EHR Phenotyping for Determining Patient Follow-up on Genetic Variants

Jennifer A. Pacheco, BA1,2, Laura J. Rasmussen-Torvik, PhD 2, Adam S. Gordon, PhD 2, Tess D. Pottinger, MS2, Megan J. Puckelwartz, PhD 2, Christin Hoell, MS2, Beth Wilmot, PhD1, Michael Mooney, PhD1
1Oregon Health and Sciences Univ., Portland, OR; 2Northwestern Univ., Chicago, IL

Introduction

Many genetic variants are of unknown significance (VUS). Efficient and accurate electronic health record (EHR) phenotyping, having facilitated genome-wide association studies (gwas.org), could identify patients with genetic variants who exhibit phenotypic features that might indicate pathogenicity of those variants. Identifying and following up with these patients could improve their healthcare, and assist in improving genetic variant categorization.

Methods

Subjects (N=3860) were recruited at Northwestern Medicine for 2 studies and sequenced on 2 platforms with the 3 genes (LDLR, APOB, PCSK9) containing variants that collectively explain ~40% of diagnosed cases of familial hypercholesterolemia (HC) (FH). Rare variants in these genes were queried in ClinVar (ncbi.nlm.nih.gov/clinvar) for pathogenic/likely pathogenic (P/LP), conflicting interpretations of pathogenicity (CP), and VUS classifications; and unreported non-synonymous variants (URV) were noted. Four EHR phenotype algorithms of varying complexity were implemented (Figure 1) to find the most efficient and effective algorithm(s) for identifying patients with HC: 2 algorithms, for primary HC (PH) and FH (PheKB.org/node/602); a subset of the PH algorithm: maximum low-density lipoprotein (LDL) without recurring high triglycerides (high LDL); and ICD diagnosis codes, grouped into phecodes (maps 1.2, 1.2b1, PheWAScatalog.org) for HC phenotypes. Overlap of subjects with both the genotypes and phenotypes was assessed, and also, if a VUS was found in >1 subject, we calculated % of those with HC.

Results, Discussion, and Conclusion

Of the 23 patients with P/LP FH variants, 21 were found by any algorithm to have a record of HC in the EHR. Furthermore, 324 out of 570 patients with CP variants were found to have HC, 120 out of 174 patients with VUSs but no P/LP nor CP variants were found to have HC, and 178 out of 308 with only URVs were found to have HC. As expected the phecode algorithm (the simplest), found the most patients with HC with P/LP variants (21), or other queried variant types (599). Both the high LDL and PH algorithms found a similar number of HC patients with P/LP variants or other variants, although overall PH found more than high LDL algorithm. As expected, the FH algorithm (the strictest), found the least number of patients with HC with P/LP variants (3) or other variants (15). Since none of the phenotyping algorithms found evidence of HC in the EHR for all patients having P/LP variants, and nor did a single phenotype algorithm identify most of them, for patients with FH genetic variants, both phecode and PH algorithms are needed to identify patients for diagnostic evaluation. Of the 128 subjects with VUSs in FH genes in cohort b, 71% had evidence of HC in their EHR, indicating those 25 VUSs may be P/LP. With further assessment, these methods, combined with other data, could be used to identify phenotypes in patients with VUSs and similar variants for categorizing VUSs, URVs, or CP variants as either P/LP or benign/likely benign.

References