

Medication Diversification Tool

Team CodeRx

Contact:

Joseph LeGrand, PharmD, MS
115C Creighton Ave
Nashville, TN, 37206
joseph.legrand@vumc.org

Entity (for any funds distribution): CodeRx, LLC

Solution Name: Medication Diversification Tool

Challenge Category I: Enhancement to Synthea (Pediatrics)

GitHub Repo: <https://github.com/coderxio/medication-diversification>

YouTube Presentation: <https://youtu.be/9u5BSCWQtJM>

Phase II Submission

Medication Diversification Tool

Abstract:

Our medication diversification tool (MDT) enhances Synthea datasets by increasing variation of medication orders in Synthea's Generic Module Framework (GMF). Synthea modules often have hard-coded, singular medication products and strengths, which produce medication profiles that fail to represent the real-life complexity of pharmaceutical therapy. This narrow selection of medications is a major limitation to the application of Synthea-generated data for research and development. Increasing variation of medication profiles will improve the usability of Synthea data for researchers and technology developers by better reflecting real-world medicine.

The MDT leverages three publicly-available and government-maintained data sets: RxClass, RxNav, and Medical Expenditure Panel Survey (MEPS). To help Synthea better represent medication use in the US, the MDT uses these sources to enable the creation of new submodules and improvements to existing modules. The MDT provides developers options to select different medications for treatment based on disease state, medication mechanism of action, or pharmacological class. It then filters these medications to commonly-used products based on MEPS data.

Synthea would maintain its accuracy and validity by integrating the MDT submodule, which leverages data from regularly maintained RxClass (monthly) and MEPS (yearly). This setup allows for dynamic adjustments to Synthea modules and prevents medication therapy modules from becoming outdated.

The MDT can generate medication orders for all current disease or future modules within Synthea that are treated with a medication. For this challenge, we applied MDT to a use case of pediatric asthma to improve Synthea's representation of the diverse prescribing patterns in the US. The application of the MDT transformed the current asthma module from prescribing a single medication for all pediatric patients to prescribing a validated distribution across seven medications, conforming to current US asthma guidelines. Overall, MDT increased the diversity of medications prescribed based on real-world data, and thus created more practical synthetic medication profiles in Synthea.

Introduction

Looking at the Synthea data through a pharmacist's lens, we observed the medication selection for many of the Synthea modules does not represent real-world diversity of pharmaceutical disease management. In several different modules, every synthetic patient was being prescribed the same dosages of the same medications; when in reality, patient medication lists are often messy and complex. The complexity of patients' medication lists is the most challenging aspect of medication management for clinicians (Sellappans et al., 2015).

This narrow view of medications will result in downstream applications that fail in the rigor of real-world application (Chen et al., 2019). For example, the standard of care for asthma consists of one inhaled corticosteroid maintenance medication with an albuterol rescue inhaler, there are at least half a dozen different medication products commonly prescribed to patients (GINA, 2021; NAEPP, 2012) but Synthea only prescribes one. Additionally, prescribing patterns may differ based on patient-specific factors such as age. The tool we developed aids Synthea developers by creating diverse and validated medication profiles without requiring significant clinical knowledge or complex module build.

Solution

Our solution -- the medication diversification tool (MDT) -- is a python package that generates Synthea lookup tables and a submodule JSON file. These enhance the medication diversity of Synthea by

combining multiple data sources mixed with patient-specific (e.g., age, gender, location) to generate medication orders. Synthea developers choose specific medication ingredients, medication mechanisms of action, treatable/preventable disease states, or pharmacological classes, and MDT outputs a submodule representative of prescribing patterns of those medication products in the US -- obtaining a validated distribution of prescription volume across the nation, in a certain state, and/or within a certain age range.

We applied the MDT to a use case of pediatric asthma for Phase II because asthma is a complex condition with clearly defined guidelines and different prescription therapies. Using guidelines-based medication classifications and age specific distributions, the MDT programmatically generates Synthea JSON that enhances the existing medication orders in the asthma module. While we focused on pediatric asthma for this challenge, our solution has applications in all disease states within Synthea that are treated with medication. We have uploaded our open source code to GitHub (<https://github.com/coderxio/medication-diversification>) and our presentation video to YouTube (<https://youtu.be/9u5BSCWQtJM>).

Methodology

Our MDT increased variation of medication orders in Synthea by leveraging the open source medication classification hierarchies aggregated in RxClass and descriptive statistics from Medical Expenditure Panel Survey (MEPS) data sources (RxClass, n.d.; Medical Expenditure Panel Survey Home, n.d.; Hill et al., 2011). Integration of RxClass is done via the National Library of Medicine’s (NLM) supported application programming interface (API). The MDT enhanced the selection of RxNorm medications within Synthea by utilizing the knowledge base of complex pharmaceutical relationships provided by RxClass.

Prescription related MEPS data was made available to MDT from two of the downloadable files: Prescribed Medicines and Household Component Full-Year. This data was used to assign clinically relevant and validated medication utilization information to the medications retrieved from RxClass, stratified by US state, patient age, and patient gender. MEPS and RxClass data were cross-referenced to provide distributions based on real-world data. To illustrate the process on pediatric asthma, we outlined the steps below in Table 1.

Table 1. Steps and Examples for the Medication Diversification Tool (MDT) - Pediatric Asthma Maintenance Inhaler Example

Steps	Pediatric Asthma Maintenance Inhaler Example
1. Use existing online National Library of Medicine (NLM) graphical user interfaces (GUIs) to find the right medications to match a clinical guideline - either based on class or individual ingredient(s). <ol style="list-style-type: none"> a. RxClass - class of medications based on disease state, mechanism of action, pharmacological class, etc. b. RxNav - individual medication ingredients 	Per pediatric asthma guidelines (Hogan & Mahrm, 2021), a long-term (chronic) maintenance inhaler should contain a single ingredient inhaled corticosteroid (ICS). User searches RxClass: <ul style="list-style-type: none"> ● Navigate to https://mor.nlm.nih.gov/RxClass/ ● Search for “corticosteroids” and select RESPIRATORY SYSTEM > ... > Corticosteroids (RxClass ID = R01AD)
2. Input medication class / ingredient IDs into MDT and define other settings and filters:	User inputs settings into MDT:

<ul style="list-style-type: none"> a. Include vs exclude medication IDs b. Dose forms / dose form groups c. Ingredient term type (TTY) filter d. Module settings e. Demographic distribution flags f. Prescription settings 	<ul style="list-style-type: none"> ● Include RxClass ID = R01AD (corticosteroids) ● Dose forms = “Metered Dose Inhaler”, “Dry Powder Inhaler”, “Inhalation Suspension” ● Ingredient term type (TTY) filter = “IN” ● Age ranges = 0-5, 6-103 														
<p>3. Using the RxClass ID, MDT will use the RxClass API to generate a list of RXCUIs which represent medication ingredients that belong to that specific class, and will also cross-reference data from RxNorm to:</p> <ul style="list-style-type: none"> a. Combine RxClass and RXCUI “includes” and then exclude any RXCUIs in the RxClass/RXCUI “exclude” settings - can potentially create complex queries to target or exclude specific medications b. Keep only dose forms / dose form groups specified in settings (optional) c. Keep only ingredient term types (TTYs) specified in settings (optional) 	<p>Ingredients in the “corticosteroids” class:</p> <table border="1" data-bbox="831 470 1349 804"> <thead> <tr> <th>Ingredient</th> <th>RXCUI</th> </tr> </thead> <tbody> <tr> <td>beclomethasone</td> <td>285155</td> </tr> <tr> <td>budesonide</td> <td>19831</td> </tr> <tr> <td>ciclesonide</td> <td>274964</td> </tr> <tr> <td>flunisolide</td> <td>25120</td> </tr> <tr> <td>fluticasone</td> <td>41126</td> </tr> <tr> <td>mometasone</td> <td>108118</td> </tr> </tbody> </table> <p>*NOTE: Some ingredients filtered out because they don’t have any products with one of the specified dose forms OR aren’t available as single ingredient products.</p>	Ingredient	RXCUI	beclomethasone	285155	budesonide	19831	ciclesonide	274964	flunisolide	25120	fluticasone	41126	mometasone	108118
Ingredient	RXCUI														
beclomethasone	285155														
budesonide	19831														
ciclesonide	274964														
flunisolide	25120														
fluticasone	41126														
mometasone	108118														
<p>4. MDT will cross-reference these ingredient-level RXCUIs with data from the Medical Expenditure Panel Survey (MEPS) to create a distribution of the most commonly prescribed medications for each age range (and gender / state if desired) at the ingredient level (i.e. budesonide).</p> <p>NOTE: this step will generate the first level of table transitions in the Synthea module JSON (Prescribe_Inгредиent).</p>	<p>MEPS ingredient-level distribution for the above ingredients (for age range = 0-5 y.o.):</p> <table border="1" data-bbox="831 1037 1349 1262"> <thead> <tr> <th>Ingredient</th> <th>MEPS Distribution</th> </tr> </thead> <tbody> <tr> <td>beclomethasone</td> <td>3.1%</td> </tr> <tr> <td>budesonide</td> <td>47.7%</td> </tr> <tr> <td>fluticasone</td> <td>49.2%</td> </tr> </tbody> </table> <p>*NOTE: Some ingredients filtered out because the MEPS population data shows 0% utilization of these ingredients.</p>	Ingredient	MEPS Distribution	beclomethasone	3.1%	budesonide	47.7%	fluticasone	49.2%						
Ingredient	MEPS Distribution														
beclomethasone	3.1%														
budesonide	47.7%														
fluticasone	49.2%														
<p>5. MDT will map the ingredient-level RXCUIs to product-level RXCUIs and cross-reference data from MEPS to create a distribution of the most commonly prescribed medications for each age range (and gender / state if desired) at the product level (i.e. budesonide 0.25 mg/mL inhalation suspension).</p> <p>NOTE: this step will generate the second level of table transitions in the Synthea module JSON (Prescribe_<<Ingredient_Name>>) as well as all of the MedicationOrder states (Prescribe_<<Product_Name>>).</p>	<p>MEPS product-level distribution for budesonide ingredient (for age range = 0-5 y.o.):</p> <table border="1" data-bbox="831 1449 1349 1808"> <thead> <tr> <th>Product</th> <th>MEPS Distribution</th> </tr> </thead> <tbody> <tr> <td>budesonide 0.25 mg/mL inhalation suspension</td> <td>55.3%</td> </tr> <tr> <td>budesonide 0.125 mg/mL inhalation suspension</td> <td>29.5%</td> </tr> <tr> <td>budesonide 0.125 mg/mL inhalation suspension [Pulmicort] (brand name)</td> <td>15.2%</td> </tr> </tbody> </table> <p>*NOTE: Distributions are calculated for <i>each</i> ingredient - budesonide is just one example.</p>	Product	MEPS Distribution	budesonide 0.25 mg/mL inhalation suspension	55.3%	budesonide 0.125 mg/mL inhalation suspension	29.5%	budesonide 0.125 mg/mL inhalation suspension [Pulmicort] (brand name)	15.2%						
Product	MEPS Distribution														
budesonide 0.25 mg/mL inhalation suspension	55.3%														
budesonide 0.125 mg/mL inhalation suspension	29.5%														
budesonide 0.125 mg/mL inhalation suspension [Pulmicort] (brand name)	15.2%														

<p>6. MDT will then take the distributions of these medication ingredients and products at these different levels of demographic granularity and programmatically generate lookup table transition CSV files and the JSON file for a Synthea module, which will result in a variety of prescribed medications based on realistic prescribing data for the desired medications.</p>	<p>The MDT-generated module JSON and lookup table transition CSV files can be used in place of existing hard-coded individual medication products and strengths within Synthea submodules and modules by replacing existing MedicationOrder states with CallSubmodule states.</p>
--	---

Validation

Medication distributions generated by the Medication Diversification Tool (MDT) leverage MEPS as a source of truth for medication prescribing. Usage of MEPS provides a dual function. First, it creates a distribution of medications which mimics US prescribing habits. Second, it provides a way to validate medication usage in a synthetic patient population. MEPS has been shown to be a valid representation of real-world care with high accuracy in prescription use (Hill, 2007; Hill et al., 2011).

To measure the MDT's performance we compared two synthetic populations, generated with 500,000 patients each, against MEPS using a chi-squared goodness of fit test. Both populations use the Synthea asthma module, however, the first population uses the current asthma module and the second relies on MDT to generate medication orders. The synthetic populations yielded 7,213 and 6,961 pediatric asthma patients, respectively.

Table 2 displays the patient count and medication distributions for each different ICS product found in MEPS for pediatric patients. Under the current Synthea asthma module 100% of patients received Flovent 0.044 MG inhaler. Compared to the Synthea + MDT which yielded a distribution over seven different products with a range of 0.01% (Budesonide 0.5 Mg/MI) to 34.16% (Flovent 0.044 Mg).

Table 2. MEPS, Current Synthea, Synthea + MDT Distributions

Drug	MEPs	Current Synthea	Synthea + MDT
Beclomethasone	3.10%	0.00%	2.92%
Beclomethasone Dipropionate 0.04 Mg - Qvar	3.10%	0.00% (0)	2.92% (203)
Budesonide	47.70%	0.00%	47.55%
Budesonide 0.125 Mg/MI	14.07%	0.00% (0)	14.04% (977)
Budesonide 0.125 Mg/MI - Pulmicort	7.02%	0.00% (0)	7.37% (513)
Budesonide 0.25 Mg/MI	26.38%	0.00% (0)	26.13% (1819)
Budesonide 0.5 Mg/MI	0.00%	0.00% (0)	0.01% (1)
Fluticasone	49.20%	100.00%	49.53%
Fluticasone Propionate 0.044 Mg - Flovent	33.51%	100.00% (7213)	34.16% (2378)
Fluticasone Propionate 0.11 Mg - Flovent	15.69%	0.00% (0)	15.37% (1070)
* No other ICS prescribed to peds			
* Brand names provided for brand products			

A chi-squared goodness of fit test was performed on both populations against MEPS. The current Synthea population did not fit the MEPS distribution ($X^2 = 7168.52$, $df = 5$, $N = 14410$, $p = 0.00$), however, our Synthea + MDT model fit well ($X^2 = 2.73$, $df = 6$, $N = 13906$, $p = 0.84$). This validates both that MDT improved the degree to which Synthea asthma module represents prescribed medication in the US and that the randomness involved in Synthea + MDT produced a realistic medication distribution. For additional details on the validation and for a reproducible Jupyter Notebook, view the GitHub repo in the docs/validation folder (<https://github.com/coderxio/medication-diversification/tree/main/docs/validation>).

It is also worth noting MDT improved the clinical relevance of prescription generated Synthea data in the Synthea + MDT population, as orders in this population were more comprehensive of the medications recommended by asthma treatment guidelines (GINA, 2021; NAEPP 2012).

Discussion/Challenge Requirements

The Medication Diversification Tool (MDT) is a novel solution that greatly enhances Synthea by addressing the overly simplistic medication prescribing currently found in Synthea modules. The ability to include/exclude medication based on clinical concepts allows the MDT to be useful not only in pediatric asthma, but any disease state module in Synthea treated with medications. This encourages developers to build more realistic medication prescribing into current and future modules, as it allows developers to generate complex medication distributions which would not be possible manually through entering single RxNorm codes. Our MDT gives validity to Synthea with regards to medication profiles in synthetic patients since it mimics real-world data.

This tool will encourage Synthea use by researchers as created modules will contain more realistic synthetic medication data that have been validated against US patient data into Synthea. This enables researchers to identify trends, test hypotheses, and increase significance of machine learning models around medication prescribing leading to new insights that would not have been possible with lack of access to real patient data.

Additionally, MDT helps developers as it provides data that matches the complexities of real life which can improve software prior to release. For example, medication adherence or clinical decision support applications require a database of medications from which to select the prescribed medication. By adding a diverse set of medications to Synthea, developers may build and test their applications with a more robust sample of medications that are most commonly used in practice.

The groundbreaking methods used in MDT to integrate RxClass, RxNav, and MEPS paves the way for the following updates to address other limitations in Synthea:

1. Automatically generating realistic medication instructions for prescriptions (partially working in submission - code is commented out).
2. Establishing a process for enhancing Synthea with MEPS data -- which is rich with other care and patient-level details such as insurance type, visit type, race/ethnicity, family, education, employment status, health status, language(s), and medical conditions.
3. Modeling prescription fill and medication adherence over time from MEPS data.
4. Preventing prescribed medication before the earliest FDA marketing start date or after drugs get removed from the market.
5. Adding features to check allergy or drug-drug interactions by leveraging data from NLM.
6. Adding user settings for low/medium/high dose range of an ingredient.
7. Improving medication costs modeling through publicly available datasets such as NADAC, Medicare or Medicaid.
8. Creating models to base dosing on lab values. For example, levothyroxine dosing based on T3/T4 lab values.

References

- Chen, J., Chun, D., Patel, M., Chiang, E., & James, J. (2019). The validity of synthetic clinical data: a validation study of a leading synthetic data generator (Synthea) using clinical quality measures. *BMC Medical Informatics and Decision Making*, 19(1).
<https://doi.org/10.1186/s12911-019-0793-0>
- Dube, K., & Gallagher, T. (2016). *Approach and Method for Generating Realistic Synthetic Electronic Healthcare Records for Secondary Use*. <https://doi.org/10.1007>
- GINA. Global Initiative for Asthma (2021). *Global Strategy for Asthma Management and Prevention*
<https://ginasthma.org/gina-reports/>
- NAEPP. National Asthma Education and Prevention Program. (2012). *Guidelines from the National Asthma Education and Prevention Program*.
https://www.nhlbi.nih.gov/files/docs/guidelines/asthma_qrg.pdf
- Hill, S. C. (2007). The Accuracy of Reported Insurance Status in the MEPS. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 44(4), 443–468.
https://doi.org/10.5034/inquiryjrnl_44.4.443
- Hill, S. C., Zuvekas, S. H., & Zodet, M. W. (2011). Implications of the Accuracy of MEPS Prescription Drug Data for Health Services Research. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 48(3), 242–259. https://doi.org/10.5034/inquiryjrnl_48.03.04
- Hogan, A. D., & Mahr, T. A. (2021). Update on pediatric asthma treatment options, doses, label changes. *AAP News*.
<https://www.aappublications.org/news/2020/07/01/focusasthma070120>
- MEPS. *Medical Expenditure Panel Survey Home*. (n.d.). www.meps.ahrq.gov. Retrieved February 25, 2021, from <https://www.meps.ahrq.gov/mepsweb/>
- RxClass*. (n.d.). mor.nlm.nih.gov. Retrieved February 25, 2021, from <https://mor.nlm.nih.gov/RxClass/#>
- Sellappans, R., Lai, P. S. M., & Ng, C. J. (2015). Challenges faced by primary care physicians when prescribing for patients with chronic diseases in a teaching hospital in Malaysia: a qualitative study: Figure 1. *BMJ Open*, 5(8), e007817. <https://doi.org/10.1136/bmjopen-2015-007817>
- Zuvekas, S. & Olin, G. (2009). Validating Household Reports of Health Care Use in the Medical Expenditure Panel Survey. *Health Serv Res*, 44(5), 1679–1700.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754554/>