



Transforming the Psychiatric Hospital Database to the OMOP Common Data Model in China

J. Li¹, Y.S. Zhang², G.M. Xu¹, L.L. Zhang², X.Q. Zhu³, W.K. Zou¹, L. Feng³, X.W. Zhang⁴, M.V. Zandt^{5,6}, X.L. Wang⁵, J. Li⁵, G. Wang³

¹ Tianjin Anding hospital, Tianjin, China

² Hebei the 6th hospital, Baoding, China

³ Beijing Anding hospital affiliated to Capital Medical University, Beijing, China

⁴ SmindU Medical Science & Technology Co., Ltd., Beijing, China

⁵ Real World Solutions, IQVIA, Durham, North Carolina, USA

⁶ Observational Health Data Sciences and Informatics, New York, New York, USA

Background

The medical data from electronic health record (EHR) has been transformed to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). However, the psychiatric data has its own characteristics, which constitute barriers to convert the data to OMOP CDM, e.g. the large amount of free text from psychiatric records is difficult to extract. Here, we describe the process of converting clinical data from 3 geographically adjacent psychiatric hospitals in China to the OMOP CDM.

Methods

The standard psychiatric diseases and psychotic drugs were mapped in OMOP version 5.3 Vocabulary. All data required by the CDM were extracted, transformed, loaded (ETL) into the CDM. Large amount of unstructured and semi-structured data from psychiatric records was extracted from text using natural language programming (NLP) which is an approach to analyzing and collecting the meaningful data from the free text. To evaluate and improve the data quality of the OMOP database, a rule-based assessment “Data Quality Dashboard (DQ)” implemented in OMOP database was applied.

Results

The successful transformed OMOP CDM was listed in 8 core domains (person, visit, condition, observation, procedure, measurement, drug and cost), which cover all of the elements from health care process. 829,000 patients (female: 56.4%; male: 43.6%) with total 6,961,000 visit records (inpatient: 1.7% ; outpatient: 98.3%) were included in the database. There are 576 various diagnosis of mental disorder, 429 drugs, 189 procedures, 141 measurements, 124 observations and 4 kind of costs were recorded. For the data quality, the total value of DQD is 99%.

Conclusions

The conversion of the psychiatric data into the OMOP CDM was successful. It provides the possibility to increase the quality of psychiatric research. The next steps are extension of the psychiatric items range in the database and increasing the accuracy of mappings between the EHR and OMOP CDM.



Transforming the Psychiatric Hospital Database to the OMOP Common Data Model in China

J. Li¹, Y.S. Zhang², G.M. Xu¹, L.L. Zhang², X.Q. Zhu³, W.K. Zou¹, L. Feng³, X.W. Zhang⁴, M.V. Zandt^{5,6}, X.L. Wang⁵, J. Li⁵, G. Wang³

¹ Tianjin Anding hospital, Tianjin, China

² Hebei the 6th hospital, Baoding, China

³ Beijing Anding hospital affiliated to Capital Medical University, Beijing, China

⁴ SmindU Medical Science & Technology Co., Ltd., Beijing, China

⁵ Real World Solutions, IQVIA, Durham, North Carolina, USA

⁶ Observational Health Data Sciences and Informatics, New York, New York, USA

Abstract

Objectives: The medical data from electronic health record (EHR) has been transformed to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)(1). Here, we describe the process of converting clinical data from 3 geographically adjacent psychiatric hospitals in China to the OMOP CDM.

Methods: The standard psychiatric diseases and psychotic drugs were mapped in OMOP version 5.3 Vocabulary. All data required by the CDM were extracted, transformed, loaded (ETL) into the CDM. Large amount of unstructured and semi-structured data from psychiatric records was extracted from text using natural language programming (NLP) which is an approach to analyzing and collecting the meaningful data from the free text. To evaluate and improve the data quality of the OMOP database, a rule-based assessment “Data Quality Dashboard (DQ)” implemented in OMOP database was applied.

Results: The successful transformed OMOP CDM was listed in 8 core domains (person, visit, condition, observation, procedure, measurement, drug and cost), which cover all of the elements from health care process. 829,000 patients (female: 56.4%; male:43.6%) with total 6,961,000 visit records (inpatient: 1.7% ; outpatient : 98.3%) were included in the database. There are 576 various diagnosis of mental disorder, 429 drugs, 189 procedures, 141 measurements, 124 observations and 4 kind of costs were recorded. For the data quality, the total value of DQD is 99%.

Conclusion: The conversion of the psychiatric data into the OMOP CDM was successful. It provides the possibility to increase the quality of psychiatric research. The next steps are extension of the psychiatric items range in the database and increasing the accuracy of mappings between the EHR and OMOP CDM.

Background

The medical data from electronic health record (EHR) has been transformed to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). However, the psychiatric data has its own characteristics, which constitute barriers to convert the data to OMOP CDM, e.g. the large amount of free text from psychiatric records is difficult to extract, Here, we describe the process of converting clinical data from 3 geographically adjacent psychiatric hospitals in China to the OMOP CDM.

Methods

The standard psychiatric diseases and psychotic drugs were mapped in OMOP version 5.3 Vocabulary. All data required by the CDM were extracted, transformed, loaded (ETL) into the CDM. Large amount of unstructured and semi-structured data from psychiatric records was extracted from text using natural language programming (NLP) which is an approach to analyzing and collecting the meaningful data from the free text. To evaluate and improve the data quality of the OMOP database, a rule-based assessment “Data Quality Dashboard (DQ)” implemented in OMOP database was applied.

Results

The successful transformed OMOP CDM was listed in 8 core domains (person, visit, condition, observation, procedure, measurement, drug and cost), which cover all of the elements from health care process (Tab. 1). 829,000 patients (female: 56.4%; male:43.6%) with total 6,961,000 visit records (inpatient: 1.7% ; outpatient : 98.3%) were included in the database. There are 576 various diagnosis of mental disorder, 429 drugs, 189 procedures, 141 measurements, 124 observations and 4 kind of costs were recorded. For the data quality, the total value of DQD is 99%.

Results

Data Sources	Beijing Anding Hospital (BJH)	Tianjin Anding Hospital (TJH)	Hebei Sixth Hospital (HBH)	Unity Database
Duration	2013-2019	2013-2019	2018-2019	2013-2019
Persons	435,000	336,000	58,000	829,000
Visits	3,277,000	3,471,000	213,000	6,961,000
-Inpatient	57,000	46,000	15,000	118,000
-Outpatient	3,220,000	3,426,000	198,000	6,844,000
Condition	9,243,000	4,454,000	313,000	14,010,000
Measurement	36,868,000	8,605,000	2,068,000	47,541,000
procedure	1,553,000	1,241,000	396,000	3,190,000
Drug	8,123,000	6,034,000	438,000	14,595,000
Observation	4,200,000	440,000	900,000	5,540,000
Cost Records	10,000,000		300,000	

Conclusions

The conversion of the psychiatric data into the OMOP CDM was successful. It provides the possibility to increase the quality of psychiatric research. The next steps are extension of the psychiatric items range in the database and increasing the accuracy of mappings between the EHR and OMOP CDM.