MANAGEMENT
PERFORMANCEHUB

Demystifying Medical Records: A crash-course in healthcare data analysis

Amy Ann Tylenda
Data Scientist, Indiana MPH

What are we talking about?

1. Data collection
2. Frequently used data and classifications
3. Caveats and how to handle them
4. HIPAA
5. De-identification and Identifiers
6. Example Project
7. Key Takeaways

MANAGEMENT PERFORMANCEHUB

What are some common examples of healthcare data?

MANAGEMENT
PERFORMANCEHUB

What is healthcare data?

- Medical History
- Medical Codes
- Demographics
- Social Determinants:

  healthcare access, education, social and community context, economic stability, and home environment

MANAGEMENT PERFORMANCEHUB

Who collects data and when?

MANAGEMENT
PERFORMANCEHUB

Who collects data and when?

Who

- Doctors and Nurses
- Pharmacists and Pharm Techs
- Lab Techs and Surgeons
- Billing/Coding Specialists

When

- Scheduling appointments
- Going to the doctor
- Getting blood work done

MANAGEMENT
PERFORMANCEHUB

Where is data collected and stored?

MANAGEMENT
PERFORMANCEHUB

Who collects data and when?

Collected

- Hospitals
- Clinics
- Pharmacies
- Outpatient Centers
- Laboratories

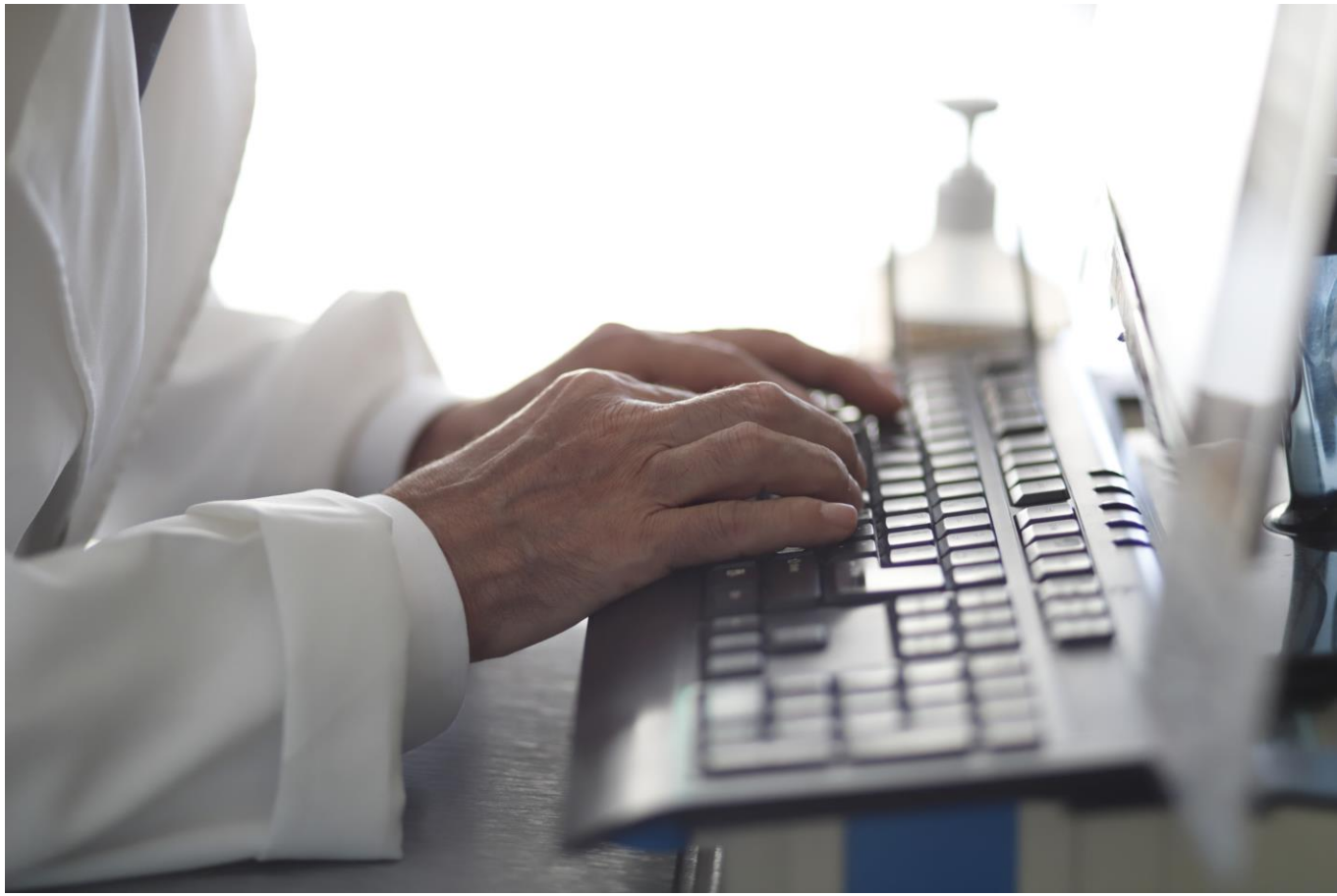
Stored

- Electronic Health/Medical Records
- Paper Records
- Imaging and Laboratory results

MANAGEMENT
PERFORMANCEHUB

Types of EHRs

- Epic
- Cerner
- Athena Health



MANAGEMENT
PERFORMANCEHUB

Frequently Utilized Healthcare Data

- Unique Identifiers (DOB, Demographics, Location of encounter)
- Diagnosis Codes
- Procedures
- Medications
- Vaccinations

MANAGEMENT PERFORMANCEHUB

Common Medical Classifications

- **ICDs are patient diagnoses: International classification of Diseases**
  - *Created by the World Health Organization*
- **CPTs are services rendered: Current Procedural Terminology**
  - *Five-digit codes identify a procedure or service*
- **NDCs are drug product identifiers: National Drug Code**
  - *Created by the FDA as a ten digit three-segment number*

MANAGEMENT
PERFORMANCEHUB

Example: Common Medical Classifications

Diagnosis code (ICD-10)

- S93.421A: Sprain of the deltoid ligament of the right ankle, initial encounter

Procedure code (CPT)

- 29425: Lower Extremity Application of Cast

Drug code (NDC)

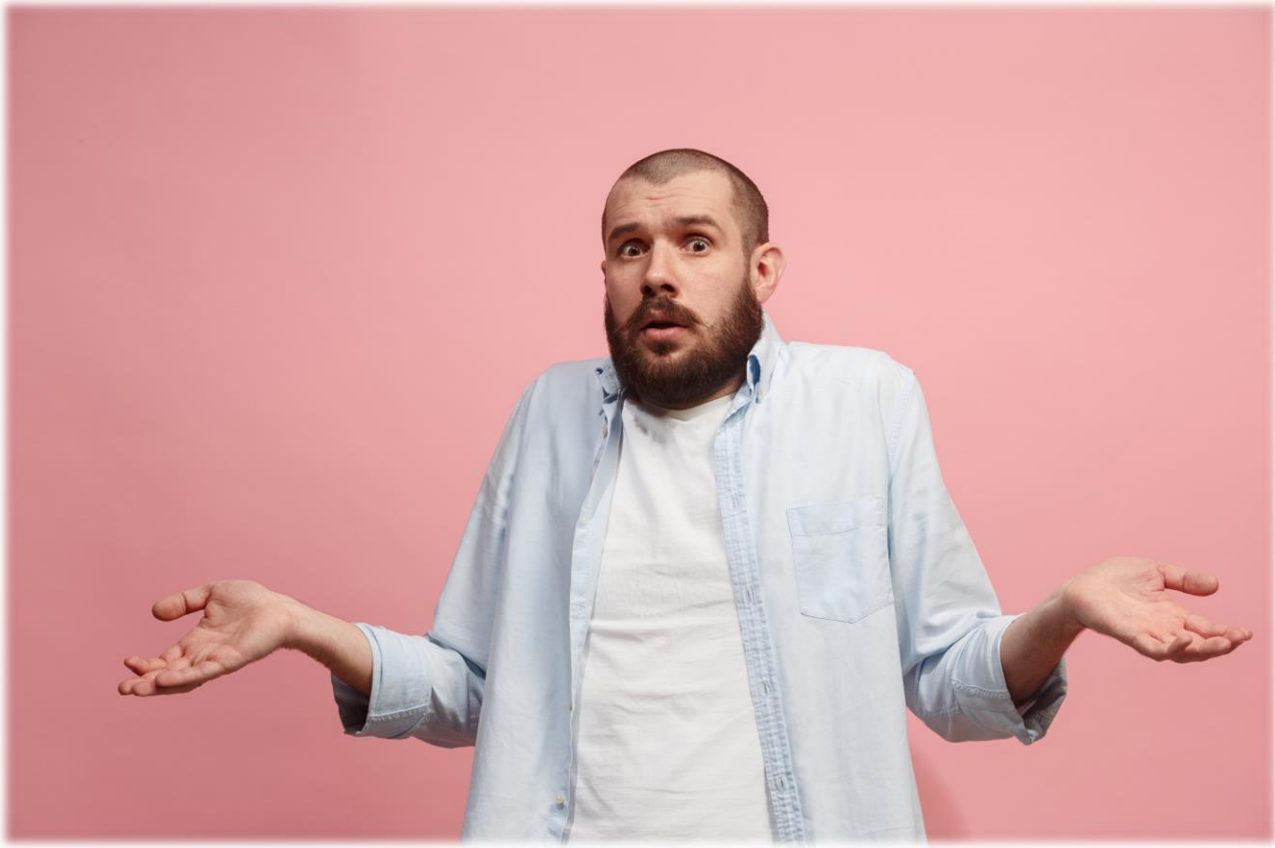
- 51655-035-50: Ibuprofen 600MG tablet 60 tablet bottle

What are some data caveats?

MANAGEMENT
PERFORMANCEHUB

Caveats in Healthcare Data

- It’s messy and often incomplete
- It ramps up in the **2010s**
- It's complex
- It’s ***hard*** to find what you are looking for
- *Every institution has nuances*



MANAGEMENT PERFORMANCEHUB

How to Combat the Caveats

- **Knowledge is power!**
- Pick date ranges with sufficient data
- Utilize medical coding crosswalks
- Ask for standardized codes
- Spot check the incoming data every so often

MANAGEMENT PERFORMANCEHUB

What is HIPAA?

MANAGEMENT
PERFORMANCEHUB

HIPAA: What is it and why it matters?

- Health Insurance Portability and Accountability Act of 1996
- HIPAA Privacy Rule regulates the use and disclosure of PHI
  PHI stands for Protected Health Information
- Business associate of covered entities, must be HIPAA compliant with HIPAA-protected data sources
  IDOH data, FSSA-Medicaid, etc.
- When in doubt, act like it's protected

MANAGEMENT
PERFORMANCEHUB

De-identification

- Remaining HIPAA compliant often requires de-identification of PHI
- Data should be thoroughly de-identified prior to hand-off
- Safe Harbor and Expert Determination are viable methods of de-identification
- The Safe Harbor method is the most straightforward form of de-identification.

MANAGEMENT
PERFORMANCEHUB

Identifiers

- Utilize the Safe Harbor method which includes the following identifiers:
  - *Names, geographic subsidiaries smaller than a state, anything after the first 3 digits of a zip code, telephone numbers, email addresses, social security numbers, medical record numbers*
- Go to hhs.gov/hipaa/ to see the complete list of identifiers.
- Other identifiers may need to be removed per state laws/regulations before releasing the data.

Getting into the Data Space

- Dig into public data
- Learn common coding languages
- Practice building dictionaries

MANAGEMENT PERFORMANCEHUB

Example Project

Hello,

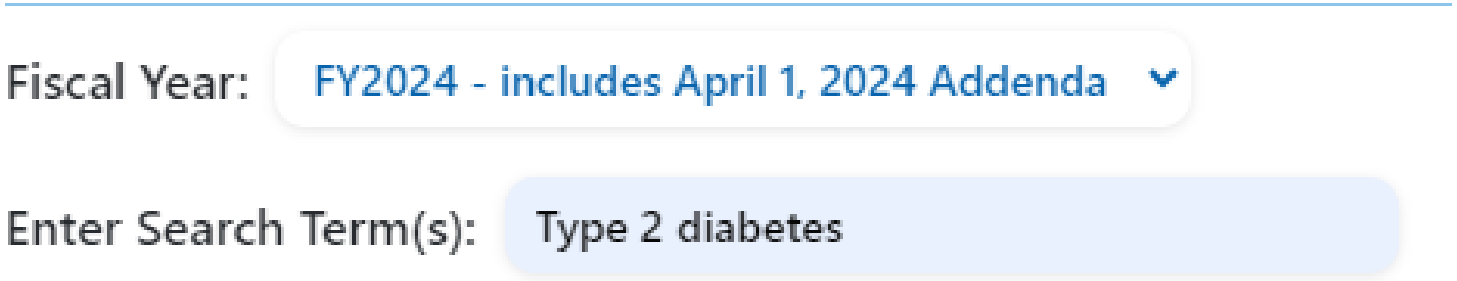
Please provide me with a list of patients who were diagnosed with Type 2 Diabetes between 2017 and 2020. Of these patients, please indicate which ones have taken metformin.

Dr. Anderson

Example: Find the diagnosis codes

- What medical classification are we using? ***ICD***
- Which version of ICD are we using? (Hint: Check the years to include)

  ICD-10
- Check ICD-10 codes from https://icd10cmtool.cdc.gov/

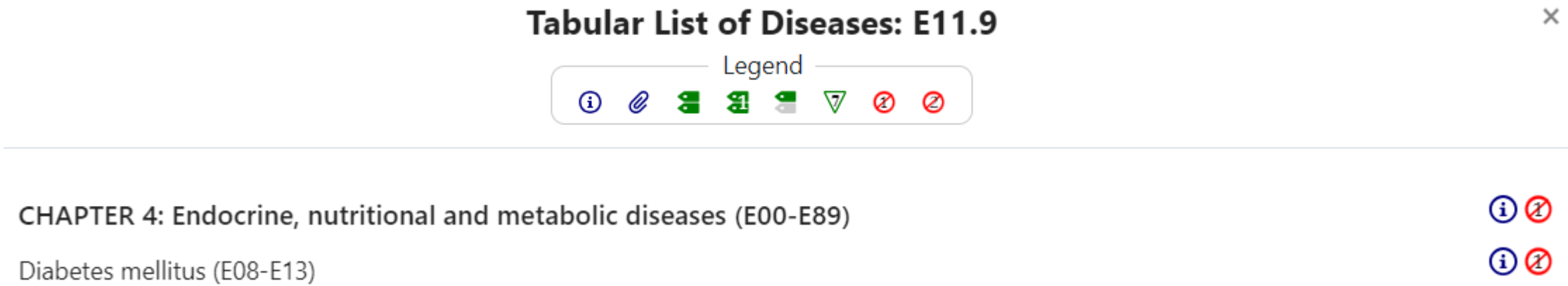


MANAGEMENT PERFORMANCEHUB

Example: Find the diagnosis codes

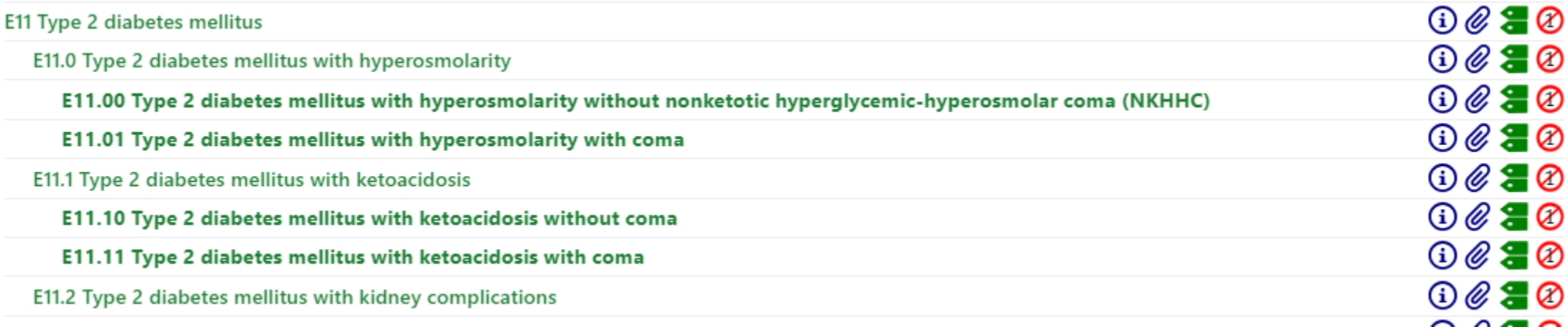
- Acidosis (lactic) E87.20
  - **in Type 1 diabetes E10.10**
  - in Type 1 diabetes E10.10 **with coma E10.11**
- Delivery (childbirth) (labor)
  - complicated O75.9 by diabetes O24.92 pre-existing O24.32 **type 1 O24.02**
  - complicated O75.9 by diabetes O24.92 pre-existing O24.32 **type 2 O24.12**
- Diabetes, diabetic (mellitus) (sugar) E11.9
  - specified type NEC E13.9 **with chronic kidney disease E13.22**
  - specified type NEC E13.9 **with dermatitis E13.620**
  - specified type NEC E13.9 **with foot ulcer E13.621**
  - specified type NEC E13.9 **with gangrene E13.52**

Example: Find the diagnosis codes



MANAGEMENT PERFORMANCEHUB

Example: Find the diagnosis codes



MANAGEMENT
PERFORMANCEHUB

Example: Find the drug codes

- What medical classification are we using? ***NDC***
- What NDC's are included?
  - https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory

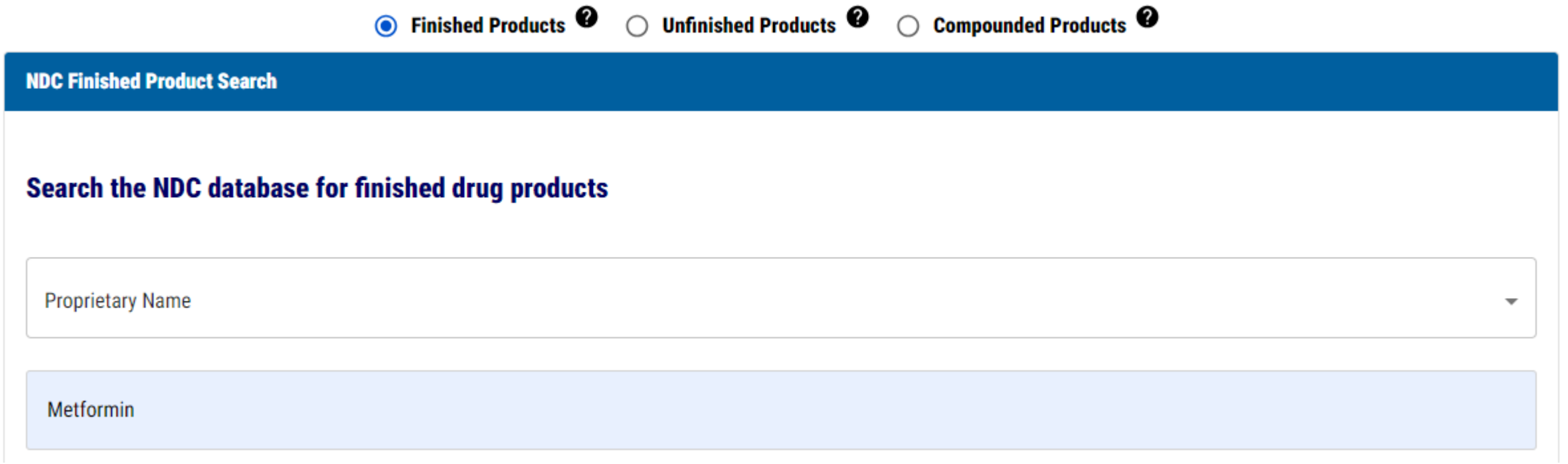
MANAGEMENT
PERFORMANCEHUB

Example: Find the drug codes

Additional References

- Search National Drug Code Directory
- NDC database file - Text Version (zip format)
- NDC database file - Excel version (zip format)
- NDC unfinished drugs database file (zip format)
- NDC compounded drugs database file (zip format)
- NDC database excluded drugs database file (zip format)
- NDC product file definitions
- NDC package file definitions
- NDC Application Programming Interface (Firefox and Chrome recommended)

Example: Find the drug codes



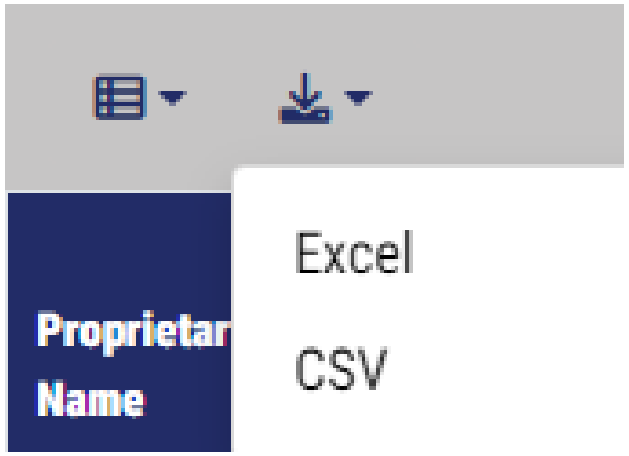
MANAGEMENT
PERFORMANCEHUB

Example: Find the drug codes

Keyword Search | Reset | Showing 1 to 50 of 1,000 entries | Show 50 entries

Proprietary Name	NDC Package Code	Strength	Dosage Form	Route	Application Number / Monograph ID	Labeler Name	Product NDC	Non Proprietary Name	Substance Name	Product Type Name	Start Marketing Date	End Marketing Date	Market Category	Package Description	Pharm Class	DEA
alogliptin and metformin hydrochloride	45802-169-72	12.5mg/1, 500 mg/1	TABLET, FILM COATED	ORAL	NDA203414	Padagis Israel Pharmaceuticals Ltd	45802-169	alogliptin and metformin hydrochloride	ALOGLIPTIN BENZOATE; METFORMIN HYDROCHLORIDE	HUMAN PRESCRIPTION DRUG	04/08/2016	N/A	NDA AUTHORIZED GENERIC	60 TABLET, FILM COATED in 1 BOTTLE (45802-169-72)	Biguanide [EPC], Biguanides [CS], Dipeptidyl Peptidase 4 Inhibitor [EPC], Dipeptidyl Peptidase 4 Inhibitors [MoA]	N/A
alogliptin and metformin hydrochloride	45802-211-72	12.5mg/1, 1000 mg/1	TABLET, FILM COATED	ORAL	NDA203414	Padagis Israel Pharmaceuticals Ltd	45802-211	alogliptin and metformin hydrochloride	ALOGLIPTIN BENZOATE; METFORMIN HYDROCHLORIDE	HUMAN PRESCRIPTION DRUG	04/08/2016	N/A	NDA AUTHORIZED GENERIC	60 TABLET, FILM COATED in 1 BOTTLE (45802-211-72)	Biguanide [EPC], Biguanides [CS], Dipeptidyl Peptidase 4 Inhibitor [EPC], Dipeptidyl Peptidase 4 Inhibitors [MoA]	N/A

Example: Find the drug codes



MANAGEMENT
PERFORMANCEHUB

Key Takeaways

- Humans have collected this data. It's **not** going to be perfect.
- Get familiar with what the data looks like. This will help you better understand and remember the caveats.
- Communicate caveats to data requestors up front and make suggestions.
- Use the Safe Harbor method when dealing with row-level data sets and stay in compliance with HIPAA

Questions?

MANAGEMENT
PERFORMANCEHUB

Contact Info

Amy Ann Tylenda
Data Scientist

Email: atylenda@mph.in.gov

MANAGEMENT PERFORMANCEHUB

Indiana Management Performance Hub
Web: data.IN.gov
Data Hub: hub.mph.in.gov