Poster Code: RWD123

One Concept, Many Cohorts: Impact of Cohort Phenotype Definitional Variance in Depression, Major Depressive Disorder and Treatment-Resistant Depression on Real-World Research

Aaron Kamauu MD MS MPH[1] Jared Kamauu BFA[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?

Choose depression definitions by ***which*** codes drive the cohort, not how many, **and publish complete code lists!**

BACKGROUND

Context

Depression, MDD and TRD are operationalized in real-world research through **computable operational definitions (CODefs)** — most often ICD-10-CM code lists in claims.

Our companion literature review found:

24 references reviewed

33 distinct depression / MDD / TRD CODefs

7 distinct depression / MDD code lists

The Gap

How much code-list choices change **who enters a depression cohort** in real-world data had not been quantified.

Minor CODef differences "may have a large impact on study results." **— FDA, 2023**

OBJECTIVES

Primary: *Quantify how variance across 10 published depression CODefs changes cohort size and overlap.*

- Compare code-list content across definitions
- Quantify patients shared by all 10 vs. definition-specific
- Map where definition-specific patients come from

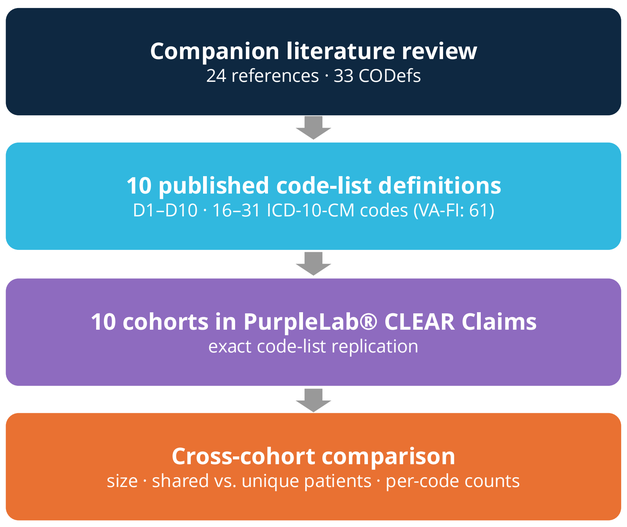
METHODS

Study Design & Data Source

Cross-sectional replication in **PurpleLab® CLEAR Claims** (U.S. all-payer). Each definition used its **exact published code list** (ICD-9-CM translated; ICD-10 adapted to ICD-10-CM).

Study period: January 2015 – May 2026

Figure 1: From literature to cohorts



Analysis

Patients per cohort, and patient overlap across every combination of definitions (**UpSet** analysis).

LIMITATIONS

- Code lists applied as published; timing, setting and claim-count criteria were not varied.
- No reference standard in claims: results show *who* is selected, not which definition is most valid for a given use case.
- U.S. claims; patterns may differ in other RWD and in other geographies.

REFERENCES

1. FDA. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. 2023.
2. Kamauu JH, et al. Variation in Conceptual and Operational Definitions of Depression in RWR. ISPOR EU 2026 (companion).
3. Märtens C, et al. J Psychosom Res. 2026;204:112579.
4. Pilon D, et al. PLoS One. 2019;14(10):e0223255.
5. Engels A, et al. BMC Psychiatry. 2020;20(1):591.
6. Xu R, et al. medRxiv. 2024 (PMID 39211891).
7. Solberg LI, et al. Am J Med Qual. 2006;21(4):238-45.
8. Hwang S, et al. Am J Geriatr Psychiatry. 2015;23(7):726-34.
9. Cheng D, et al. (VA-FI) J Gerontol A Biol Sci Med Sci. PMID 34272940.
10. CMS Chronic Conditions Data Warehouse. Chronic Condition Algorithms: Depression.

DISCLOSURES & ACKNOWLEDGMENTS

A Kamauu, J Kamauu, M Buck, C Parker, A Kamauu: employees of Navidence Inc.; A Kamauu & C Parker: co-founders. SL DuVall: employee of PurpleLab Inc., which provided CLEAR Claims data.

RESULTS

10 — Published depression CODefs replicated

16–31 — Distinct ICD-10-CM codes per definition (VA-FI: 61)

87M–110M — Patients per cohort (119M in any definition)

86M — Patients in all 10 definitions (72% of 119M)

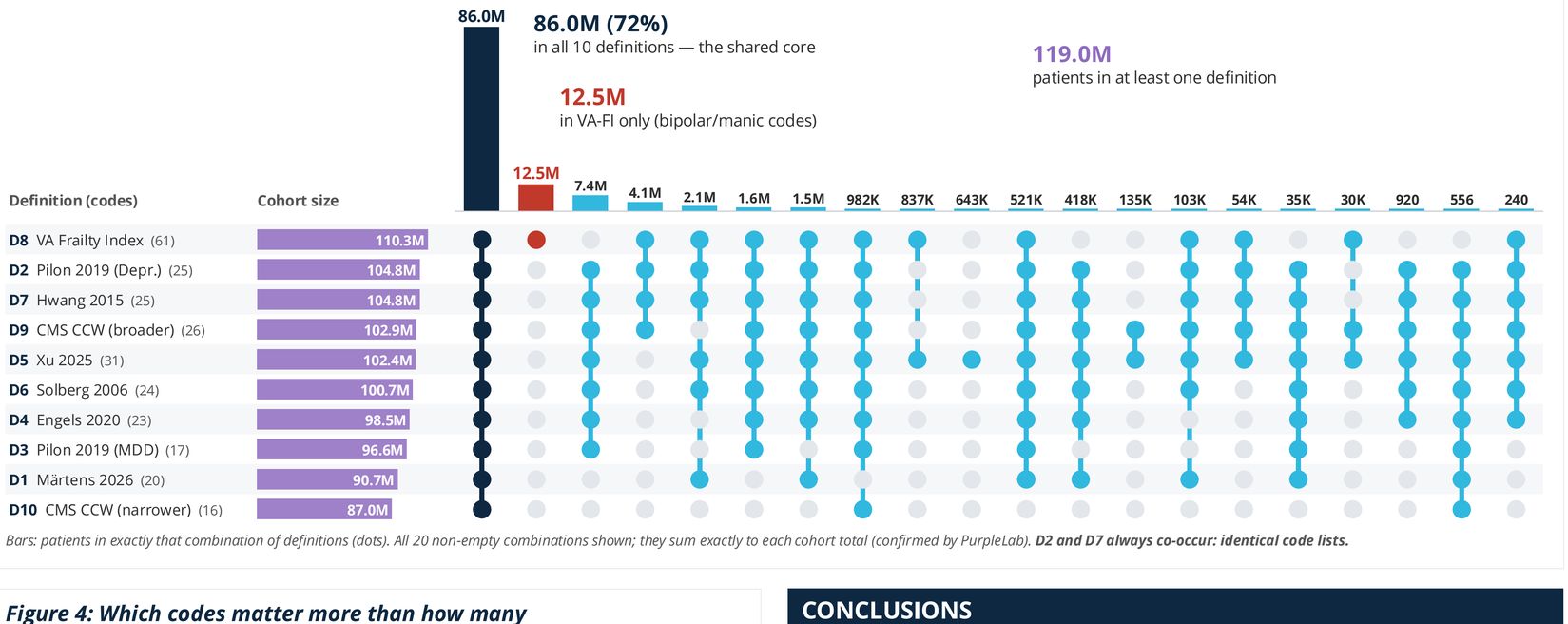
Figure 2: Code groups included by each definition — a shared MDD core, divergent edges

ICD-10-CM	Code group	D1 Märtens 2026	D2 Pilon 2019 (Depr.)	D3 Pilon 2019 (MDD)	D4 Engels 2020	D5 Xu 2025	D6 Solberg 2006	D7 Hwang 2015	D8 VA Frailty Index	D9 CMS CCW (broader)	D10 CMS CCW (narrower)	# of 10
SHARED CORE · included in all 10 definitions												
F32.0–F32.3	MDD, single episode (mild → severe)	■	■	■	■	■	■	■	■	■	■	10
F32.9	MDD, single episode, unspecified	■	■	■	■	■	■	■	■	■	■	10
F33.0–F33.3, F33.9	MDD, recurrent (incl. unspecified)	■	■	■	■	■	■	■	■	■	■	10
F33.4x	MDD, recurrent, in remission	■	■	■	■	■	■	■	■	■	■	10
VARIABLE EDGES · where definitions diverge												
F32.4–F32.5	MDD, single episode, in remission		■	■	■	■	■	■	■	■	■	9
F32.A	Depression, unspecified		■	■	■	■	■	■		■		7
F32, F32.8x, F33, F33.8	Other / header depressive codes	5/6	■		■	■	■	■	3/6	■		8
F34.1	Dysthymic disorder	■	■			■	■	■	■			6
F34.8x, F34.9, F39	Other persistent / unspecified mood					3/5			■			2
F06.3x	Mood disorder due to physiological cond.					■				2/3		2
F43.21, F43.23	Adjustment disorder w/ depressed mood		1/2					1/2	■	1/2		4
F30.x, F31.x	Manic episode & bipolar disorder								■			1
N94.3	Premenstrual tension syndrome					■						1
Distinct ICD-10-CM codes per definition*		**20**	**25**	**17**	**23**	**31**	**24**	**25**	**61**	**26**	**16**	

■ All codes in group included ▭ Partial (codes included / codes in group) Not included

** Codes as implemented in PurpleLab.*

Figure 3: Patient overlap across all 10 definitions (UpSet plot)



Bars: patients in exactly that combination of definitions (dots). All 20 non-empty combinations shown; they sum exactly to each cohort total (confirmed by PurpleLab). ***D2 and D7 always co-occur: identical code lists.***

Figure 4: Which codes matter more than how many



D3 (17 codes) has 5.9M more patients than D1 (20 codes); D5 (31 codes) has fewer than D2/D7 (25).

CONCLUSIONS

Key Findings

- All 10 published definitions share the same core MDD codes; **86.0M patients** qualify under this shared core set of codes.
- Yet cohorts ranged **87.0M–110.3M**: a 23M-patient swing from code-list choice alone.
- VA-FI's bipolar/manic codes add **12.5M** patients no other definition captures.
- Code count does not predict size; two published definitions are identical (D2 = D7).

Implications for RWE

- Select depression CODefs **empirically**, knowing what each code adds.
- Publish complete code lists so cohorts can be reproduced.
- Report results across alternative CODefs.

Contact: aaron@navidence.com