Harnessing the potential of natural-language processing and interconnected data streams for complex diseases in the hospital setting: Lupus case study in France (Lupus REAL)

PT9

AUTHORS

Marc Jouve[1]
marc.jouve@sancare.fr

Batourou Sano[3]
batourou.sano@alirahealth.com

Eymeric Lagonotte[2]
elagonotte@chu-reims.fr

Jean-François Ricci[4]
jf.ricci@alirahealth.com

Marion Valette[1]
marion.valette@sancare.fr

Pr Vincent Vuiblet[5]
vvuiblet@chu-reims.fr

[1]Sancare, Paris, Paris, France, [2]URCA, Reims, France, [3]Alira Health, Paris, France, [4]Alira Health, Basel, BS, Switzerland, [5]CHU de Reims, Reims, Grand Est, France

RATIONALE AND OBJECTIVES

Efficiently identifying and characterizing with precision patients suffering from a heterogeneous condition with diverse manifestations with traditional data sources is complex and resource intensive.

This study assesses the ability of an interconnected hospital-specific natural language processing (NLP)-powered solution (REALLI) to identify and characterize patients with lupus in France.

METHODOLOGY

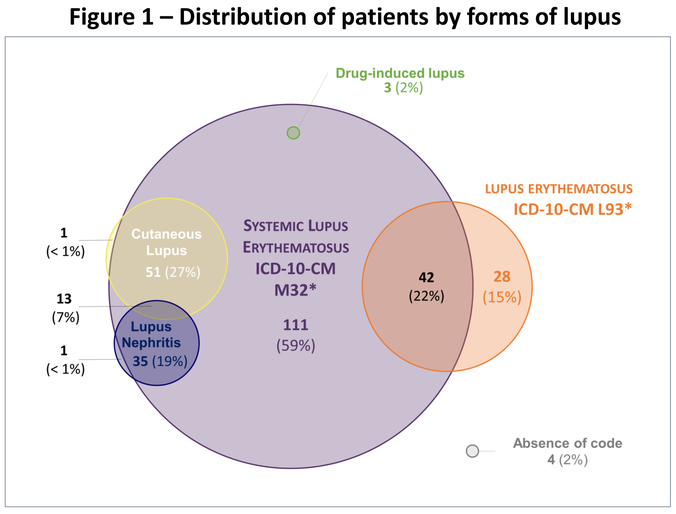
- **Study design**: a retrospective single-center study based on claims data and Electronic Medical Records (EMRs). Data were extracted via the "REALLI solution," a hospital-native health research information system which provides near real-time access to full content of hospital clinical and administrative data.
- **Study period**: Hospital patient-level data were retrieved from January 2018 through March 2023. Patient follow-up began at first admission in study period until March 2023 or death, whichever came first.
- **Study population**: Upon Ethics and Scientific Committee approval of the study protocol, eligible adult patients were retrieved by searching the native hospital patient EMRs for "lupus" or "lupique" text strings and diagnosis codes (ICD-10-CM L93.X or M32.X). We mailed eligible adult patients a study specific patient information letter and posted the study on Alira Health portal (https://alirahealth.com/cnil-and-transparency/) and CHU de Reims research portal (https://www.iias.fr/pages/tous-les-projets-iias).

RESULTS

> Lupus classification

Overall, 187 adult patients were identified in the University Hospital of Reims (France) with a diagnosis of Lupus. Patients may have presented with non-mutually exclusive forms of lupus but were segmented in mutually exclusive groups: lupus erythematosus (LE; or L93.X only), systemic lupus erythematosus (SLE; M32.X only, or in combination with L93.X. SLE patients were then classified into non-mutually exclusive sub-groups based on different lupus manifestations identified in text searches in patients' EMRs, i.e., lupus nephritis (LN) and cutaneous lupus erythematosus (CLE). See Figure 1.

Figure 1 – Distribution of patients by forms of lupus



> Demographics

Average age was 51.4 years (14.4), 88% females. Among all sub-groups, patients with CLE were the youngest (mean 49.7 years) and patients with CLE+LN the oldest (mean 58.9 years). The proportion of active smokers varied from 21% in LN patients to 31% in the CLE+LN group. See Table 1.

Table 1 - Socio-demographic characteristics of lupus patients

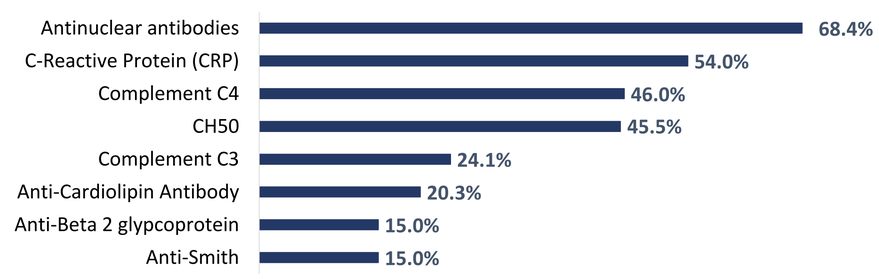
	Overall	LE	SLE	CLE	LN	CLE + LN
Number of patients (%)	187 (100%)	70 (37%)	153 (82%)	64 (34%)	48 (26%)	13 (7%)
Mean age (SD)	51.4 (14.4)	53.6 (14.4)	50.9 (14.1)	49.7 (14.0)	55.1 (15.8)	58.9 (14.3)
Female (%)	164 (88%)	63 (90%)	133 (87%)	59 (92%)	44 (92%)	12 (92%)
Active Smoker (%)	38 (20%)	18 (26%)	35 (23%)	18 (28%)	10 (21%)	4 (31%)

CLE: Cutaneous Lupus Erythematosus; LE: Lupus Erythematosus; LN: Lupus Nephritis; SLE: Systemic Lupus Erythematosus

> Biomarkers

History of lupus-specific and other biomarker testing was retrieved. For instance, antinuclear antibodies, CRP, complement 4, and CH50 tests were found in 68.4%, 54.0%, 46.0% and 45.5% of patients, respectively (lab values not reported here). See Figure 2.

Figure 2 – Proportion of patients with biomarker tests



> Disease activity index

The Systemic Lupus Erythematosus Disease Activity (SLEDAI) was reported in 9.6% of patients EMRs, with an average score of 11.7 (9.6) and a median score of 10.0 ([4.0 ; 17.0]).

> Other conditions

Lupus patients tend to suffer from concurrent comorbidities and/or develop lupus-induced complications. In our study, the most prevalent conditions were cardiovascular (excluding hypertension, 63% of patients), metabolic (40%), polyarthritis (39%), hypertension (39%) and gastrointestinal (37%). Renal manifestations were identified in 35% of patients, including 92% and 100% in LN and CLE+LN subgroups, respectively.

Overall, 165 (88%) of patients suffered from at least one of the conditions listed above. See Figure 3. Overall, 10% and 12% of patients presented with metabolic diseases only and polyarthritis only, respectively, while 63% patients with cardiovascular diseases presented with at least one other condition.

Figure 3 – Distribution of patients across disease combinations*

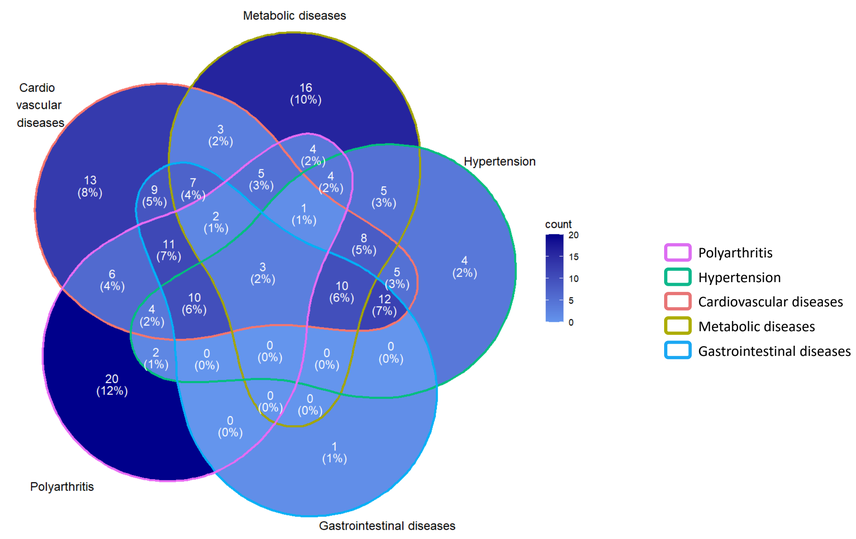


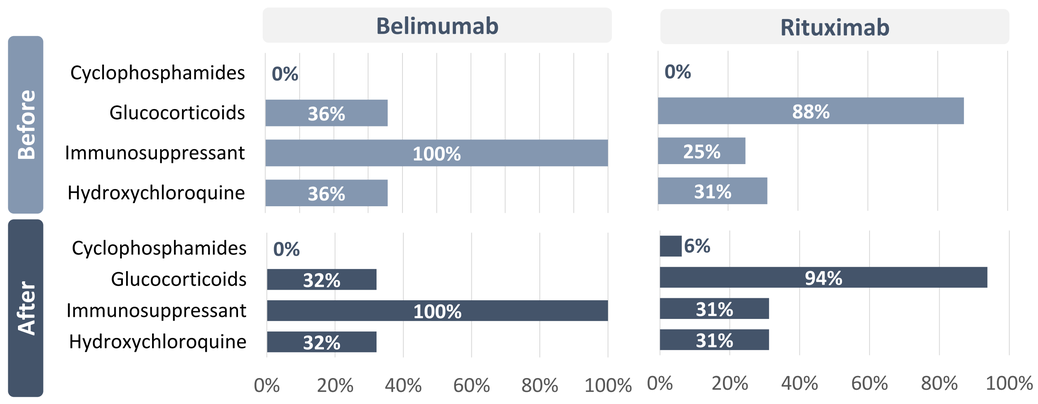
*Figure 3 is based on patients identified with at least one of the conditions above (N=165).
Percentages calculated on the total number of patients with at least one of the conditions above.

> Drug treatments

Nearly half of the patients were treated with glucocorticoids and/or hydroxychloroquine during the study period (51.3% and 47.1%, respectively). Biologics were prescribed less frequently, with 15.0% (n=28) of patients who received belimumab and 8.6% (n=16) rituximab during the study period, including one patient who received both biologics.

All patients treated with belimumab (n=28) received immunosuppressants before and after use of belimumab. 88% (14) and 94% (15) of rituximab-treated patients were treated with glucocorticoids before/after initiation of rituximab, respectively (Figure 4).

Figure 4 – Drug treatments before and after initial of biologics (rituximab and/or belimumab)



CONCLUSION

A NLP-powered solution connected to multidata hospital sources is effective in characterizing complex pathologies, with a breadth and granularity of information, such as biomarkers, disease activity index, and text string, not available in traditional RWE data sources. Methods and NLP developed for this single-center study will be deployed to multihospital systems in France to test the robustness of the findings.

SANCARE HEALTH & DATA

AliraHealth

INSTITUT DE L'INTELLIGENCE ARTIFICIELLE EN SANTÉ REIMS - CHAMPAGNE ARDENNE