

DR. PETER IZMIRLY (Orcid ID : 0000-0001-5445-2182)

DR. S. SAM LIM (Orcid ID : 0000-0003-2361-0787)

DR. CRISTINA DRENKARD (Orcid ID : 0000-0002-6832-7291)

DR. ELIZABETH D FERUCCI (Orcid ID : 0000-0003-0478-8205)

PROF. CAROLINE GORDON (Orcid ID : 0000-0002-1244-6443)

Article type : Full Length

**Title:** Prevalence of Systemic Lupus Erythematosus in the United States: Estimates from a Meta-Analysis of the Centers for Disease Control and Prevention National Lupus Registries

**Running Title:** Systemic Lupus Erythematosus Prevalence Meta-Analysis

**Authors:** Peter M. Izmirly\* MD, MSc, Hilary Parton† MPH, Lu Wang‡ PhD, W. Joseph McCune§ MD, S. Sam Lim§§ MD, MPH, Cristina Drenkard§§ MD, PhD, Elizabeth D. Ferucci¶ MD, MPH, Maria Dall'Era|| MD, Caroline Gordon†† MD, Charles G. Helmick‡‡ MD, and Emily C. Somers\*\* PhD ScM

**Affiliations:** \*Division of Rheumatology, Department of Medicine, New York University School of Medicine, New York, NY; † New York City Department of Health and Mental Hygiene, Long Island City, New York, NY; ‡ Department of Biostatistics, University of Michigan Schools of Medicine & Public Health, Ann Arbor, MI; § Division of Rheumatology, Department of Medicine, University of Michigan Schools of Medicine, Ann Arbor, MI; §§ Department of Medicine, Division of Rheumatology, Emory University, Atlanta, GA; ¶ Division of Community Health Services, Department of Clinical and

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/ART.41632](https://doi.org/10.1002/ART.41632)

This article is protected by copyright. All rights reserved

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

Research Services, Alaska Native Tribal Health Consortium, Anchorage, AK; <sup>||</sup>  
Department of Medicine, University of California at San Francisco, San Francisco, CA;  
<sup>††</sup>Rheumatology Research Group, Institute of Inflammation and Ageing, College of  
Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom;  
<sup>‡‡</sup>Division of Population Health, Centers for Disease Control and Prevention, Atlanta,  
GA; <sup>\*\*</sup>Division of Rheumatology, Department of Internal Medicine, Department of  
Environmental Health Sciences, Department of Obstetrics and Gynecology, University  
of Michigan, Ann Arbor MI.

**Funding:** Support for this analysis was provided by cooperative agreements between the New York City Department of Health and Mental Hygiene and New York University School of Medicine, and the Centers for Disease Control and Prevention U01DP006489 and National Institutes of Health UL1TR002240 to the University of Michigan. Data collection was funded under the following cooperative agreements: Georgia Lupus Registry: DP08806, Michigan Lupus Epidemiology and Surveillance Program: DP001441, California Lupus Surveillance Program: A114297, Manhattan Lupus Surveillance Program: DP002827, and Indian Health Service: Interagency agreement award IAA10FED1003070.

**Conflicts of interest:** Gordon (Centers for Disease Control and Prevention, MGP UCB, grants from UCB and Sandwell and West Birmingham Hospitals NHS Trust), Izmirly (GlaxoSmithKline)

**Corresponding author:** Peter M. Izmirly, MD, NYU School of Medicine  
550 First Avenue, MSB 625 New York, NY 10016 Email: [Peter.Izmirly@nyumc.org](mailto:Peter.Izmirly@nyumc.org)  
Phone: 212-263-5802 Fax: 646-501-5208

**Abstract word count:** 240. **Manuscript word count:** 1657. **Figure:** 1. **Tables:** 1.  
**Supplemental Figure:** 1.

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

**Keywords:** Systemic lupus erythematosus, Meta-Analysis, Epidemiology, public health surveillance

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Author Manuscript

## ABSTRACT

**Objective:** Epidemiologic data for systemic lupus erythematosus (SLE) is limited, particularly for racial/ethnic subpopulations in the United States (U.S.). Leveraging data from the Centers for Disease Control and Prevention (CDC) National Lupus Registry network of population-based SLE registries, a meta-analysis estimating U.S. SLE prevalence was performed.

**Methods:** The CDC National Lupus Registry network included four registries in unique states and a fifth in the Indian Health Service (IHS). All registries used the 1997 revised American College of Rheumatology (ACR) classification criteria for the SLE case definition. Case finding spanned either 2002-2004 or 2007-2009. A random effects model was employed given heterogeneity across sites. Applying sex/race-stratified estimates to the 2018 Census population, an estimate for the number of SLE cases in the U.S. was generated.

**Results:** 5,417 cases fulfilled the ACR SLE classification criteria. Pooled prevalence from the four state-specific registries was 72.8/100,000 (95%CI:65.3,81.0), 9 times higher for females than males (128.7 vs 14.6), and highest among Black females (230.9), followed by Hispanic (120.7), white (84.7) and Asian/Pacific Islander females (84.4). Male prevalence was highest in Black males (26.7) followed by Hispanic (18.0), Asian/Pacific Islander (11.2), and white males (8.9). The American Indian/Alaska Native had the highest race-specific SLE estimates for females (270.6/100,000) and males (53.8/100,000). In 2018, 204,295 persons (95% CI:160,902,261,725) in the U.S. fulfilled ACR SLE classification criteria.

**Conclusions:** A coordinated network of population-based SLE registries provided more accurate estimates for SLE prevalence and numbers affected in the U.S.

Author Manuscript

## INTRODUCTION

The heterogeneity of the clinical manifestations of systemic lupus erythematosus (SLE) and lack of a singular diagnostic test make SLE difficult for epidemiologists to study (1). Previous estimates for SLE rates in the United States (U.S.) have been predominantly derived from tertiary care settings and relatively small, homogeneous patient populations that contain limited data on key demographic groups in the U.S. (1). Other explanations for the varied estimates, which range from 19-241 per 100,000, include

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

racial/ethnic disparities in SLE susceptibility and mortality, differing case definitions, heterogeneous sources for case ascertainment, small populations, possible inaccuracy of self-report, unreliability in coding in health system databases, and variable access to health care for high-risk populations (2, 3).

The Centers for Disease Control and Prevention (CDC) funded a network of five population-based SLE registries, each using similar active surveillance methods, to determine SLE incidence and prevalence in populations reflecting a broad distribution of racial/ethnic demographics in the U.S. Data from these five registries have provided overall prevalence and incidence rates of SLE, as well as estimates that focused on the major U.S. demographic groups, including whites and Blacks (3, 4), Asians/Pacific Islanders and Hispanics (5, 6) and American Indians/Alaska Natives (AI/AN) (7). Leveraging these data, we performed a meta-analysis to estimate the overall prevalence of SLE and to provide an estimate of the number of SLE cases in the U.S. in 2018.

### **METHODS**

#### ***Data Sources and Study Selection***

The CDC-supported and SLE-dedicated registries were based in the following source populations which contained a mix of urban and rural areas: Georgia (Fulton & DeKalb Counties – Georgia Lupus Registry (GLR)) (3) and Michigan (Washtenaw & Wayne Counties – Michigan Lupus Epidemiology and Surveillance Program (MILES)) (4), both of which have a large Black population (approximately 50%); California (San Francisco County – California Lupus Surveillance Program, CLSP)) (5) and New York (New York County – Manhattan Lupus Surveillance Program (MLSP)) (6), which have populations with substantial representation of Asian/Pacific Islanders and Hispanics; and the Indian Health Service (IHS) (IHS facilities from Alaska, Phoenix, and Oklahoma City areas) (7), which provided estimates for the AI/AN population. Active surveillance for these registries was performed at various times between 2003-2015 using the surveillance exemption to HIPAA and public health authorization by respective State or City Health

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

Departments, which allowed access to medical records without individual consent. The case definitions for SLE prevalence varied slightly by the time period evaluated in each registry with all taking place between 2002-2009 (3-7). All registries used the 1997 revised American College of Rheumatology (ACR) classification criteria for SLE as the primary case definition for SLE (8, 9). The registries employed harmonized methods, including the utilization of a variety of case-finding sources and screening for potential SLE cases using the same core set of ICD-9 codes. Registries used a consistent approach to capture the relevant clinical and demographic information and core definitions from a standardized data dictionary. Trained medical abstractors who underwent routine quality assurance monitoring collected the data. Population denominators were based on intercensal population estimates for the respective source populations. Sex- and race/ethnicity-specific prevalence estimates were calculated per 100,000 person-years and age-adjusted to the 2000 U.S. Standard Population, Supplementary Figure 1 (10). Data were extracted from the published manuscripts by two authors independently (PI & HP) who agreed on all data used.

### ***Data Synthesis and Analysis***

A meta-analysis was conducted to derive pooled prevalence estimates using data from four similar CDC-funded registries based in states (GLR, MILES, CLSP and MLSP) for the age-standardized prevalence (adjusted to the 2000 U.S. Standard Population) (10) and for rates stratified by sex and race/ethnicity categories other than AI/AN (3-6). In contrast to the 4 state-based registries, the IHS-based registry (7) was different, focusing on one demographic (American Indian/Alaska Natives; AI/AN) and thus was handled separately.

For the meta-analysis, heterogeneity across sites was tested by Cochran's Q and  $I^2$  statistic (11, 12). Due to significant heterogeneity, we used a random effects model, weighted by the population denominator for each site, to calculate pooled prevalence (13). Such random effects models allow an underlying distribution of the effect sizes across different studies. Pooled race- and ethnicity-specific estimates were calculated,

except for the AI/AN estimate, which was based solely on the IHS registry that covered multiple states.

In the MLSP publication (6), rates were presented as combined race and ethnicity categories (e.g., non-Hispanic white). For this meta-analysis, race and ethnicity rates were calculated separately for consistency across the registries based in states.

Hispanic ethnicity and race categories overlap, so Hispanic estimates include all races and each race category includes Hispanics (i.e., race and Hispanic ethnicity are not mutually exclusive).

To estimate the number of SLE cases in the U.S., the pooled age-adjusted sex- and race-specific prevalence rates from the four states and the AI/AN prevalence from the Indian Health Service were separately extrapolated to 2018 U.S. Census population data; these stratum-specific estimates were then summed for the total population count. The pooled prevalence estimates do not incorporate the Hispanic rates because that would lead to duplicate counting.

## RESULTS

### *Prevalence*

The five registries contributed 5,417 SLE cases fulfilling ACR classification criteria from diverse areas across the country. The random effects model for the meta-analysis of estimates from the four registries based in states yielded an overall SLE prevalence of 72.8 per 100,000 (95% CI: 65.3, 81.0); Figure 1A. The prevalence among females was about 9 times higher than males (128.7 vs 14.6); Table 1. From the race- and ethnicity-specific pooled estimates from the four state-specific registries, prevalence was the highest among Black females (230.9, 95% CI: 178.2, 299.2), followed by Hispanic females (120.7, 95% CI: 84.0, 173.4), white females (84.7, 95% CI 68.4, 104.8), and Asian/Pacific Islander females (84.4, 95% CI: 48.3, 147.4); Figure 1B, Table 1. Among males, prevalence followed a similar pattern, with the highest rates among Black males (26.7, 95% CI: 19.6, 36.4), followed by Hispanic males (18.0, 95% CI 15.6, 20.8), Asian/Pacific Islander males (11.2, 95% CI: 5.7, 21.9), and white males (8.9, 95% CI: 8.0, 10.1), Figure 1C, Table 1.



## Systemic Lupus Erythematosus Prevalence Meta-Analysis

The prevalence estimates for AI/AN from the IHS Registry (not included in the pooled meta-analysis estimates from the four registries based in states) were the highest for all races for both females (270.6, 95% CI: 237.5, 307.0) and males (53.8, 95% CI: 36.2, 77.1), Figures 1B and 1C, Table 1.

### ***Numbers of SLE cases in the United States***

Applying the sex- and race-specific prevalence estimates to the corresponding stratum-specific population denominators from 2018 U.S. Census data, we estimated that 204,295 persons (95% CI: 160,902, 261,725) in the U.S. fulfilled ACR SLE classification criteria; Table 1.

## **DISCUSSION**

Using the ACR classification criteria to clinically define SLE, overall SLE prevalence in the U.S. was estimated to be 72.8 per 100,000 (95% CI: 65.3, 81.0) during calendar years 2002-2009. Prevalence was approximately 9 times higher in females compared to males, and highest among American Indians/Alaska Natives and Black females.

Extrapolating sex- and race-specific estimates to 2018 U.S. Census data, we estimated that 204,295 (95% CI: 160,902, 261,725) persons (184,323 females and 19,972 males) in the U.S. fulfilled ACR SLE classification criteria.

Limitations and strengths for each of the five component registries have been previously described (3-7). There are several limitations. First, although the registries were designed to employ similar methods, there were subtle differences in the comprehensive case finding sources that were approached by the different registries and the ICD-9 criteria used to identify possible cases. Second, case finding may have missed some true cases meeting the ACR criteria, so the actual numbers may be slightly higher than these estimates as demonstrated by the capture-recapture analyses conducted by the state-based registries (3-6). Third, data on race and ethnicity were abstracted from medical records that may not accurately represent the patient's own racial or ethnic identification. Hispanic ethnicity and the different races encompass several heterogeneous groups, and SLE rates among them may differ. Fourth, AI/AN prevalence was based on one registry, although 3 geographic regions with different population characteristics were encompassed in the IHS registry, improving

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

generalizability of results (7). Due to significant heterogeneity, the IHS registry data were not used in the meta-analysis pooled prevalence calculations; however, the IHS-derived AI/AN estimates were used for the national estimate calculation for number of SLE cases. Fifth, secondary case definitions used by the five registries ((3-7), results not reported in these analyses) resulted in slightly higher estimates in most instances, although that greater sensitivity may have occurred with lower specificity. Sixth, these analyses did not include other forms of lupus such as “early” or “incomplete” lupus, drug-induced lupus or primary cutaneous lupus (14, 15). Seventh, our prevalence estimates from 2002-2004 and 2007-2009 were applied to the 2018 census population. That provides a more relevant estimate of numbers with SLE, but it might be slightly affected if lupus prevalence changed significantly during that period.

These analyses also have several strengths. First, case finding likely captured a wider spectrum of SLE than previous studies because of the HIPAA surveillance exemption and case finding that facilitated data collection that extended beyond academic medical centers. Second, cases were validated through standardized and quality-controlled abstracting and rigid reviews of all available medical records. Third, the standard 1997 revised ACR classification criteria (8,9) were used. Fourth, the registries used harmonized methods and data dictionaries, and included a large number of SLE cases from diverse populations across the country with substantial representation of males and the major racial and ethnic groups found in the U.S. Fifth, employing these estimates allowed us to estimate the numbers affected with SLE in the U.S. This estimate approaches the 1983 definition of a rare disease used in the U.S. (i.e., 200,000) (16) and is lower than the widely used estimate of 1.5 million people (17). In summary, using estimates from a large, coordinated network of population-based registries that conducted active surveillance for SLE, a more accurate prevalence estimate for the U.S. was obtained. This likely represents a lower bound for SLE prevalence in the U.S.

**Acknowledgements:** The authors would like to acknowledge Benjamin Wainwright and Ruth Fernandez-Ruiz for their assistance in preparing the manuscript.

## References

1. Lim SS, Drenkard C, McCune WJ, Helmick CG, Gordon C, Deguire P, et al. Population-based lupus registries: advancing our epidemiologic understanding. *Arthritis Rheum.* 2009;61(10):1462-6.
2. Drenkard C, Lim SS. Update on lupus epidemiology: advancing health disparities research through the study of minority populations. *Curr Opin Rheumatol.* 2019;31(6):689-96.
3. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: The Georgia Lupus Registry. *Arthritis Rheumatol.* 2014;66(2):357-68.
4. Somers EC, Marder W, Cagnoli P, Lewis EE, DeGuire P, Gordon C, et al. Population-based incidence and prevalence of systemic lupus erythematosus: the Michigan Lupus Epidemiology and Surveillance program. *Arthritis Rheumatol.* 2014;66(2):369-78.
5. Dall'Era M, Cisternas MG, Snipes K, Herrinton LJ, Gordon C, Helmick CG. The Incidence and Prevalence of Systemic Lupus Erythematosus in San Francisco County, California: The California Lupus Surveillance Project. *Arthritis Rheumatol.* 2017;69(10):1996-2005.
6. Izmirly PM, Wan I, Sahl S, Buyon JP, Belmont HM, Salmon JE, et al. The Incidence and Prevalence of Systemic Lupus Erythematosus in New York County (Manhattan), New York: The Manhattan Lupus Surveillance Program. *Arthritis Rheumatol.* 2017;69(10):2006-17.
7. Ferucci ED, Johnston JM, Gaddy JR, Sumner L, Posever JO, Choromanski TL, et al. Prevalence and incidence of systemic lupus erythematosus in a population-based

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

registry of American Indian and Alaska Native people, 2007-2009. *Arthritis Rheumatol.* 2014;66(9):2494-502.

8. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40(9):1725.
9. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;25(11):1271-7.
10. Standard Populations (Millions) for Age-Adjustment. National Cancer Institute, National Institutes of Health; 2019. p. <https://seer.cancer.gov/stdpopulations/>.
11. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods.* 2006;11(2):193-206.
12. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539-58.
13. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health.* 2019;22(4):153-60.
14. Drenkard C, Parker S, Aspey LD, Gordon C, Helmick CG, Bao G, et al. Racial Disparities in the Incidence of Primary Chronic Cutaneous Lupus Erythematosus in the Southeastern US: The Georgia Lupus Registry. *Arthritis Care Res (Hoboken).* 2019;71(1):95-103.
15. Izmirly P, Buyon J, Belmont HM, Sahl S, Wan I, Salmon J, et al. Population-based prevalence and incidence estimates of primary discoid lupus erythematosus from the Manhattan Lupus Surveillance Program. *Lupus Sci Med.* 2019;6(1):e000344.
16. FAQs about Rare Diseases 2020. National Institutes of Health Genetic and Rare Diseases Information Center. [cited; Available from: <https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases>
17. Bruskin-Goldring Research Study conducted through telephone survey for the Lupus Foundation of America. 1994.

**Figure 1 Legend:**

**Figure 1A: Meta-Analysis Results of SLE Prevalence Based on Four CDC Population-Based Registries, Overall and by Registry Site.** Overall prevalence estimates for the SLE registry sites in Michigan (MI), Georgia (GA), New York (NY), and California (CA) are represented by circles, with the size of the circle corresponding to the weight of the contribution to the meta-analysis, and the diamond representing the results from the meta-analysis and the lines corresponding to 95% confidence intervals.

**Figure 1B & C: Meta-Analysis Results of SLE Prevalence Based on Four CDC Population-Based Registries, Overall and by Race and Hispanic Ethnicity Among Females (B) and Males (C)**

The overall female and male meta-analysis estimate is based on results from the lupus registry sites in Michigan, Georgia, New York, and California.

<sup>a</sup> Estimates for Blacks and whites are based on pooled estimates from the four state-based registries; Asian/Pacific Islanders (Asian/PI) are based on pooled estimates from Michigan, California and New York; American Indian/Alaska Native (AI/AN) estimates are from the Indian Health Service registry data previously published.

<sup>b</sup> Estimates for Hispanics are based on pooled estimates from Michigan, California and New York.

Systemic lupus erythematosus (SLE) cases were defined according to the American College of Rheumatology (ACR) criteria.

**Table 1: Estimated Number of Persons with Systemic Lupus Erythematosus Living in the United States in 2018**

Race/Ethnicity (Number of sites in analysis)	FEMALE		
	SLE Prevalence <sup>a</sup> per 100,000 (95% CI)	Population Denominator	Estimated # SLE Cases in United States (95% CI)

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

### Race

Black (4)	230.9	(178.2, 299.2)	24,880,722	57,450	(44,337, 74,443)
White (4)	84.7	(68.4, 104.8)	130,137,989	110,227	(89,014, 136,437)
Asian/PI (3)	84.4	(48.3, 147.4)	12,544,896	10,588	(6,059, 18,491)
AI/AN (1)	270.6	(237.5, 307.0)	2,238,966	6,059	(5,318, 6,874)
<b>Total<sup>b</sup></b>	<b>128.7</b>	<b>(113.3, 146.2)</b>	<b>169,802,573</b>	<b>184,323</b>	<b>(144,729, 236,245)</b>

### Ethnicity

Hispanic <sup>c</sup> (3)	120.7	(84.0, 173.4)	30,689,083	37,042	(25,779, 53,215)
---------------------------	-------	---------------	------------	--------	------------------

### MALE

Race	Prevalence <sup>a</sup>		Population	Estimated # SLE Cases	
	per 100,000 (95% CI)		Denominator	in United States (95% CI)	
Black (4)	26.7	(19.6, 36.4)	22,961,129	6,131	(4,500, 8,358)
White (4)	8.9	(8.0, 10.1)	127,942,583	11,387	(10,235, 12,922)
Asian/PI (3)	11.2	(5.7, 21.9)	11,660,533	1,306	(665, 2,554)
AI/AN (1)	53.8	(36.2, 77.1)	2,134,870	1,149	(773, 1,646)
<b>Total<sup>b</sup></b>	<b>14.6</b>	<b>(12.2, 17.5)</b>	<b>164,699,115</b>	<b>19,972</b>	<b>(16,173, 25,480)</b>
<b>Ethnicity</b>					
Hispanic <sup>c</sup> (3)	18.0	(15.6, 20.8)	31,281,605	5,631	(4,880, 6,507)

Systemic lupus erythematosus cases were defined according to the 1997 revised American College of Rheumatology criteria.

<sup>a</sup> Estimates for Blacks and whites are based on pooled estimates from the four state-based registries; Asian/Pacific Islanders (Asian/PI) and Hispanics are based on pooled estimates from Michigan, California and New York; American Indian/Alaska Native (AI/AN) estimates are based on the Indian Health Service Registry.

<sup>b</sup> The pooled 'total' prevalence estimate includes Black, white and Asian/Pacific Islanders (Asian/PI). Since the American Indian/Alaska Native (AI/AN) prevalence was

## Systemic Lupus Erythematosus Prevalence Meta-Analysis

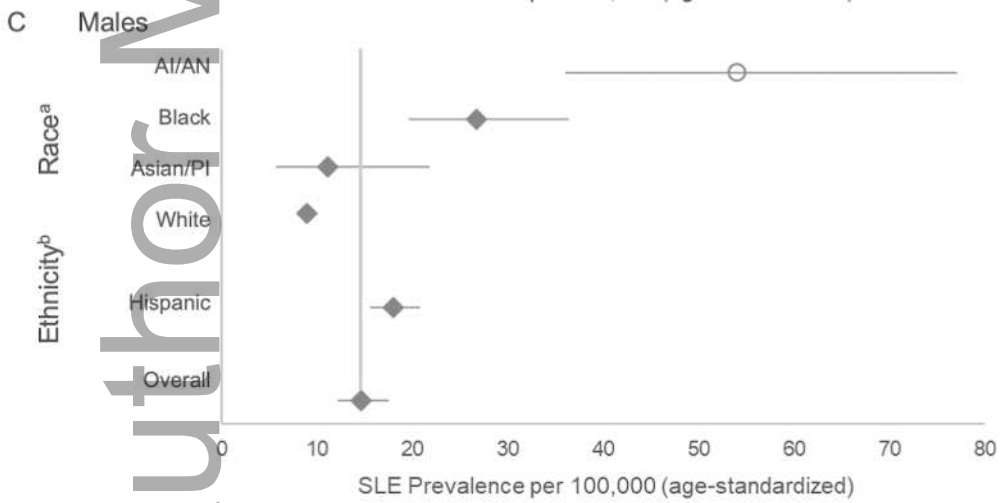
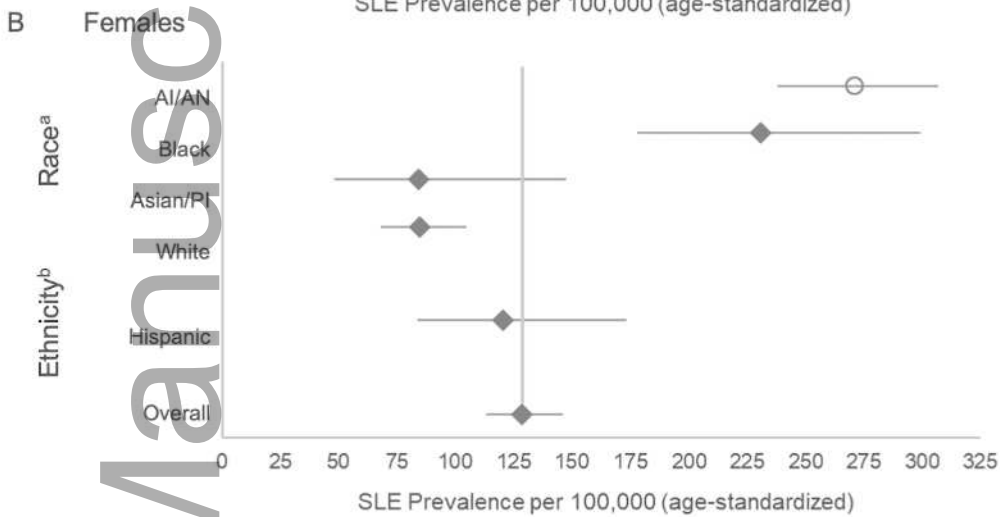
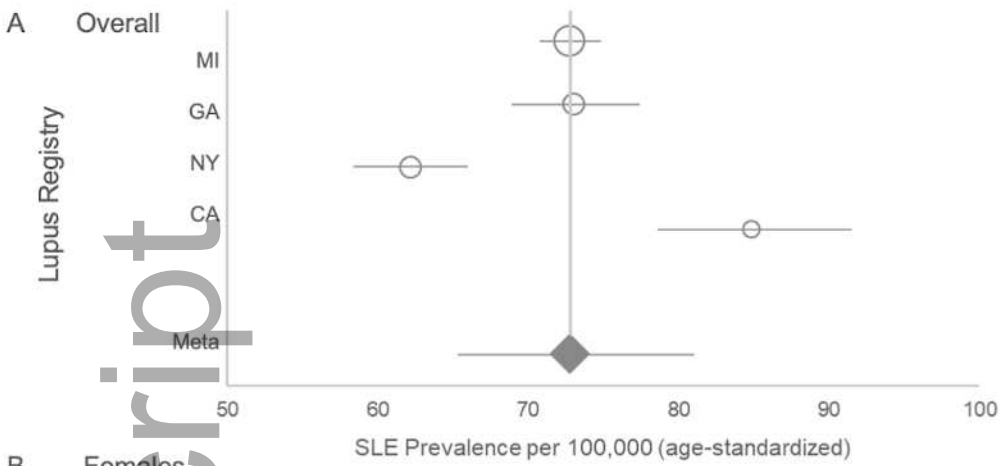
based on one registry and was significantly higher, it was not included in the pooled prevalence per 100,000.

<sup>c</sup> Hispanic ethnicity is not mutually exclusive from the race categories, i.e., all Hispanic persons are included in one of the race categories. Thus, the pooled estimates do not incorporate the Hispanic rates since that would lead to duplicate counting. Estimates for Hispanics are based on pooled estimates from Michigan, California and New York.

### Supplemental Figure 1 Legend

#### Prevalence of Systemic Lupus Erythematosus from Five CDC Lupus Registries Using ACR Classification Criteria

A) female and B) male age-adjusted rates. Systemic lupus erythematosus cases were defined according to the 1997 revised American College of Rheumatology criteria.



art\_41632\_f1.tiff