Association of Interleukin 6 Receptor Variant With Cardiovascular Disease Effects of Interleukin 6 Receptor Blocking Therapy A Phenome-Wide Association Study

Tianxi Cai, ScD; Yichi Zhang, PhD; Yuk-Lam Ho, MPH; Nicholas Link, BA; Jiehuan Sun, PhD; Jie Huang, MS; Tianrun A. Cai, MD; Scott Damrauer, MD; Yuri Ahuja, BS; Jacqueline Honerlaw, RN, BSN, MPH; Jie Huang, PhD; Lauren Costa, MPH; Petra Schubert, MPH; Chuan Hong, PhD; David Gagnon, MD, MPH, PhD; Yan V. Sun, PhD; J. Michael Gaziano, MD, MPH; Peter Wilson, MD; Kelly Cho, PhD, MPH; Philip Tsao, PhD; Christopher J. O’Donnell, MD, MPH; Katherine P. Liao, MD, MPH; for the VA Million Veteran Program

IMPORTANCE Electronic health record (EHR) biobanks containing clinical and genomic data on large numbers of individuals have great potential to inform drug discovery. Individuals with interleukin 6 receptor (IL6R) single-nucleotide polymorphisms (SNPs) who are not receiving IL6R blocking therapy have biomarker profiles similar to those treated with IL6R blockers. This gene–drug pair provides an example to test whether associations of IL6R SNPs with a broad range of phenotypes can inform which diseases may benefit from treatment with IL6R blockade.

OBJECTIVE To determine whether screening for clinical associations with the IL6R SNP in a phenome-wide association study (PheWAS) using EHR biobank data can identify drug effects from IL6R clinical trials.

DESIGN, SETTING, AND PARTICIPANTS Diagnosis codes and routine laboratory measurements were extracted from the VA Million Veteran Program (MVP); diagnosis codes were mapped to phenotype groups using published PheWAS methods. A PheWAS was performed by fitting logistic regression models for testing associations of the IL6R SNPs with 1342 phenotype groups and by fitting linear regression models for testing associations of the IL6R SNP with 26 routine laboratory measurements. Significance was reported using a false discovery rate of 0.05 or less. Findings were replicated in 2 independent cohorts using UK Biobank and Vanderbilt University Biobank data. The Million Veteran Program included 332 799 US veterans; the UK Biobank, 408 455 individuals from the general population of the United Kingdom; and the Vanderbilt University Biobank, 13 835 patients from a tertiary care center.

EXPOSURES IL6R SNPs (rs2228145; rs4129267).

MAIN OUTCOMES AND MEASURES Phenotypes defined by International Classification of Diseases, Ninth Revision codes.

RESULTS Of the 332 799 veterans included in the main cohort, 305 228 (91.7%) were men, and the mean (SD) age was 66.1 (13.6) years. The IL6R SNP was most strongly associated with a reduced risk of aortic aneurysm phenotypes (odds ratio, 0.87-0.90; 95% CI, 0.84-0.93) in the MVP. We observed known off-target effects of IL6R blockade from clinical trials (eg, higher hemoglobin level). The reduced risk for aortic aneurysms among those with the IL6R SNP in the MVP was replicated in the Vanderbilt University Biobank, and the reduced risk for coronary heart disease was replicated in the UK Biobank.

CONCLUSIONS AND RELEVANCE In this proof-of-concept study, we demonstrated application of the PheWAS using large EHR biobanks to inform drug effects. The findings of an association of the IL6R SNP with reduced risk for aortic aneurysms correspond with the newest indication for IL6R blockade, giant cell arteritis, of which a major complication is aortic aneurysm.

Published online August 8, 2018.

© 2018 American Medical Association. All rights reserved.
Naturally occurring variants in the human genome can serve as experiments of nature to study potential drug targets. Individuals with genetic variants detected as single-nucleotide polymorphisms (SNPs) can have profiles similar to individuals receiving a treatment. An example of a gene–drug pair is the interleukin 6 receptor (IL6R) genetic variant Asp358Ala (rs2228145) and the IL6R antagonists tocilizumab and sarilumab. Both are indicated for the treatment of rheumatoid arthritis (RA), and tocilizumab is indicated for giant cell arteritis (GCA). Individuals with the IL6R variant not taking IL6R blockade have biochemical parameters similar to individuals taking the drug. For example, patients initiating IL6R blockade experience a significant reduction in C-reactive protein (CRP) levels. Among individuals with the Asp358Ala IL6R genetic variant, carriers had an 8.3% lower CRP level compared with those without the variant. Interleukin 6 is a proinflammatory cytokine that triggers inflammation by binding IL6R on the cell membrane. Functional studies of the Asp328Ala genetic variant showed that carriers have reduced expression of membrane-bound IL6R, leading to an impaired response to IL6. Similarly, tocilizumab and sarilumab impair response to IL6 by blocking its ability to bind to IL6R.

The significance of this gene–drug relationship suggests that a large-scale screen of phenotypes or a phenome-wide association study (PheWAS) of the IL6R genetic variant may uncover potential therapeutic targets for IL6R antagonists. A phenome-wide association study can also identify potential adverse effects. This concept and approach has been discussed in the literature. However, it is only recently that large biobanks with linked EHR data, such as the Veterans Affairs Million Veteran Program (MVP) and the UK Biobank, have become available to fully test this hypothesis.

In this proof-of-concept study, we performed a PheWAS on the Asp235Ala IL6R genetic variant to determine whether cations of a genetic variant of interest with a broad range of phenotypes available in the electronic health record (EHR).

**Figure. Phenome-Wide Association Plot of IL6R in the Million Veteran Program**

Phenome-wide association study plot for IL6R showing phenotypes with significant associations in the Million Veteran Program. The blue line indicates the significance threshold for Bonferroni correction as a reference. Up-facing triangles indicate an increased risk while down-facing triangles indicate a reduced risk for a phenotype or laboratory measurement.

**Key Points**

**Question** Can a phenome-wide association study enable the use of genetics to inform drug development?

**Findings** In this phenome-wide association study using electronic health record and genetic data from 332,799 US veterans, the association between a genetic variant of interleukin 6 receptor (IL6R) with potential effects of IL6R blocker therapy was assessed. The study identified a recently approved indication for IL6R blocker therapy associated with aortic aneurysm and identified off-target effects observed from clinical trials.

**Meaning** The phenome-wide association study approach using large biobanks and genetics is a promising tool to assess potential beneficial and adverse effects of therapeutic agents with known pathways and related genes.
known effects for IL6R blockade as well as known off-target effects from clinical trials can be detected using data from a large US-based biobank study. Additionally, results were replicated in 2 independent biobank cohorts using freely available online data.

Methods

Study Populations
The Million Veteran Program served as the main cohort for this PheWAS. The Million Veteran Program is a longitudinal cohort study with clinical EHR data containing inpatient and outpatient data linked with genomic data. The Million Veteran Program recruits from approximately 50 Veterans Affairs facilities across the United States. Inclusion criteria in the MVP include age of 18 years or older, having a valid mailing address, and having the ability to provide informed consent. At recruitment, individuals completed baseline and lifestyle questionnaires, including self-reported race/ethnicity, and provided blood samples for genotyping and biomarker studies. All individuals in the study provided written informed consent as part of the MVP. This study was approved through the Veterans Affairs central institutional review board as part of the MVP.

The UK Biobank is a prospective study of the effects of lifestyle, environmental, and genomic factors on disease outcomes. The study recruited approximately 500,000 volunteers from the general population of the United Kingdom aged 40 to 69 years from 2006 to 2010. The phenotypes available in the UK Biobank are derived from diverse sources, including inpatient International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes. Currently online, only grouped inpatient ICD-10 codes are available. Data for individual ICD-10 and any outpatient ICD-10 codes were not available. Estimates of the genome-phenome associations along with their P values from the UK Biobank were obtained using Gene ATLAS.

The Vanderbilt University Biobank (BioVU) is a DNA biobank at Vanderbilt University Medical Center linked to a copy of their EHR containing inpatient and outpatient data with a goal to explore the connection between genetics and health outcomes. All estimates of the genome-phenome associations along with their P values (<.05) based on 13,835 BioVU participants are freely available online. Deidentified data from the UK Biobank and BioVU are available freely online, and informed consent was waivered.

Statistical Methods
The phenome-wide association study analysis included both ICD-9–based phenotypes and a list of routine laboratory measurements that were available in 75% or more of patients in the MVP. The ICD-9–based phenotypes were defined by mapping ICD-9 codes to PheWAS codes, as published by Denny et al. Using the standard approach, a participant was defined as having a phenotype if they had 2 or more PheWAS codes. We excluded PheWAS codes with a prevalence of 0.1% or less from the analysis. The laboratory measurements, defined by the average of all available measurements for each patient, consisted of complete blood count, including white blood cell count, hemoglobin level, platelet count, creatinine level, estimated glomerular filtration rate, liver function tests, and lipid levels, as well as total cholesterol level, high-density lipoprotein cholesterol level, low-density lipoprotein cholesterol level, and triglyceride level.

The primary study screened for associations of rs2228145 (risk allele A; Asp358Ala) with each individual phenotype defined by PheWAS codes by fitting logistic regression models. Linear regression models were fitted to test for associations of the log-transformed laboratory measurements with the PheWAS codes. We applied standard quality control pipelines, such as discordance between sex inferred by genotyping vs self-report. We also excluded related individuals (half-siblings, first-degree relatives, or second-degree relatives) as measured by the Kinship-Based Inference for GWAS software. All models were adjusted for age at the last visit, sex, and the 20 leading principle components to adjust for population stratification, follow-up time in months, and total number of ICD-9 codes. The follow-up time and total number of ICD-9 codes was included to adjust for the density of EHR data for each patient; both variables were log-transformed.

To adjust for multiple testing, we defined statistical significance as a P value less than a threshold controlling for a false discovery rate (FDR) of 5% using the Benjamini–Hochberg procedure. This Benjamini–Hochberg FDR control ensured that among the associations considered significant, at most 5% of the associations were false-positive. For reference, we also reported the threshold for Bonferroni correction at a family-wise error rate of 5%. As a sensitivity analysis, we additionally performed the PheWAS using 1 or more PheWAS codes to define a phenotype.

Based on prior knowledge regarding the hypothesized effects of IL6R treatment on inflammatory and cardiovascular phenotypes, we additionally extracted laboratory data on CRP levels, cardiac troponin levels, creatine kinase–MB (CK-MB) level, and brain natriuretic peptide level. We then tested the associations of rs2228145 with these additional laboratory phenotypes using linear regression models.

Validation of Significant Outcomes
With Medical Record Review
For phenotypes considered significant after FDR control, we validated the accuracy of each phenotype through medical record review; 20 participants were randomly selected among those who had 1 or more PheWAS codes and reviewed for evidence of the phenotype in the narrative notes. All reviewers were clinically trained health care professionals (T.A.C., J.H., S.D., and K.P.L.). We reported the positive predictive value (PPV) of participants with 1 or more PheWAS codes and 2 or more PheWAS codes. The PPV in general was calculated as the number of confirmed phenotypes based on medical record review divided by participants with either 1 or more or 2 or more PheWAS codes.

Replication Using UK Biobank and BioVU Online Data
The phenotypes with significant associations with the IL6R genetic variant in MVP were further examined in the BioVU
and UK Biobank. Since the exact SNP used for the MVP PheWAS was not available in these data sets, we compiled a list of SNPs in high linkage disequilibrium with our SNP of interest, rs2228145. Fourteen SNPs in high linkage disequilibrium (ie, $R^2 > 0.9$) with rs2228145 were identified from the African and European populations obtained from the LDlink website using the LDproxy function.15 From this list, rs4129267 ($R^2 = 0.99$) was available in both the BioVU and UK Biobank and was used to replicate findings from the MVP.

In the online UK Biobank data, phenotypes were grouped by inpatient $ICD$-10 codes, and analyses could not be performed with individual $ICD$-10 data. To replicate significant findings from the MVP, mapping from $ICD$-10 groups in UK Biobank were mapped back to $ICD$-9 codes and then to PheWAS codes. In some cases, the UK Biobank $ICD$-10 groups could not be mapped to a PheWAS code because the group was too broad. Using the closest match based on phenotype description, we then extracted data on the association of rs4129267 with the available phenotypes of interest using the Gene ATLAS website.14 Since phenotypes from BioVU were defined by PheWAS code, a direct look-up online of the associations of rs4129267 with the phenotypes of interest was performed.8 For the replication studies, a $P$ value less than .05 was considered significant. All analyses were implemented in R, version 3.2.2 (the R Foundation).

**Results**

The *$IL6R$* PheWAS in the MVP studied 332 799 participants, of whom 305 228 (91.7%) were men with a mean (SD) age of 66.1 (13.6) years and a mean (SD) follow-up time of 11.9 (4.9) years. General characteristics of the population, including the most common conditions based on PheWAS codes and representation by region of the United States, are shown in Table 1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>66.1 (13.6)</td>
</tr>
<tr>
<td>Male</td>
<td>305 228 (91.7)</td>
</tr>
<tr>
<td>Duration of EHR follow-up, mean (SD), y</td>
<td>11.94 (4.92)</td>
</tr>
</tbody>
</table>

Abbreviation: EHR, electronic health record.

*Groupings based on phenome-wide association study codes.

Twenty-two significant phenotypes were associated with the *$IL6R$* genetic variant, of which 13 (59%) were associated with vascular and cardiac disease; the threshold for significance was $P < 6.6 \times 10^{-4}$. The phenotypes with the strongest association with *$IL6R$* were aortic aneurysm (odds ratio [OR], 0.90; 95% CI, 0.87-0.93) as well as a specific type of aortic aneurysm, abdominal aortic aneurysm (AAA) (OR, 0.87; 95% CI, 0.84-0.90), and coronary atherosclerosis and ischemic heart disease (OR, 0.95; 95% CI, 0.94-0.97) (Figure; Table 2) (eTable 1 in the Supplement). Based on medical record review, the PPV of the PheWAS codes ranged from 55% to 100% (eTable 2 in the Supplement).

**Association of $IL6R$ Variant With Laboratory Findings**

Each allele of the *$IL6R$* variant was associated with higher levels of hemoglobin and albumin (Figure). These 2 laboratory measurements are known to increase with *$IL6R$* antagonist therapy.5,22-26 The known reduction in CRP level was observed in the PheWAS, as were findings consistent with a reduced risk of ischemic heart disease, lower levels of troponin I, and lower CK-MB levels (Table 3).

The sensitivity analysis showed similar associations of the *$IL6R$* variant with reduced risk for aortic aneurysms and cardiovascular disease phenotypes (eFigure in the Supplement). Additionally, in a secondary analysis, we examined the association of the *$IL6R$* variant with procedure codes for aortic rupture repair (n = 1667), as these patients may be considered to have a more severe form of AAA. We observed a 16% reduction in risk for AAA rupture repair among individuals with the *$IL6R$* variant (OR, 0.84; 95% CI, 0.78-0.90).

**Replicating MVP PheWAS Findings in the UK Biobank and BioVU**

From UK Biobank, we replicated an association of the *$IL6R$* variant with coronary heart disease phenotypes, including chronic ischemic heart disease (OR, 0.99; 95% CI, 0.9968-0.9991; $P = .002$) among the 408 455 individuals with genomic and inpatient $ICD$-10 billing group data (Table 2) (eTable 3 in the Supplement). From BioVU, we performed the replication on 13 835 individuals and confirmed associations of the *$IL6R$* variant with abdominal aortic aneurysm (OR, 0.83; 95% CI, 0.7082-0.9726; $P = .02$) and atopic or contact dermatitis (OR, 1.08; 95% CI, 1.0020-1.1554; $P = .04$).

**Discussion**

This study demonstrated an application of large EHR biobanks as a research platform using genetics to inform drug development pipelines using a PheWAS approach. From previous studies, the *$IL6R$* genetic variant was known to have similar biochemical effects as *$IL6R$* antagonist therapy; a higher number of *$IL6R$* alleles (0, 1, or 2) is associated with effects of a higher dose of *$IL6R$* blockers. This phenome-wide association study identified several potential associations. The discussion will focus on phenotypes associated with the *$IL6R$* variant in the MVP with replication in the UK Biobank or BioVU.
Abbreviations: NA, not applicable; OR, odds ratio.

sarilumab, including associations with higher hemoglobin levels and atopic dermatitis. Data from the IL6R clinical trials showed associations of the IL6R variant with lower CRP levels and atopic dermatitis, broadly characterized as a rash, is also a common adverse effect of IL6R blockers. Conjunctivitis and pleuritis, though uncommon, have also been reported in clinical trials. The association with gout was interesting; however, there is a paucity of data regarding a link with IL6R in the literature.

The IL6R PheWAS confirmed findings from a mendelian randomization study and a meta-analysis of a reduced risk for a range of coronary heart disease phenotypes among carriers of the IL6R variant in both the MVP and the UK Biobank. Further, studies of cardiac biomarkers in the MVP showed associations of the IL6R variant with lower CK-MB and troponin levels in a dose-dependent manner. These findings agree with results from a randomized clinical trial of tocilizumab compared with placebo administered to patients with RA within 2 days after a non–ST-elevation myocardial infarction. Patients treated with tocilizumab post–ST

Table 2. Significant Associations of IL6R With Phenome-Wide Association Study (PheWAS) Codes in the Million Veteran Program (MVP) and Replication Results From UK Biobank and Vanderbilt University Biobank (BioVU)*

<table>
<thead>
<tr>
<th>Clinical Phenotype</th>
<th>MVP PheWAS Code Description</th>
<th>Prevalence, %</th>
<th>OR (95% CI)</th>
<th>P Value</th>
<th>Replication UK Biobanka</th>
<th>BioVu c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aortic aneurysm</td>
<td>Abdominal aortic aneurysm</td>
<td>2.4</td>
<td>0.87 (0.84-0.90)</td>
<td>3.73 × 10⁻⁵</td>
<td>Yes²</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Aortic aneurysm</td>
<td>3.0</td>
<td>0.90 (0.87-0.93)</td>
<td>1.02 × 10⁻¹¹</td>
<td>Yes²</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Other aneurysm</td>
<td>3.4</td>
<td>0.92 (0.89-0.94)</td>
<td>2.57 × 10⁻⁹</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Aneurysm of iliac artery</td>
<td>0.3</td>
<td>0.83 (0.75-0.92)</td>
<td>2.73 × 10⁻⁴</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>Coronary atherosclerosis</td>
<td>20.5</td>
<td>0.95 (0.94-0.97)</td>
<td>3.43 × 10⁻¹²</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Ischemic heart disease</td>
<td>25.4</td>
<td>0.95 (0.94-0.97)</td>
<td>3.97 × 10⁻¹²</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Other chronic ischemic heart disease</td>
<td>14.9</td>
<td>0.95 (0.93-0.97)</td>
<td>6.04 × 10⁻¹²</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Myocardial infarction</td>
<td>5.5</td>
<td>0.94 (0.92-0.97)</td>
<td>1.69 × 10⁻⁶</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>Vascular disease</td>
<td>Atherosclerosis of native arteries of the extremities with intermittent claudication</td>
<td>1.6</td>
<td>0.90 (0.87-0.94)</td>
<td>3.92 × 10⁻⁶</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Peripher al vascular disease, unspecified</td>
<td>6.5</td>
<td>0.95 (0.93-0.97)</td>
<td>5.02 × 10⁻⁶</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Atherosclerosis of the extremities</td>
<td>2.3</td>
<td>0.92 (0.89-0.96)</td>
<td>9.73 × 10⁻⁶</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Atherosclerosis</td>
<td>3.1</td>
<td>0.93 (0.91-0.96)</td>
<td>1.80 × 10⁻⁵</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Peripher al vascular disease</td>
<td>6.9</td>
<td>0.96 (0.94-0.98)</td>
<td>3.87 × 10⁻⁵</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Skin conditions</td>
<td>Degenerative skin conditions</td>
<td>19.4</td>
<td>1.03 (1.01-1.04)</td>
<td>1.72 × 10⁻⁴</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Seborrheic dermatitis</td>
<td>4.5</td>
<td>1.05 (1.02-1.07)</td>
<td>3.30 × 10⁻⁴</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Atopic/contact dermatitis</td>
<td>11.9</td>
<td>1.03 (1.01-1.05)</td>
<td>3.72 × 10⁻⁴</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Erythematousquamous dermatosis</td>
<td>4.6</td>
<td>1.04 (1.02-1.07)</td>
<td>5.10 × 10⁻⁴</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>Acquired deformities of finger</td>
<td>0.3</td>
<td>1.19 (1.09-1.31)</td>
<td>2.57 × 10⁻⁴</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Gouty arthropathy</td>
<td>1.5</td>
<td>1.08 (1.03-1.13)</td>
<td>5.23 × 10⁻⁴</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>Pulmonary</td>
<td>Pleurisy/pleural effusion</td>
<td>1.8</td>
<td>1.07 (1.03-1.12)</td>
<td>3.47 × 10⁻⁴</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>Renal</td>
<td>Disorders resulting from impaired renal function</td>
<td>1.1</td>
<td>1.10 (1.04-1.16)</td>
<td>4.92 × 10⁻⁴</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Eye</td>
<td>Conjunctivitis, non-infectious</td>
<td>1.4</td>
<td>1.08 (1.03-1.13)</td>
<td>6.60 × 10⁻⁴</td>
<td>Yes</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not applicable; OR, odds ratio.

* Details on mapping and association study in UK Biobank and BioVU can be found in Table 3 in the Supplement.

© 2018 American Medical Association. All rights reserved.
Association of Interleukin 6 Receptor Variant With CVDEffects of IL6R Blocking Therapy

Table 3. Association of the IL6R Variant With Biomarkers Associated With Inflammation and Cardiovascular Disease Risk

<table>
<thead>
<tr>
<th>Laboratory Measurement</th>
<th>β (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-reactive protein</td>
<td>-0.06 (-0.08 to -0.04)</td>
<td>2.76 × 10⁻¹¹</td>
</tr>
<tr>
<td>Troponin I</td>
<td>-0.04 (-0.06 to -0.02)</td>
<td>7.12 × 10⁻⁵</td>
</tr>
<tr>
<td>CK-MB</td>
<td>-0.01 (-0.02 to -0.003)</td>
<td>.01</td>
</tr>
<tr>
<td>Pro-B-type natriuretic peptide</td>
<td>-0.02 (-0.05 to 0.01)</td>
<td>.18</td>
</tr>
</tbody>
</table>

Abbreviation: CK-MB, creatine kinase–MB.

Other limitations include the accuracy of ICD-9 codes for defining phenotypes. Studies are underway to improve the accuracy of phenotypes, which in turn would improve the power of PheWAS.58 Because the PheWAS is a hypothesis-generating tool, we applied an FDR of 5% to determine significant associations. Whether the threshold should be more or less stringent or if an alternate approach is needed is also under investigation. Additionally, the phenotypes used in the standard PheWAS mapping6,16 are not independent and can be highly correlated (eg, myocardial infarction and coronary atherosclerosis). Both the Benjamini-Hochberg FDR control procedures applied in this study are considered robust in accounting for the presence of positive correlations between outcomes of interest.59 As with all hypothesis-generating approaches, the results require validation in independent cohorts.

The lack of replication for more phenotypes from the MVP in UK Biobank and BioVU may also be because of a lack of power from differences in sample size and depth of data. The online data currently available for UK Biobank are inpatient ICD-10 code groups only. Therefore, the association analysis performed in UK Biobank was between the IL6R variant and an inpatient ICD-10 code group for aortic aneurysm and dissection, a broader phenotype. The prevalence of this broad ICD-10 group in UK Biobank was 0.34%, significantly lower than in the MVP or BioVU. Although BioVU contains both inpatient and outpatient data, it also has a significantly smaller population than the MVP or UK Biobank. We believe the ability to replicate the association of a single IL6R variant with AAA further supported the strength of the association and this approach. Lastly, while the PheWAS may identify potential associations of IL6R with potential drug targets, the correlation between the effect size of a genetic variant with a phenotype can differ from the actual treatment effect.4,60

Conclusions

In conclusion, our study highlighted the PheWAS as a promising approach leveraging clinical EHR data to query potential effects of new therapeutics in cases in which the drug mechanism of action has a clear link with a genetic variant. As a proof-of-concept study, we performed a PheWAS using a genetic variant with biochemical effects similar to a known therapy, tocilizumab and sarilumab. The strongest association observed was a reduced risk of aortic aneurysms, which was in line with the newest indication for IL6R blockade in GCA, in which aortic aneurysm can be a presenting sign. The findings of associations with different subphenotypes of aneurysms, such as AAA and aneurysm of the iliac artery, suggest that the beneficial effects of IL6R antagonist therapy may extend beyond treatment for aneurysms associated with large vessel vasculitis. Additionally, the IL6R PheWAS identified expected drug effects, such as reduction in CRP level, and off-target effects, including atopic dermatitis. Our findings support previous studies of reduced risk for coronary heart disease with IL6R blockade, and ongoing clinical trials may validate these findings. Importantly, this study also demonstrated the important role of biobanks with freely available data, such as UK Biobank and BioVU, as resources that can help to catalyze studies for the research community.
Association of Interleukin 6 Receptor Variant With CVD Effects of IL6R Blocking Therapy

Research Original Investigation

One Million Veteran Program participants and staff.

REFERENCES


Association of Interleukin 6 Receptor Variant With CVD Effects of IL6R Blocking Therapy

Original Investigation Research

Rheum Dis of a randomised, placebo-controlled study of interleukin-6 receptor blockade on surrogates of 40 54(9):2817-2829. doi:10.1002/art.22033

response to methotrexate. A double-blind randomised controlled clinical trial of interleukin-6 receptor antagonist, tocilizumab, in European patients with rheumatoid arthritis who had an incomplete response to methotrexate. Arthritis Rheum. 2006; 54(9):2817-2829. doi:10.1002/art.22033


