Identifying Genetic Pleiotropy through a Literature-wide Association Study (LitWAS) and a Phenotype Association Study (PheWAS) in the Age-related Eye Disease Study 2 (AREDS2)

THE UNIVERSITY OF ARIZONA COLLEGE OF MEDICINE PHOENIX

Michael Simmons, UA; Freekje Van Asten, PhD, NIH; Ayush Singhal, PhD; Tiarnan Keenan, FRCO PhD, NIH; NIH, Emily Chew, MD, NIH; Zhiyong Lu, PhD NCBI, NLM, NIH

NIH

Introduction

Genetic pleiotropy refers to how single genetic polymorphisms relate to multiple phenotypes. Identifying genetic pleiotropy may provide valuable insight about the mechanisms behind gene-disease associations.

In this thesis, we explore such genetic pleiotropy with a Phenome-wide Association Study (PheWAS) using data from the Age-Related Eye Study 2 (AREDS2).

We also employ a novel text mining approach (a Literature-wide Association Study (LitWAS)) to extract pleiotropic associations from the published literature as a hypothesis generation mechanism.

Methods

The AREDS2 trial was a double-masked, multi-center, randomized clinical trial that investigated multivitamins as a preventative treatment for age-related macular degeneration (AMD). Data from AREDS2 covers 123 phenotypes including AMD features, other ocular conditions, cognitive function and cardiovascular, neurological, gastrointestinal and endocrine disease.

LitWAS: A previously validated relationship extraction algorithm was used to isolate descriptions of genetic associations with these phenotypes in MEDLINE abstracts. Results were filtered to exclude negated findings and normalize variant mentions.

PheWAS: Variant associations were tested by correlating SNPs selected from literature with literature-supported phenotypes. A number of SNPS were then selected for correlation testing with all available phenotypes. Associations that reached Bonferroni-corrected significance were replicated in AREDS.

Results

Table 1. AREDS2 Baseline Characteristics	
With genetic testing	1826
Female	1071
White	1776
Never smoked	792
Some college education	881
Mean Age (st dev)	73.2 years (7.7 years)
Mean follow-up (st dev)	4.8 years (0.5 years)
AREDS2 treatment	
Control	446
Lutein/Zeaxanthin	449
DHA/EPA	481
Combination Therapy (L+Z+DHA+EPA)	450

Table 1. Only AREDS2 patients who consented for genetic testing were included in this study.

LitWAS

Table 2. Literature-wide Association Study Results	
Total number of SNPs	9372
Total Unique Genes	2884
Minimum number of pleiotropic associations shared by one SNP	2*
Average phenotypes per SNP	3.1
Maximum number of pleiotropic associations shared by one SNP	38

Table 2: Single-nucleotide polymorphisms (SNPs) with at least two shared phenotypes were included in the LitWAS.

Table 3. Literature-mined Associations with the ARMS2 Variant, rs1040924				
Phenotype	PMID	OR	p-value	Impact Phrase
Exudative AMD	24362810	11.7	<0.0001	"ARMS2 … polymorphisms were associated with very high risk for exudative AMD"
Non-exudative AMD	20381870	6.01	<0.0001	"there was a significant association between genotype and presence of [geographic atrophy] for … rs1040924"
Tobacco	26067391	8.33	---	"smoking synergistically increased the susceptibility of AMD for variants of [rs1040924]"
Vitreous hemorrhage	22509112	6.52	0.07	"Association of rs10490924 with … vitreous hemorrhage was reported in two studies"
Subretinal hemorrhage	21397333	12.4	0.0001	"[rs10494924] increased the likelihood for hemorrhagic PED by 12.4-fold [in polypoidal choroidal vasculopathy]"
Reticular pseudodrusen	24595987	3.23	0.008	"Logistic regression analysis revealed that … [the] T-allele at ARMS2 A69S … were risk factors for RPD"

Table 3: A selection of results from the LitWAS. Values source from either the abstract or article text. PMID = PubMed Unique Identifier

PheWAS

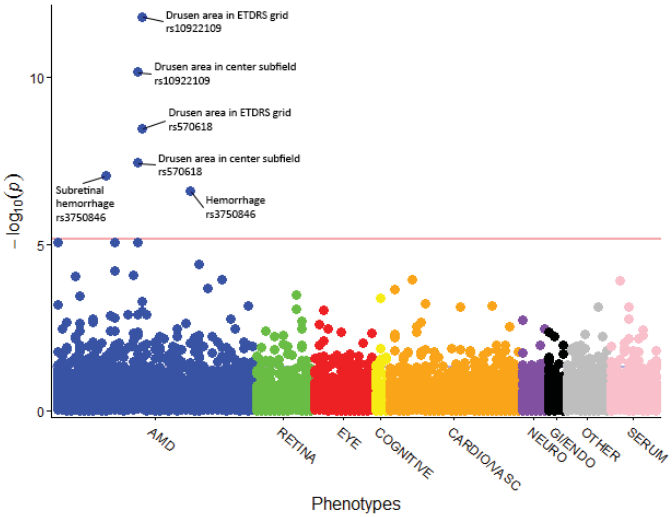

Figure 1: A Manhattan plot displaying the strengths of association between all tested SNPs and the AREDS2 phenotypes. Significant associations exist between variants of the ARMS2 gene and hemorrhagic phenotypes of the retina.

Table 4. ARMS2 (rs3750846) and Hemorrhage Characteristic of AMD in AREDS and AREDS2				
Patient Population	Hemorrhage	%	OR (CI)	p-value
AREDS				
Large drusen at baseline and CNV at any time (n=1001)	691	69.0	1.74 (1.2-2.6)	0.004
Large drusen at baseline and CNV developed during follow-up (n=475)	284	59.8	1.57 (1.0-2.5)	0.068
CNV at any time (n=1285)	848	66.0	1.78 (1.3-2.5)	<0.001
CNV developed during follow-up (n=560)	329	58.8	1.69 (1.1-2.6)	0.019
AREDS2				
Large drusen at baseline and CNV at any time (n=615)	191	31.1	1.58 (0.9-2.7)	0.097
Large drusen at baseline and CNV developed during follow-up (n=480)	142	29.6	2.0 (1.0-3.8)	0.038
CNV at any time (n=1946)	969	49.8	1.75 (1.3-2.3)	<0.001
CNV developed during follow-up (n=520)	155	29.8	1.94 (1.1-3.5)	0.031

Table 4. The association between the ARMS2 variant, rs3750846, and sub-retinal hemorrhage is consistent in prevalent and incident populations in both AREDS and AREDS2.

Discussion and Conclusions

The rs3750846 variant of the ARMS2-HTRA1 locus is associated with subretinal hemorrhage.

Automatic literature mining, when paired with clinical data, is a promising method for exploring genotype-phenotype relationships.

Acknowledgements

This research was supported by the NIH Intramural Research Program, National Library of Medicine, and the NIH Medical Research Scholars Program, a public-private partnership supported jointly by the NIH and generous contributions to the Foundation for the NIH from the Doris Duke Charitable Foundation, the Howard Hughes Medical Institute, the American Association for Dental Research, the Colgate-Palmolive Company, and other private donors. No funds from the Doris Duke Charitable Foundation were used to support research that used animals.