## Supplemental Materials:

### Table of Contents:

**Supplemental Methods:** Derivation of case-finding algorithms 2

**Supplemental Table 1:** Characteristics of the manually-assembled single-center cohort for pilot algorithm review at the discovery site, to determine data elements for inclusion in the computable phenotype. 5

**Supplemental Table 2:** Diagnostic and provider code sets used in computable phenotypes. 6

**Supplemental Table 3:** Procedural code sets used in computable phenotypes. 8

**Supplemental Table 4:** Kidney disease/glomerular disease code set used in computable phenotypes. 9

**Supplemental Table 5:** Performance characteristics of SLE computable phenotype algorithm across PEDSnet population with two or more rheumatology or nephrology encounters, stratified by presence of absence of kidney involvement. 11

**Supplemental Table 6:** Performance metrics of SLE and algorithms compared to historical controls. 12

### Supplemental Methods: Derivation of case-finding algorithms

Computable phenotypes do not exist for the individual classification criteria from the American College of Rheumatology (ACR), European League Against Rheumatism (EULAR), or SLE International Clinic (SLICC) systems. Many epidemiologic studies have used the International Classification of Diseases, Ninth Revision (ICD-9) clinical modification billing code data (specifically 2 or 3 counts of the SLE ICD-9-CM code 710.0) to identify patients with SLE within administrative databases ([3](#_ENREF_3), [20](#_ENREF_20), [31-36](#_ENREF_31)), but used alone as diagnostic criteria, single ICD-9-CM codes perform poorly with low positive predictive values ([37](#_ENREF_37)). Case-finding algorithms combining ICD-9-CM codes with keywords, laboratory values, and medication data improve performance in adult cohorts ([17-19](#_ENREF_17), [25](#_ENREF_25), [26](#_ENREF_26)) but there is a paucity of data analyzing algorithms for cSLE or pediatric lupus nephritis.

The discovery cohort included 219 SLE patients, 80 of whom had lupus nephritis (Supplemental Table 1). Algorithm SLE\_A1, which required 2 or more SLE diagnosis codes 60 or more days apart, showed good sensitivity, but lacked specificity. Due to the widespread use of hydroxychloroquine in patients after SLE diagnosis, the combination of diagnosis and hydroxychloroquine medication codes was tested as an alternative means for entry into the SLE cohort. While sensitivity of SLE\_A3, which required the same criteria as SLE\_A1 OR 1 or more SLE codes plus hydroxychloroquine exposure, improved to 98%, the specificity dropped significantly. Further modification of the algorithm to include 2 or more SLE diagnosis codes 60 or more days apart AND hydroxychloroquine exposure (SLE\_A5) resulted in a trade-off between specificity and sensitivity, with comparable positive predictive value to algorithm SLE\_A1. Requiring 3 or more diagnosis codes, with a shorter interval between visits associated with these codes, resulted in improvements in specificity and positive predictive value in all variations tested (SLE\_A2, SLE\_A4, SLE\_A6). Finally, the addition of a kidney biopsy procedure (Supplemental Table 3) resulted in optimal performance, and ensured that the lupus nephritis cohort was a cusbet of the SLE cohort. This algorithm SLE\_A7 was therefore selected for validation.

For identification of lupus nephritis patients, algorithm LN\_A1 and LN\_A2 were tested in parallel using the same approach as SLE\_A1 and SLE\_A2, but with lupus nephritis diagnosis codes (a subset of the SLE diagnosis codes, Supplemental Table 2). LN\_A1 and LN\_A2 were highly specific but lacked sensitivity (Table 2). In response to lower than anticipated counts for lupus nephritis and low sensitivity across sites, algorithms LN\_A3 and LN\_A4 included patients with at least one diagnosis of SLE accompanied by either a nephritis-related diagnosis or a kidney biopsy procedure. LN\_A3 and LN\_A4 were substantially more sensitive than LN\_A1 and LN\_A2. Removal of the nephritis-related diagnosis criterion in LN\_A5 and LN\_A6 increased specificity and positive predictive value. Finally, LN\_A7 incorporated an LN diagnosis combination which included an SLE diagnosis accompanied by a kidney disease or glomerular disease code on the same date. This code combination was treated as a lupus nephritis diagnosis for the algorithm. This was motivated by observations of how ICD9-CM codes for lupus nephritis in source data were mapped to the SNOMED-CTcodes across sites in the PEDSnet learning health system.

The addition of a hydroxychloroquine exposure criterion, and requirement for ≥3 lupus nephritis codes with ≥30 days to LN\_A7 ensured that the lupus nephritis cohort was a subset of the SLE cohort (SLE\_A7). Considering performance against the discovery cohort, lupus nephritis diagnosis representation across sites, and a requirement that the lupus nephritis cohort be a subset of the SLE cohort, LN-A7 was selected for validation.

### Supplemental Table 1: Characteristics of the manually-assembled single-center cohort for pilot algorithm review at the discovery site, to determine data elements for inclusion in the computable phenotype

|  |  |  |
| --- | --- | --- |
| **Characteristic \*** | **SLE Cases** | **Lupus Nephritis Cases** |
| N | 219 | 80 |
| Age at initial in-person visit (y) | 11 (6, 15) | 11 (7, 144) |
| Female gender | 181 (83%) | 61 (76%) |
| Asian race | 20 (9%) | 8 (10%) |
| Black race | 83 (38%) | 30 (38%) |
| Multiracial | 4 (2%) | 3 (3.8%) |
| Other race | 1 (0.5%) | 0 |
| Unknown race | 40 (18%) | 16 (20%) |
| White race | 71 (32%) | 23 (29%) |
| Hispanic ethnicity | 31 (14%) | 12 (15%) |
| Follow-up time (y) | 8 (5, 14) | 9 (5, 14) |

\* Continuous data presented as median (interquartile range) and categorical data as N (%)

\*\* There were 318 patients in the cohort, of whom 219 were validated as SLE and 80 were validated as lupus nephritis through manual chart review. The other 99 cases included incomplete lupus or non-lupus cases that had at least one billing code for lupus (such as Sjogren’s and mixed connective tissue disease).

### Supplemental Table 2: Diagnostic and provider code sets used in computable phenotypes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **OMOP Concept ID** | **Code in Vocabulary** | **Description** | **Vocabulary** | **Flag** |
| 257628 | 55464009 | Systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 4295179 | 403486000 | Acute systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 36676444 | 773333003 | Autosomal systemic lupus erythematosus  | SNOMED | SLE inclusion |
| 4346976 | 239889005 | Bullous systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 37110504 | 724767000 | Chorea co-occurent and due to systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 37110517 | 724781003 | Demyelination of central nervous system co-occurent and due to SLE (disorder)  | SNOMED | SLE inclusion |
| 4296502 | 403487009 | Fulminating systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 4055640 | 196138005 | Lung disease with systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 4057084 | 19682006 | Lupus hepatitis (disorder) | SNOMED | SLE inclusion |
| 4101469 | 25380002 | Pericarditis co-occurrent and due to systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 4319305 | 95332009 | Rash of systemic lupus erythematosus (disorder) | SNOMED | SLE inclusion |
| 44784527 | 698694005 | Systemic lupus erythematosus in remission (disorder) | SNOMED | SLE inclusion |
| 4301051 | 403488004 | Systemic lupus erythematosus of childhood (disorder) | SNOMED | SLE inclusion |
| 4344400 | 239890001 | Systemic lupus erythematosus with multisystem involvement (disorder) | SNOMED | SLE inclusion |
| 4344158 | 239887007 | Systemic lupus erythematosus with organ/system involvement (disorder) | SNOMED | SLE inclusion |
| 4318863 | 95644001 | Systemic lupus erythematosus encephalitis | SNOMED | SLE inclusion |
| 4299106 | 77753005 | Lupus disease of the lung | SNOMED | SLE inclusion |
| 4316215 | 95408003 | Systemic lupus erythematosus arthritis | SNOMED | SLE inclusion |
| 4149913 | 95644001 | Systemic lupus erythematosus encephalitis | SNOMED | SLE inclusion |
| 46273369 | 309762007 | Systemic lupus erythematosus with pericarditis | SNOMED | SLE inclusion |
| 4269448 | 72181000119109 | Endocarditis due to SLE | SNOMED | SLE inclusion |
| 37117740 | 36471000 | Dilated cardiomyopathy secondary to SLE | SNOMED | SLE inclusion |
| 4217054 | 732960002 | Secondary autoimmune hemolytic anemia co-occurrent and due to SLE | SNOMED | SLE inclusion |
| 37016279 | 417303004 | Retinal vasculitis due to SLE | SNOMED | SLE inclusion |
| 4285717 | 308751000119106 | Glomerular disease due to systemic lupus erythematosus | SNOMED | LN or SLE inclusion |
| 4250483 | 68815009 | Systemic lupus erythematosus glomerulonephritis syndrome (disorder) | SNOMED | LN or SLE inclusion |
| 4186940 | 73286009 | SLE glomerulonephritis syndrome, WHO class I | SNOMED | LN or SLE inclusion |
| 4297164 | 4676006 | SLE glomerulonephritis syndrome, WHO class II | SNOMED | LN or SLE inclusion |
| 4267801 | 76521009 | SLE glomerulonephritis syndrome, WHO class III | SNOMED | LN or SLE inclusion |
| 4178133 | 36402006 | SLE glomerulonephritis syndrome, WHO class IV | SNOMED | LN or SLE inclusion |
| 4002526 | 52042003 | SLE glomerulonephritis syndrome, WHO class V | SNOMED | LN or SLE inclusion |
| 37399735 | 11013005 | SLE glomerulonephritis syndrome, WHO class VI | SNOMED | LN or SLE inclusion |
| 37395585 | 295121000119101 | Nephrosis co-occurrent and due to systemic lupus erythematosus (disorder) | SNOMED | LN or SLE inclusion |
| 4145240 | 295111000119108 | Nephrotic syndrome co-occurrent and due to systemic lupus erythematosus | SNOMED | LN or SLE inclusion |
| 46270384 | 307755009 | Renal tubulo-interstitial disorder in systemic lupus erythematosus | SNOMED | LN or SLE inclusion |
| 4316373 | 295101000119105 | Nephropathy co-occurrent and due to systemic lupus erythematosus | SNOMED | LN or SLE inclusion |
| 257628 | 95609003 | Neonatal lupus erythematosus (disorder) | SNOMED | LN and SLE exclusion |
|  |  |  |  |  |
| **OMOP Concept ID** | **Code in Vocabulary** | **Description** | **Vocabulary** | **Flag** |
| 38003880 | 207RN0300X | Allopathic & Osteopathic Physicians, Internal Medicine, Nephrology | NUCC | Provider code |
| 38003882 | 207RR0500X | Allopathic & Osteopathic Physicians, Internal Medicine, Rheumatology | NUCC | Provider code |
| 38003955 | 2080P0210X | Allopathic & Osteopathic Physicians, Pediatrics, Pediatric Nephrology | NUCC | Provider code |
| 38003957 | 2080P0216X | Allopathic & Osteopathic Physicians, Pediatrics, Pediatric Rheumatology | NUCC | Provider code |
| 44777772 | 259 | Pediatric Nephrology | HES Specialty | Provider code |
| 44777783 | 361 | Nephrology | HES Specialty | Provider code |
| 38004479 | 39 | Nephrology | Medicare Specialty | Provider code |
| 44777791 | 410 | Rheumatology | HES Specialty | Provider code |
| 38004491 | 66 | Rheumatology | Medicare Specialty | Provider code |
| 45756818 | OMOP4822021 | Pediatric Rheumatology | ABMS | Provider code |
| 45756813 | OMOP4822026 | Pediatric Nephrology | ABMS | Provider code |

### Supplemental Table 3: Procedural code sets used in computable phenotypes.

|  |  |  |  |
| --- | --- | --- | --- |
| **OMOP Concept ID** | **Code in Vocabulary** | **Term** | **Vocabulary** |
| 2003588 | 55.23 | Closed [percutaneous] [needle] biopsy of kidney | ICD-9-CM Procedure |
| 2109566 | 50200 | Renal biopsy; percutaneous, by trocar or needle | CPT4 |
| 2211783\* | 76942 | Ultrasonic guidance for needle placement (eg, biopsy, aspiration, injection, localization device), imaging supervision and interpretation | CPT4 |
| 35608187 | 769246001 | Closed renal biopsy | SNOMED |
| 36717689 | 719033000 | Evaluation of kidney biopsy specimen | SNOMED |

\* OMOP Concept ID accompanied by string search in source values for “renal” or “kidney.”

### Supplemental Table 4: Combination code sets used in computable phenotypes.

|  |
| --- |
| **Kidney disease/glomerular disease** |
| **OMOP ID** | **Code in Vocabulary** | **Description** | **Vocabulary** |
| 198124 | 90708001 | Kidney disease | SNOMED |
| 4059452 | 197679002 | Glomerular disease | SNOMED |
| **Nephritis-related diagnosis** |
| **OMOP ID** | **Code in Vocabulary** | **Description** | **Vocabulary** |
| 192362 | 197591002 | Nephrotic syndrome with membranoproliferative glomerulonephritis | SNOMED |
| 192364 | 197603000 | Nephrotic syndrome associated with another disorder | SNOMED |
| 193253 | 52845002 | Nephritis | SNOMED |
| 194405 | 197590001 | Nephrotic syndrome with membranous glomerulonephritis | SNOMED |
| 195314 | 52254009 | Nephrotic syndrome | SNOMED |
| 196464 | 236392004 | Rapidly progressive glomerulonephritis | SNOMED |
| 197319 | 197589005 | Nephrotic syndrome with proliferative glomerulonephritis | SNOMED |
| 199071 | 266549004 | Nephrotic syndrome with minimal change glomerulonephritis | SNOMED |
| 252365 | 77182004 | Membranous glomerulonephritis | SNOMED |
| 259070 | 197579006 | Acute proliferative glomerulonephritis | SNOMED |
| 312358 | 20917003 | Chronic glomerulonephritis | SNOMED |
| 433257 | 80321008 | Mesangiocapillary glomerulonephritis | SNOMED |
| 435003 | 197580009 | Acute nephritis with lesions of necrotizing glomerulitis | SNOMED |
| 435308 | 19351000 | Acute glomerulonephritis | SNOMED |
| 435320 | 197582001 | Acute glomerulonephritis associated with another disorder | SNOMED |
| 442074 | 197626007 | Focal membranoproliferative glomerulonephritis | SNOMED |
| 442075 | 197616000 | Chronic glomerulonephritis associated with another disorder | SNOMED |
| 442076 | 197614002 | Chronic rapidly progressive glomerulonephritis | SNOMED |
| 4027120 | 197598008 | Nephrotic syndrome, diffuse mesangiocapillary glomerulonephritis | SNOMED |
| 4030513 | 236403004 | Focal segmental glomerulosclerosis | SNOMED |
| 4056346 | 197593004 | Nephrotic syndrome, minor glomerular abnormality | SNOMED |
| 4056478 | 197707007 | Chronic nephritic syndrome | SNOMED |
| 4058840 | 197594005 | Nephrotic syndrome, focal and segmental glomerular lesions | SNOMED |
| 4058843 | 197600002 | Nephrotic syndrome, diffuse crescentic glomerulonephritis | SNOMED |
| 4125958 | 236395002 | Post-infectious glomerulonephritis | SNOMED |
| 4128055 | 236381000 | Steroid-resistant nephrotic syndrome | SNOMED |
| 4222610 | 83866005 | Focal AND segmental proliferative glomerulonephritis | SNOMED |
| 4241966 | 59479006 | Mesangiocapillary glomerulonephritis, type II | SNOMED |
| 4260398 | 44785005 | Minimal change disease | SNOMED |
| 4263367 | 36171008 | Glomerulonephritis | SNOMED |
| 4286024 | 68544003 | Acute post-streptococcal glomerulonephritis | SNOMED |
| 4294813 | 75888001 | Mesangiocapillary glomerulonephritis, type I | SNOMED |
| 4298809 | 7724006 | Nephritic syndrome | SNOMED |
| 36712846 | 12511000132108 | Persistent proteinuria | SNOMED |
| 45769862 | 79131000119100 | Kidney lesion | SNOMED |

### Supplemental Table 5: Performance characteristics of SLE computable phenotype algorithm across PEDSnet population with two or more rheumatology or nephrology encounters, stratified by presence of absence of kidney involvement. Confidence intervals calculated using exact binomial test.

|  |
| --- |
| **SLE – with nephritis** |
| **Center** | **Sensitivity (95% CI)** | **Specificity (95% CI)** | **Positive predictive value (95% CI)** | **Negative predictive value (95% CI)** |
| 1 | 100% (84, 100) | 9%3 (82, 98) | 100% (86, 100) | 100% (93, 100) |
| 2 | 100% (82, 100) | 89% (78, 96) | 100% (86, 100) | 100% (93, 100) |
| 3 | 100% (86, 100) | 98% (90, 100) | 96% (80, 99) | 100% (93, 100) |
| 4 | 100% (83, 100) | 89% (78, 96) | 96% (80, 99) | 100% (93, 100) |
| 5 | 100% (82, 100) | 88% (76, 95) | 100% (86, 100) | 100% (93, 100) |
| 6 | 100% (82, 100) | 88% (76, 95) | 96% (80, 99) | 100% (93, 100) |
| 7 | 100% (86, 100) | 98% (90, 100) | 100% (86, 100) | 100% (93, 100) |
| **All** | 100% **(98, 100)** | **92% (88, 94)** | **98% (95, 99)** | **100% (99, 100)** |
| **SLE – without nephritis** |
| 1 | 100% (84, 100) | 93% (82, 98) | 84% (64, 96) | 1**00%** (93, 100) |
| 2 | 100% (82, 100) | 89% (78, 96) | 76% (55, 91) | 1**00%** (93, 100) |
| 3 | 100% (86, 100) | 98% (90, 100) | 100% (86, 100) | 1**00%** (93, 100) |
| 4 | 100% (83, 100) | 89% (78, 96) | 80% (59, 93) | 1**00%** (93, 100) |
| 5 | 100% (82, 100) | 88% (76, 95) | 72% (51, 88) | 1**00%** (93, 100) |
| 6 | 100% (82, 100) | 88% (76, 95) | 76% (55, 91) | 1**00%** (93, 100) |
| 7 | 100% (86, 100) | 98% (90, 100) | 96% (80, 99) | 1**00%** (93, 100) |
| **All** | 100% **(98, 100)** | **92% (88, 94)** | **83% (77, 89)** | **100% (99, 100)** |

### Supplemental Table 6: Performance metrics of SLE and lupus nephritis algorithms compared to historical controls ([17-20](#_ENREF_17))

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SLE** | **Sensitivity** | **Specificity** | **Positive predictive value** | **Negative predictive value** |
| Hanly *et al.* #1 | 82% |  97% | 86% | 95% |
| Hanly *et al. #2* | 52% |  98% | 87% | 89% |
| Hanly *et al. #3* | 80% |  97% | 85% | 95% |
| Hanly *et al. #4* | 41% | 100% | 99% | 87% |
| Hanly *et al. #5* | 83% |  99% | 95% | 96% |
| Hanly *et al. #6* | 85% |  98% | 91% | 96% |
| Hanly *et al. #7* | 87% |  92% | 74% | 97% |
| Barnado *et al. #1* | 40% |  NA | 95% | NA |
| Barnado *et al. #2* | 38% |  NA | 95% | NA |
| Barnado *et al. #3* | 70% |  NA | 92% | NA |
| Barnado *et al. #4* | 61% |  NA | 92% | NA |
| Barnado *et al. #5* | 66% |  NA | 91% | NA |
| Barnado *et al. #6* | 34% |  NA | 91% | NA |
| Barnado *et al. #7* | 33% |  NA | 91% | NA |
| Barnado *et al. #8* | 86% |  NA | 89% | NA |
| **SLE\_A7 (discovery)** | **89%** |  **71%** | **87%** | **75%** |
| **SLE\_A7 (6 site validation)** a | **100%** |  **92%** | **91%** | **100%** |
| **SLE with LN** |  |  |  |  |
| Chibnik *et al. #1* |  97% b |  NA c | 89% | NA |
| Chibnik *et al. #2* |  54% b |  NA | 92% | NA |
| Chibnik *et al. #3* |  48% b |  NA | 91% | NA |
| Chibnik *et al. #4* | 100% |  NA | 89% | NA |
| **SLE\_A7 (6 site validation)** d | **100%** |  **92%** | **98%** | **100%** |
| **Lupus Nephritis** |  |  |  |  |
| Chibnik *et al. #1* | 98%\* |  NA | 80% | NA |
| Chibnik *et al. #2* | 56%\* |  NA | 86% | NA |
| Chibnik *et al. #3* | 55%\*  |  NA | 88% | NA |
| Chibnik *et al. #4* | 100% |  NA | 79% | NA |
| Li *et al. #1* | 87% |  95% | 63% | 99% |
| Li *et al. #2* | 62% |  97% | 68% | 96% |
| Li *et al. #3* | 64% |  97% | 69% | 96% |
| Li *et al. #4* | 82% |  95% | 64% | 98% |
| Li *et al. #5* | 53% |  98% | 77% | 95% |
| Li *et al. #6* | 63% |  95% | 59% | 96% |
| Li *et al. #7* | 88% |  93% | 58% | 99% |
| Li *et al. #8* | 83% |  94% | 59% | 98% |
| Li *et al. #9* | 33% | 100% | 94% | 93% |
| **LN\_A7 (discovery)** | **80%** |  **98%** | **94%** | **94%** |
| **LN\_A7 (6 site validation)** | **89%** |  **92%** | **93%** | **89%** |

a performance using 6-institution PEDSnet data (did not include PEDSnet discovery site, see Table 3)

b compared to algorithm 4, the most sensitive at detecting SLE

c NA, data not available

d performance for identifying SLE patients with nephritis (see supplemental table 5)