Journal of Clinical Lipidology (2016) 10, 1230–1239

Journal of Clinical Lipidology

Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study

Maya S. Safarova, MD, PhD, Hongfang Liu, PhD, Iftikhar J. Kullo, MD*

Department of Cardiovascular Diseases, Mayo Clinic, Rochester, MN, USA (Drs Safarova and Kullo); and Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA (Dr Liu)

KEYWORDS:
Familial hypercholesterolemia;
Screening;
Prevalence;
Awareness;
Control;
Hypercholesterolemia;
Electronic phenotyping;
eEpidemiology;
Electronic health records;
Informatics

BACKGROUND: Little is known about prevalence, awareness, and control of familial hypercholesterolemia (FH) in the United States.

OBJECTIVE: To address these knowledge gaps, we developed an ePhenotyping algorithm for rapid identification of FH in electronic health records (EHRs) and deployed it in the Screening Employees And Residents in the Community for Hypercholesterolemia (SEARCH) study.

METHODS: We queried a database of 131,000 individuals seen between 1993 and 2014 in primary care practice to identify 5992 (mean age 52 ± 13 years, 42% men) patients with low-density lipoprotein cholesterol (LDL-C) ≥190 mg/dL, triglycerides <400 mg/dL and without secondary causes of hyperlipidemia.

RESULTS: Our EHR-based algorithm ascertained the Dutch Lipid Clinic Network criteria for FH using structured data sets and natural language processing for family history and presence of FH stigmata on physical examination. Blinded expert review revealed positive and negative predictive values for the SEARCH algorithm at 94% and 97%, respectively. The algorithm identified 32 definite and 391 probable cases with an overall FH prevalence of 0.32% (1:310). Only 55% of the FH cases had a diagnosis code relevant to FH. Mean LDL-C at the time of FH ascertainment was 237 mg/dL; at follow-up, 70% (298 of 423) of patients were on lipid-lowering treatment with 80% achieving an LDL-C ≤100 mg/dL. Of treated FH patients with premature CHD, only 22% (48 of 221) achieved an LDL-C ≤70 mg/dL.

CONCLUSIONS: In a primary care setting, we found the prevalence of FH to be 1:310 with low awareness and control. Further studies are needed to assess whether automated detection of FH in EHR improves patient outcomes.

Dr. Safarova is supported by American Heart Association Postdoctoral Fellowship Award 16POST27280004. Dr. Kullo is funded by the National Human Genome Research Institute's electronic Medical Records and Genomics Network through grants HG04599 and HG006379 to Mayo Clinic. The National Human Genome Research Institute and American Heart Association had no role in the design and conduct of the work; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

* Corresponding author. Department of Cardiovascular Diseases, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
E-mail address: Kullo.Iftikhar@mayo.edu

Submitted March 4, 2016. Accepted for publication August 1, 2016.

1933-2874/
http://dx.doi.org/10.1016/j.jacl.2016.08.001

Introduction

An important paradigm of precision medicine is to screen individuals for disease before overt clinical manifestations particularly when treatment is available and beneficial. An example is familial hypercholesterolemia (FH) where major adverse events such as sudden cardiac death, myocardial infarction, and stroke can be prevented by timely initiation of lipid-lowering therapy. A relatively common genetic disorder, FH is associated with dramatically increased lifetime risk of premature atherosclerotic cardiovascular disease (ASCVD) due to elevated plasma low-density lipoprotein cholesterol (LDL-C) levels.[1,2] FH has been labeled a Tier 1 public health genomics condition[3] and is one of the few genetic diseases that meets the World Health Organization criteria for population-based large-scale screening programs aimed at early disease detection and timely treatment.[4] Increased attention has focused on FH recently as advances in genomics are providing important insights into the genetic architecture of lipid disorders,[5,6] and novel classes of lipid-lowering drugs are opening up new avenues of therapy in these high-risk patients.[7–10]

However, substantial gaps in our knowledge of prevalence, awareness, and control of FH remain. Few studies have specifically addressed the prevalence of heterozygous FH. Almost half a century ago, the prevalence of heterozygous FH was estimated at 1:500 among relatives of survivors of myocardial infarction.[11] Excluding specific populations with a "founder effect," the reported estimates of prevalence of FH vary widely.[12] In the Danish population, prevalence was reported to be 1:137 (0.7%),[13] whereas in a study from neighboring Finland, the prevalence was 1:600 (0.2%).[14] Based on genetic screening, the prevalence of FH in the Netherlands' population was 1:200.[15] In the US National Health and Nutrition Examination Survey (NHANES), the prevalence of FH diagnosed using clinical criteria was estimated to be 1:250 (0.4%).[16] This significant difference in reported prevalence rates motivates study of the prevalence of FH in a community-based setting in the United States. Furthermore, little is known about the extent to which FH is underdiagnosed and undertreated in the United States. Indeed, some have projected that <10% of prevalent cases are diagnosed and treated.[17,18]

To address these knowledge gaps, we undertook the Mayo Screening Employees And Residents in the Community for Hypercholesterolemia (SEARCH) study in the Mayo Employee and Community Health (ECH) system that delivers primary care to residents of Olmsted County and southeastern Minnesota. A unified electronic data trust that includes comprehensive clinical records of ECH patients enabled the research described in this report. We developed an electronic phenotyping algorithm to mine electronic health records (EHR) to identify patients who met the Dutch Lipid Clinic Network (DLCN) criteria for FH with the long-term goal of addressing knowledge gaps in prevention, awareness, control of FH, and the prevention of premature ASCVD.

Material and methods

Study population and settings

This cross-sectional study was approved by the Institutional Review Board of Mayo Clinic, Rochester, Minnesota. Individuals in the Mayo ECH system who had given permission for their medical records to be used for research and had clinical data available in the EHR were considered eligible for the study. Lipid levels were extracted from structured laboratory databases from June 21, 1993, to December 31, 2014. The index date was defined as the date of the earliest LDL-C level ≥190 mg/dL. Race was categorized as "white," "black or African American," "Asian," "other," and "choose not to disclose." Initially, EHRs of a random sample of 115 patients with severe hypercholesterolemia were manually reviewed to inform development of the electronic phenotyping algorithm for FH. Because a variety of medical conditions (hypothyroidism, cholestatic liver diseases, nephrotic syndrome, severe renal failure, and pregnancy)[19] can increase LDL-C levels, patients who had such conditions within 1 year before the index date were excluded. We identified these conditions using Logical Observation Identifiers Names and Codes (LOINC) for renal, thyroid, and liver function laboratory parameters as well as the *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* codes for pregnancy.

Defining FH case status

To identify cases of FH, we used a modified numerical score system of the DLCN criteria.[13] From the original set of clinical criteria,[20] we excluded two criteria: a first-degree relative with tendon xanthomas or corneal arcus, and children aged <18 years with severe hypercholesterolemia. This decision was based on initial manual review of EHR of patients with LDL-C ≥190 mg/dL, which revealed that these variables were not recorded by providers. Variables that were incorporated in the SEARCH ePhenotyping algorithm (version 1.0) are listed in Table 1. FH case status was assigned as "definite" if the score was >8 points or "probable" when the score was 6–8, with the remaining individuals identified as "possible" FH (scored 3–5 points) serving as controls. The number of definite and probable FH cases was combined to estimate the prevalence of FH.

Ascertaining FH criteria from the EHR

We mined both structured data sets and unstructured clinical text to ascertain FH criteria from the EHR (Table 1