Published in final edited form as:

Arthritis Care Res (Hoboken). 2018 November ; 70(11): 1630–1636. doi:10.1002/acr.23553.

Phenome-wide association studies uncover a novel association of increased atrial fibrillation in males with systemic lupus erythematosus

April Barnado, MD, MSCI¹, Robert J. Carroll, PhD², Carolyn Casey, DO³, Lee Wheless, MD, PhD⁴, Joshua C. Denny, MD, MS^{1,2}, and Leslie J. Crofford, MD¹

¹Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

²Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

³Department of Medicine, Lehigh Valley Health Network, Allentown, Pennsylvania, USA

⁴Department of Dermatology, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Abstract

Objective—Phenome-wide association studies (PheWAS) scan across billing codes in the electronic health record (EHR) and re-purpose clinical EHR data for research. We examined if PheWAS could function as an EHR discovery tool for SLE and uncover novel clinical associations in male versus female SLE patients.

Methods—We used a de-identified version of the Vanderbilt University Medical Center EHR with over 2.8 million subjects. We performed PheWAS comparing 1) SLE cases vs. age, sex, and race-matched controls and 2) male vs. female SLE cases and controlled for multiple testing using a false discovery rate (FDR) of 0.05.

Results—We identified 1097 SLE subjects and 5735 matched controls. Comparing SLE cases to matched controls, SLE cases were more likely to have codes for SLE disease criteria. In the male vs. female SLE case-only PheWAS adjusting for age and race, males were more likely to have atrial fibrillation (AF) OR = 4.50 (FDR $p = 3.23 \times 10^{-3}$). Chart reviewed confirmed AF with the majority of subjects developing AF after SLE diagnosis and having multiple risk factors for AF. Adjusting for age, sex, race, and coronary artery disease (CAD), SLE disease was significantly associated with AF ($p = 0.002$).

Conclusion—Using PheWAS to compare males vs. females with SLE, we uncovered a novel association of increased AF in SLE males. SLE disease status was independently associated with AF even after adjusting for age, sex, race, and CAD. These results demonstrate the utility of PheWAS as an EHR discovery tool for SLE.

Corresponding Author: April Barnado, MD, MSCI, Address: 1161 21st Avenue South, T3113 MCN, Nashville, TN 37232, Telephone: (615) 322 – 4746, Fax: (615) 322 – 6248, april.barnado@vanderbilt.edu.

Conflict of Interest: none

Keywords

systemic lupus erythematosus; electronic health records; phenome-wide association study

Introduction

Electronic health records (EHR) serve as an efficient and cost-effective discovery tool to complement cohort and administrative database studies (1–3). One method that repurposes clinical EHR data for research is called phenome-wide association studies (PheWAS). PheWAS scan across billing codes in the EHR similar to a genome-wide association study (GWAS) scanning across the genome. PheWAS have replicated hundreds of known phenotype genotype associations (4–7) and also uncovered novel genetic associations in multiple autoimmune diseases (4, 5). PheWAS have been expanded to look at associations of phenotypes with autoantibodies in RA (8–10). PheWAS findings have been validated using multiple EHRs and with orthogonal methods (5–7, 11, 12)

Systemic lupus erythematosus (SLE) studies typically involve a single-center cohort with 100 to 1,000 SLE patients. These studies can take years to conduct and can be expensive to retain patients. SLE is also studied using administrative databases with limitations in identifying SLE patients by relying solely on billing codes, which may not accurately capture SLE patients (13, 14). Further, dense data on a patient's labs and SLE disease course are often lacking. These limitations have also impacted studies comparing the disease course of male and female SLE patients. While SLE is more common in females, males may have a more accelerated disease course with increased renal disease and mortality (15–22). Studies are limited with low numbers of male SLE subjects and conflicting findings (23–26). Further, many studies only examine for differences in ACR SLE criteria and not other important comorbidities such as cardiovascular disease. Whether these comorbidities impact the disease course in males with SLE is not well understood.

To the best of our knowledge, an EHR-based PheWAS has not been conducted in SLE. In this proof-of-concept study, we sought to demonstrate that PheWAS could function as a discovery tool for SLE. We also used PheWAS to examine differences in comorbidities in males compared to females with SLE.

Patients and Methods

Study Population

We identified potential SLE subjects in Vanderbilt's Synthetic Derivative (SD) (3) after approval from the Institutional Review Board of Vanderbilt University Medical Center (VUMC). VUMC is a regional, tertiary care medical center. The SD is a de-identified version of the EHR that contains over 2.8 million subjects with longitudinal data over several decades. The SD contains all available information in the EHR such as diagnostic and procedure codes, demographics, text from inpatient and outpatient notes (including both subspecialty and primary care), laboratory values, radiology reports, and medication orders.

The SD is composed of approximately equal numbers of males and females who are predominantly Caucasian (81%), reflecting the patient population of VUMC.

We used our validated EHR algorithm to identify SLE patients within the SD. This previously described algorithm (14) uses 4 counts of the SLE ICD-9 code (710.0) and a positive anti-nuclear antibody (ANA) with a titer of 1:160 while excluding ICD-9 codes for systemic sclerosis (SSc) (710.1) and dermatomyositis (DM) (710.3). This algorithm has a positive predictive value (PPV) of 89% and a sensitivity of 86%.

Non-SLE controls were defined as subjects within the SD who did not have ICD-9 codes under the 710.* heading of “Diffuse diseases of connective tissue” or 714.* heading of “Rheumatoid Arthritis and other inflammatory polyarthropathies” or ICD-10 codes under M05.* (“Rheumatoid Arthritis with rheumatoid factor”), M06.* (“Other rheumatoid arthritis”), M32 (“SLE”), M33.* (“Dermatopolymyositis”), M34.* (“Systemic sclerosis”), M35.* (“Other systemic involvement of connective tissue”), and M36.* (“Systemic disorders of connective tissue in diseases classified elsewhere”). Controls were matched on age (± 5 years), race, and sex in a 5:1 ratio to the SLE cases to maximize power while maintaining close matching. Controls were selected as “medical home patients” who received their longitudinal care at VUMC (27), defined as at least 3 outpatient visits within 5 years, to ensure adequate density of records similar to the SLE cases. Controls were not matched on number of visits to ensure controls were healthy. Clinical variables included current age, defined as age at time of data analysis, age at first use of the SLE ICD-9 code, sex, and race. Race was a mixture of self-report and administrative entry. Prior studies have validated that these EHR race assignments reflect self-report and genetic ancestry (28). As there were very low numbers for Hispanic and Asian subjects, these 2 groups were combined as “other” and counted as a third race group in addition to Caucasians and African Americans for the PheWAS.

Phenome-wide association studies and statistics

In PheWAS, the 18,000 ICD-9 codes are collapsed into 1,800 PheWAS codes that represent distinct clinical diagnoses. The listing of ICD-9 codes that are mapped to PheWAS codes is available at <http://phewascatalog.org>. To be considered a case, a subject has to have at least 2 instances of the PheWAS code on different days, at least 1 day apart. A subject is a control if there are no instances of the ICD-9 code. Subjects having only 1 instance of the code are excluded to eliminate the possibility of coding errors or preliminary diagnoses that are ultimately ruled out (29). For a PheWAS code to be used, there must be at least 20 cases (4, 5). For each PheWAS code, an unconditional logistic regression model is created with odds ratios (ORs) and 95% confidence intervals (95% CI) reported. For the most significant PheWAS code identified by this screening, conditional logistic regression models of SLE cases and matched controls were then constructed to calculate an OR and 95% CI for the association between the code and SLE disease status adjusting for age, sex, and race. Analyses were performed and graphed in the PheWAS package (29) in R version 3.2.5. We performed PheWAS comparing 1) SLE cases vs. matched controls and 2) SLE males vs. females. For the SLE male vs. female PheWAS, we ran adjusted models for current age and race. To adjust for multiple hypothesis testing, we used a false discovery rate (FDR) p of

0.05. There were 900 testable phenotypes for the SLE cases vs. controls and 398 for the SLE male vs. female PheWAS. We assessed for demographic differences in males vs. females with SLE using the Mann-Whitney U test for continuous variables, as there were non-normal distributions in the data, and chi-square or Fisher's exact test for categorical variables. Two-sided p values < 0.05 were considered to indicate statistical significance. Analyses were conducted using IBM SPSS software, version 24.0 (SPSS).

Chart Review

PheWAS does not account for the time relationship between ICD-9 codes, so the timing between codes has to be assessed on chart review. Manual chart review was performed for the most significant association found in the SLE male vs. female PheWAS to confirm that the PheWAS code for atrial fibrillation (AF) captured a true clinical diagnosis of AF. Both inpatient and outpatient notes and electrocardiogram (ECG) reports were reviewed, and a subject was classified as a case if a physician or an ECG documented AF. We also reviewed SLE disease manifestations, AF risk factors, timing of AF relative to SLE diagnosis, and corticosteroid use.

Results

Study Population

Using our validated algorithm (14), we identified 1097 SLE cases. These SLE cases have been described previously (14). Briefly, the SLE cases were predominantly female (90%) and Caucasian (69%) with mean age at first SLE ICD-9 code of 40 years, a mean 9 years of EHR follow-up, and a mean 20 counts of the SLE ICD-9 code. We then identified 5735 control subjects who were matched in a 5:1 ratio with age (± 5 years), sex, and race to the 1097 SLE cases. SLE cases and controls had similar mean current age (51 ± 17 vs. 50 ± 17 , $p = 0.84$), sex (90% female, $p = 1.00$), and race (65% Caucasian, 25% African American, 3% Hispanic, 2% Asian, and 5% Unknown, $p = 1.00$) with SLE cases having longer years of follow-up in the EHR compared to controls (9.2 ± 6.0 vs. 7.6 ± 6.2 , $p < 0.001$).

SLE case vs. control PheWAS

As a proof-of-concept study, we compared the 1097 SLE cases to the 5735 matched controls to validate PheWAS as a discovery tool in the EHR for SLE. As expected, SLE cases were more likely to have codes related to ACR SLE criteria (28). Specifically, SLE cases had codes for nephritis odds ratio (OR) = 44.18 (95% CI 30.88 – 63.20, FDR $p = 7.97 \times 10^{-93}$), pain in joint OR = 2.89 (2.50 – 3.35, FDR $p = 2.20 \times 10^{-44}$), thrombocytopenia OR = 7.45 (5.58 – 9.96, FDR $p = 2.13 \times 10^{-40}$), pleurisy; pleural effusion OR = 4.80 (3.80 – 6.06, FDR $p = 3.20 \times 10^{-38}$), and pericarditis OR = 16.61 (10.73 – 25.73, FDR $p = 7.49 \times 10^{-35}$) (Supplemental Figure 1, Supplemental Table 1). Compared to matched controls, SLE cases were also more likely to have codes related to cardiovascular disease and infections (Supplemental Table 1, Supplemental Figure 2).

Male vs. Female SLE PheWAS

Using PheWAS, we compared males ($n = 111$) vs. females ($n = 986$) with SLE. Both males and females had similar mean current age (52 ± 18 vs. 50 ± 17 , $p = 0.42$) and age at first

SLE ICD-9 code (43 ± 18 vs. 40 ± 17 , $p = 0.09$) with similar race breakdown (Table 1). Males and females also had similar mean years of EHR follow-up (8 ± 5 vs. 9 ± 5 , $p = 0.23$) and number of counts of the SLE ICD-9 code (17 ± 15 , 19 ± 22 , $p = 0.73$). Figure 1 shows the male vs. female PheWAS adjusting for current age and race. Compared to females, males were more likely to have 4 codes that met the FDR p of 0.05 including atrial fibrillation OR = 4.50 (95% CI 2.32 – 8.72, FDR $p = 3.23 \times 10^{-3}$), other chronic ischemic heart disease unspecified OR = 4.40 (2.24 – 8.64, FDR $p = 3.23 \times 10^{-3}$), atrial fibrillation and flutter OR = 4.08 (2.12 – 7.83, FDR $p = 3.23 \times 10^{-3}$), and coronary atherosclerosis OR = 3.19 (1.73 – 5.89, FDR $p = 0.02$). In a sensitivity analysis, excluding 4 SLE subjects who developed AF > 1 year prior to SLE did not change these results.

Atrial fibrillation chart review

Of the 51 SLE subjects with an atrial fibrillation (AF) code, all 51 subjects on chart review had AF documented by a physician or on an ECG. Of the 51, 34 were female and 17 male. Comparing females and males, there was a similar mean age at first SLE code (49 ± 18 vs. 56 ± 20 , $p = 0.14$) and mean age at first AF code (54 ± 18 vs. 57 ± 20 , $p = 0.40$). Both females and males with SLE and AF were predominantly Caucasian (74% vs. 71%, $p = 0.82$). When comparing Caucasian ($n = 37$) and African Americans ($n = 14$) with SLE and AF, the African Americans had a younger mean age at first SLE code (35 ± 16 vs. 57 ± 16 , $p < 0.001$) and age at first AF code (38 ± 14 vs. 61 ± 16 , $p < 0.001$).

For timing of AF relative to SLE, the majority of SLE subjects developed SLE before AF ($n = 28$). Of the remaining subjects, 19 developed SLE and AF within the same year, and 4 developed SLE after AF (Figure 2). On average, SLE subjects had their first SLE code 3.5 ± 4.9 years before their first AF code. Females had on average their first SLE code 4.9 ± 4.9 years prior to their first AF code compared to 0.65 ± 3.2 years for males ($p = 0.005$). On chart review of the 19 subjects who had a SLE and AF code within the same year, 5 had AF during an initial presentation of SLE with 4 having serositis and 1 neuropsychiatric manifestations. Other variables examined included chronic vs. paroxysmal AF, corticosteroid use, hyperthyroidism, and risk factors for AF. Of the 51 SLE AF subjects, only 2 had hyperthyroidism, which developed after AF. Only 3 subjects (6%) had chronic AF. There were 5 patients with a history of antiphospholipid antibody syndrome with 1 patient having active arterial thromboses at time of AF. Of the 41 with medication data at time of AF, 21 (51%) were on corticosteroids with only 2 receiving high doses of corticosteroids, specifically methylprednisolone 1000 mg. For risk factors for AF, 12% (6/51) were current smokers, 37% (19/51) had a history of CAD at AF diagnosis, 86% (44/51) hypertension, 27% (14/51) diabetes mellitus, and 67% (34/51) a body mass index (BMI) ≥ 25 . On average, SLE subjects had 2 ± 1 risk factors. When comparing these risk factors in the SLE males vs. females, there was a higher proportion of males with CAD at AF diagnosis compared to females (59% vs. 27%, $p = 0.02$) with other risk factors not significantly different.

When comparing SLE cases to the matched controls, SLE cases were more likely to have codes for AF with an OR of 1.86 (95% CI 1.35 – 2.57, FDR $p = 6.52 \times 10^{-4}$). As SLE cases had significantly longer EHR follow-up compared to controls, we conducted a sensitivity

analysis adjusting for years of EHR follow-up to see if the longer follow-up could account for the higher risk of AF. Adjusting for years of follow-up, SLE cases were still more likely to have AF compared to matched controls (OR = 1.86, 95% CI 1.34 – 2.58, FDR $p = 7.16 \times 10^{-4}$).

In a conditional logistic regression of SLE cases and matched controls adjusting for age, sex, and race, SLE disease status was significantly associated with AF with an OR of 1.63 (95% CI 1.34 – 1.97, $p = 9.0 \times 10^{-7}$). CAD was also significantly associated with AF with an OR of 10.28 (95% CI 7.96 – 13.28, $p < 0.001$). After adjusting for CAD, SLE disease status remained significantly associated with AF with an OR of 1.45 (95% CI 1.15 – 1.83, $p = 0.002$). HTN was also significantly associated with AF (OR = 2.85, 95% CI 2.19 – 3.69, $p = 3.4 \times 10^{-15}$). SLE remained associated with AF even after adjusting for HTN, CAD, age, sex, and race ($p = 0.007$).

In a sensitivity analysis adjusting for years of follow-up, SLE remained significantly associated with AF (OR = 1.46, 95% CI 1.16 – 1.85, $p = 0.002$). A sensitivity analysis was also performed where subjects having AF > 1 year prior to SLE diagnosis were excluded. Both SLE (OR = 1.44, 95% CI 1.14 – 1.81, $p = 0.002$) and CAD (10.42, 95% CI 7.58 – 14.31, $p < 0.001$) remained significantly associated with AF.

Discussion

These data represent the first use of phenome-wide association studies (PheWAS) in SLE. PheWAS provide a robust, comprehensive approach to scan the entire EHR and systemically evaluate dense, longitudinal data. Using PheWAS to compare SLE cases to matched controls, SLE cases, as expected, had more codes for ACR SLE criteria (30) validating that PheWAS can function as an EHR discovery tool for SLE. We then used PheWAS to compare comorbidities in males and females with SLE and found a novel clinical association of increased atrial fibrillation (AF) in males. We demonstrated that PheWAS could not only identify SLE disease manifestations and important comorbidities in the EHR but also uncover novel clinical associations.

In the PheWAS comparing males vs. females with SLE, we uncovered a new clinical association of increased AF in males with SLE. In the United States, AF is more common in males compared to females with a strong age effect. For example, the OR for males vs. females at age 15 to 44 years is approximately 3 vs. an OR of 1 for age ≥ 85 years (31). Our unadjusted OR for SLE males vs. females with AF was 4.88, very similar to our age and race adjusted OR of 4.50. Therefore, age and race were not the sole drivers of our findings. We used chart review to further investigate the increased AF in males. Males and females with SLE and AF were not significantly different in terms of duration of EHR follow-up, age at first AF code, or race. Comparing Caucasians vs. African Americans with SLE and AF, African Americans were significantly younger at time of first AF and SLE codes. African Americans with SLE and AF had a strikingly younger mean age of onset of AF of 38 years compared to 68 years in the general population (32). This younger age of AF onset could result from an earlier accumulation of risk factors for AF including hypertension and CAD compared to the general population.

To investigate the increased AF in males vs. females with SLE, we examined the background signal for AF in our EHR. SLE cases were more likely to have codes for AF compared to matched controls with an OR of 1.86 (FDR $p = 6.52 \times 10^{-4}$). We hypothesize that this increased risk of AF in SLE is due to comorbidities that may be more common in SLE patients compared to matched controls due to SLE disease activity, damage, and side effects from immunosuppressants. Looking at AF risk factors in our 51 SLE AF subjects, the majority of subjects had multiple risk factors such as hypertension, obesity, and CAD. Notably, compared to females, males with SLE were more than twice as likely to have CAD at time of AF diagnosis suggesting that CAD may account, at least partly, for the increased AF in the males. SLE case status was significantly associated with AF in a conditional logistic regression model adjusting for age, sex, and race (OR = 1.63, 95% CI 1.34 – 1.97, $p = 9.0 \times 10^{-7}$). This association was largely unchanged even after adding CAD and HTN to the model, suggesting that while traditional AF risk factors such as CAD and HTN play a role in AF, SLE-related factors may also independently contribute.

For SLE-related factors, case reports show that SLE patients may develop AF secondary to pericarditis or myocarditis (33). Our results support these observations with 4 of our 51 SLE AF subjects developing AF at the time of pericarditis. Case reports also show that SLE patients develop AF after high-dose methylprednisolone therapy (33). In our 51 SLE AF subjects, only 2 were receiving high-dose methylprednisolone. One study hypothesized that elevated norepinephrine could play an etiologic role in a wide range of diseases including both SLE and AF (34). While studies have found elevated norepinephrine signaling separately in SLE and AF, there are no studies looking at norepinephrine signaling in patients with both SLE and AF.

Our study has important clinical implications highlighting that clinicians should not only be aware of the increased risk of CAD in SLE but also AF, particularly in patients with a history of CAD. Recognizing this increased risk of AF could prompt more expedited referrals and testing that result in earlier diagnoses. Specifically, recognizing this increased risk could prompt clinicians to evaluate for AF in common clinical scenarios such as palpitations or dyspnea where clinicians may not normally test for AF or may instead attribute these symptoms to SLE manifestations such as serositis. As AF is associated with increased risks of CVD, heart failure, and death, earlier diagnosis and treatment may reduce these risks (31). Further studies are needed to confirm this increased risk of AF in SLE and its association with CAD. This study also highlights that males with SLE and CAD may represent a particularly vulnerable patient group. Studies have mainly focused on increased CAD in SLE females (35, 36) but very few have focused on SLE males (37), as they are often underrepresented in studies. Larger studies are needed to assess CAD in SLE males.

While PheWAS have confirmed both known epidemiologic associations and uncovered novel clinical associations in SLE, there are limitations to our study. We used an internally validated algorithm to identify SLE patients with a PPV of 89% and a sensitivity of 86%. Although this algorithm has strong test characteristics, there is a possibility that some of our SLE patients may not have a SLE diagnosis. Another limitation of the clinical EHR data is that disease activity and damage measures such as the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) (38) and Systemic Lupus International Collaborating

Clinics Damage Index (SDI) (39) are not routinely collected in clinical settings and thus we cannot adjust for disease activity or damage. Next, PheWAS codes only capture billing codes at Vanderbilt. With healthcare fragmentation, patients can receive care in multiple systems, which may not be documented at Vanderbilt. These potential missed diagnoses, however, would generally bias us to the null result, unless missing data was non-randomly distributed such that more missing data occurred in the controls, who had less follow-up time in the EHR compared to the SLE cases. We did estimate the AF prevalence in our controls was similar to, if not higher than, the general population suggesting, that under-ascertainment of AF in our controls was unlikely. Lastly, our study was performed at a single institution's EHR potentially limiting generalizability of our results to other groups of SLE patients. Using an EHR-based cohort to study SLE, however, captures a wider net of SLE patients that may more accurately reflect on community patients compared to SLE patients that are recruited into a cohort.

EHR data is a rich data source for conducting discovery research in relatively rare diseases such as SLE. PheWAS is a robust tool that can comprehensively scan across the EHR. We used the EHR to assemble a large SLE cohort and perform the first PheWAS in SLE. The PheWAS comparing males vs. females with SLE uncovered a new association of increased atrial fibrillation in males. The PheWAS comparing SLE cases to matched controls demonstrated that PheWAS could identify SLE disease manifestations and capture the burden of cardiovascular disease in SLE patients. Our results demonstrate the feasibility and potential of PheWAS for discovery research in SLE.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by grants NIH/NICHD 5K12HD043483-12 (Barnado), NLM R01-LM010685 (Denny), NCRR/NIH UL1 RR024975, NCATS/NIH ULTR000445

References

1. Pathak J, Kho AN, Denny JC. Electronic health records-driven phenotyping: challenges, recent advances, and perspectives. *J Am Med Inform Assoc.* 2013; 20(e2):e206–11. [PubMed: 24302669]
2. Ritchie MD, Denny JC, Crawford DC, Ramirez AH, Weiner JB, Pulley JM, et al. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. *Am J Hum Genet.* 2010; 86(4):560–72. [PubMed: 20362271]
3. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther.* 2008; 84(3):362–9. [PubMed: 18500243]
4. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics.* 2010; 26(9):1205–10. [PubMed: 20335276]
5. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol.* 2013; 31(12):1102–10. [PubMed: 24270849]
6. Denny JC, Crawford DC, Ritchie MD, Bielinski SJ, Basford MA, Bradford Y, et al. Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical

records for genome- and phenome-wide studies. *Am J Hum Genet.* 2011; 89(4):529–42. [PubMed: 21981779]

7. Hebring SJ, Schrodi SJ, Ye Z, Zhou Z, Page D, Brilliant MH. A PheWAS approach in studying HLA-DRB1*1501. *Genes Immun.* 2013; 14(3):187–91. [PubMed: 23392276]
8. Doss J, Mo H, Carroll RJ, Crofford LJ, Denny JC. Phenome-Wide Association Study of Rheumatoid Arthritis Subgroups Identifies Association between Seronegative Disease and Fibromyalgia. *Arthritis Rheumatol.* 2016
9. Liao KP, Kurreeman F, Li G, Duclos G, Murphy S, Guzman R, et al. Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls. *Arthritis Rheum.* 2013; 65(3):571–81. [PubMed: 23233247]
10. Liao KP, Sparks JA, Hejblum BP, Kuo IH, Cui J, Lahey LJ, et al. Phenome-wide association study of autoantibodies to citrullinated and non-citrullinated epitopes in rheumatoid arthritis. *Arthritis Rheumatol.* 2016
11. Denny JC, Bastarache L, Roden DM. Phenome-Wide Association Studies as a Tool to Advance Precision Medicine. *Annu Rev Genomics Hum Genet.* 2016; 17:353–73. [PubMed: 27147087]
12. Bush WS, Oetjens MT, Crawford DC. Unravelling the human genome-phenome relationship using phenome-wide association studies. *Nat Rev Genet.* 2016; 17(3):129–45. [PubMed: 26875678]
13. Moores KG, Sathe NA. A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data. *Vaccine.* 2013; 31(Suppl 10):K62–73. [PubMed: 24331075]
14. Barnado A, Casey C, Carroll RJ, Wheless L, Denny JC, Crofford LJ. Developing Electronic Health Record Algorithms that Accurately Identify Patients with Systemic Lupus Erythematosus. *Arthritis Care Res (Hoboken).* 2016
15. Jonsson H, Nived O, Sturfelt G. Outcome in systemic lupus erythematosus: a prospective study of patients from a defined population. *Medicine (Baltimore).* 1989; 68(3):141–50. [PubMed: 2785628]
16. Pistiner M, Wallace DJ, Nessim S, Metzger AL, Klinenberg JR. Lupus erythematosus in the 1980s: a survey of 570 patients. *Semin Arthritis Rheum.* 1991; 21(1):55–64. [PubMed: 1948102]
17. Reveille JD, Bartolucci A, Alarcon GS. Prognosis in systemic lupus erythematosus. Negative impact of increasing age at onset, black race, and thrombocytopenia, as well as causes of death. *Arthritis Rheum.* 1990; 33(1):37–48. [PubMed: 2302266]
18. Prete PE, Majlessi A, Gilman S, Hamideh F. Systemic lupus erythematosus in men: a retrospective analysis in a Veterans Administration Healthcare System population. *J Clin Rheumatol.* 2001; 7(3):142–50. [PubMed: 17039119]
19. Garcia MA, Marcos JC, Marcos AI, Pons-Estel BA, Wojdyla D, Arturi A, et al. Male systemic lupus erythematosus in a Latin-American inception cohort of 1214 patients. *Lupus.* 2005; 14(12): 938–46. [PubMed: 16425573]
20. Crosslin KL, Wiginton KL. Sex differences in disease severity among patients with systemic lupus erythematosus. *Gend Med.* 2011; 8(6):365–71. [PubMed: 22153881]
21. Schwartzman-Morris J, Putterman C. Gender differences in the pathogenesis and outcome of lupus and of lupus nephritis. *Clin Dev Immunol.* 2012; 2012:604892. [PubMed: 22690240]
22. Andrade RM, Alarcon GS, Fernandez M, Apte M, Vila LM, Reveille JD, et al. Accelerated damage accrual among men with systemic lupus erythematosus: XLIV. Results from a multiethnic US cohort. *Arthritis Rheum.* 2007; 56(2):622–30. [PubMed: 17265497]
23. Mok CC, Lau CS, Chan TM, Wong RW. Clinical characteristics and outcome of southern Chinese males with systemic lupus erythematosus. *Lupus.* 1999; 8(3):188–96. [PubMed: 10342711]
24. Stefanidou S, Benos A, Galanopoulou V, Chatziyannis I, Kanakoudi F, Aslanidis S, et al. Clinical expression and morbidity of systemic lupus erythematosus during a post-diagnostic 5-year follow-up: a male:female comparison. *Lupus.* 2011; 20(10):1090–4. [PubMed: 21700658]
25. Voulgari PV, Katsimbri P, Alamanos Y, Drosos AA. Gender and age differences in systemic lupus erythematosus. A study of 489 Greek patients with a review of the literature. *Lupus.* 2002; 11(11): 722–9. [PubMed: 12475002]

26. Tan TC, Fang H, Magder LS, Petri MA. Differences between male and female systemic lupus erythematosus in a multiethnic population. *J Rheumatol*. 2012; 39(4):759–69. [PubMed: 22382348]
27. Schildcrout JS, Denny JC, Bowton E, Gregg W, Pulley JM, Basford MA, et al. Optimizing drug outcomes through pharmacogenetics: a case for preemptive genotyping. *Clin Pharmacol Ther*. 2012; 92(2):235–42. [PubMed: 22739144]
28. Dumitrescu L, Ritchie MD, Brown-Gentry K, Pulley JM, Basford M, Denny JC, et al. Assessing the accuracy of observer-reported ancestry in a biorepository linked to electronic medical records. *Genet Med*. 2010; 12(10):648–50. [PubMed: 20733501]
29. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics*. 2014; 30(16):2375–6. [PubMed: 24733291]
30. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1997; 40(9):1725.
31. Writing Group M, Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, et al. Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation*. 2016; 133(4):e38–360. [PubMed: 26673558]
32. Golwala H, Jackson LR 2nd, Simon DN, Piccini JP, Gersh B, Go AS, et al. Racial/ethnic differences in atrial fibrillation symptoms, treatment patterns, and outcomes: Insights from Outcomes Registry for Better Informed Treatment for Atrial Fibrillation Registry. *Am Heart J*. 2016; 174:29–36. [PubMed: 26995367]
33. Aslam AK, Vasavada BC, Sacchi TJ, Khan IA. Atrial fibrillation associated with systemic lupus erythematosus and use of methylprednisolone. *Am J Ther*. 2001; 8(4):303–5. [PubMed: 11441330]
34. Fitzgerald PJ. Elevated norepinephrine may be an etiological factor in a wide range of diseases: age-related macular degeneration, systemic lupus erythematosus, atrial fibrillation, metabolic syndrome. *Med Hypotheses*. 2013; 80(5):558–63. [PubMed: 23410497]
35. Manzi S, Meilahn EN, Rairie JE, Conte CG, Medsger TA Jr, Jansen-McWilliams L, et al. Age-specific incidence rates of myocardial infarction and angina in women with systemic lupus erythematosus: comparison with the Framingham Study. *Am J Epidemiol*. 1997; 145(5):408–15. [PubMed: 9048514]
36. Bruce IN, Urowitz MB, Gladman DD, Ibanez D, Steiner G. Risk factors for coronary heart disease in women with systemic lupus erythematosus: the Toronto Risk Factor Study. *Arthritis Rheum*. 2003; 48(11):3159–67. [PubMed: 14613278]
37. Riveros Frutos A, Casas I, Rua-Figueroa I, Lopez-Longo FJ, Calvo-Alen J, Galindo M, et al. Systemic lupus erythematosus in Spanish males: a study of the Spanish Rheumatology Society Lupus Registry (RELESSER) cohort. *Lupus*. 2017; 26(7):698–706. [PubMed: 27799439]
38. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. *Arthritis Rheum*. 1992; 35(6):630–40. [PubMed: 1599520]
39. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum*. 1996; 39(3):363–9. [PubMed: 8607884]

Significance and Innovations

- This is the first study to use phenome-wide association studies (PheWAS) to perform discovery research in systemic lupus erythematosus (SLE) using the electronic health record (EHR).
- In this proof-of-concept study, we demonstrate that PheWAS could identify SLE disease manifestations in the EHR validating PheWAS as a discovery tool.
- Using PheWAS to compare males and females with SLE, we report the first association of increased atrial fibrillation (AF) in males with SLE.
- SLE disease status was independently associated with AF even after adjusting for age, sex, race, and coronary artery disease.

Figure 1. PheWAS of males vs. females with SLE

The x axis represents the PheWAS codes that are mapped to ICD-9 codes organized and color-coded by organ system. The y axis represents the level of significance. The PheWAS is adjusted for current age and race. Each triangle represents a PheWAS code. Females are the reference group. Triangles pointing up represent codes more common in males. Triangles pointing down represent codes more common in females. The horizontal red line represents the false discovery rate (FDR) of 0.05. The four codes that met this level of significance are shown above including atrial fibrillation, other chronic ischemic heart disease, atrial fibrillation and flutter, and coronary atherosclerosis.

Figure 2. Histogram displaying number of years between first systemic lupus erythematosus (SLE) code and first atrial fibrillation (AF) code

The x axis shows the difference in years from the first SLE code to first AF code, and the y axis shows the number of SLE AF subjects. Years < 0 represent number of years SLE was coded before AF. Years > 0 represent number of years SLE was coded after AF. There are 51 SLE AF subjects represented. Subjects on average had a first SLE code 3.5 ± 4.9 years before the first AF code.

Table 1Characteristics of males vs. females with SLE[†].

Demographics	Males (n = 111)	Females (n = 986)	p value [*]
Current age, mean $\pm$ SD	52 $\pm$ 18	50 $\pm$ 17	p = 0.42
Age at first SLE ICD-9 code, mean $\pm$ SD	43 $\pm$ 18	40 $\pm$ 17	p = 0.09
Race (%)			
Caucasian	69%	69%	p = 0.32
African American	27%	26%	
Asian	4%	3%	
Hispanic	0%	2%	
Years of follow-up in the EHR, mean $\pm$ SD	8 $\pm$ 5	9 $\pm$ 5	p = 0.23
Number of counts of the SLE ICD-9 code (710.0), mean $\pm$ SD	17 $\pm$ 15	19 $\pm$ 22	p = 0.73

[†]Used 4 or more counts of the SLE ICD-9 code (710.0) and ANA positive 1:160 while excluding ICD-9 codes for systemic sclerosis (710.1) and dermatomyositis (710.3) to identify potential patients with SLE.

^{*}Mann-Whitney U for continuous variables, Fisher's exact test for categorical variables.