Using a Large Electronic Health Records Database to Define Representative Patient Populations for Lupus Nephritis Trials

Andrew Bevan, MRes[1]; Davide Garrisi, MSc, MBA[2]; Sarah Stump, MS[3], Carmichael Angeles, MD[3]

Corresponding Author Email: Andrew.Bevan@ppd.com

[1]PPD, Part of Thermo Fisher Scientific, Cambridge, United Kingdom; [2]PPD, Part of Thermo Fisher Scientific, Milan, Italy; [3]PPD, Part of Thermo Fisher, Wilmington, NC, United States

Background

- Lupus nephritis (LN) is a severe renal manifestation of systemic lupus erythematosus (SLE) that disproportionately affects women and people of African descent. Furthermore, those of African descent have worse renal outcomes.
- There is intensive interest in developing new treatments for LN, including more than 18 drugs currently in development, and approvals by the FDA of 3 drugs in the last 3 years, voclosporin (Lupkynis®), belilumab (Benlysta®) and . anifrolumab-fnia (Saphnelo®).
- An analysis by the authors of 10 LN trials completed in clinicaltrials.gov, revealed low representation of Black or African-Americans (Mdn 9.5% [IQR 4.9-13.1%]).
- FDA guidance places emphasis on studying clinically relevant trial populations but does not address how such populations might be defined.

Objective

- To assess the demographic characteristics of individuals diagnosed with LN in the USA from an electronic health record (EHR) database and propose statistical parameters for cohort sample sizes for given demographic variables to support planning for inclusion of clinically relevant populations in LN trials.

Methods

- We conducted a study of data from the TriNetX Analytics Network, which contains EHRs from >140 million individuals in the US.
- Demographics assessed for living individuals with an ICD-10-CM code M32.14 Glomerular Disease in SLE in the last 5 years that had received healthcare services and did not also have an ICD-10-CM code N18.6 ESRD.
- Binomial confidence intervals (CIs) were calculated for each demographic cohort to define parameters for hypothetical sample sizes for clinical trials.

Results

Demographic Characteristics

- 16,400 individuals met eligibility criteria.
- Race and ethnicity was available for all.
- Gender was available for 15,810, 84% were female.
- Demographic characteristics are summarized in **Table 1.**

Gender Cohort Models

- Based on an analysis of binomial CIs, a US trial of 100 subjects would be considered "statistically representative" ($p < 0.05$) of the gender distribution of the TriNetX LN population if it included a range of 75-90 females and 9 - 25 males.
- 95% CIs for a hypothetical US trial of 300 subjects is provided in **Table 2**.

Race and Ethnicity Cohort Models

- Race and ethnicity cohort models are provided in **Table 3** and **Table 4** respectively for hypothetical US LN trials of 100 and 300 subjects.
- The same hypothetical US trial of 100 subjects would be considered representative of the race and ethnic distribution of the TriNetX LN population if it included a range of 29 - 49 white, 24 - 43 Black or African American, 2 - 13 Asian and 8 - 22 Hispanic or Latino patients.

Table 1. Demographic Characteristics of Lupus Nephritis Patients

Number of Patients Identified	16,400
Age (Median ± Standard Deviation)	45 ± 18
Gender (%)	
Number of Patients with Gender Available	15,810
Female	84
Male	16
Race (%)	
White	39
Black or African American	33
Asian	6
Other, Mixed or Unknown	22
Ethnicity (%)	
Not Hispanic or Latino	62
Unknown	24
Hispanic or Latino	14

Table 2. Sample Size and Gender Cohort Models

Demographic Cohorts	Expected # of Cohort Subjects	95% Confidence Intervals
(A) Sample Size = 100		
Female	84	75 - 90
Male	16	9 - 25
(B) Sample Size = 300		
Female	252	232 – 264
Male	48	36 – 62

Table 3. Sample Size and Race Cohort Models

Demographic Cohorts	Expected # of Cohort Subjects	95% Confidence Intervals
(A) Sample Size = 100		
White	39	29 – 49
Black or African American	33	24 – 43
Asian	6	2 – 13
Other, Mixed or Unknown	22	14 - 31
(B) Sample Size = 300		
White	117	100 – 134
Black or African American	99	83 – 116
Asian	18	11 – 28
Other, Mixed or Unknown	66	52 - 81

Table 4. Sample Size and Ethnicity Cohort Models

Demographic Cohorts	Expected # of Cohort Subjects	95% Confidence Intervals
(A) Sample Size = 100		
Not Hispanic or Latino	62	52 – 72
Unknown	24	16 – 34
Hispanic or Latino	14	8 – 22
(B) Sample Size = 300		
Not Hispanic or Latino	186	169 – 203
Unknown	72	58 – 88
Hispanic or Latino	42	31 – 55

Conclusions

- As emphasized by the FDA, therapeutics should be investigated in trial populations that are representative of those likely to use the treatment if approved.
- People of African descent have been unrepresented in LN trials completed to date as demonstrated by the disparity between demographic data from CT.gov vs EHR.
- Large EHR databases enable rapid identification of "real world" demographics of patients with rare renal conditions such as LN, and thus **are a viable tool to model and support** the planning of "representative cohorts" for inclusion into clinical trials, especially when suitable epidemiological data is scarce.

Presented at the National Kidney Foundation Spring Clinical Meetings, Long Beach 14 - 18 May 2024

Supported by funding from PPD, a part of Thermo Fisher Scientific