아주대학교 AJOU UNIVERSITY

OHDSI

How OMOP CDM Enables Nursing Research

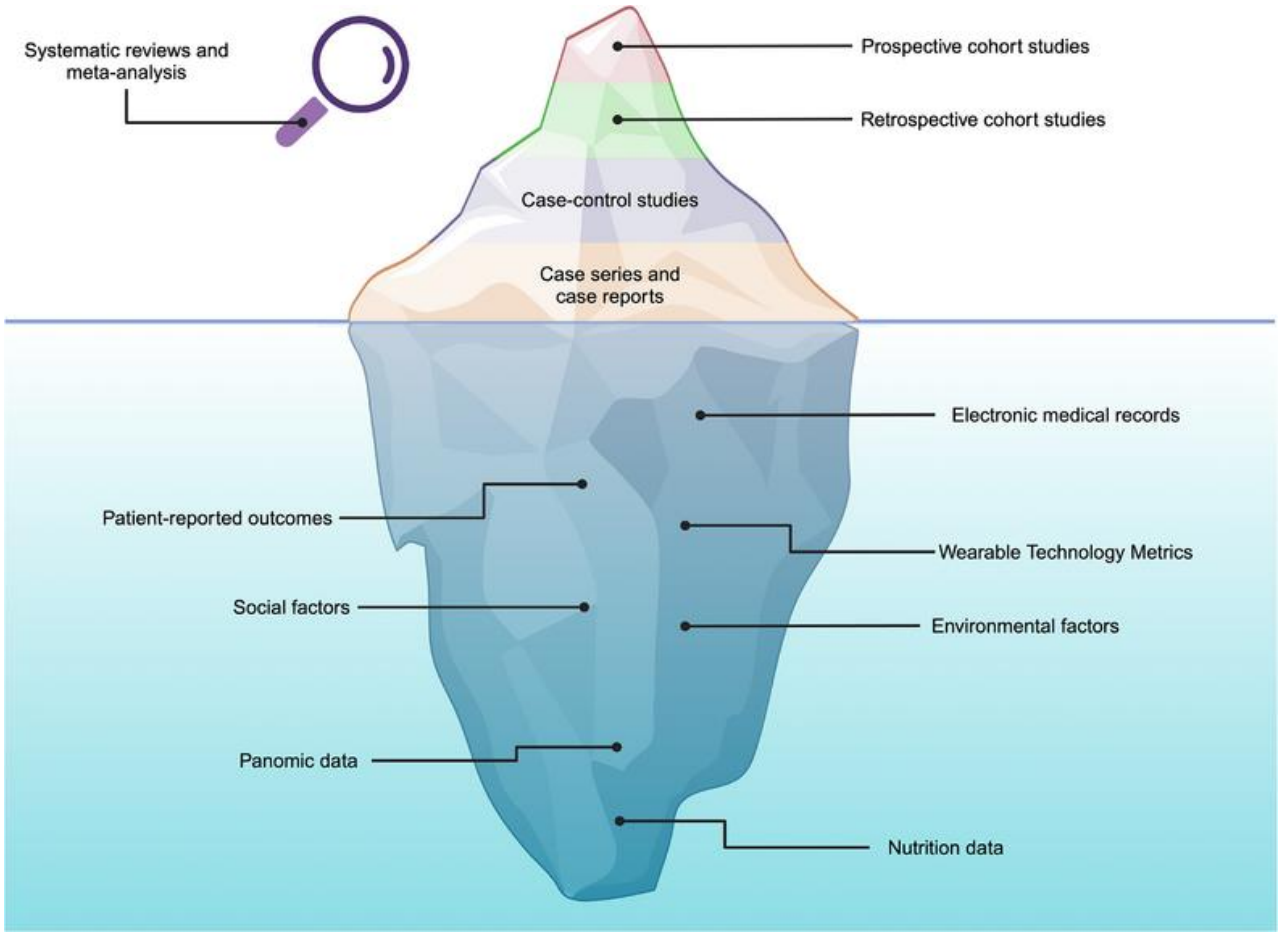
Standardizing and Operationalizing Nursing Narratives for Research

2026.03.05

Department of Biomedical Informatics,
Ajou University
Sujin Gan

Background

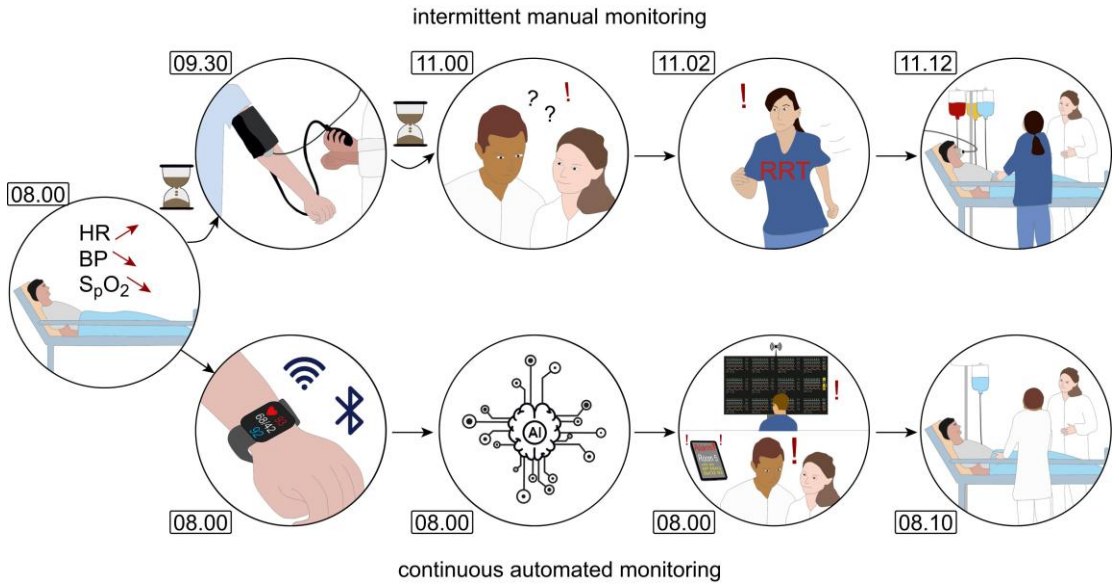
Real-World Data in Healthcare Research



- Electronic health records contain large amounts of clinical information generated during routine care.
- Not all clinically meaningful information is captured in structured data fields.

Why Nursing Data Matters

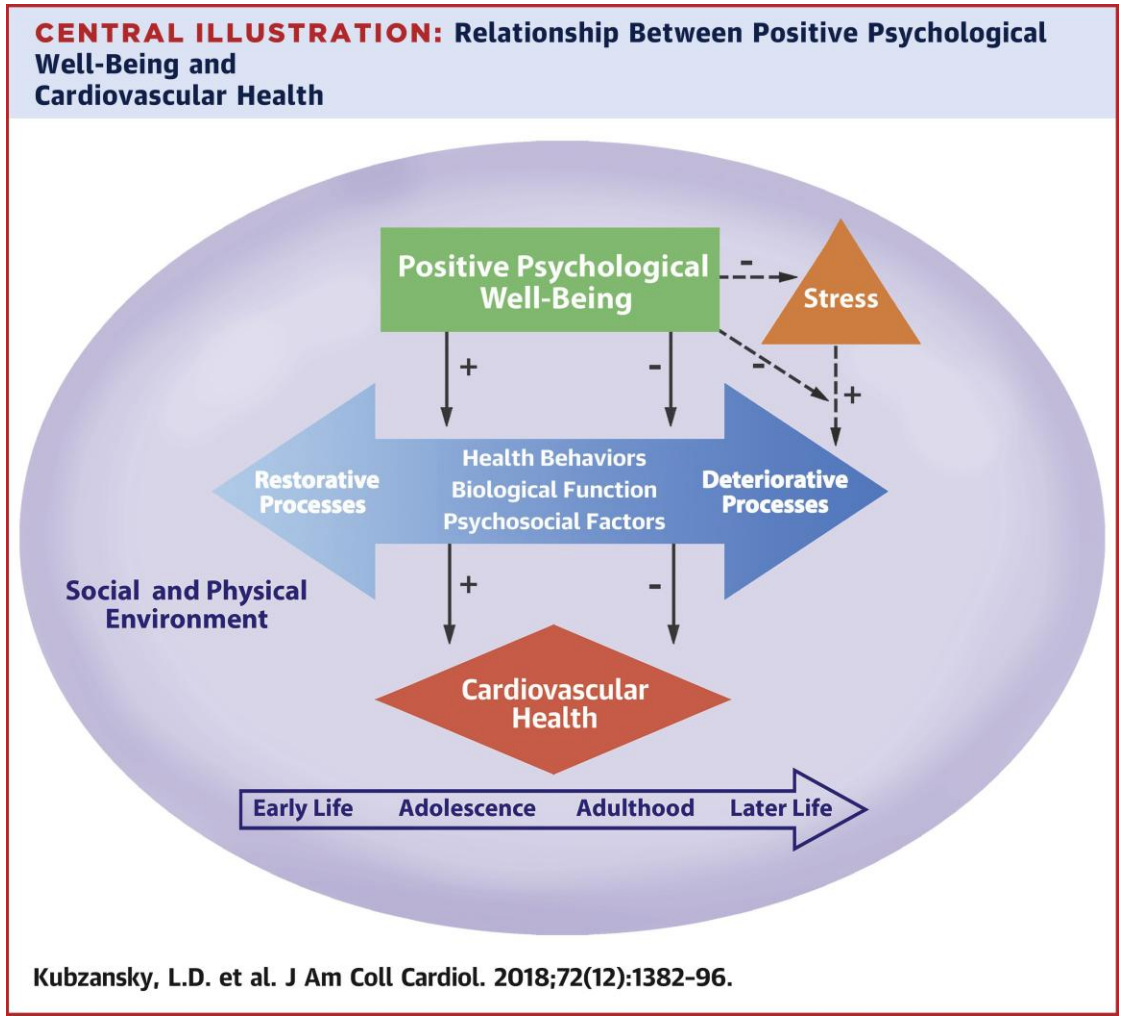
Nursing Documentation captures patient-centered information



Characteristic	Description	Example
Patient Info	Name, age, visit date	N/A
Chief Complaint	Main reason for visit	Cough for 2 weeks
History	Relevant medical background	No fever, non-smoker
Symptoms	Patient's experiences	N/A
Examination	Doctor's observations	Clear lungs, no wheezing
Diagnosis	Provider's belief	Suspected mild viral bronchitis
Treatment Plan	Medication, tests, follow-ups	Rest, fluids, no antibiotics
Provider Notes	Additional observations	N/A

Reflection of subtle changes, continuous

Why Nursing Data Matters



- Psychosocial and behavioral factors influence health outcomes
- These signals are often captured in nursing documentation

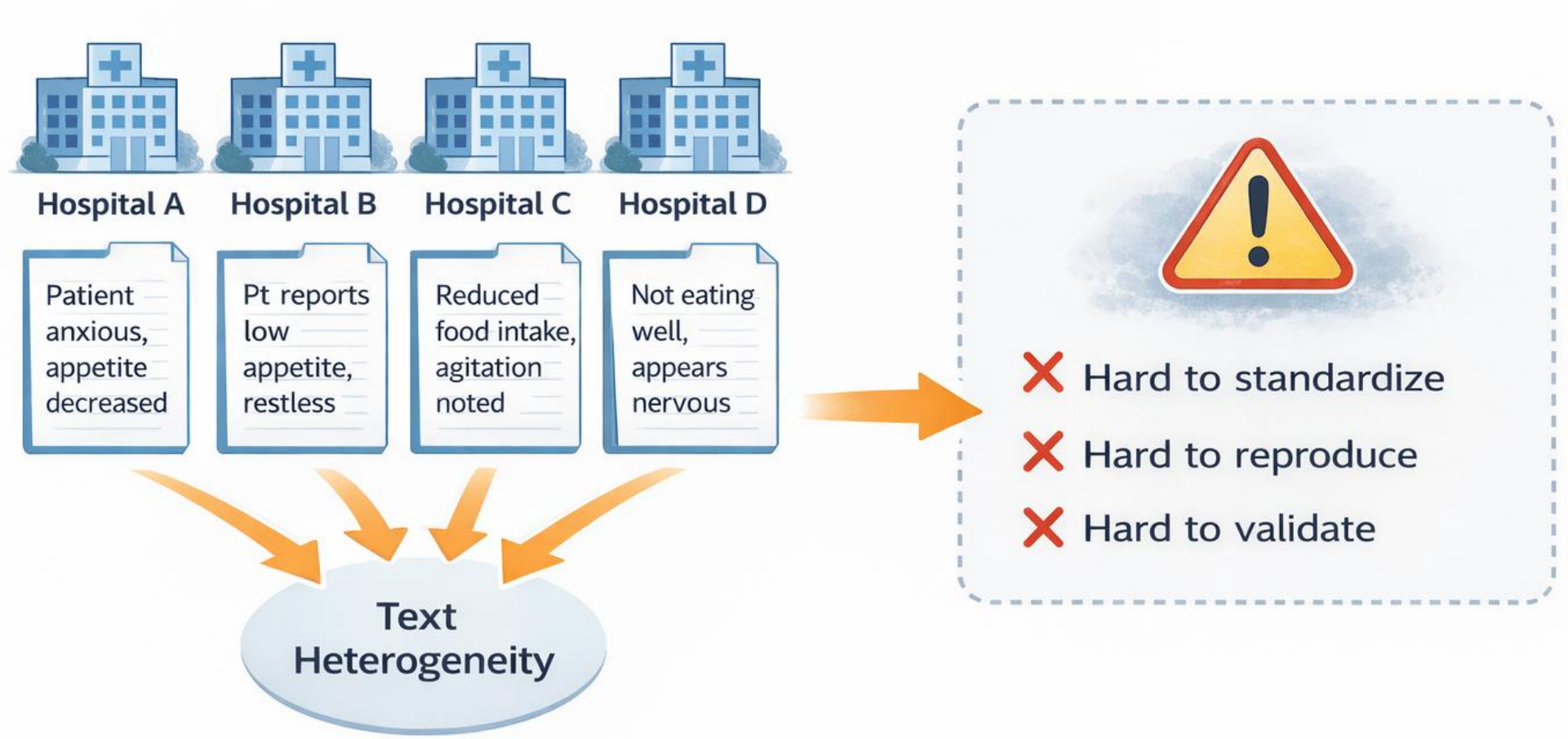
Growing Interest — But Limited Reproducibility

- Natural language processing has increasingly been used to extract early clinical signals from nursing documentation that are not captured in structured EHR data.
- However, models trained on one institution’s nursing notes often do not generalize well to other hospitals.

Challenges in Using Nursing Narratives

아주대학교 AJOU UNIVERSITY

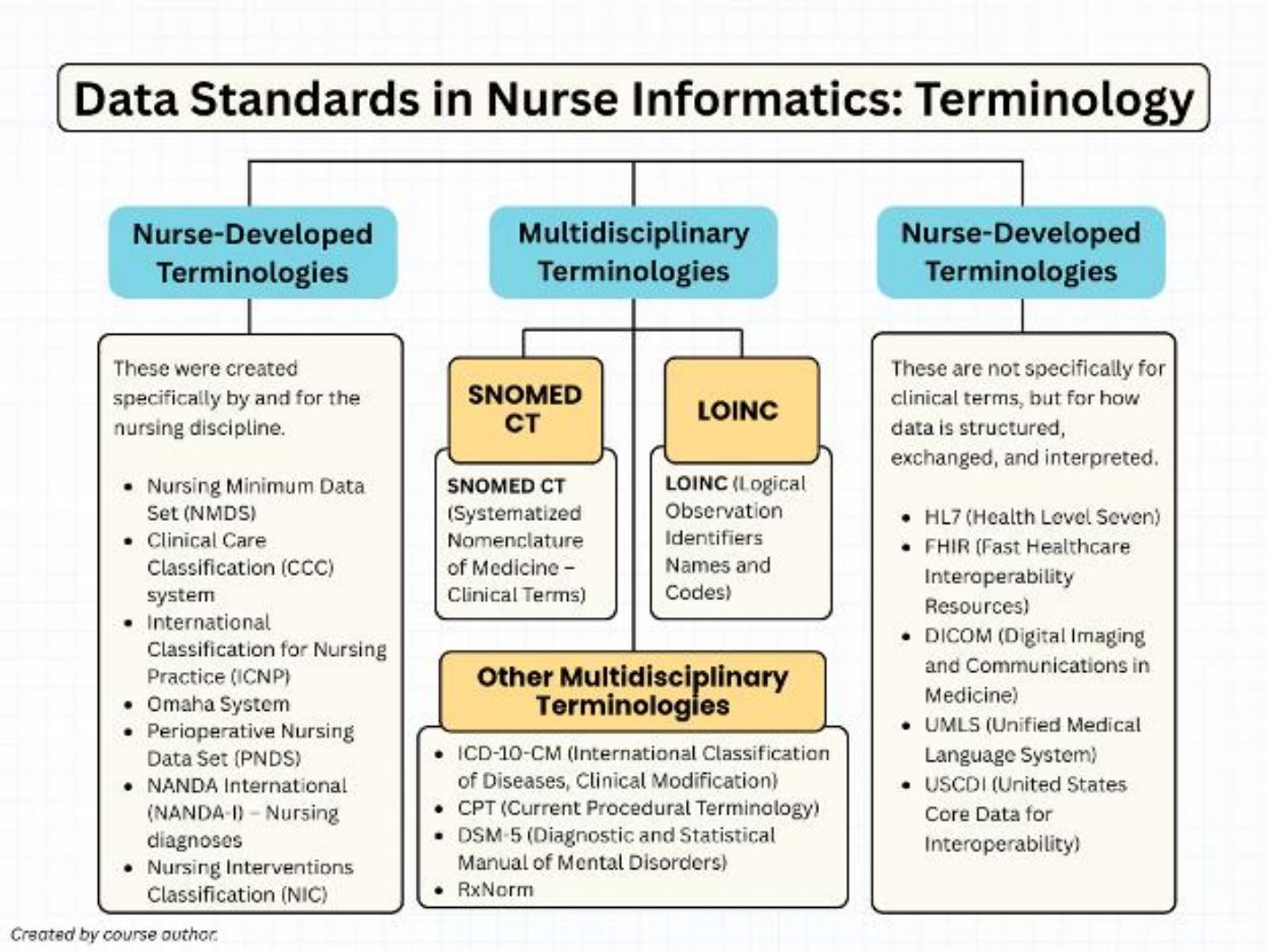
Stored as unstructured free text



Opportunity

Nursing already has standardized concepts

Nurse-Developed Terminology (NDT) Databases	
Clinical Care Classification (CCC)	A standardized, coded terminology designed for documenting nursing care across the care continuum.
NANDA International (NANDA-I)	Standardized nursing diagnoses used worldwide to describe patient problems and responses
International Classification for Nursing Practice (ICNP)	Comprehensive terminology developed by the International Council of Nurses (ICN).
Nursing Interventions Classification (NIC)	Standardized language for *interventions* (nursing treatments and actions).
Nursing Outcomes Classification (NOC)	Standardized language for evaluating patient *outcomes* of nursing care.
Perioperative Nursing Data Set (PNDS)	A standardized nursing language specific to perioperative (surgical) care.
(11)	



Created by course author.

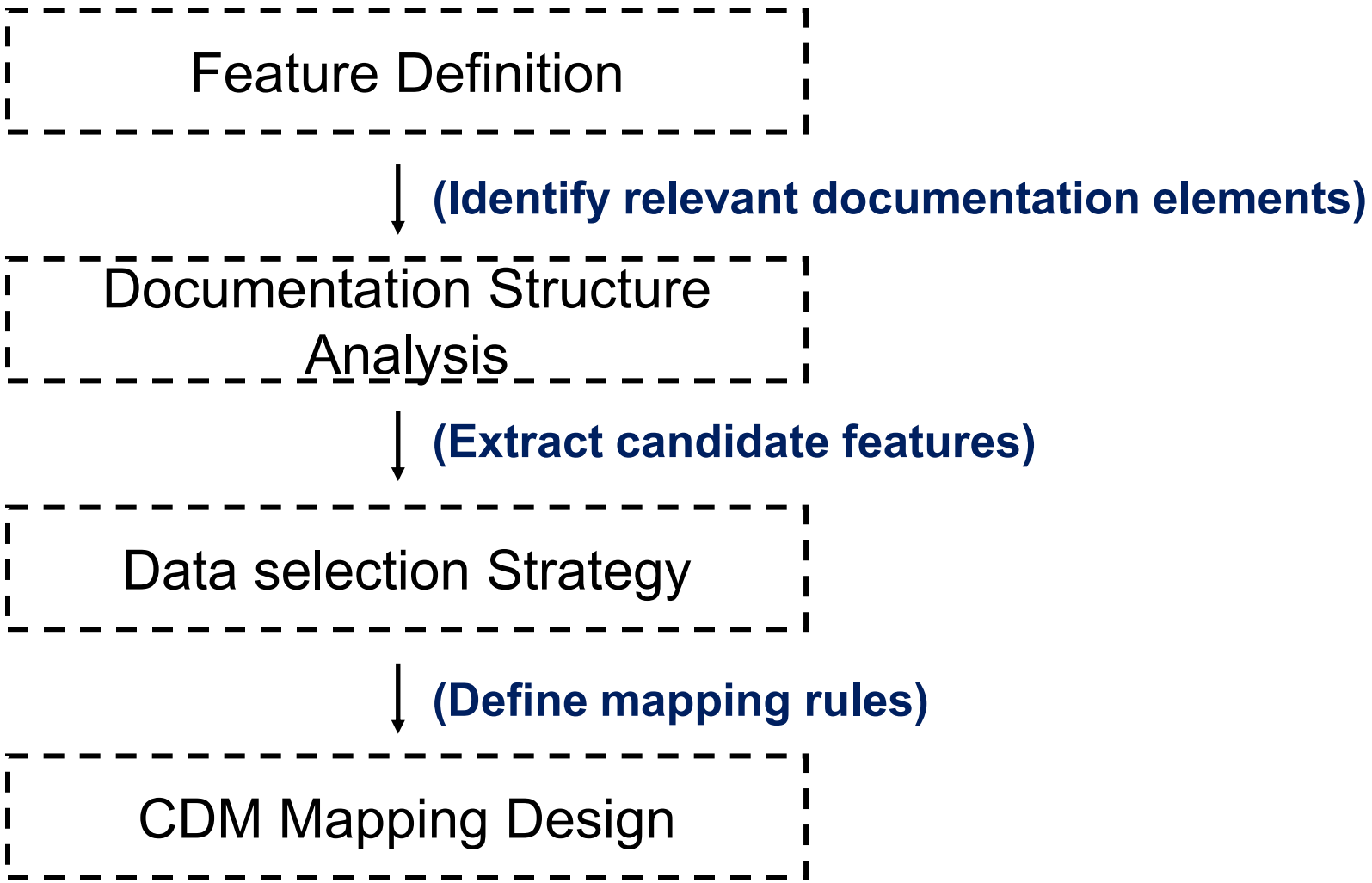
Research objective

Standardizing nursing narratives for research

- Extract clinically meaningful signals from nursing narratives
- Convert free-text nursing documentation into structured features
- Represent these features within the OMOP CDM framework

Proposed Framework

Framework overview



Feature Dictionary Development

아주대학교 AJOU UNIVERSITY

Defining standardized nursing features

- A feature dictionary was developed to systematically capture clinically meaningful signals in nursing documentation
- Features were derived from **nursing standard terminologies and clinical documentation practices**
- This dictionary defines how nursing narratives can be transformed into structured research features

Nursing Feature Dictionary (Derived for Clinical Signal Extraction)

아주대학교 AJOU UNIVERSITY

Feature Category	Feature Examples
Patient-reported symptoms	abdominal pain, chest pain, headache, musculoskeletal pain, dizziness, nausea, dyspnea complaint, pain score
Observed clinical signs	abnormal mental status, abnormal heart rhythm, abnormal respiratory pattern, abnormal body temperature, circulation deficit, fluid imbalance, renal dysfunction indicators
Infection and clinical condition indicators	infection-related diagnosis, infection-related medication, fever observation, systemic inflammatory concern
Nursing monitoring behavior	heart rate measurement, respiratory rate measurement, blood pressure measurement, temperature measurement, SpO_2 monitoring
Monitoring pattern signals	clustered vital measurements, frequent reassessment, limited vital measurement
Nursing documentation signals	nursing note written, narrative clinical description, repeated documentation, vital sign comments
Clinical escalation indicators	physician notification, urgent reassessment, abnormal finding reported
Medication-related nursing actions	PRN medication administered, scheduled medication withheld, symptom-relief medication administration

Nursing Documentation Structure Analysis

아주대학교 AJOU UNIVERSITY

Understanding the structure of nursing documentation

- Nursing documentation structures were examined within the hospital CDW.
- Section headers and templates were analyzed.
- Narrative patterns were evaluated to assess extraction feasibility.

Data scope

- Institution: Ajou University Medical Center (CDW)
- Data collection period: **2020.01.01 ~ 2024.12.31**
- Data types used:
  - Nursing records, including flow sheets including flowsheets, fall risk assessments, and pain evaluations
- Total number of records: **Approximately 600 million**

Data Selection Rules for CDW → CDM Migration

아주대학교 AJOU UNIVERSITY

Only clinically relevant and valid nursing records are migrated to CDM

1. Concept relevance: Records must correspond to predefined feature domains
2. Value validation: Entries must contain meaningful clinical values
3. Metadata completeness: Records must include essential metadata (e.g., timestamp)

Nursing Documentation Structure Analysis

- Total Record Count: 674,841,231
- Attributes: nursrec_key, **stmnt_gb**, stmnt_key, stmnt_val, property, act_dt, ptnt_no, med_dept
  - stmnt_gb: **ACT**, REL, **QUT**, **FRE, TET,** GOL, **STE**

Code	Full Term	Concern features	Record Counts
ACT	**Action**	**→ Medication Administration, GCS, Pain**	**183,408,479**
REL	Related Factor		714,648
QUT	**Quote**	**→ Vital sign frequency**	**57,013,792**
FRE	**Free Text**	**→ Medication Administration**	**112,661,691**
TET	**Test / Evaluation**	**→ GCS**	**73,330,940**
GOL	Goal		5,471,184
STE	**Status Evaluation**	**→ Nursing note content**	**166,881,974**

Data Selection Rules for CDW → CDM Migration

아주대학교 AJOU UNIVERSITY

Assessing CDW data compatibility with the predefined features

Metric	Definition
Match	Data element in the EHR is the same as the data element in Reference model
Partial Match	Data element in the EHR is different but intended to capture the same type of data as the data element in Reference model
Conflicting	Data element in EHR is not equal to the data element from Reference model
Extra	Data element is in the EHR but does not exist in Reference model
Missing	Data element is not in EHR but does exist in Reference model

Only clinically meaningful and structurally compatible elements were migrated to OMOP CDM.

Example nursing narrative structure

아주대학교 AJOU UNIVERSITY

These mappings are preliminary and may require further review.

Features	Nursing statemnets
General concern	Patient reports discomfort due to gas
Pain level	Patient reports intermittent pain
Pain level	Patient reports intermittent generalized pain
Fluid volume alteration	Increased thirst reported
Communication problem	Sensory aphasia present
Mood disorder	Emotional instability observed
Fall risk	Unable to go down stairs
Fall risk	Unable to climb stairs
Violence gesture	Aggressive behavior observed
Violence gesture	Patient shows aggressive behavior
Mood disorder	Patient expresses feelings of emptiness
Abnormal mental state	Grandiose delusion present
Abnormal respiration	Hyperventilation observed
Abnormal respiration	Hyperventilation confirmed
Abnormal mental state	Delusion of reference present

Preliminary Results

Nursing narrative analysis

Metric	Result
Nursing narrative entries analyzed	1,716
Entries containing predefined feature signals	786
Feature detection rate	~46%

Mapping to OMOP CDM

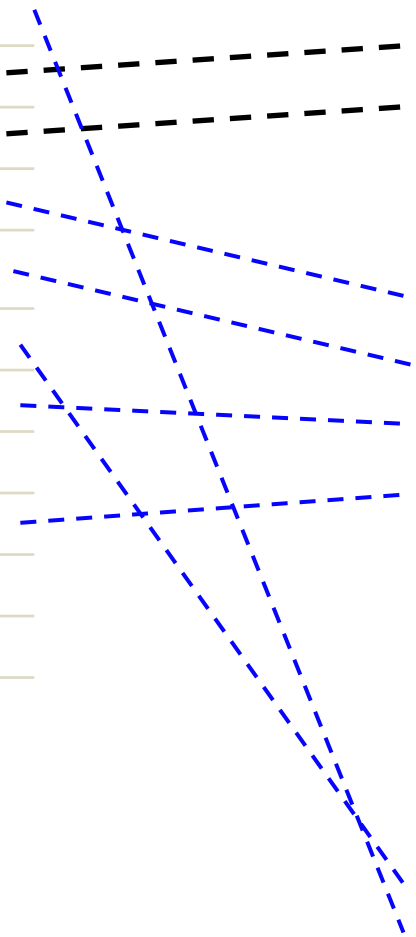
Mapping nursing features into OMOP CDM

- Nursing features were mapped to appropriate CDM domains.
- The Mapping process involved collaboration with the nursing department in Ajou University Hospital
- Terminology alignment was performed with domain experts
- Mapping rules were refined iteratively during development.

Example: Mapping Nursing Features to OMOP CDM

아주대학교 AJOU UNIVERSITY

Columns
_feature_note_id
_feature_note_date
_feature_note_datetime
_feature_note_title
_model_factor
concer_feature_note_source_value
concer_feature_source
_feature
_feature_note_text
med_dept
provider_id



CDM Field	User Guide
note_id	1
person_id	12345678
note_date	2020-01-01
note_datetime	2020-01-01 18:00
note_type_concept_id	44814645
note_class_concept_id	0
note_title	Vital Sign
concer_feature_note_source_value	All 5 measurements taken at the same time
concer_feature_source	활력징후 측정함
note_text	BT(℃):36.8, PR(회/분):70, RR(회/분):20, BP(mmHg):123/79
encoding_concept_id	0
language_concept_id	4175771
provider_id	32581
visit_occurrence_id	3
visit_detail_id	3
note_source_value	QUT
_feature_note_id	1

Example: Transforming CDW Nursing Data into CDM Format

아주대학교 AJOU UNIVERSITY

CDW

ptnt_no	act_dt	stmnt_gb	stmnt_key	stmnt_val	property
		STE	13915	발열 있음	BT(℃):37.9
		QUT	21529	활력징후 측정함	BT(℃):36.8, PR(회/분):70, RR(회/분):20, BP(mmHg):123/79
		QUT	21529	활력징후 측정함	BT: 36.8℃
		TET	-1	열 없이 자고있음	36.5℃
		TET	NULL	미열있음	37.5

CDM

note_id	person_id	...	note_title	_feature	_feature_source	note_text	...	visit_occurrence_id	visit_detail_id	note_source_value	concer_feature_note_id
		...	Vital Sign	Temperature measurement	발열 있음	BT(℃):37.9	...	NULL	NULL	STE	
		...	Vital Sign	All 5 vital measurements taken at the same time	활력징후 측정함	BT(℃):36.8, PR(회/분):70, RR(회/분):20, BP(mmHg):123/79	...	3	3	QUT	
		...	Vital Sign	Temperature measurement	활력징후 측정함	BT: 36.8℃	...	7	7	QUT	
		...	Vital Sign	Temperature measurement	발열 있음	BT(℃):37.9	...	NULL	NULL	STE	
			Vital Sign	Temperature measurement	열 없이 자고있음	36.5℃	...	NULL	NULL	TET	
			Vital Sign	Temperature measurement	미열있음	37.5	...	NULL	NULL	TET	

Example: Extracting Clinical Features from Nursing Narratives

아주대학교 AJOU UNIVERSITY

CDW

ptnt_no	act_dt	stmnt_gb	stmnt_key	stmnt_val	property
		STE	259	시간당 소변량 감소함	10-20ml/hr
		TET	-1	통증 감소함	NPIS: 3점
		OUT	20531	의식상태 확인함	의식수준: Alert
		FRE		Dr. 김재중 splint 재고정함. splint 재고정 후 통증은 가라 앉았다함.	

Q. Can it be considered a feature ‘pain level’?

CDM

note_id	...	note_title	_feature	_feature_source	note_text	...	visit_occurrence_id	visit_detail_id	note_source_value	concer_feature_note_id
	...	Nursing Assessment		시간당 소변량 감소함	10-20ml/hr	...	4	4	STE	
	...	Nursing Assessment	Pain level	통증 감소함	NPIS: 3점	...	64	64	TET	
	...	Nursing Assessment	Pain level	통증 감소함	NPIS: 3점	...	64	64	TET	

Conclusion

Conclusion

아주대학교 AJOU UNIVERSITY

Standardizing nursing narratives for research

- Nursing documentation can be structured using the OMOP CDM framework
- Future goal: interoperable CDM mapping across institutions

Integrating nursing documentation into CDM could enable richer patient insights

- Patient-reported symptoms, Social Determinants of Health
- Subtle clinical changes

감사합니다.

sujin0114@ajou.ac.kr