Poster Code: B-271

# Impact of Variance in Cohort Phenotype Definitions on Real-World Research: An Assessment of ASCVD Definitions on Real-World Data

Scan for poster + more

**Aaron Kamauu MD MS MPH,[1] Jared Kamauu BFA,[1] Aimee Harrison MFA,[1] Michael Buck PhD,[1] Craig Parker MD MS,[1] Allise Kamauu MS,[1] Scott L DuVall PhD[2]**

*[1]Navidence Inc., USA; [2]PurpleLab Inc., USA*

**So What?**

*Different ASCVD definitions produce dramatically different study cohorts — across 6 replicated definitions: total number of distinct codes varied 202-705, and patient counts ranged dramatically with only 29-37% shared across all cohorts.*

*Researchers should publish complete code lists and use data-driven computable operational definition (CODef) selection.*

## BACKGROUND

### Context

In real-world research (RWR), complex clinical concepts like Atherosclerotic Cardiovascular Disease (ASCVD) are defined by combining multiple composite medical conditions into cohort phenotypes. Consistent conceptual and computable operational definitions (CODefs) are essential for meaningful, reproducible results.

### The Gap / Problem

Prior work identified significant variability in ASCVD conceptual and operational definitions. However, the direct impact on actual patient cohorts in real-world data (RWD) had not been quantified.

*See companion posters **B-269, B257 & B-256***

### HEOR / RWE Relevance

FDA Guidance (2023) states: "small differences in the choice of operational definition … may have a large impact on study results." This study provides empirical evidence quantifying that impact, demonstrating that definition choice directly shapes study cohorts.

## OBJECTIVES

### Primary Objective

*To assess the impact of definition variability on cohort phenotypes by replicating published ASCVD definitions in two real-world data sources and comparing the resulting patient cohorts.*

### Secondary Objectives

- Quantify overlap of patient coverage across replicated ASCVD cohort definitions
- Identify significant differences that impact study outcomes
- Demonstrate the relationship between code count and cohort size

## METHODS

### Study Design & Data Source

In this cross-sectional cohort replication study, six published ASCVD cohort definitions, with ICD-10-CM / ICD-10-CA / ICD-9-CM code lists, were identified from a targeted literature review. All six were replicated in PurpleLab® CLEAR Claims (US national claims) and MIMIC-IV (EHR from 28 hospitals).

***Figure 1: Literature Review → Cohort Definition Selection***

**Targeted Literature Review**
**ASCVD Studies**
(n = 11)

↓ *Published Code Lists Required*

**ASCVD Cohort Definitions**
**w/ Published Code Lists**
(n = 6)

↓ *202–705 ICD Codes Each*

**Replicated Cohorts in PurpleLab® CLEAR Claims and MIMIC IV**
(n = 21,929,791 | n = 14,782)

↓ *Exact Code List Replication*

**Compared Cohort Overlap**
(Patient Coverage & Differences)

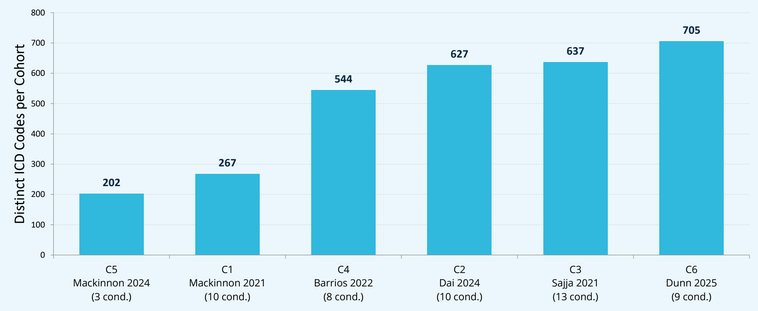
## METHODS (CONT.)

### Computable Operational Definitions (CODefs™)

Study elements — exposures, outcomes, and covariates — were defined using Navidence CODefs™: rigorous, data source-agnostic algorithms that ensure precision and reproducibility across RWR.

## RESULTS

***Figure 2: # of Distinct Codes per ASCVD Cohort Definition***




| 6 | 14.8K, 21.9M | 29-37% |
|---|---|---|
| ASCVD Cohort Definitions | Qualifying Patients per Data Source | Patients Common Across *All* Cohorts |

***Figure 3: Included Conditions by ASCVD Cohort Definition***

| Condition | C1 | C2 | C3 | C4 | C5 | C6 | Count |
|---|---|---|---|---|---|---|---|
| Peripheral Artery Disease (PAD) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6/6 |
| Ischemic Stroke | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6/6 |
| Myocardial Infarction (MI) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6/6 |
| Stable Angina | ✓ | ✓ | ✓ | ✓ | | ✓ | 5/6 |
| Unstable Angina | ✓ | ✓ | ✓ | ✓ | | ✓ | 5/6 |
| Abdominal Aortic Aneurysm | ✓ | ✓ | ✓ | ✓ | | | 4/6 |
| Descending Thoracic Aneurysm | ✓ | ✓ | ✓ | ✓ | | | 4/6 |
| Transient Ischemic Attack (TIA) | ✓ | | ✓ | ✓ | | ✓ | 4/6 |
| Coronary Revascularization | ✓ | ✓ | ✓ | | | | 3/6 |
| Other Arterial Revascularization | ✓ | ✓ | ✓ | | | | 3/6 |
| Acute Coronary Syndrome (ACS) | | ✓ | ✓ | | | ✓ | 3/6 |
| Coronary Artery Disease (CAD) | | | ✓ | | | ✓ | 2/6 |
| Cerebrovascular Disease | | | ✓ | | | ✓ | 2/6 |

***Figure 4: Patient Count % across ASCVD Cohort Definitions Replicated in RWD***

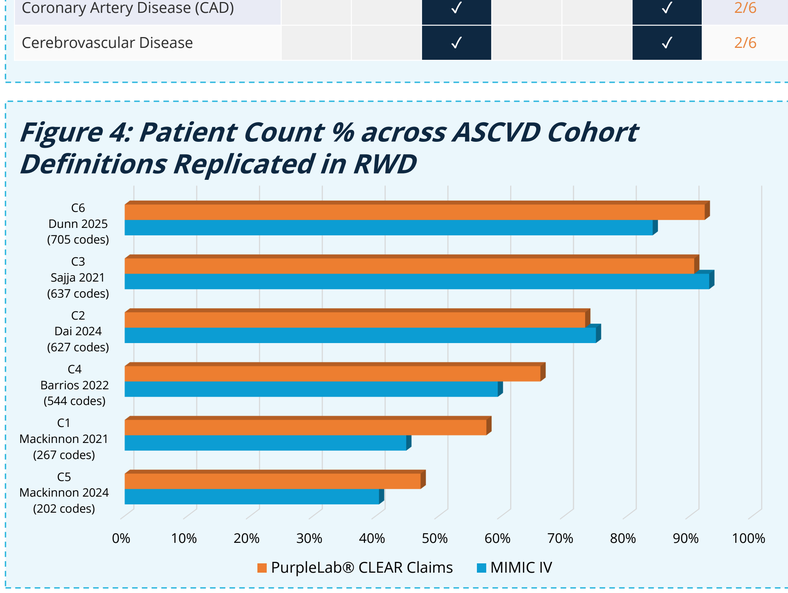
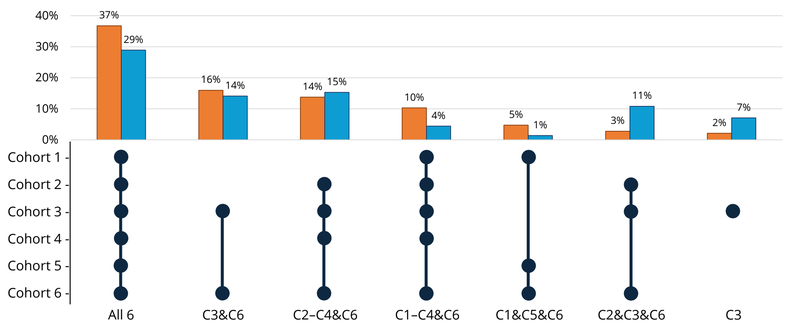


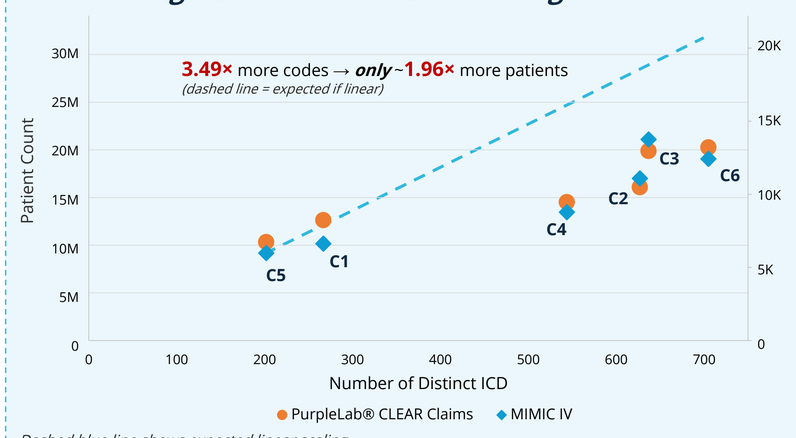

***Figure 5: Top 7 Patient Overlaps across ASCVD Cohort Definitions Replicated in RWD***




*UpSet plot visualizing the overlaps/intersections, offering a scalable alternative to Venn diagrams. How to read it: Each column represents a specific overlap (dots), and the graph show the results of that intersection.*

## RESULTS (CONT.)

***Figure 6: # of Distinct Codes per ASCVD Cohort Definition vs. Resulting Patient Counts: Diminishing Returns***




*Dashed blue line shows expected linear scaling.*

## CONCLUSIONS

### Key Findings

- Six published ASCVD definitions used 202-705 distinct codes covering 3-13 conditions — variance compounds.
- Despite >3x more codes, the largest cohort had ~2x patients — impact is driven by which codes, not how many.
- Only 29-37% of patients were shared across all 6 definitions, demonstrating that cohorts are not interchangeable.
- Data-driven CODef selection is essential for research cohorts given the significant impact on health outcomes and endpoints.

### Clinical Research Implications

- Researchers, regulators, and payers ***expect*** published, complete CODefs and code lists for all RWR studies.
- While study needs ultimately determine final decisions, definition selection should be ***informed by data*** with transparent reporting to enable proper cross-study comparison.
- Publishing complete definitions supports the FDA's call for rigor in operational definitions for RWE-based regulatory decision-making.

## LIMITATIONS

*Patient overlap analysis based on PurpleLab® CLEAR Claims and MIMIC-IV; results may vary across other data sources.*

*Cross-cohort patient characteristic comparisons planned as next step.*

## REFERENCES

1. FDA. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. 2023.
2. Kamauu J, et al. Defining ASCVD in RWR: An Analysis of Conceptual and Operational Definition Variability. ISPOR 2026; **Poster B-269**.
3. Mackinnon ES, et al. Increasing Prevalence and Incidence of ASCVD in Adult Patients in Ontario, Canada. CJC Open. 2021;4(2):206-213.
4. Dai D, et al. Multimorbidity in ASCVD and Its Associations With Adverse CV Events and Healthcare Costs. J Health Econ Outcomes Res. 2024;11(1):75-85.
5. Sajja A, et al. A simplified approach to identification of risk status in patients with ASCVD. Am J Prev Cardiol. 2021;7:100187.
6. Barrios V, et al. Treatment patterns and use of healthcare resources of patients with ASCVD. Front Cardiovasc Med. 2022;9:966049.
7. Mackinnon ES, et al. Real-World Risk of Recurrent CV Events in ASCVD Patients. Cardiol Ther. 2024;13(1):205-220.
8. Dunn TJ, et al. Comparing characteristics and costs among patients with T2D and established ASCVD. J Manag Care Spec Pharm. 2025;31(2):117-126.

## DISCLOSURES & ACKNOWLEDGMENTS

*A Kamauu, J Kamauu, M Buck & C Parker: Employees of Navidence Inc. A Kamauu, C Parker: Co-Founders of Navidence Inc. A Harrison: Employee of Navidence Inc. during this work. S DuVall: Employee of PurpleLab.*
*No conflicts of interest to disclose.*



Contact: aaron@navidence.com