

Development and Assessment of a Portability Strategy for Computable Operational Definitions

James R. Rogers¹; Amanda Shields²; Vladimir Turzhitsky¹; Shefali Haldar¹; Aimee Harrison²; Aaron Kamau²

¹Merck & Co., Inc., Rahway, NJ, USA; ²Navidence, Aurora, CO, USA

Background and objective

- Computable operational definitions (CODEfs) are representations of medical concepts in a format that supports precision in real-world data (RWD) analyses
- Portability, defined as applying CODEfs to different data sources or across different time periods, is important for establishing consistency and reproducibility across RWD analyses, but is difficult to formally evaluate, given heterogeneity in CODEf development environments
- **The objective of this study is to compare different validated literature review CODEfs by using weakly supervised machine learning as a possible benchmark, using type 2 diabetes mellitus (T2DM) as an example**
- This study consists of 3 parts: (1) identify validated T2DM CODEf from the literature, (2) build the silver standard for benchmarking, and (3) benchmark the literature review definitions against the silver standard

Targeted literature review: Methods

- A targeted literature review was conducted using PubMed to identify published scientific articles that developed CODEfs for T2DM
- Articles were included if they presented at least one validated CODEf for T2DM that was developed using a RWD source
- Extracted data included CODEf characteristics and operational setup, validation metrics, RWD used, and other study-related details. Relevant covariates were also extracted to inform model-building step

Targeted literature review: Results

- Per **Table 1**, a total of 12 CODEfs were identified that could be feasibly used for this study (ie, fit-for-use with corresponding data elements)
- The most common validated CODEf was “at least 1 T2DM dx” with 8 references
- The majority of CODEfs were derived in the US and used EHR data, as the most common form of validation was manual chart review

Table 1. Literature review CODEfs derived for targeted literature review

Literature review CODEf	Number of references	Geography of data	Data source(s) leveraged
At least 1 T2DM dx and at least 1 T2DM rx	2	USA, Japan	EHR, administrative claims
At least 1 T2DM rx	3	USA, Australia	EHR
At least 1 T2DM rx (does not include pramlintide)	1	USA	EHR
At least 1 T2DM rx (does not include insulin and pramlintide)	1	USA	EHR
At least 1 T2DM dx or at least 1 T2DM rx	1	USA	EHR
At least 1 T2DM dx	8	USA, Japan, Spain	EHR, administrative claims
At least 1 T2DM dx and no T1DM dx	1	USA	EHR
At least 1 inpat T2DM dx or at least 2 T2DM dx	1	USA	EHR
At least 2 T2DM dx	2	USA	EHR
At least 1 T2DM dx, no T1DM dx, and no insulin	1	USA	EHR
At least 3 T2DM dx, no T1DM dx, at least 1 T2DM rx (does not include insulin and pramlintide)	1	USA	EHR
At least 1 inpat T2DM dx	1	Italy	Registry

USA, United States of America; EHR, electronic health record.

Silver standard: Methods

- The silver standard for this project was of outputs from machine learning models in which literature review results were benchmarked against, all using data from Optum’s deidentified Clinformatics® Data Mart Database
- To build the silver standard, we leverage work from Swerdel JN et al., 2019
- The training cohort consisted of patients from Jan 2017 to Dec 2018, while the testing cohort consisted of patients entering Jan 2019 to Dec 2019
 - A case was defined as a patient with at least 5 T2DM dx
 - Noncases were a randomly sampled set of individuals with no T2DM dx, based on the estimated T2DM prevalence per Xu G et al., 2018
- Random forest, LASSO, and XGBoost models were built, with model inputs being T2DM covariates identified from the literature review process
- Best-performing models were based on highest F1-scores
- Silver standards were determined based on majority vote from the best-performing models
- These standards, along with the literature review CODEfs, were then applied to the testing set for benchmarking, focusing on sensitivity, positive predictive value (PPV), and F1-score

Silver standard: Results

- Testing data contained 32,334 cases and 301,006 noncases, with the silver standards leading to 26,878 cases to be benchmarked against
- **Table 2** presents validation metrics comparing the silver standard against the literature review CODEfs
- Relative to the silver standard, best-performing CODEfs per F1-score were “at least 1 T2DM dx and at least 1 T2DM rx,” “at least 1 T2DM rx (does not include pramlintide),” and “at least 1 T2DM rx”
- Validation statistics were inconsistent between what was benchmarked with the silver standard and what was found in the publications for the selected CODEfs

Table 2. Comparison of literature review CODEfs against silver standard

		Benchmarked against silver standard			Reported in literature	
Literature review CODEf	Cases identified	F1-score	Sensitivity	PPV	Sensitivity ^a	PPV ^a
At least 1 T2DM dx and at least 1 T2DM rx	25,259	0.93	0.91	0.96	0.10	0.60 to 0.96
At least 1 T2DM rx	28,319	0.93	0.95	0.90	0.24 to 0.97	0.80
At least 1 T2DM rx (does not include pramlintide)	28,310	0.93	0.95	0.90	0.62	0.80
At least 1 T2DM rx (does not include insulin and pramlintide)	20,632	0.79	0.70	0.91	0.58	0.81
At least 1 T2DM dx or at least 1 T2DM rx	42,031	0.77	0.99	0.63	0.55 to 0.94	0.81 to 0.84
At least 1 T2DM dx	38,971	0.77	0.94	0.65	0.53 to 0.99	0.21 to 0.92
At least 1 T2DM dx and no T1DM dx	37,571	0.75	0.90	0.65	0.55	0.64
At least 1 inpat T2DM dx or at least 2 T2DM dx	23,852	0.71	0.67	0.75	0.84	0.78
At least 2 T2DM dx	22,246	0.70	0.64	0.78	0.42 to 0.83	0.28 to 0.73
At least 1 T2DM dx, no T1DM dx, and no insulin	31,560	0.63	0.69	0.58	0.47	0.70
At least 3 T2DM dx, no T1DM dx, at least 1 T2DM rx (does not include insulin and pramlintide)	4,887	0.30	0.18	0.99	0.26	1.00
At least 1 inpat T2DM dx	6,606	0.27	0.17	0.68	0.12	-

^aSome include ranges because of multiple references; a “dash” indicates just one reference without the validation metric performed.

Comparing cohort characteristics: Results

- **Table 3A, 3B,** and **3C** present baseline demographic (at cohort entry), comorbidity (within past year), and medication (within past year) characteristics related to T2DM for the silver standard and the CODEfs applied to the testing data
- Per **Table 3A**, demographics were generally consistent across all cohorts except for slightly higher capture of non-Hispanic patients in the “at least 1 inpat T2DM dx” algorithm relative to the others
- Per **Table 3B**, the silver standard cohort and “at least 1 T2DM dx and at least 1 T2DM rx” cohort were fairly consistent, though the latter had slightly lower prevalence of baseline conditions. When comparing the silver standard cohort to the “at least 1 T2DM rx” cohort, the former had a higher prevalence of cardiovascular conditions and neuropathy
- Per **Table 3C**, the silver standard cohort, “at least 1 T2DM dx and at least 1 T2DM rx” cohort, and “at least 1 T2DM rx” cohort had similar prevalence of medication capture

Conclusions

- When compared against the silver standard, the best-performing literature review CODEfs were “at least 1 T2DM dx and at least 1 T2DM rx” and “at least 1 T2DM rx” and were reflected in similar baseline characteristics
 - CODEfs involving medications performed better, demonstrating that involvement of a medication component is important for appropriate T2DM CODEf construction
 - Other components, such as increased count of components (eg, multiple dx) and inpatient status, led to noticeable decreases in sensitivity, though those aspects could reflect a different T2DM CODEf of interest (eg, severe cases) that would require a change in the silver standard focus
- As T2DM is well-studied and prevalent in the US, the underlying models for the silver standard were largely reflective of common diagnostic journeys for those patients
 - Despite trivial representation from the models, the silver standard provides a common benchmark for varying CODEfs to be measured against, thus assisting in CODEf selection
 - As an extended option, the underlying models may serve as a CODEf construction tool for less-studied or poorly understood conditions

Table 3. Highlighted baseline demographics*

Table 3A. Demographics

CODEf	N	Age	Sex		Race				Ethnicity		
		Mean (SD)	Female	Male	Asian	Black	White	Unknown	Hispanic	Non-Hispanic	Unknown
Silver standard	26878	65.50 (12)	13341 (50)	13537 (50)	921 (3)	4029 (15)	15380 (57)	6548 (24)	3381 (13)	18496 (69)	5001 (19)
At least 1 T2DM dx and at least 1 T2DM rx	25259	64.45 (12)	12689 (50)	12570 (50)	898 (4)	3848 (15)	14260 (56)	6253 (25)	3266 (13)	17250 (68)	4743 (19)
At least 1 T2DM rx	28319	62.38 (14)	14785 (52)	13534 (48)	987 (4)	4045 (14)	15834 (56)	7453 (26)	3534 (12)	19101 (67)	5684 (20)
At least 1 T2DM rx (does not include pramlintide)	28310	62.39 (14)	14779 (52)	13531 (48)	987 (4)	4045 (14)	15828 (56)	7450 (26)	3534 (12)	19094 (67)	5682 (20)
At least 1 T2DM rx (does not include insulin and pramlintide)	20632	62.82 (14)	10804 (52)	9828 (48)	842 (4)	2808 (14)	11391 (55)	5591 (27)	2651 (13)	13727 (66)	4254 (21)
At least 1 T2DM dx or at least 1 T2DM rx	42031	64.63 (14)	21998 (52)	20033 (48)	1450 (3)	5987 (14)	23903 (57)	10691 (25)	4779 (11)	28843 (69)	8409 (20)
At least 1 T2DM dx	38971	66.14 (13)	19902 (51)	19069 (49)	1361 (4)	5790 (15)	22329 (57)	9491 (24)	4511 (12)	26992 (69)	7468 (19)
At least 1 T2DM dx and no T1DM dx	37571	66.34 (12)	19170 (51)	18401 (49)	1334 (4)	5594 (15)	21499 (57)	9144 (24)	4395 (12)	26005 (69)	7171 (19)
At least 1 inpat T2DM dx or at least 2 T2DM dx	23852	67.10 (12)	12039 (50)	11813 (50)	785 (3)	3749 (16)	13685 (57)	5633 (24)	3013 (13)	16639 (70)	4200 (18)
At least 2 T2DM dx	22246	67.04 (12)	11217 (50)	11029 (50)	745 (3)	3536 (16)	12688 (57)	5277 (24)	2870 (13)	15459 (70)	3917 (18)
At least 1 T2DM dx, no T1DM dx, and no insulin	31560	66.72 (13)	16060 (51)	15500 (49)	1208 (4)	4505 (14)	18039 (57)	7808 (25)	3621 (12)	21744 (69)	6195 (20)
At least 3 T2DM dx, no T1DM dx, at least 1 T2DM rx (does not include insulin and pramlintide)	4887	67.57 (11)	2424 (50)	2463 (50)	209 (4)	754 (15)	2731 (56)	1193 (24)	815 (17)	3342 (68)	730 (15)
At least 1 inpat T2DM dx	6606	68.14 (12)	3301 (50)	3305 (50)	148 (2)	1089 (16)	3995 (60)	1374 (21)	655 (10)	4918 (74)	1033 (16)

Table 3B. Conditions

CODEf	N	Cardiovascular conditions	Cerebrovascular conditions	Nephropathy	Neuropathy	Ophthalmic conditions	Peripheral vascular
Silver standard	26878	10337 (38)	2711 (10)	7052 (26)	8758 (33)	4699 (18)	5469 (20)
At least 1 T2DM dx and at least 1 T2DM rx	25259	9156 (36)	2348 (9)	6007 (24)	7556 (30)	4273 (17)	4683 (18)
At least 1 T2DM rx	28319	9396 (33)	2404 (8)	6099 (22)	7715 (27)	4383 (16)	4732 (17)
At least 1 T2DM rx (does not include pramlintide)	28310	9392 (33)	2404 (8)	6098 (22)	7713 (27)	4381 (16)	4731 (17)
At least 1 T2DM rx (does not include insulin and pramlintide)	20632	6276 (30)	1511 (7)	3572 (17)	4467 (22)	2268 (11)	2802 (14)
At least 1 T2DM dx or at least 1 T2DM rx	42031	15806 (38)	4177 (10)	9723 (23)	10732 (26)	5957 (14)	7331 (17)
At least 1 T2DM dx	38971	15566 (40)	4121 (11)	9631 (25)	10573 (27)	5847 (15)	7282 (19)
At least 1 T2DM dx and no T1DM dx	37571	14938 (40)	3909 (10)	9172 (24)	9975 (26)	5380 (14)	6906 (18)
At least 1 inpat T2DM dx or at least 2 T2DM dx	23852	11007 (46)	3059 (13)	8105 (34)	9249 (39)	5054 (21)	6122 (26)
At least 2 T2DM dx	22246	10042 (45)	2736 (12)	7855 (35)	9071 (41)	4965 (22)	5925 (27)
At least 1 T2DM dx, no T1DM dx, and no insulin	31560	12258 (39)	3165 (10)	6975 (22)	7209 (23)	3706 (12)	5252 (17)
At least 3 T2DM dx, no T1DM dx, at least 1 T2DM rx (does not include insulin and pramlintide)	4887	2281 (47)	617 (13)	2047 (42)	2582 (53)	1344 (28)	1777 (36)
At least 1 inpat T2DM dx	6606	4559 (69)	1565 (24)	2700 (41)	2611 (40)	1192 (18)	2012 (30)

Table 3C. Medications

CODEf	N	AGI	DDP4i	GLP-1 agonists	Insulin	Meglitinides	Metformin	Pramlintide	SGLT2i	Sulfonylureas	TZD
Silver standard	26878	80 (<1)	3527 (13)	2812 (10)	6849 (26)	171 (1)	20137 (75)	8 (<1)	2311 (9)	7392 (28)	1638 (6)
At least 1 T2DM dx and at least 1 T2DM rx	25259	85 (<1)	3514 (14)	2910 (12)	6939 (28)	170 (1)	19513 (77)	8 (<1)	2361 (9)	7339 (29)	1649 (6)
At least 1 T2DM rx	28319	143 (<1)	3538 (12)	3103 (11)	7687 (27)	172 (1)	21621 (76)	9 (<1)	2386 (8)	7427 (26)	1683 (6)
At least 1 T2DM rx (does not include pramlintide)	28310	143 (<1)	3538 (12)	3100 (11)	7678 (27)	172 (1)	21620 (76)	0 (0)	2384 (8)	7427 (26)	1683 (6)
At least 1 T2DM rx (does not include insulin and pramlintide)	20632	121 (1)	2642 (13)	1757 (8)	0 (0)	110 (<1)	18030 (87)	0 (0)	1671 (8)	5864 (28)	1308 (6)
At least 1 T2DM dx or at least 1 T2DM rx	42031	143 (<1)	3538 (8)	3103 (7)	7687 (18)	172 (<1)	21621 (51)	9 (<1)	2386 (6)	7427 (18)	1683 (4)
At least 1 T2DM dx	38971	85 (<1)	3514 (9)	2910 (8)	6939 (18)	170 (<1)	19513 (50)	8 (<1)	2361 (6)	7339 (19)	1649 (4)
At least 1 T2DM dx and no T1DM dx	37571	82 (<1)	3418 (9)	2796 (7)	6011 (16)	161 (<1)	19098 (51)	2 (<1)	2271 (6)	7175 (19)	1614 (4)
At least 1 inpat T2DM dx or at least 2 T2DM dx	23852	65 (<1)	2628 (11)	2283 (10)	5824 (24)	144 (1)	12772 (54)	4 (<1)	1717 (7)	5452 (23)	1228 (5)
At least 2 T2DM dx	22246	63 (<1)	2582 (12)	2252 (10)	5751 (26)	143 (1)	12249 (55)	4 (<1)	1695 (8)	5339 (24)	1204 (5)
At least 1 T2DM dx, no T1DM dx, and no insulin	31560	63 (<1)	2578 (8)	1554 (5)	0 (0)	106 (<1)	15781 (50)	0 (0)	1629 (5)	5695 (18)	1262 (4)

*All characteristics are presented as N (%), except for age
AGI, alpha-glucosidase inhibitors; DDP4i, dipeptidyl peptidase 4 inhibitors; GLP-1, glucagon-like peptide-1; SGLT2i, sodium-glucose cotransporter 2 inhibitors; TZD, thiazolidinedione.

References

1. Swerdel JN, et al. *J Biomed Inform.* 2019;97:103258.
2. Xu G, et al. *BMJ.* 2018;362:k1497.

Conflict of interest statement

JRR, VT, and SH are employees of Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA, who may own stock and/or hold stock options in Merck & Co., Inc., Rahway, NJ, USA.
AS, AH, and AK are employees of Navidence Inc, Aurora, CO, USA who may own stock and/or hold stock options in Navidence Inc.