APPLICATION OF TEXT MINING METHODS TO IDENTIFY LUPUS NEPHRITIS FROM ELECTRONIC HEALTH RECORDS

Background Lupus nephritis (LN), or chronic inflammation of the kidneys, is a frequent complication of SLE and associated with higher overall morbidity and mortality. Previous studies have found that LN affects individuals of Hispanic, African American and Asian descent more than those of European ancestry. Accurate estimates of LN in the population remain limited due to the inability to capture this information through structured data fields in electronic health records (EHRs). We aimed to develop a text mining pipeline to extract information on LN diagnosis within clinical notes in EHRs. We sought to validate this administrative algorithm in comparison to SLEDAI scores at clinical visits.

Methods Individuals with a single diagnosis code for SLE (710.0) during June 1, 2012 November 5, 2016 from the EHR of a university health system were included (n=2,509). All clinical notes for patients were extracted and annotated using a clinical text-mining tool, the Clinical Event Recognizer (CLEVER), and a custom-built dictionary that included lupus nephritis and associated terms. Positive and negative mentions of LN were tagged and evaluated for performance.

Results Over 1300 positive and 400 negative mentions of LN were detected from clinical notes. Manual review of 50 note mentions (25 positive and 25 negative) determined that our text-mining tool detected LN with 79% sensitivity and 86% specificity.

Conclusions We conducted the first text-mining strategy to extract LN status from clinical notes. Additional evaluation of the LN text-mining pipeline on a gold-standard chart-reviewed cohort of SLE patients (n=332) is ongoing. Further refinement of the pipeline will allow us to classify stage of LN severity vs. highest SLEDAI-2K.

Abstract 187 Table 1 Number of positive and negative mentions of lupus nephritis identified through the text-mining tool and manual review.

<table>
<thead>
<tr>
<th>Text-Mining Tool</th>
<th>Manual Review</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>22</td>
</tr>
<tr>
<td>Negative</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
</tr>
</tbody>
</table>

Follow-up studies to determine factors associated with this important disease outcome.

Funding Source(s): NIH-NIAMS F32 AR070585 (Gianfrancesco)

VALIDATION OF AN ADMINISTRATIVE ALGORITHM TO ASSESS SLE DISEASE SEVERITY

Abstract 188 Table 1 Garris administrative algorithm for SLE severity vs. highest SLEDAI-2K

<table>
<thead>
<tr>
<th>Highest SLEDAI-2K</th>
<th>Mild vs. Moderate/Severe</th>
<th>Mild/Moderate vs. Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>63.9%</td>
<td>94.3%</td>
</tr>
<tr>
<td>Specificity</td>
<td>82.8%</td>
<td>36.7%</td>
</tr>
<tr>
<td>C-statistic</td>
<td>0.740</td>
<td>0.755</td>
</tr>
</tbody>
</table>

Background SLE severity is heterogeneous: some patients have mild disease with rashes and arthritis, while others have severe multi-organ system involvement. It is challenging to study SLE in administrative databases given this heterogeneity. Garris et al developed an administrative claims-based SLE severity algorithm derived from elements of the SLEDAI, SLAM and BILAG instruments (Garris, J Med Econ 2013). It employs ICD-9, CPT and NDC codes over a 1 year period and classifies patients as having mild, moderate or severe disease. We sought to validate this algorithm in comparison to SLEDAI scores at clinical visits.

Methods We identified 100 SLE patients followed at the Brigham and Women’s Hospital (BWH) Lupus Center (2008–2010) with SLEDAI-2K (Gladman, J Rheumatol 2002) data at each visit over a 1 year period per person. We also obtained ICD-9, CPT and NDC codes for the Garris algorithm items (e.g. codes for glucocorticoids, ICD-9 codes for periarticular) for the same year per subject. We compared Garris SLE severity to the highest SLEDAI-2K in that year. We defined the SLEDAI-2K categories of mild <3, moderate 3–6, and severe >6 as in the literature (Polacheck, Arthritis Care Res 2017). We compared classification in binary categories of mild vs. moderate/severe and mild/moderate vs. severe. For each, we calculated sensitivity, specificity, and C-statistics.

Results We analyzed 377 SLEDAI-2K assessments on 100 subjects (mean 3.77 [SD 2.63]) in the BWH Lupus Cohort. For the Garris vs. highest SLEDAI-2K model, 56 of 100 subjects were classified similarly by Garris and highest SLEDAI-2K (23/36 mild, 22/34 moderate, and 11/36 severe by SLEDAI-2K). The performance characteristics compared to the highest SLEDAI-2K of the year were: C-statistics were 0.755 for mild vs. moderate vs. severe SLE severity and 0.740 for mild vs. moderate/severe (table). Sensitivity of the Garris algorithm compared to the highest SLEDAI-2K were 63.9% for mild vs. moderate/severe and 94.3% for mild/moderate vs. severe.

Abstract 188 Table 1 Garris administrative algorithm for SLE severity vs. highest SLEDAI-2K

1Cameron B Speyer*, 1Daniel Li, 3Hongshu Guan, 1Kazuki Yoshida, 1April Jorge, 3Candace H Feldman, 2Karen H Costenbader. 1Division of Rheumatology, Immunology and Allergy; Brigham and Women’s Hospital; 2Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital

Funding Source(s): NIH-NIAMS F32 AR070585 (Gianfrancesco)
Background The epidemiology of definite antiphospholipid syndrome (APS) in the general population has not been described. A recent meta-analysis concluded that it was difficult to determine the frequency of a clinically significant antiphospholipid (aPL) profile in patients with aPL-related clinical outcomes due to the lack of robust data; only 4% of the studies had the current cutoff values for antiphospholipid antibodies (aCL) and less than one fifth of them had confirmation after 6–12 weeks. This study aimed to characterize the epidemiology of definite APS based on the 2006 updated international consensus (Sydney) classification criteria.

Methods An inception cohort of patients with incident APS in 2000–2015 in a geographically well-defined population were identified based on comprehensive individual medical record review. All cases met the definite 2006 Sydney consensus APS criteria, including the laboratory and clinical criteria as well as laboratory confirmation after 12 weeks. Lupus anticoagulant, IgM and IgG aCL and anti-2 glycoprotein-1 antibodies were tested in a centralized lab (cutoff >40 GPL/MPL). Incidence rates were age and sex adjusted to the US white 2010 population. Prevalence estimates were obtained from the incidence rates assuming no increased mortality associated with APL and assuming migration in/out of the area was independent of disease status.

Results In 2000–2015, 33 cases of incident APS by the Sydney criteria were identified (mean age 54.2 years, 55% female; 97% Caucasian). The annual incidence of APS was 2.1 (95% confidence interval [CI]: 1.4–2.8) per 1 000 000 population aged 18 years. Incidence rates were similar in both sexes. The estimated prevalence of APS was 50 per 1 000 000 (95% CI: 42–58) and was similar in both sexes. Six (18%) patients had a concurrent diagnosis of systemic lupus erythematosus (SLE). The most frequent clinical manifestation was deep venous thrombosis. The overall mortality of patients with APS was not significantly different from the general population (standardized mortality ratio: 1.61; 95% CI: 0.74–3.05).

Conclusions Results from this first ever population based study revealed that definite APS occurred in about 2 persons per 1 000 000 per year. The estimated prevalence is 50 per 1 000 000. Overall mortality was not different from the general population. The incidence and prevalence of APS in the same population was at least as common as SLE.

Funding Source(s): None

OPPOSITE PROFILES OF COMPLEMENT IN ANTIPHOSPHOLIPID SYNDROME (APS) AND SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) AMONG PATIENTS WITH ANTIPHOSPHOLIPID ANTIBODIES

1Chack-Yung Yu*, 1Stephanie Sawelli, 2Robert Rouby, 3Kathryn Kitzmiller, 1Danielle Zhou, 4Haikady Nagaraja, 1Evan Mulhern, 1Fatima Barbar-Smilej, 1Stacy P Arton, 1Vee-Ling Wu, 1Nationwide Children’s Hospital and The Ohio State University, 2Division of Rheumatology, Allergy and Immunology, The University of North Carolina at Chapel Hill, Chapel Hill, NC; 3Grand Canyon University (and The Ohio State University); 4The Ohio State University; 1Nationwide Children’s Hospital; 2Nationwide Children’s Hospital and The Ohio State University and the Loyola University Chicago

Background Antiphospholipoid syndrome (APS) is the association of antiphospholipid antibodies (aPL) with thromboses and/or recurrent pregnancy loss (RPL). Among patients with SLE, one-third have aPL and 10%–15% have a manifestation of secondary APS. Although mouse models suggested complement drives the pathogenesis of APS, we have little knowledge on how complement proteins and genes contribute to the pathology of human APS, and the concurrence of SLE and APS.

Methods We performed a cross-sectional study on complement proteins and genes in 525 patients with aPL. Among them, 237 experienced thromboses and 293 had SLE; 111 had both SLE and thromboses, and 106 had neither SLE nor thrombosis. Complement protein levels were determined by radial immunodiffusion for C4, C3 and factor H; and by functional ELISA for mannan binding lectin (MBL). Total C4, C4A and C4B gene copy numbers (GCN) were measured by TaqMan-based realtime PCR.

Results Two to six copies of C4 genes are frequently present in a diploid genome, and each copy may code for an acidic C4A or a basic C4B protein. We observed significantly (a) higher protein levels of total C4, C4A, C4B, C3 and anticoagulant (ACLA) IgG, (b) increased frequencies of lupus anticoagulant and males, and (c) depressed levels of complement factor H, MBL and ACLA-IgM among patients with thrombosis than those without thrombosis (n=288). We also observed significantly lower GCNs of total C4 and C4A among aPL-positive patients with both SLE and thrombosis than others. By contrast, aPL-positive subjects with SLE had significantly reduced protein levels of C3, total C4, C4A, C4B and ACLA-IgG, and higher frequency of females than those without SLE. Patients with thrombosis but without SLE (n=126), and patients with SLE but without thrombosis (n=182) had the greatest differences in mean protein levels of C3 (p=2.6×10–6), C4 (p=2.2×10–9) and ACLA-IgG (p=1.2×10–5). RPL occurred in 23.7% of female patients and thrombotic SLE patients had the highest frequency of RPL (41.0%; p=3.8×10–10). Compared with non-RPL females, RPL had significantly higher frequency of thrombosis and elevated C4 protein levels. Female patients with homozygous C4A deficiency all experienced RPL (p=0.0001) but the opposite was true for patients with homozygous C4B deficiency (p=0.017).

Conclusions There are substantial differences for complement protein concentrations and genetic diversities among aPL-positive patients with thromboses, recurrent pregnancy loss and SLE. These results are relevant for diagnosis and management of APS and SLE.

Funding Source(s): This work was supported by NIH grants 1R01 AR050078, 1R21070905, 1R01 AR073311 (CYY), and RR00046 (UNC).