

**Title:** Common Data Elements for Maternal Health Research: An OMOP-CDM Concept Mapping Study

**Authors:** Andreea Creanga,<sup>1,2</sup> Elizabeth Stierman,<sup>1</sup> Carrie Wolfson,<sup>1</sup> Benjamin Martin,<sup>2</sup> Khyzer Aziz,<sup>2</sup> Meighan Mary,<sup>1</sup> Sarah Clifford,<sup>1</sup> Amanda Burgess,<sup>1</sup> Paul Nagy<sup>1,2</sup>

**Affiliations:** <sup>1</sup> Johns Hopkins Bloomberg School of Public Health; <sup>2</sup> Johns Hopkins School of Medicine.

**Background:** Globally, about 15% of pregnant women experience life-threatening complications and, every day, >1,000 women die from pregnancy-related causes.<sup>1</sup> Despite considerable progress over the past three decades to improve maternal outcomes around the world,<sup>2</sup> there is still much to learn about the effectiveness of new and promising interventions, and how best to implement effective interventions in various settings. To enable such learning, it is critical for maternal health researchers to utilize common data elements (CDEs), therefore ensure consistency in data collection and promote interoperability and collaboration across research endeavors. In April 2024, with funding from and in close collaboration with the National Institutes of Health (NIH), the [Maternal Health Data Innovation & Coordination Hub](#) at Johns Hopkins University (JHU) convened a panel of experts (n=67) to support the development of CDEs for use in maternal health research. The panel includes representatives from academic research centers (e.g., NIH-funded Maternal Health Research Centers of Excellence and Maternal-Fetal Medicine Units Network), professional organizations (e.g., American College of Obstetricians and Gynecology, Society for Maternal Fetal Medicine), various NIH institutes, and other federal health agencies in the United States. In parallel, the Maternal Health Data Innovation & Coordination Hub works to facilitate the use of real-life data structured with the OMOP-CDM across 10 Maternal Health Research Centers of Excellence in the United States<sup>3</sup> through the newly established *OHDSI Maternal Health Data Science Fellowship*.<sup>4</sup> To support both areas of work, this analysis examines the extent to which maternal health constructs prioritized by the expert panel are available in the OMOP-CDM.

**Methods:** Two modified Delphi exercises are being undertaken to derive a minimum set of priority CDEs for maternal health. The purpose of Delphi 1 (completed in May 2024; response rate 83.6% for both rounds) was to prioritize constructs (i.e., topics of interest grouped as biomedical and psychosocial) that should be captured in a standardized manner across all maternal health research. Two voting rounds were used to reach consensus on construct prioritization (defined as >75% “yes, include” votes), with the first round allowing panelists to “write-in” constructs that others could vote on during the second round. The purpose of Delphi 2 (planned for July 2024) is to recommend specific CDEs for priority constructs identified during Delphi 1.

Using the search function in ATLAS and the JHU OMOP-CDM instance, we aimed to identify concepts that correspond to the priority constructs from Delphi 1. We quantified the number of standard and valid concepts (+ if <10; ++ if 10-99; +++ if ≥100) identified for each priority construct and, among these, examined the common types of domains (condition, drug, measurement, observation, or procedure) and standard vocabularies (SNOMED, LOINC, INCD10CM, CPT4, RxNorm, OMOP Extension) used for mapping. We also noted whether the identified standard and valid concepts yield (yes/no) descendant person counts (DPC) in the full JHU OMOP-CDM instance with approximately 3.1 million records between July 1, 2016 and May 31, 2024. Subsequently, through reviews of distance=1 parent and children concepts, we qualified the perceived ease of deriving accurate concept sets in ATLAS for our Delphi 1 priority constructs as follows: *easy*, if an exact standard and valid concept match was identified in OMOP; *moderate*, if related standard and valid concept matches were identified but they would require de-duplication or an OMOP extension; and *difficult*, if related standard and valid concept matches were not identified or they would require derivation via cohort definition terms.

**Results:** Delphi 1 panelists voted on 267 biomedical and 194 psychosocial constructs. They prioritized 36 biomedical and 23 psychosocial constructs for inclusion in a minimum NIH-endorsed dataset to be used by all maternal health researchers. All but one biomedical and 2 psychosocial constructs did not have a standard and valid concept in our OMOP instance. Overall, among priority constructs with standard and valid concepts in OMOP, 31 (55%) mapped to <10 and only 4 (7%) to  $\geq 100$  such concepts. Nine in 10 priority constructs mapped to observation domain concepts in OMOP, with more biomedical than psychosocial priority constructs mapping to concepts in other domains. SNOMED and LOINC were the most frequent source vocabularies for our priority constructs. OMOP mapping for 21 (58%) biomedical and 6 (26%) priority constructs had DPCs for one or more standard and valid concepts identified. In our assessment, currently, derivation of an accurate concept set will be relatively easy for 24 biomedical and 17 psychosocial priority constructs (70% total), moderate for 11 biomedical and 4 psychosocial priority constructs (25% total), and difficult for at least one biomedical and 2 psychosocial priority constructs (5% total).

**Conclusions:** This concept mapping study provides support for the use of OMOP-CDM data to conduct research in maternal health. Except for the specific timing of key events in the postpartum period, health insurance status prior to pregnancy, and everyday experience of discrimination, all other 33 priority constructs identified by our Delphi 1 exercise are available in OMOP-CDM. Prior work conducted by the OHDSI community in other medical fields demonstrated that the use of a CDM and standardized medical terminology mitigates limitations with various data sources, and that the more precise the mapping, the more reliable the evidence generated is.<sup>5,6</sup> Considering our study findings, the OHDSI community, which now includes an increasing number of maternal health scientists, would benefit from support with deduplication of standard and valid concepts in OMOP vocabularies and from phenotyping work for key maternal health constructs, including and going beyond those included in this analysis.

**Disclaimer:** The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Funding/Support:** Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number U24HD113136.

## References:

1. World Health Organization. Trends in maternal mortality 2000 to 2020: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO. Available at: <https://www.who.int/publications/i/item/9789240068759>
2. Souza JP, Day LT, Rezende-Gomes AC, Zhang J, Mori R, Baguiya A, Jayaratne K, Osoti A, Vogel JP, Campbell O, Mugerwa KY, Lumbiganon P, Tunçalp Ö, Cresswell J, Say L, Moran AC, Oladapo OT. A global analysis of the determinants of maternal health and transitions in maternal mortality. *Lancet Glob Health*. 2024;12(2):e306-e316.
3. Maternal Health Data Innovation and Coordination Hub. Available at: <https://maternalhealthhub.jhu.edu>
4. OHDSI Maternal Health Data Science Fellowship. Information available at: <https://www.ohdsi.org/ohdsi-news-updates/>
5. Matcho A, Ryan P, Fife D, Gifkins D, Knoll C, Friedman A. Inferring pregnancy episodes and outcomes within a network of observational databases. *PLoS One*. 2018;13(2):e0192033.
6. Weaver J, Hardin JH, Blacketer C, Krumme AA, Jacobson MH, Ryan PB. Development and evaluation of an algorithm to link mothers and infants in two US commercial healthcare claims databases for pharmacoepidemiology research. *BMC Med Res Methodol*. 2023;23(1):246.

**Table 1. Availability and characteristics of OMOP-CDM concepts for biomedical constructs prioritized during Delphi exercise**

Biomedical Domains	Biomedical Constructs	Delphi Vote	Available in OMOP-CDM								Ease of deriving accurate concept set <sup>4</sup>	
			Standard & valid concepts <sup>1</sup>	Domains <sup>2</sup>	Vocabularies					DPC <sup>3</sup>		
					SNOMED	LOINC	ICD10CM	CPT4	RxNorm			OMOP Extension
Pregnancy Episode	Pregnancy status	97%	+	O, P	x	x					0	
	Gestational age at time of event	97%	++	C, M, O, P	x	x	x	x			✓	
	Plurality	97%	+	M, O		x					0	
	Pregnancy outcome	94%	++	C, M, O	x	x	x				✓	
	Estimated due date	94%	+	C, O	x	x		x			✓	
	Days postpartum at time of event	81%										*Can be derived w/ cohort definition
Delivery Episode	Mode of delivery	94%	++	C, M, O, P	x	x	x	x			✓	
	Date of delivery/end of pregnancy	93%	+	O	x						✓	
Maternal Health Conditions and Outcomes	Maternal death	97%	+	C, O	x	x	x				0	
	Causes of maternal death	94%	+	M, O		x	x				0	
	Gestational diabetes	93%	++	C, D, M, O, P	x	x	x	x	x		✓	
	Severe maternal morbidity	90%	+++	C, P	x	x	x	x			0	
	Gestational hypertension	90%	++	C, D, M, O, P	x	x	x	x	x	x	✓	
	Preeclampsia	90%	++	C, D, M, O, P	x	x	x	x	x	x	✓	
	Obstetric hemorrhage	90%	+++	C, D, M, O, P	x	x	x	x	x	x	✓	
	Eclampsia	87%	++	C, M, O, P	x	x	x	x		x	✓	
	Date of maternal death	84%	+	O	x						0	
	HELLP syndrome	81%	+	C, M, O	x	x					✓	
	Sepsis	81%	++	C, D, M, O, P	x	x	x	x	x		✓	
Placental complications	77%	+++	C, D, M, O, P	x	x	x	x	x		✓		
Neonatal Characteristics and Outcomes	Neonatal death	97%	++	O	x						0	
	Causes of neonatal death	90%	+	C, O	x	x	x				0	
	Date of birth	84%	+	O	x						0	
	Neonatal birthweight	84%	++	C, M, O, P	x	x	x	x			✓	
	Neonatal sex assigned at birth	84%	+	O	x						0	
	Timing of neonatal death	77%	++	O	x						0	
Maternal Health History	Pregnancy history (GPA status)	97%	++	O	x						✓	
	Chronic (pre-gestational) diabetes	94%	++	C, D, M, O, P	x	x	x	x	x		✓	
	Chronic (pre-gestational) hypertension	94%	++	C, D, M, O, P	x	x	x	x	x	x	✓	
	Prior cesarean	87%	+	C, M, O, P	x	x	x	x			✓	
	Comorbidities	81%	+++	C, P	x	x	x	x			0	
Maternal Health Status Assessments	Pre-pregnancy weight	90%	+	M, O		x					0	
	Weight (current)	87%	+	M, O		x					✓	
	Height	81%	+	M, O		x					✓	

	Gestational weight gain	77%	+	C, M	x	x	x				0	
Care Encounters	ICU admission	77%	+	M, O, P	x	x		x			✓	

**Notes for both Tables:** <sup>1</sup>Number of standard and valid OMOP concepts denoted: + if <10; ++ if 10-99; +++ if ≥100; <sup>2</sup>Domains assessed include: condition (C); drug (D); measurement (M); observation (O); and procedure (P); <sup>3</sup>DPC, descendant person counts; <sup>4</sup>Perceived ease of deriving accurate concept sets in ATLAS categorized as follows: *easy (shown in green)*, if an exact standard and valid concept match was identified in OMOP; *moderate (shown in yellow)*, if related standard and valid concept matches were identified but they would require de-duplication or an OMOP extension; and *difficult (shown in red)*, if related standard and valid concept matches were not identified or they would require derivation via cohort definition terms.

**Table 2. Availability and characteristics of OMOP-CDM concepts for psychosocial constructs prioritized during Delphi exercise**

Psychosocial Domains	Psychosocial Constructs	Delphi Vote	Available in OMOP-CDM									Ease of deriving accurate concept set <sup>4</sup>
			Standard & valid concepts <sup>1</sup>	Domains <sup>2</sup>	Vocabularies						DPC <sup>3</sup>	
					SNOMED	LOINC	ICD10CM	CPT4	RxNorm	OMOP Extension		
Mental Health	Depressive disorders	84%	+	C, M, O, P	x						✓	
Substance Use	Smoking/tobacco use	92%	+	C, D, M, O, P	x	x	x	x	x		✓	
	Alcohol use	88%	+	C, D, M, O, P	x	x	x	x	x		✓	
	Substance/drug use	84%	++	C, D, M, O, P	x	x	x	x	x		✓	
Infant Feeding	Human milk or breastfeeding	84%	++	C, M, O, P	x	x					0	
Violence	Intimate partner violence	80%	++	C, M, O, P	x	x					0	
Access to Medical Care	Access to health care	92%	++	O	x	x					0	
	Health insurance status prior to pregnancy	88%										*Potentially derived w/ cohort definition
	Health insurance status (current)	80%	++	O	x	x					0	
	Health insurance type	76%	++	O	x	x					0	
Economic Stability	Food security	88%	+	M, O, P	x	x					0	
	Transportation	80%	++	O	x	x					✓	
Patient Experience	Everyday experiences of discrimination	88%										
Demographics	Age	96%	+	O	x						✓	
	Educational attainment	96%	+	O		x					0	
	Ethnicity and race	92%	+	O		x					0	
	Sex assigned at birth	88%	+	O		x					0	
	Partnership/marital status	84%	+	O	x	x					0	
	Primary language	84%	+	O	x	x					0	
	Current place of residence	84%	+	O	x	x					0	
	Gender identity	80%	+	O		x					0	
	Birthplace (country)	77%	+	O	x	x					0	
Disability status	76%	+	O		x					0		