

OHDSI内では、実名での活動になります。
Zoom参加時も「名前は実氏名で」お願いします。

OHDSI Japan evening conference #7

イブニング カンファレンス(第7回)
2020.6.30



OHDSI チュートリアル報告

- 2020/6/5の日本医療情報学会 第24回春季学術大会で開催。
 - ・ **なぜOHDSIか**
木村 映善 (愛媛大学医学部 医療情報学講座)
 - ・ **OHDSI Japan 日本での取り組み**
岡田 昌史 (IQVIAソリューションズジャパン株式会社)
 - ・ **医療機関におけるOHDSI(OMOP CDM)の利用**
青柳 吉博 (国立がん研究センター東病院 臨床研究支援部門)
 - ・ **88 hours: OHDSIのCOVID-19への取り組み**
平松 達雄 (国際医療福祉大学 未来研究支援センター)
- ✓ 資料はOHDSI Japan サイトに掲載
- 500名近い方のご参加
⇒コミュニティメンバー登録も423人に。
- 秋の連合大会にも申し込み済。審査待。
“Inside OHDSI” OHDSI/OMOPの中身に踏みこんだもの。



2020 OHDSI Symposium

- Global&USのシンポジウム（メインのもの）
- 10月18日～21日（4日間）、WEB開催。
- 参加費無料～（寄付も受け付けます）
- 参加登録受付中↓

<https://www.ohdsi.org/2020-symposium-registration/>

■ 「コラボレーターショーケース」

アブストラクト締め切り **7月16日** 1～2ページ

- ◆ トピック: 医療観察データの標準と管理、方法論的研究、オープンソースのアナリティクス開発、OHDSIの分析的ユースケースからの臨床研究、臨床的特徴抽出、人口レベルの推定、患者レベルの予測
- ◆ 発表形式: ポスター、口頭、ソフトウェアデモ(ビデオ)



Asia-Pacific Symposium 2020

- Asia-Pacificのシンポジウム（今回初めて）
- 12月5日～6日(半日2日間)、WEB開催。
- 参加費無料～のはず。

■ 内容検討中

- Welcome to OHDSI AP 2020
- Panel discussion
- Asia Pacific Study – Hypertension & Diabetes
- Region Chapter Introduction
- Individual country breakout



OHDSI Japanの活動場所

ここしばらくは
ネット上での活動。
そのうちリアルも
復活させたい。

- OHDSI Japanサイト: 情報提供
www.ohdsi-japan.org
- 本家Forumの日本コーナー
Forumに登録すると、書き込める。世界中だれでも読める。
<https://forums.ohdsi.org/c/ohdsi-in-japan>
- 日本コミュニティ メンバー登録
Doorkeeperのサイト。ODJで一番大きな枠組み。案内メールが
流れてくる。発言場所はない。イベントの参加登録ができる。
- 日本のコラボレータML
メーリングリスト。ML内で発言できる。流量は多くない。登録希
望者は平松まで（サイト/DKのお問い合わせから）
- 日本のSlack
登録希望は岡田先生まで（お問い合わせでも転送します）。
OHDSI Book翻訳のために作られたが、それ以外でもgenealチャ
ンネルにどうぞ。WS内で発言できる。

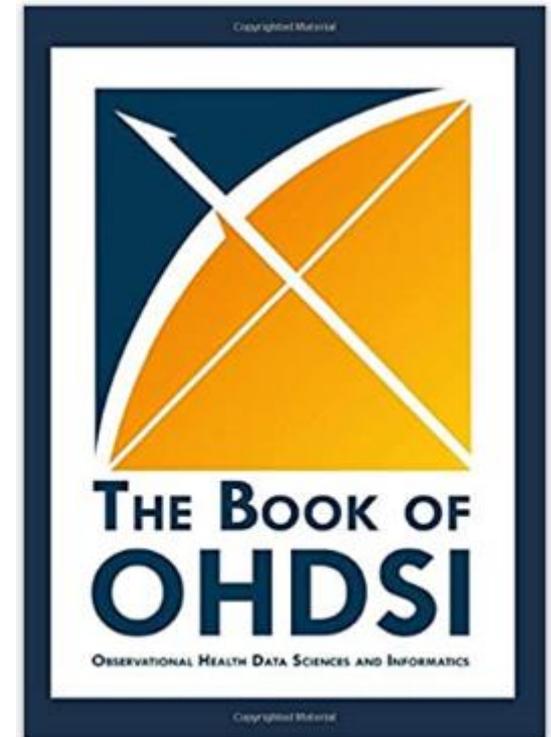


The Book of OHDSI紹介

- OHDSIのすべて（？）が書かれた本
- 無料で読める

<https://ohdsi.github.io/TheBookOfOhdsi/>

- Amazonで物理本も買える
(\$10.82+送料 / ¥2,007)
- 日本語版はOHDSI Japan
メンバーで翻訳進行中





The Book of OHDSI 紹介

第 1 部 The OHDSI Community

1. The OHDSI Community
2. Where To Begin
3. Open Science

第 2 部 Uniform Data Representation

4. The Common Data model
5. Standardized Vocabularies
6. Extract Transform Load

第 3 部 Data Analytics

7. Data Analytics Use Cases
8. OHDSI Analytics Tools
9. SQL and R
10. Defining Cohorts
11. Characterization
12. Population-Level Estimation
13. Patient-Level Prediction

第 4 部 Evidence Quality

14. Evidence Quality
15. Data Quality
16. Clinical Validity
17. Software Validity
18. Method Validity

第 5 部 OHDSI Studies

19. Study Steps
20. OHDSI Network Research

Appendix

- A. Glossary
- B. Cohort definitions
- C. Negative controls
- D. Protocol template
- E. Suggested Answers

References



1. The OHDSI Community

- The Journey from Data to Evidence
様々な医療データからエビデンスをつくる道のりの課題
- Observational Medical Outcomes Partnership
OMOPプロジェクトの紹介とOHDSIへのつながり
- OHDSI as an Open-Science Collaborative
OHDSIのミッション、ビジョン、目標
 - ◆ ミッション: より良い健康上の意思決定とより良いケアを促進するためのエビデンスをコミュニティが共同で生成できるようにすることで、人々の健康を改善する。
 - ◆ ビジョン: 観察研究により健康と病気の総合的な理解が生み出される世界。
- OHDSI's Progress
OHDSIはここまで達成しました。
- Collaborating in OHDSI
コラボレーションへのお誘い。



2. Where To Begin

産官学入り混じったコミュニティで、誰でも参加できる。

- Join the Journey

参加するにはこんな手段がある。

(「入会届を出す」ような手続きはありません。)

- ◆ OHDSI Forums
- ◆ OHDSI Events
- ◆ OHDSI Community Calls
- ◆ OHDSI Workgroups
- ◆ OHDSI Regional Chapters
- ◆ OHDSI Research Network

- Where You Fit In

こんな方は、こんなやりかたでいかが。



3. Open Science

- Open Science
OHDSIはオープンサイエンスの価値観に基づいている
- Open-Science in Action: the Study-a-Thon
「Study-a-Thon」はじめました
- Open Standards
OMOP共通データモデル(OMOP-CDM)と標準化語彙
- Open Source
オープンソースのソフトウェア
- Open Data
通常はないが、可能ではある。シミュレートデータはある。
- Open Discourse
シンポジウム、共同研究等
- OHDSI and the FAIR Guiding Principles
オープンデータの公開に関するFAIR原則への対応



4. The Common Data model

CMDテーブルの説明

- Design Principles
設計の考え方

- Data Model Conventions

CMDの規則/習慣。人中心、スキーマ、データテーブル、「ドメイン」、「コンセプト」、フィールド命名規則、コンセプトと元データ値

- CDM Standardized Tables

臨床イベントテーブル16個、語彙テーブル10個、メタデータテーブル2個、医療システムデータテーブル4個、医療経済データテーブル2個、派生要素テーブル3個、結果スキーマテーブル2個

(子宮内膜症を発症したローレンの例で説明)

PERSON Table、OBSERVATION_PERIOD Table、
VISIT_OCCURRENCE、CONDITION_OCCURRENCE、
DRUG_EXPOSURE、PROCEDURE_OCCURRENCE



5. Standardized Vocabularies(1)

- Why Vocabularies, and Why Standardizing

Vocabulary(語彙)、標準化が必要な理由、OHDSI語彙チーム Athena、既存の語彙の利用

- Concepts

コンセプト(概念)という概念。コンセプトID、コンセプト名
ドメイン(大)、語彙集(用語集)、コンセプトクラス(小)

標準コンセプト(S) 変換後データで使う(マッピングの目的地)

非標準コンセプト: 元データに対して使われる

分類コンセプト: データにはつかない。Queryで指定するのに使う。

コンセプトコード: データの値(コード)。元データで使う。

コンセプトの追加と廃止

- Relationships 2つの概念の関係

マッピング: 非標準から標準の概念へ maps to / mapped from

複数先に分割されるマッピングもある。階層関係、複数の親もある。

異なる語彙間での関係::「~相当」関係。同じ語彙間での関係: : 解剖学的,etc..



5. Standardized Vocabularies(2)

- Hierarchy (CONCEPT_ANCESTOR table)
CONCEPT_RELATIONSHIPテーブルから自動的に構築。
高品質で包括的な階層構造は、薬物と状態の2つのドメインにのみ。
- Internal Reference Tables
内部使用するTable (DOMAIN_ID、VOCABULARY_ID、
CONCEPT_CLASS_ID、CONCEPT_RELATIONSHIP_ID)
- Special Situations
 - ◆ 性別は出生時の生物学的性別。それ以外はOBSERVATIONとして。
 - ◆ OMOPの「Condition」は兆候と症状も含む。
 - ◆ デバイスは、標準がない。
 - ◆ 人間の提供者(非医療提供者も)の専門はすべてプロバイダードメインで定義
 - ◆ 腫瘍学、放射線学、ゲノミクス：開発中
 - ◆ 薬物ドメイン：RxNormとRxNorm-extension
 - ◆ NULLの問題：NULL、CID=0、“unknown”の取り扱い



6. Extract Transform Load(1)

ソースデータをCDMに再構築し、標準化された語彙にマッピングする必要がある。ソースデータが更新されるたびに再実行できるように、このETLプロセスは反復可能であることが重要。

●STEP1: ETLの設計 2つのOHDSIツール

White Rabbit : データをスキャンし、ETLの設計のための情報レポートを作成する。2次元の表を前提にしているため、レセプトやSS-MIX2には使えない。

Rabbit-In-a-Hat : 元テーブルからCDMテーブルへの対応、元フィールドからCDMフィールドへの対応を設計するときに用い、ETLプロセスのドキュメントを生成する。単なる対応ではない場合が多いので、そこは手動で記述する。

ETLするテーブルの順番。



6. Extract Transform Load(2)

●STEP2: 標準コードへのマッピング

- 非標準コードから、標準コードへ。
- OHDSI語彙にあるコーディングシステムの場合は、対応表 (CONCEPT_RELATIONSHIP table)を引く。
- OHDSI語彙にないコーディングシステムの場合は、手動で。

Usagiツール：対応付けを自動提案する。適切でないときは、ユーザーがターゲットコンセプトを検索できる。元用語が英語ではないときは、Google Translateで英語列を作ってから（それで結構OK）。

●STEP3: ETLの実装

実装は個別的なので、最善の方法について正式な推奨はない。究極のユーザーフレンドリーなETLツールの開発は結局あきらめた。データの2割に対してデータに固有の低レベルコードを記述する必要があるため。

●STEP4: ETLの品質管理

レビュー（ドキュメント、プログラム、語彙マッピング）。サンプリングしたデータを手動で比較。ETL前後で総数を比較。単体テスト用データを作ってテスト。など



6. Extract Transform Load(3)

- ETL Conventions and THEMIS

Convention⇒約束事・取り決め⇒「規約」。たとえば、生年月日に誕生年がないときどうするか？各グループが同じシナリオで対応できるように。

CDM Wikiに説明が入っている。WGもある(“THEMIS”)。Conventionが不明なときは、Forumに投げてみよう。

- CDM and ETL Maintenance

最初のCDMが構築された後にも、ETLメンテナンスの継続的なサイクルがある。

メンテナンスが必要な一般的なトリガーには、ソースデータの変更、ETLのバグ、新しいOMOP語彙のリリース、CDM自体の変更/更新など。

ボキャブラリーは月に1度以上の更新がある。