit with a primary care provider (PCP) at DUHS in 2019 (the retrospective year nearest the index date without confounding due to the COVID-19 pandemic). Patients with 2 separate alanine aminotransferase (ALT) levels checked more than 6 but less than 24 months apart during the retrospective window were included. Patients were excluded if they had a body mass index (BMI) less than 25 kg/m^2 or another form of chronic liver disease or excessive alcohol use as identified by ICD code. Full inclusion and exclusion criteria are outlined in Tables A1–3.

We developed 4 computable phenotypes for the purpose of this study: clinical NAFLD (C-NAFLD), NAFLD with diagnosis (D-NAFLD), clinical NASH (C-NASH), and NASH with diagnosis (D-NASH). The (C) phenotypes were defined by clinical indices of disease, as described below, while the (D) phenotypes were derived from the presence of the ICD-10 code in the EHR. The inclusion criteria for each phenotype are shown in Tables A2 and 3.

C-NAFLD was defined as 2 ALT measurements taken 6–24 months apart that exceeded the given threshold ($\geq$40 for men, $\geq$31 for women) plus an FSI score $\geq$ 32. The ALT cutoffs are based on criteria established by Husain et al. that have been shown to have high specificity (92.4%) and positive predictive value (80.8%) in identifying NAFLD based on EHR data.[10] The FSI uses age, sex, BMI, triglyceride levels, diagnoses of hypertension and diabetes, and aspartate aminotransferase (AST)/ALT ratio to identify individuals with hepatic steatosis.[10,11] FSI $\geq$ 32 has been shown to be a reliable predictor of the presence of NAFLD.[7] C-NASH relied on the same ALT cutoffs and a Fibrosis-4 (FIB-4) score $\geq$ 2.67, calculated using the patient's age, AST, ALT, and platelet levels.[2,12,13] The same ALT cutoffs are used for both C-NAFLD and C-NASH because in both cases, the criteria is meant to indicate NAFLD (ie, it is not meant to identify NASH in the NASH population). This is so that the FIB-4 is selectively applied to a population with likely NAFLD, as identifying the presence of NAFLD with concomitant advanced fibrosis as indicated by FIB-4 is a validated approach in predicting NASH and has been recommended as part of the clinical care pathway.[13] The American Gastroenterological Association, the American Association for the Study of Liver Diseases, the European Association for the Study of the Liver, the European Association for the Study of Diabetes, the European Association for the Study of Obesity, and the Japanese Society of Gastroenterology guidelines all include the use of FIB-4 in guiding clinical decision-making surrounding the diagnosis and/or treatment of NAFLD-NASH.[14–19]

The (C) and (D) phenotypes are mutually exclusive within a given condition, ie, within NAFLD and within NASH. In our analysis, we exclude patients from the clinical phenotype if they carry a diagnosis as defined by diagnostic phenotype criteria. By contrast, the NAFLD and NASH categories are not mutually exclusive. For example, a patient could be in included in both C-NAFLD and C-NASH, but not C-NASH and D-NASH.

We abstracted laboratory data and vital signs for individual patients from the date of their second ALT measurement or the closest available date within 6 months. We included documented comorbidities and medication prescriptions from before the second ALT measurement. Race in the EHR reflects either self-reported or administrator-assigned race. Outcome data during the 2-year prospective period used billing codes to identify patients hospitalized for a CV and/or heart failure (HF) event (Table A4).

Statistical Analysis

Chi-squared tests were used to compare the proportion of patients meeting criteria within each of the categorical phenotypes, such as that pertaining to demographics, comorbidities, medication use, and outcomes. Unpaired t-tests were used to compare continuous variables such as age, lab values, and vital signs. The Wilcoxon rank-sum test was used to test for significance in evaluating the relative number of health-care visits; these data are presented in the form of median and interquartile range due to its non-normal distribution. *P* values less than .05 were considered statistically significant. Analyses were conducted using R Bioconductor version 4.1.3.[20] This study was

Table 1. Demographic Characteristics by Phenotype

Demographic characteristic	C-NAFLD	D-NAFLD	C-NASH	D-NASH
Basic characteristics				
N (total study population)	11,011	6898	1252	642
Age (y) (median, interquartile range)	56 (46, 66)	57 (47, 67)	71[a] (62, 78)	64[a] (53, 72)
Male sex	5405[a] (49.1%)	2740[a] (39.7%)	570 (45.4%)	277 (43.1%)
25 ≤ BMI <30	1837 (16.7%)	1804 (26.2%)	585 (46.7%)	184 (28.7%)
30 ≤ BMI <40	6634 (60.2%)	3718 (53.9%)	560 (44.7%)	367 (57.2%)
BMI ≥ 40	2540 (23.1%)	1376 (19.9%)	107 (8.5%)	91 (14.2%)
Race/Ethnicity				
Hispanic	550 (5.0%)	495 (7.2%)	34 (2.6%)	43 (6.7%)
Non-Hispanic Asian	308 (2.8%)	230 (3.3%)	13 (1.0%)	17 (2.6%)
Non-Hispanic White	7168 (65.1%)	4541 (65.8%)	816 (65.2%)	475 (74.0%)
Non-Hispanic Black	2680 (24.3%)	1465 (21.2%)	357 (28.5%)	90 (14.0%)
Other	305 (2.8%)	167 (2.4%)	32 (2.6%)	17 (2.6%)

P values are not available for individual BMI, or racial/ethnic categories.
[a]$P < .05$.

approved by the Duke University Institutional Review Board. It was funded in part by research grants from AstraZeneca pharmaceuticals.

Results

Study Population

A total of 204,393 patients who had a visit with DUHS primary care in 2019 had at least 2 eligible ALT measurements between January 1, 2016, and December 31, 2020. Of these patients, 64,370 were excluded on the basis of BMI or concomitant diagnoses. After applying inclusion and exclusion criteria for the phenotypes, the sample size for each was as follows: 11,011 (5.4%) for C-NAFLD; 6898 (3.4%) for D-NAFLD; 1252 (0.6%) for C-NASH; and 4730 (0.3%) for D-NASH (see Figure A1).

C-NAFLD vs D-NAFLD

Regarding demographic characteristics (Table 1), relative to those with D-NAFLD, patients with C-NAFLD were more frequently male and had a BMI in the obese category; median age was similar. There was a greater proportion of Black patients in the C-NAFLD phenotype (C-NAFLD 24.3% vs D-NAFLD 21.2%) and a greater proportion of Hispanic patients in the D-NAFLD phenotype (C-NAFLD 5.0% vs D-NAFLD 7.2%) (Figure 1).

As displayed in Table 2, the median number of annual health-care visits was statistically slightly greater among patients with D-NAFLD. The distribution of comorbidities varied, and differences in vital signs and laboratory values were all either statistically or clinically insignificant.

Medication use differed between NAFLD phenotypes. As shown in Table 3, patients with C-NAFLD were more likely than those with D-NAFLD to have been prescribed several metabolic syndrome-modifying medications including statins, angiotensin converting enzyme inhibitors, and angiotensin receptor blockers. By contrast, beta blockers, insulin, and metformin were more commonly prescribed in the D-NAFLD phenotype. Of note, glucagon-like peptide 1 (GLP-1) receptor agonists were more commonly prescribed in patients with D-NAFLD (C-NAFLD 11.8% vs D-NAFLD 13.9%, $P < .001$).

As seen in Table 4 and Figure 2, CV hospitalizations (C-NAFLD 30.3% vs D-NAFLD 34.1%, $P < .001$) were more common in D-NAFLD, while rates of HF hospitalizations were similar (C-NAFLD 13.9% vs D-NAFLD 14.5%, $P = .23$).

C-NASH vs D-NASH

The median age of the C-NASH population (Table 1) was greater than the D-NASH population. The C-NASH phenotype had a greater proportion of Black patients (C-NASH 28.5% vs D-NASH 14.0%) and the D-NASH population had a greater proportion of Hispanic patients (C-NASH 2.6% vs D-NASH 6.7%) and White patients (C-NASH 65.2% vs D-NASH 74.0%) (Figure 1).

Patients with C-NASH had more health-care encounters in both the inpatient/emergency department and outpatient settings (Table 2). Most diagnosed comorbidities (Table 2) were more common in C-NASH, with the exception of diabetes, which was more common in D-NASH (C-NASH 42.3% vs D-NASH 55.6%, $P < .001$), and hypertension, peripheral artery disease, and HF with preserved ejection fraction, which were diagnosed at similar rates across phenotypes. In terms of lab values, AST and ALT were higher among those with C-NASH.

Regarding medication use (Table 3), there was not a significant difference the prescription rates of angiotensin converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers, alphaglucosidase inhibitors, insulin, or thiazolidinediones. Statins were more prevalent in C-NASH while dipeptidyl peptidase-4 inhibitors, sodium-glucose cotransporter-2 inhibitors, and metformin were more common in D-NASH. Once again, GLP-1 receptor agonists were more common in the diagnostic phenotype (C-NASH 5.0% vs D-NASH 16.7%, $P < .001$).

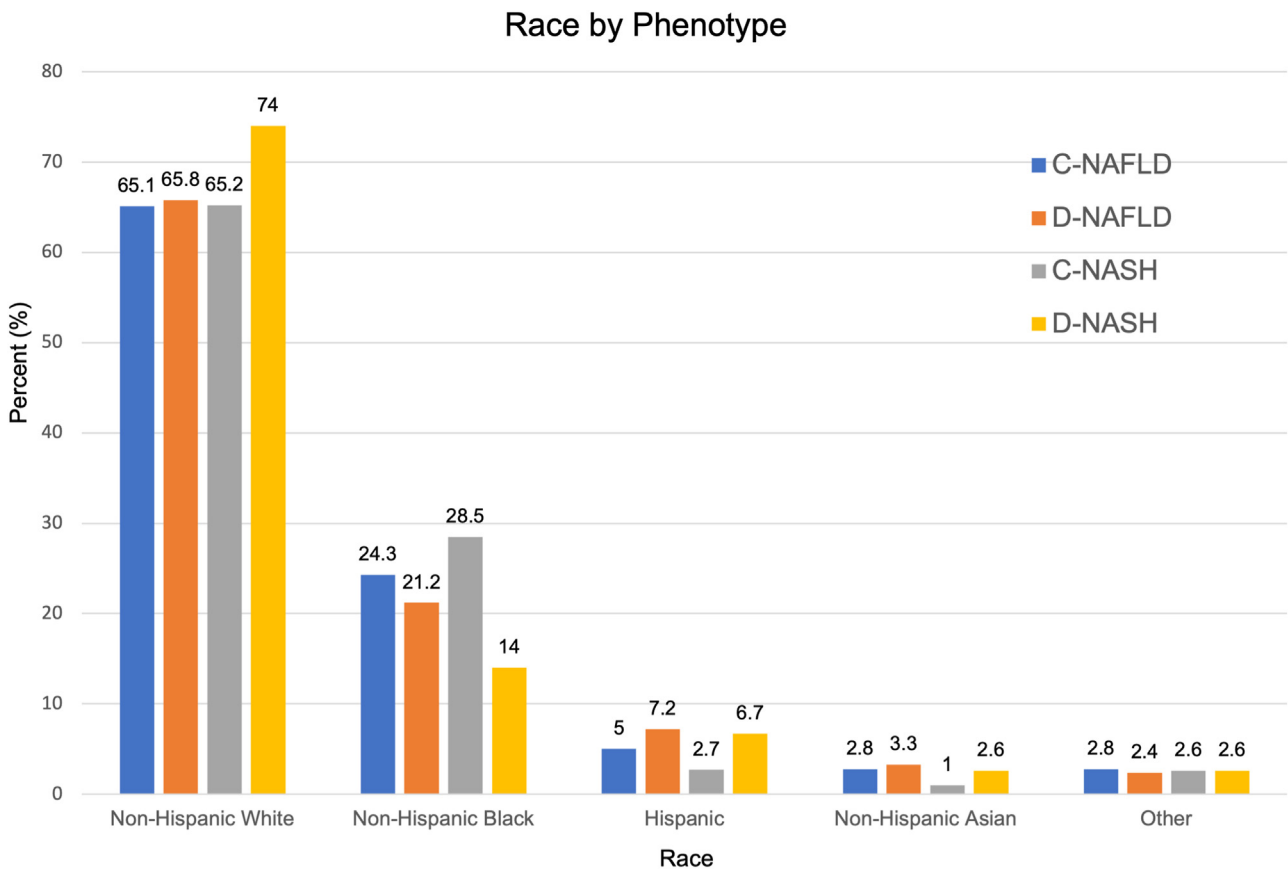

Figure 1. Race by phenotype. Non-Hispanic White patients and Hispanic patients are over-represented in the D-NAFLD/NASH groups while Black patients are over-represented in the C-NAFLD/NASH groups.

Outcomes (Table 4, Figure 2) indicate that patients with C-NASH were significantly more likely to experience CV (C-NASH 58.0% vs D-NASH 46.3%, $P < .001$) and HF (C-NASH 33.9% vs D-NASH 24.8%, $P < .001$) hospitalizations than those with D-NASH.

Discussion

More than 30% of the US adult population has NAFLD while 3%–5% have NASH.[21] Despite this, few individuals carry diagnoses, even among those with evidence of disease.[6] Given the American Gastroenterological Association's recently published NAFLD clinical care pathway and the expected Food and Drug Administration approval of medications to treat NASH, developing improved methods to easily identify patients with NAFLD and NASH from the EHR remains a critical need.[13,22] We therefore used EHR data to determine whether computable phenotypes could identify patients at the population level that remain underdiagnosed with NAFLD and NASH. Indeed, our findings suggest that population-level studies and clinical trial recruitment efforts that rely on EHR diagnoses may miss potentially eligible patients in a systematically biased fashion.

We confirmed that even in a large academic medical center, fewer than 5% of patients who met study inclusion criteria carried a diagnosis of NAFLD in the EHR. Nearly double that number of patients met clinical criteria for NAFLD. This was similar for NASH. In this context, we report the following major findings. First, clinical indices of fatty liver disease that have been applied to individuals may have a role in identifying patients with or at risk for NAFLD on the population level. Second, a significant number of patients *who have existing clinical evidence of NAFLD* do not carry a diagnosis despite being seen in primary care and, further, the characteristics of this population of patients differ significantly from those of populations with a diagnosis. Third, patients with D-NASH are more commonly prescribed disease-modifying treatment for NAFLD-spectrum disease (ie, GLP-1 receptor agonists) than those with C-NASH, indicating that diagnosis may be associated with appropriate treatment. And fourth, patients who meet clinical index criteria for NASH but are not diagnosed with disease are hospitalized more frequently than those who have an ICD diagnostic code for NASH, indicating that high-risk patients are potentially being underdiagnosed.

Recent studies have evaluated the utility of various approaches including genome-wide association studies, machine learning, and clinical characteristics such as BMI in identifying NAFLD/NASH based on EHR data.[23] We report an approach for using pre-existing clinical data to identify

Table 2. Health-Care Encounters, Comorbidities, Vital Signs, and Lab Values by Phenotype

Relevant clinical feature	C-NAFLD	D-NAFLD	*P* value	C-NASH	D-NASH	*P* value
Annual health-care encounters (median, interquartile range)						
IP + ED encounters	0 (0, 0.4)	0.2 (0, 0.6)	$P < .001$[a]	0.4 (0, 1)	0.2 (0, 0.6)	$P < .001$[a]
Outpatient visits	7 (4, 14)	10 (6, 18)	$P < .001$[a]	13 (6, 25)	10 (6, 19)	$P < .001$[a]
Comorbidities						
Hypertension	8453 (76.8%)	4904 (71.7%)	$P < .001$	1044 (83.4%)	512 (79.8%)	$P = .058$
Hyperlipidemia	4569 (41.5%)	3118 (45.2%)	$P < .001$	690 (55.1%)	331 (51.6%)	$P = .156$
Diabetes	4553 (41.3%)	3015 (43.7%)	$P < .01$	529 (42.3%)	357 (55.6%)	$P < .001$
Myocardial infarction	593 (5.4%)	350 (5.1%)	$P = .382$	156 (12.5%)	55 (8.6%)	$P < .05$
Coronary artery disease	1529 (13.9%)	1080 (15.7%)	$P < .001$	376 (30.0%)	143 (22.3%)	$P < .001$
Stroke/Transient ischemic attack	755 (6.9%)	522 (7.6%)	$P = .077$	204 (16.3%)	75 (11.7%)	$P < .01$
Peripheral artery disease	423 (3.8%)	292 (4.2%)	$P = .207$	130 (10.4%)	50 (7.8%)	$P = .082$
Cancer	581 (5.3%)	440 (6.4%)	$P < .01$	233 (18.6%)	38 (5.9%)	$P < .001$
Heart failure with reduced ejection fraction	539 (4.9%)	256 (3.7%)	$P < .001$	190 (15.2%)	60 (9.3%)	$P < .001$
Heart failure with preserved ejection fraction	596 (5.4%)	395 (5.7%)	$P = .390$	181 (14.5%)	78 (12.1%)	$P = .138$
Labs and vitals (mean, SD)						
Systolic blood pressure (mmHg)	131 (17)	130 (17)	$P < .01$	127 (23)	127 (19)	$P = .992$
Diastolic blood pressure (mmHg)	79 (11)	77 (11)	$P < .001$	72 (15)	73 (13)	$P = .189$
Total cholesterol (mg/dL)	180 (45)	179 (47)	$P = .39$	162 (48)	166 (45)	$P = .138$
HDL cholesterol (mg/dL)	44 (13)	45 (13)	$P = .289$	48 (18)	43 (14)	$P < .001$
LDL cholesterol (mg/dL)	102 (38)	101 (39)	$P = .052$	86 (38)	92 (36)	$P < .005$
AST (IU/L)	47 (68)	45 (13)	$P < .001$	84 (200)	43 (68)	$P < .001$
ALT (IU/L)	40 (176)	41 (328)	$P = .853$	138 (540)	50 (146)	$P < .001$
Platelet count (10^9/L)	247 (74)	249 (80)	$P < .05$	145 (63)	203 (92)	$P < .001$

ED, emergency department; HDL, high-density lipoprotein; IP, inpatient hospitalizations; LDL, low-density lipoprotein; SD, standard deviation.
[a]Non-normal distribution of data.

populations with a clinical phenotype for NAFLD or NASH which may be more accessible than machine learning methods yet more precise than models that use less clinical information. This represents a potential solution to a major limitation of current EHR-based studies of NAFLD-NASH; that is, these diseases are known to be significantly underdiagnosed.

The present study reveals important demographic and clinical differences in patient populations with clinical fatty liver disease relative to those with diagnosed disease. Given that the number of outpatient health-care visits is similar or higher in (C) groups relative to (D) groups, the lack of diagnosis in (C) patients is not likely to be due to these patients not presenting to care. Further, the ability to identify these patients as meeting clinical criteria per our methodology suggests that clinical evidence pointing to disease is present in these patients, and yet, they are not assigned diagnoses. Importantly, we highlighted the significant lack of NAFLD and NASH diagnoses among Black patients despite meeting clinical risk factors. Our findings that Black populations suffer from the same if not more of the comorbidities seen among NAFLD and NASH patients suggest that despite epidemiological evidence suggesting lower rates of NAFLD among Black patients,[24] providers must make an additional effort to address these health disparities and bias within the health-care system to ensure proper suspicion for disease, diagnosis, and referral for treatment.

Table 3. Medication Use by Phenotype

Medication	C-NAFLD	D-NAFLD	*P* value	C-NASH	D-NASH	*P* Value
Statin	6319 (57.4%)	3694 (53.6%)	$P < .001$	842 (67.3%)	388 (60.4%)	$P < .005$
Angiotensin converting enzyme inhibitor	4214 (38.3%)	2384 (34.6%)	$P < .001$	492 (39.3%)	250 (38.9%)	$P = .92$
Angiotensin II receptor blocker	3333 (30.3%)	1832 (26.6%)	$P < .001$	407 (32.5%)	189 (29.4%)	$P = .191$
Beta blocker	4837 (43.9%)	3414 (49.5%)	$P < .001$	796 (63.6%)	378 (58.9%)	$P = .052$
Alphaglucosidase inhibitors	27 (0.2%)	21 (0.3%)	$P = .550$	2 (0.2%)	2 (0.3%)	$P = .879$
Insulin	2650 (24.1%)	1952 (28.3%)	$P < .001$	511 (40.8%)	273 (42.5%)	$P = .506$
Dipeptidyl Peptidase-4 inhibitor	973 (8.8%)	629 (9.1%)	$P = .538$	118 (9.4%)	85 (13.2%)	$P < .05$
Sodium-glucose Cotransporter-2	760 (6.9%)	514 (7.5%)	$P = .173$	45 (3.6%)	68 (10.6%)	$P < .001$
Glucagon-like peptide 1 receptor agonist	1302 (11.8%)	959 (13.9%)	$P < .001$	62 (5.0%)	107 (16.7%)	$P < .001$
Metformin	3972 (36.1%)	2617 (37.9%)	$P < .05$	340 (27.2%)	260 (40.5%)	$P < .001$
Thiazolidinediones	228 (2.1%)	458 (2.4%)	$P = .169$	22 (1.8%)	18 (2.8%)	$P = .183$

Table 4. Hospitalizations by Phenotype

Hospitalizations	C-NAFLD	D-NAFLD	*P* value	C-NASH	D-NASH	*P* value
Cardiovascular	3330 (30.3%)	2348 (34.1%)	$P < .001$	737 (58.0%)	296 (46.3%)	$P < .001$
Heart failure	1497 (13.9%)	978 (14.5%)	$P = .281$	416 (33.9%)	155 (24.8%)	$P < .001$

Additionally, higher rates of diabetes in the (D) relative to (C) populations could reflect the role of screening guidelines in facilitating diagnosis. Both the European Association for the Study of the Liver and the American Diabetes Association recommend NAFLD/NASH screening among certain patients with diabetes.[25,26] Therefore, rates of screening or index of suspicion for fatty liver disease may be higher for patients with a diagnosis of diabetes relative to other comorbidities. Notably, recent changes in the definition of NAFLD, now metabolic dysfunction–associated steatotic liver disease, recognize that metabolic fatty changes to the liver may co-occur with other chronic liver diseases including alcohol use (MetALD); it will be important to evaluate the impact of these new criteria on diagnostic and classification patterns in the future.

Clinical outcomes differed significantly between C-NASH and D-NASH phenotypes. For NASH patients, there were fewer hospitalizations in the (D) groups relative to the respective (C) groups. While conclusions regarding causality between the presence or absence of a diagnosis and outcomes in terms of hospitalizations cannot be drawn from present data, one theoretical explanation consistent with our findings could be that diagnosis facilitates treatment. This mechanism is supported by medication use data, which show that D-NASH patients are more likely to be prescribed drugs for metabolic syndrome that may have a role in the treatment of NAFLD-NASH spectrum disease, namely, GLP-1 receptor agonists. While phase III trials investigating their efficacy in the treatment of NAFLD-NASH specifically are ongoing, GLP-1 receptor agonists have been reported to

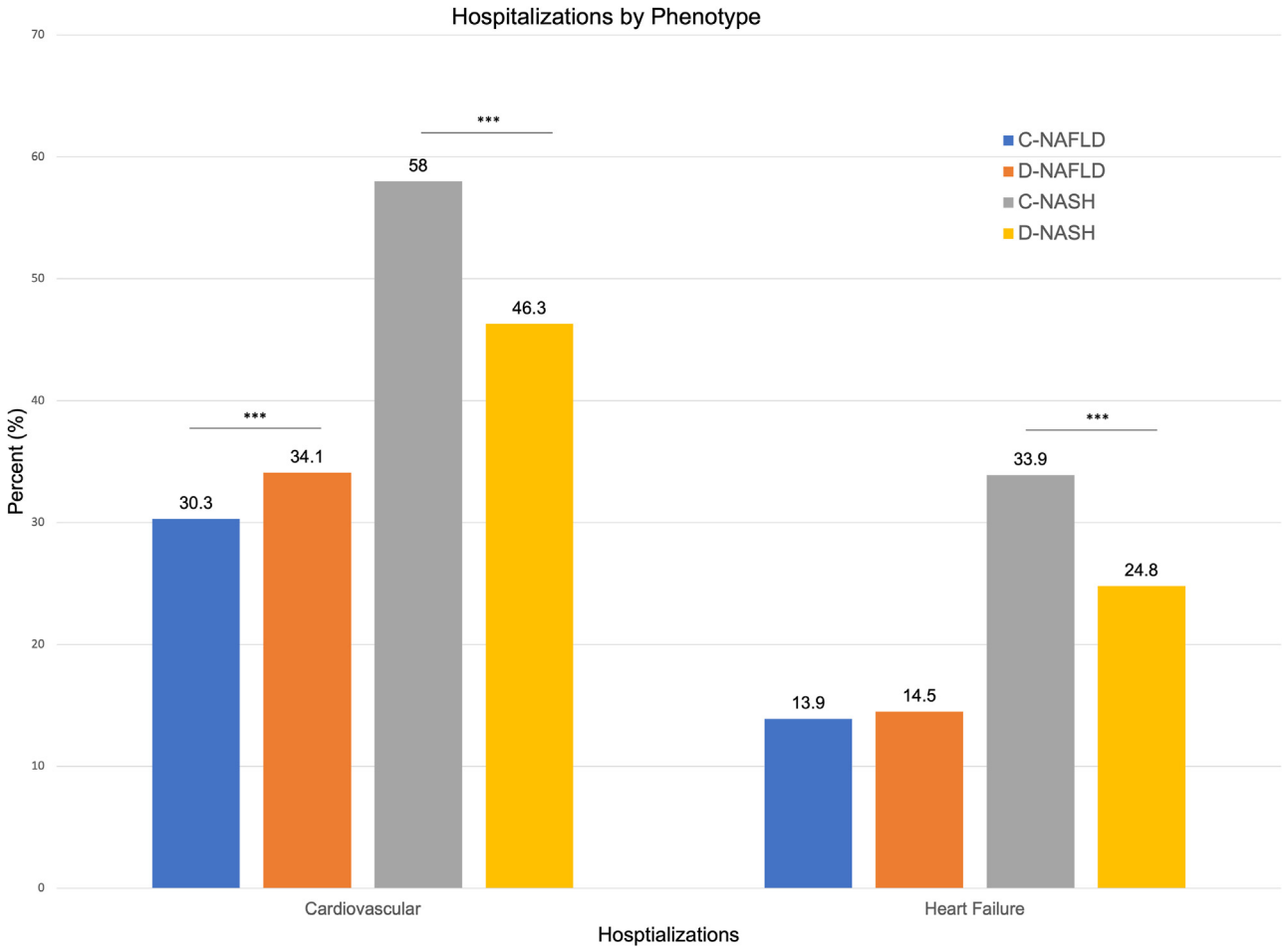

Figure 2. Hospitalizations by phenotype. Significantly fewer D-NASH patients had HF and CV hospitalizations than C-NASH patients.

"have shown a significant impact on body weight and clinical, biochemical and histological markers of fatty liver and fibrosis in patients with NAFLD".[27] Further, in a phase II clinical trial, semaglutide, a GLP-1 receptor agonist, has been shown to increase rates of NASH resolution (59% in the 0.4-mg group vs 17% in the placebo group; $P < .001$).[28] Alternatively, or additionally, those without a diagnosis may be sicker or experience additional negative social determinants of health, leading to poorer outcomes. Either way, improving outcomes and increasing equity in care requires addressing this diagnostic gap. Notably, this pattern of outcomes was not replicated for the NAFLD group. Patterns of comorbidities, laboratory findings, and medication use were more heterogeneous for NAFLD than NASH patients, potentially confounding the effect of diagnosis and management of fatty liver disease on outcomes.

Strengths and Limitations

This study has both methodologic and practical strengths. It has a large sample size with granular clinical data. The phenotypes it employs are easy to compute and therefore can be feasibly replicated. The use of clear diagnosis codes for both inclusion and exclusion criteria enables accurate identification of the target populations and is similarly replicable.

This study also has its limitations. As an EHR-based study, it shares constraints common to all studies that use this approach, including the fact that patients may receive care elsewhere that is not captured for this study.[29] However, the large sample size of the study as well as statistical measures to ensure data robustness help to mitigate such issues.

The primary limitations of this study are related that which renders it necessary; that is, the difficult of accurately identifying and classifying NAFLD-NASH spectrum disease. For instance, the use of exclusion criteria based on BMI and inclusion criteria based on elevated transaminase levels, aimed at increasing the specificity of the phenotypes, limits their sensitivity and prevents capture of certain populations known to have NAFLD-spectrum disease, ie, those without elevations in transaminase levels or BMI. Additionally, the clinical indices used have known limitations, including the inexact positive and negative predictive value of using ALT cutoffs alone as a way of defining NAFLD (ie, the Husain et al. criteria), as well as the imperfect ability of FIB-4 to define NASH, as it is an estimate of fibrosis rather than the inflammation of steatohepatitis, predicts advanced fibrosis at the given cutoff, potentially missing patients with milder fibrosis, and is less accurate at the extremes of age. Further, the phenotypes have their own limitations and could benefit from being refined. For example, in future studies, including the FSI in the C-NASH phenotype definition would likely serve to increase its specificity in identifying patients with fibrosis secondary to metabolic disease specifically. The current study does not demonstrate hepatic steatosis by imaging or biopsy and therefore cannot claim that individuals in the C-NAFLD and C-NASH phenotypes certainly have disease. However, as discussed, the measures used to define these populations clinically have been validated as having high predictive value for the presence of disease.[7,13] Likewise, there are limitations associated with reliance on ICD-10 codes in identifying populations with disease; it is possible that patients in the diagnostic phenotype do not truly have clinical NAFLD-NASH disease but rather have been inappropriately diagnosed, or that nonspecific diagnosis codes captured in the diagnostic phenotype reflect alternative etiologies of liver disease. However, the study exclusion criteria based on alternative diagnoses helps to mitigate this risk. Further, diagnostic phenotypes are used as a comparison group in understanding characteristics associated with being assigned a diagnosis rather than a gold standard representing the true presence of disease; thus, the analysis is not necessarily invalidated by a potential degree of inaccuracy with respect to the application of ICD-10 diagnosis codes.

Future Research

Additional research is required to elucidate the means by which certain populations are more likely to obtain a diagnosis (ie, studies of screening and work-up practices) to identify opportunities for intervention in increasing rates of diagnosis. Additional validation of the utility of clinical scoring systems such as the FSI and FIB-4 and the clinical phenotypes used in this study when applied at the population level will also be important in considering the potential for the application of such an approach more broadly.

Conclusion

In conclusion, clinical indices may be able to identify patients with or at risk for NAFLD-NASH at the population level, which may prove useful both in clinical care and in identifying patients for research. Further, significant differences exist between patients with clinical evidence of disease but without a diagnosis and patients with a diagnosis. These differences point to the need for physicians to maintain a high index of suspicion for NAFLD-NASH, particularly among patients from populations that may experience systematic underdiagnosis, including Black patients. Further research is necessary to validate the use of clinical markers of disease at the population level, explore mechanisms that explain the observed differences between populations with and without a diagnosis, and investigate how outcomes of patients with and without diagnoses vary according to the demographic and clinical differences identified in this study.

Supplementary Materials

Material associated with this article can be found in the online version at doi: https://doi.org/10.1016/j.gastha.2024.100611.

References

1. Cotter TG, Rinella M. Nonalcoholic fatty liver disease 2020: the state of the disease. Gastroenterology 2020; 158(7):1851–1864.
2. Younossi Z, Henry L. Contribution of alcoholic and nonalcoholic fatty liver disease to the burden of liver-related morbidity and mortality. Gastroenterology 2016; 150(8):1778–1785.
3. Dufour J-F, Anstee QM, Bugianesi E, et al. Current therapies and new developments in NASH. Gut 2022; 71(10):2123–2134.
4. Rinella ME, Sanyal AJ. Management of NAFLD: a stage-based approach. Nat Rev Gastroenterol Hepatol 2016; 13(4):196–205.
5. Alexander M, Loomis AK, Fairburn-Beech J, et al. Real-world data reveal a diagnostic gap in non-alcoholic fatty liver disease. BMC Med 2018;16(1):130.
6. Nielsen EM, Anderson KP, Marsden J, et al. Nonalcoholic fatty liver disease underdiagnosis in primary care: what are we missing? J Gen Intern Med 2022; 37(10):2587–2590.
7. Motamed N, Nikkhah M, Karbalaie Niya MH, et al. The ability of the framingham steatosis index (FSI) to predict non-alcoholic fatty liver disease (NAFLD): a cohort study. Clin Res Hepatol Gastroenterol 2021;45(6):101567.
8. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. J Hepatol 2018; 68(2):305–315.
9. Hurst JH, Liu Y, Maxson PJ, et al. Development of an electronic health records datamart to support clinical and population health research. J Clin Transl Sci 2020; 5(1):e13.
10. Husain N, Blais P, Kramer J, et al. Nonalcoholic fatty liver disease (NAFLD) in the veterans administration population: development and validation of an algorithm for NAFLD using automated data. Aliment Pharmacol Ther 2014;40(8):949–954.
11. Long MT, Pedley A, Colantonio LD, et al. Development and validation of the framingham steatosis index to identify persons with hepatic steatosis. Clin Gastroenterol Hepatol 2016;14(8):1172–1180.e2.
12. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006; 43(6):1317–1325.
13. Kanwal F, Shubrook JH, Adams LA, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. Gastroenterology 2021;161(5):1657–1669.
14. Long MT, Noureddin M, Lim JK. AGA clinical practice update: diagnosis and management of nonalcoholic fatty liver disease in lean individuals: expert review. Gastroenterology 2022;163(3):764–774.e1.
15. Loomba R, Lim JK, Patton H, et al. AGA clinical practice update on screening and surveillance for hepatocellular carcinoma in patients with nonalcoholic fatty liver disease: expert Review. Gastroenterology 2020; 158(6):1822–1830.
16. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. Hepatology 2018; 67(1):328–357.
17. EASL, EASD, EASO. EASL-EASD-EASO Clinical practice guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64(6): 1388–1402.
18. European Association for the Study of the Liver; Clinical Practice Guideline Panel; Chair; EASL Governing Board representative; Panel members. EASL Clinical practice guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. J Hepatol 2021;75(3):659–689.
19. Tokushige K, Ikejima K, Ono M, et al. Evidence-based clinical practice guidelines for nonalcoholic fatty liver disease/nonalcoholic steatohepatitis 2020. J Gastroenterol 2021;56(11):951–963.
20. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2022.
21. Noureddin M, Ntanios F, Malhotra D, et al. Predicting NAFLD prevalence in the United States using National Health and Nutrition Examination Survey 2017-2018 transient elastography data and application of machine learning. Hepatol Commun 2022;6(7):1537–1548.
22. Ganguli S, DeLeeuw P, Satapathy SK. A Review of current and upcoming treatment modalities in non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. Hepat Med 2019;11:159–178.
23. Popa SL, Ismaiel A, Cristina P, et al. Non-alcoholic fatty liver disease: implementing complete automated diagnosis and staging. A systematic review. Diagnostics (Basel) 2021;11(6):1078.
24. Samji NS, Snell PD, Singal AK, et al. Racial disparities in diagnosis and prognosis of nonalcoholic fatty liver disease. Clin Liver Dis 2020;16(2):66–72.
25. Pandyarajan V, Gish RG, Alkhouri N, et al. Screening for nonalcoholic fatty liver disease in the primary care clinic. Gastroenterol Hepatol 2019;15(7):357–365.
26. American Diabetes Association. 4. Comprehensive medical evaluation and assessment of comorbidities: standards of medical care in diabetes-2020. Diabetes Care 2020;43(Suppl 1):S37–S47.
27. Nevola R, Epifani R, Imbriani S, et al. GLP-1 receptor agonists in non-alcoholic fatty liver disease: current evidence and future perspectives. Int J Mol Sci 2023; 24(2):1703.
28. Newsome PN, Buchholtz K, Cusi K, et al. A placebo-controlled trial of subcutaneous semaglutide in nonalcoholic steatohepatitis. N Engl J Med 2020; 384(12):1113–1124.
29. Kohane IS, Aronow BJ, Avillach P, et al. What every reader should know about studies using electronic health record data but may be afraid to ask. J Med Internet Res 2021;23(3):e22219.

Received January 25, 2024. Accepted December 31, 2024.

Correspondence:
Address correspondence to: Bridget M. Geyer, MD, MSc, Duke University School of Medicine, 40 Duke Medicine Circle, Durham, North Carolina 27710. e-mail: bridget.geyer@duke.edu.

Author contributions:
Bridget M. Geyer: Writing – original draft. Cynthia A. Moylan: Investigation, Study design. Feier Chang: Statistical analysis. Peng Wu: Statistical analysis. Benjamin A. Goldstein: Statistical analysis. Kara Wegermann: Research. Sunny L. Chung: Research. Matthew Phelan: Research. Joseph Wawrzynski: Research. Jacqueline B. Henson: Research. Howard Lee: Research. Phil Ambery: Study conceptualization. Neha Pagidipati: Study conceptualization.

Conflicts of Interest:
These authors disclose the following: Phil Ambery is an employee and shareholder of AstraZeneca. Cynthia A. Moylan receives research support from: GSK, Madrigal, and Exact Sciences and is a member of consultation/advisory panels for Boehringer Ingelheim, Novo Nordisk, GLG, and Sirtex. Neha Pagidipati receives research support from Alnylam, Amgen, Bayer, Boehringer Ingelheim, Eggland's Best, Eli Lilly, Novartis, Novo Nordisk, Merck and is a member of consultation/advisory panels for Bayer, Boehringer Ingelheim, CRISPR Therapeutics, Eli Lilly, Esperion, AstraZeneca, Merck, Novartis, and Novo Nordisk, executive committees for trials sponsored by Novo Nordisk and by Amgen, DSMB for trials sponsored by J + J and Novartis, and medical advisory board for Miga Health. The remaining authors disclose no conflicts.

Funding:
This study was funded in part by AstraZeneca pharmaceuticals.

Ethical Statement:
This study was approved by the Duke University Institutional Review Board under ID Pro00102560.

Data Transparency Statement:
The data used in this study are from a private datamart but can be furnished on request to the corresponding author.

Reporting Guidelines:
Compliant with CODE-EHR best practice guidelines.