

豆知識！ HISの構造化データを利用した臨床研究と HISproの活動について

2023年10月31日
九州大学病院ARO次世代医療センター
トライアルマネージャー会議

九州大学病院メディカル・インフォメーションセンター助教
平田明恵

目次

1. HISをデータソースとする研究の概要
2. HISproの紹介

※マスタ
 ; 何の情報が
 何という名前の
 項目(変数)に紐づくか

医薬品マスタの例

- ローカルコード(ハウスコード)と標準コード
- 標準コードの意味と限界
- 利用目的とコードの粒度
- 医薬品におけるローカル名称

ローカルコード：I1000510
 先発品名称：ゾビラックス錠200mg
 一般名：アシクロビル錠200mg
 名称：アシクロビル錠200mg「ファイザー」

薬価基準	1135305020102
HOT 販売会社	1135305020102
コード 調剤包装	1135305020102
包装数量	1135305020102
YJコード	<u>6250002F1289</u>

(株式会社医薬情報研究所)

薬効分類番号：625 → 抗ウイルス剤

1. HISをデータソースとする研究の概要

✓よく使う構造化データ DWH / DPC

DWHとは

- DataWarehouse
- 一次利用のためのデータを、二次利用のために貯めておくもの
- SS-MIX2
- 二次利用のための標準規格
- 標準化ストレージと拡張ストレージからなる
- HISデータをわかりやすい形で再構築している



<http://www.ss-mix.org>

DWHで得られる情報

(ざっくり言うと)
DPC (ファイル) 以外の
HISデータ



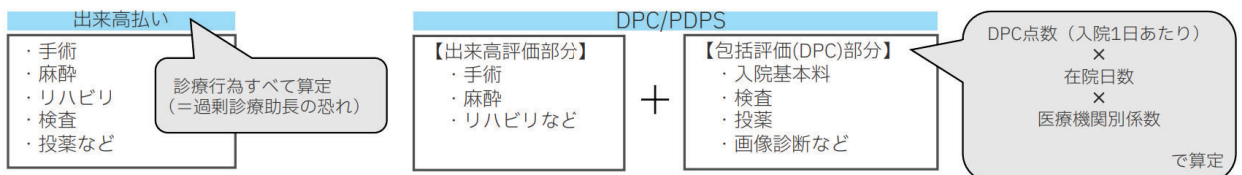
DPCとは

-DPC (DiagnosisProcedureCombination : 診断群分類)

-入院期間中に**医療資源を最も投入した**「傷病名」と「診療行為」の組み合わせによる分類。各診断群分類ごとに診療報酬点数がつけられている。

-DPC制度 (DPC/PDPS)

-従来の出来高払い方式とは異なり、「DPCに基づく包括評価部分」と「出来高評価部分」を組み合わせる診療報酬算定を行う包括医療費支払い制度。



-対象病院

-DPC対象病院・・・1757病院 (2020年4月)

-DPC準備病院・・・249病院 (2020年4月)

DPCとは

-DPCデータ

-DPC対象病院は診療報酬の請求を行うと同時にカルテ・レセプト情報を電子データとして厚生労働省に提出する義務があり、この電子データをDPCデータと呼ぶ。



-DPCデータベース

-各病院から提出されるDPCデータをデータベース化したもので、厚生労働省が保有。2019年の報告では年間約1700万件ものデータが集まっている。

△出カデータ（DPCファイル）
＝DWHの構造化データほどには整備されていない…

1. HISをデータソースとする研究の概要

✓DWH / DPC

- DWH or DPC にしかない情報
- DWH or DPC の方が信頼できそうな情報

例) 「併存症」 = DPCのみに存在

例) (入院) 主病名; DPC vs (外来) 病名; DWH

そのデータが生まれる診療場面のイメージ



より有効なデータ項目の選択

HISをデータソースとする研究の留意点

従来型 臨床研究データ 「研究の為に選定・収集」

選定済みの症例；

- irregularな背景・検査値・治療
→専門医reviewに基づく除去
- 診療科固有の「症例リスト」
→診療方針の一貫性

既存の「症例（背景）リスト」

患者ID	性別	生年月	診断	入院日	治療	既往症
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---

精度(高) / サイズ(小)

HIS構造化データ

未選定の症例；

対象症例・背景すべてHISから抽出

※ 未選定のデータから
クエスチョンに沿って
対象症例・背景情報を同定

☆リアルワールドデータ研究への発展

HIS構造化データ活用における留意点①

PECOの各要素を
どうやって「構造化データ」で「表現」するか？
～疾患・病態の「定義」～

「保険病名」の壁

- ※より信頼できそうな「病名」=そのデータが発生した臨床場面
- ※レセプト病名 vs DPC病名

疾患・病態の「定義」 同定アルゴリズム で対応

- ※他の構造化データ（検査、処方等）との組合せで「表現」
- ※特有の検討手順

目的のPECOを「表現」する為に起こるLimitation

- ※デザイン段階で把握 考慮した解析デザイン

HIS構造化データ活用における留意点②

Specificなクエスチョン・PECOの設計 ～データが取れすぎて焦点がぼやける～

データ項目（種類、期間）を容易に繰り返し取得可能

- ✓ 検定の多重性・・・？
- ✓ 結局「なに」と「なに」の関係を見たい？
- ✓ 目的（RQ）が見えなくなる → 論文の結論・考察がブレまくる

「データ整形」の壁

- ✓ とりあえず「あるだけデータ取ってきて考える」 → 無駄に複雑、挫折
- ✓ 時系列データ → 多様すぎる「現実」を忠実に反映

※綿密なデザイン & specificなPECO が不可欠

- ✓ PECOを「表現する」データ項目 = 予め計画・定義
- ✓ 得られるデータの形を想像できる臨床力、データ整形計画&スキル

そのデータが生まれる診療場面のイメージ

HIS駆動型研究習得プロジェクト（HISpro）の紹介

🕒 2022年度（初回）
CQ～デザイン策定～解析まで実施

🕒 2023年度
CQ～デザイン策定まで
生物統計事前学習導入

九州大学病院メディカル・インフォメーションセンター
九州大学病院臨床観察研究支援事業 COS3

HIS駆動型研究習得プログラム 2023年度

参加者募集

本プログラムは、臨床観察研究を主体的に実施する研究者が、電子カルテ(HIS)を活用した研究プロセスを習得することを目標としています。参加者は、ご自身の新規の研究テーマを題材に、本プログラムの手順に沿って研究デザインを完成させることで、臨床研究をより系統的に実施するための知識・技術を習得いただけます。

本プログラムの特長

- ・参加者ご本人のテーマの論文化を目指したカリキュラム
- ・疫学・統計・情報を専門とするCOS3スタッフによるメンタリング
- ・個別のテーマに適した研究デザイン・データ選択・統計解析計画
- ・参加者が主体的に取り組む演習形式
- ・医療データ研究に必要な最新のトピックを盛り込んだ疫学・統計学演習
- ・データソースの確保；九州大学病院電子カルテ情報の活用
- ・プログラム終了後、論文発表までのサポート（別途規程有り）

応募の流れ

- 1 メールにてお問い合わせください
宛先: cos3@med.kyushu-u.ac.jp
標題: HISPRO実施要項希望
本文: ・応募希望者氏名
・所属
・職種
・連絡先（電話番号、メールアドレス）
- 2 実施要項・年間スケジュールをお送りします
- 3 実施要項に記載の手順に従ってご応募ください

実施スケジュール



※詳細については、年間スケジュールをご確認ください。

臨床観察研究支援事業（COS3）

臨床観察研究支援事業 Clinical Observational Study Support System (COS3)

MIC教員・スタッフによる臨床観察研究プロセス全般の支援
(受託・共同研究)

☞ 実用化を目指した医学研究の推進体制を強化し、良質な観察研究・橋渡し研究・臨床試験を推進すること（九州大学病院第3期病院中期計画）

COS3の受託業務

① CRIN-Qシステム業務

- 「Web入力型臨床データ収集システム（CRIN-Q）」の入力画面構築
- システム及びデータの維持管理

② コンサルティング業務

- 研究デザインおよび統計解析に関する知識提供
- データベース構築に関する助言
- セミナー提供（HISpro含む）

③ 統計解析業務

- データ解析および統計解析の実行、結果図表作成

④ データベース構築業務

- 九州大学病院外部のデータ駆動型事業等への電子カルテデータの提供支援

臨床研究遂行プロセスのリスク / 専門性

学術的リスク（疫学 / 統計学）

- ⚠ クエスチョン不在
- ⚠ 先行研究確認不足
- ⚠ 関連変数の未整理
- ⚠ 実現性のない解析計画

学術的専門性・経験

- ✓ 疫学
- ✓ 統計学
- ✓ 医療情報学

研究プロセス破綻の回避

研究デザイン
データ収集
解析
論文化

研究事務局リスク（運営組織）

- ⚠ 必要となる実務の認識不足
- ⚠ 必要人員の確認不足
- ⚠ 共同研究の調整力不足
- ⚠ 継続性の想定不足

データマネジメントリスク

- ⚠ データソースの知識不足
- ⚠ データベース管理の知識不足
- ⚠ データ構造の知識不足
- ⚠ 情報関連法令・指針の認識不足

HISpro 趣旨

【参加者】

九大病院HISを活用した臨床研究デザイン手順の習得（医療RWD研究への発展）

- 疫学・統計・情報専門家との有効なコミュニケーション習得
；クエスチョンの伝え方、pit fallの収集
- デザイン思考過程の実体験
；可能な限り「後戻り・右往左往」しない、デザインに必須の手順実践
；変数選択、統計学的モデルなど概念の習得
- HIS（RWD）データの特有のデザイン、変数定義・選択の実体験
；選択登録された「レジストリ研究」との違い

【提供者】

臨床研究プロセスの専門性の啓発、臨床家（臨床研究主体）の実態・ニーズ把握

- 臨床家（臨床研究主体）との有効なコミュニケーション習得
；up-date活発な臨床疫学知識・経験の溝、聞き取り方・伝え方
- デザイン手順のup-date
- 医療RWDを活用する非医療系研究者との臨床研究手法共有
；九州大学オープンサイエンスプラットフォーム（OSP）



HISpro研究デザイン手順フロー

手順の要点

5-1. クリニカルクエスチョンの構造化

- 1. 先行研究の整理
- 2. 臨床的背景・経過の可視化
- 3. PECO, RQ, FINERの導出
- 4. &-5. DAGを用いた関連要因整理
- 6. DAGにおける主要変数のHISデータ定義

議論/修正
(打合せ)



5-2. 解析方針の策定

- 1. 研究型
- 2. 統計学的モデル, 評価項目, 検定方法, 感度分析方法
- 3. データ整形方法
- 4. &-5. データ構造定義書, 解析計画書確定

議論/修正
(打合せ)



5-3. 倫理審査申請準備

5-2. で作成したデータ構造定義書, 解析計画書をもとに倫理審査申請書類作成

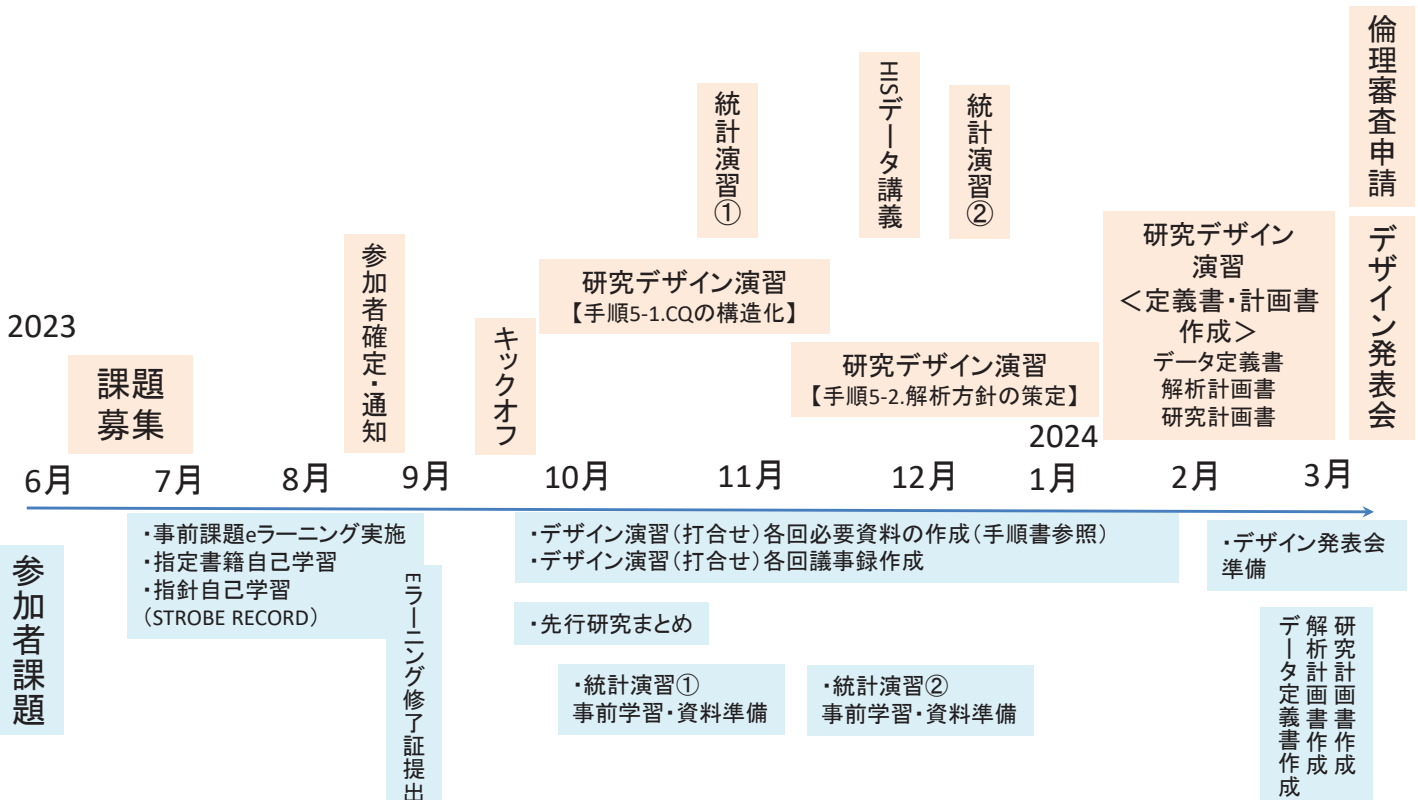
5-4. 倫理審査申請

研究デザイン打合せの要点

- FINERの検討に十分な先行研究整理が行われている
- 問題とする疾患についてPECO/RQの策定に十分な臨床背景・時系列的な要因の関係性の提示
- DAGにおけるPECO各変数および交絡因子など背景要因の網羅的な提示
- DAGにおいて特に問題とする背景およびその根拠の明確化
- 主要変数のHISデータによる実現性と定義の明確化

- 記述or分析, 横断or縦断
- 先行研究/臨床的に有益なアウトカムの差の表現方法
- 変数選択・定義, 評価項目, 統計学的モデルの観点から必要な感度分析
- 模擬データセットを用いた具体的な抽出データの形の確認
- 実現可能なデータ整形方法
- 実現可能な統計解析方法
- データ構造定義書および解析計画書による作業手順の明確化

2023年度実施スケジュール



HISpro 担当スタッフ

担当領域	人員数	所属・職種
責任者	1	MIC教授
メンター	1	MIC助教
チューター（リーダー）	1	医学研究院衛生・公衆衛生学分野学術研究員
チューター	3	MIC助教、工学研究院准教授
統計解析管理者	1	MIC共同研究員
データ管理者	3	MIC講師、MIC薬剤師、MICテクニカルスタッフ
事務連絡担当者	1	MIC事務補佐員

HISpro2022 得たもの/課題

【参加者】

- 「言い回し」に由来する、疫学統計専門家とのズレ発生源を自ら発見💡
；疫学統計領域特有の解釈をするフレーズ（例）「ケースとコントロール」「調整」「誤差」
- PECOの性質から解析方法（「モデル」）を吟味・選択する 「初めて知った」💡
- 変数の「立ち位置」「定義方法」を考える 「こうやって（変数を）決めるのか！」💡
- データ整形の実作業で実際に苦労🌀
- 自身が考えるデザイン・解析イメージを言葉で共有する難しさ🌀
- 自テーマの解析手法を、第三者に自身の言葉で説明🌀

【提供者】

- 高いモチベーションを持つ臨床家の存在を知った💡
- 臨床家（臨床研究主体）の「イメージ」を取りこぼさない難しさ🌀
；研究主体が本当に解明したいことを引き出すスキル
- デザインの必要な段階で、各専門家から情報を出し切る難しさ🌀
；とくに解析モデル選定あたり
- 臨床疫学的知見を伝える難しさ🌀

👁️ 手順の向上

臨床現場クエスチョン



HIS活用型研究



診療への還元