

Best of Both Worlds: Enhancing Claims Data With Lab Results in Real-World Evidence Generation

Xinshuo Ma, Francesca Devine, Ting-Ying Huang — Komodo Health, New York, NY, and San Francisco, CA

Introduction

- Health insurance claims have become the mainstream foundation for real-world evidence in recent decades. Yet the lack of clinical details is a well-known limitation of claims databases.
- Thanks to modern technology of tokenization, the linkage between claims and lab results is now more accessible and with patient privacy preserved.

Objective

- Using a recently developed lab dataset as an example, this study explored the data capabilities of lab results and evaluated their potential to enhance claims databases.

Methods

Study Design

- This cross-sectional study assessed the data from January 2016 to November 2023 and reviewed the cohort overlap via linking two databases.
- This study used Komodo Lab Results (KLR) and Komodo Research Dataset (KRD).

Komodo Research Dataset (KRD): Composed of administrative data and claims, KRD captures routinely collected health services utilization records and expenditures for over 330 million de-identified unique individuals in the US. Native to HIPAA-compliant, privacy-preserving tokens, KRD offers extended patient-level observations of medical encounters and outpatient pharmacy dispensings via linkage across health and pharmacy insurance plans. Data availability is as early as 2016. Specialty datasets such as genomics, laboratory test results, and electronic medical records are readily accessible via additional linkage. KRD is the optimized schema of the underlying Healthcare Map™ from Komodo Health for RWE generation and HEOR.

Komodo Lab Results (KLR): Unit-standardized laboratory results of routine and specialized tests generated by over 73 million unique individuals across care settings of hospitals, nursing facilities, and ambulatory offices in the US.

- Atherosclerotic cardiovascular disease (ASCVD) or familial hypercholesterolemia (FH) cohorts and patients identified with ICD-10-CM diagnosis, ICD-10-PCS procedure, and HCPCS procedure codes in KRD who had an LDL-C value were established to assess KLR availability of key test results for clinical features commonly required for continuous monitoring.

Results

- KLR has 33 lab tests and over 3.4 billion results recorded for over 59 million unique patients (as represented by unique tokens) in the US, including blood cell/platelets, neutrophils, immunoglobulins, eGFR, and INR (**Figure 1**).
- Over 46 million patients had serum potassium values, with 65% repeatedly tested and a median 31.2 months between the first and last values. Equivalent estimates respectively for LDL-C and HbA1c were 33 and 28 million, 60% and 58% repeatedly tested, and 35.1 and 30.1 months (**Table 1**).
- On average, 97.1% of patients were also available in KRD. Repeat testing rate was within 1% difference for all included tests between overall KLR patients and the subset who overlapped with KRD (**Figure 2**).

Figure 1. Number of Patients in Komodo Lab Results and Those Who Also Contributed to Komodo Research Dataset

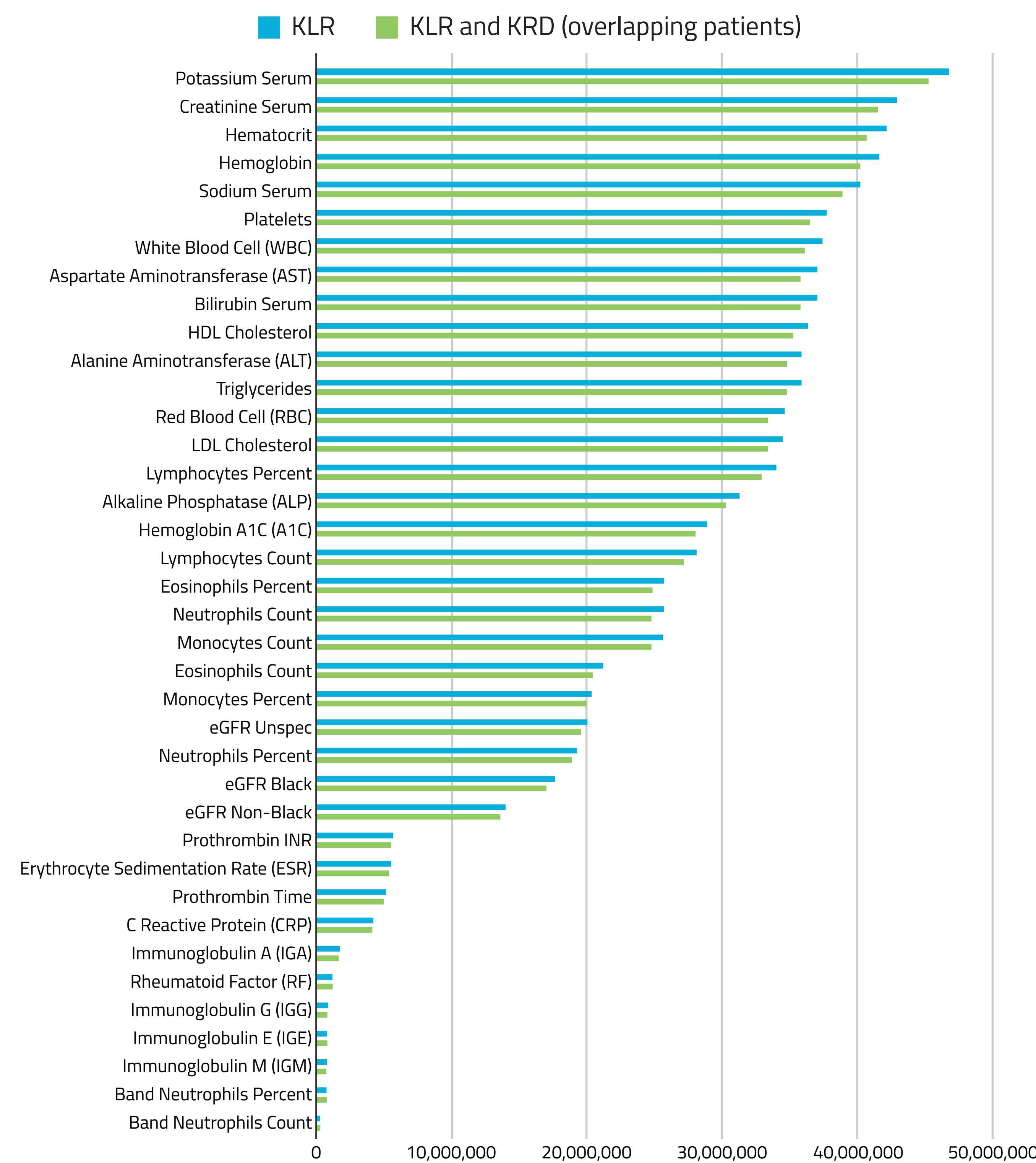


Figure 2. Percentage of Patients With at Least 2 Separate-Day Lab Tests in Komodo Lab Results and Were Available in Komodo Research Dataset

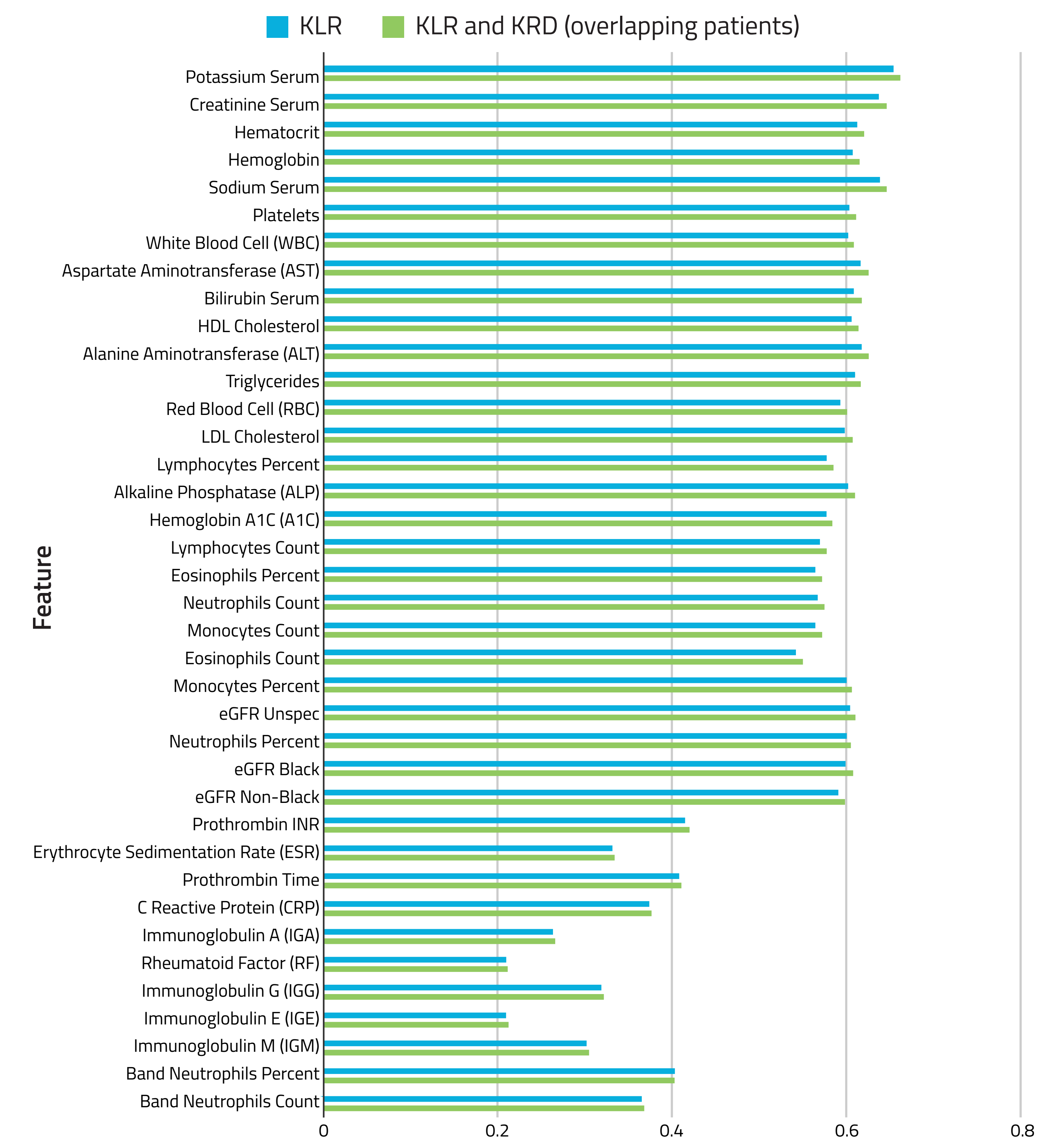


Table 1. Time Between First and Last Tests in Selected Lab Features

Feature	No. of Overlapped Patients	No. of Tests per Patient	Time Between First and Last Tests (Years)	
			Mean	Median
Aspartate Aminotransferase (AST)	35,761,584	4.1	2.9	2.4
HDL Cholesterol	35,209,572	3.3	3.4	2.9
Alanine Aminotransferase (ALT)	34,725,131	4.1	2.9	2.3
Triglycerides	34,791,443	3.3	3.3	2.9
LDL Cholesterol	33,391,600	3.3	3.4	3.0
Hemoglobin A1C (A1C)	28,014,452	3.6	3.1	2.5
eGFR Black	17,014,207	3.7	2.3	1.9
eGFR Non-Black	13,550,762	3.7	2.1	1.7

- Among ASCVD or FH patients diagnosed in 2016 or after, 9.1 million patients had a LDL-C test result in KLR. Of these patients, 45%, 26%, and 13% had their highest LDL-C value greater than or equal to clinical thresholds of 100, 130, and 160 mg/dL, respectively (**Table 2**).

Table 2. LDL-C Results in ASCVD or FH Patients

Feature	No. of Patients	% of Patients
ASCVD patients who were diagnosed after 2016 and had at least 1 LDL-C test	59,759,358	
LDL-C value ≥ 100 mg/dL	4,057,193	45%
LDL-C value ≥ 130 mg/dL	2,367,751	26%
LDL-C value ≥ 160 mg/dL	1,201,950	13%

Conclusion

- Lab datasets contain a wealth of structured clinical information.
- Together with the longitudinality strength of claims databases, lab results are a valuable resource for researchers seeking contextualization of patient health over time and granularity in outcomes assessment.

Scan here to download poster or inquire for more info.

