2019 MMA Datathon

Case Challenge: Planning a Hospital Formulary

It is December 31st, 2008 and Turing Hospital, based in New York City, is preparing to open its door to patients within the next year. The newly appointed director of pharmacy is deciding which diabetes medications to list on the hospital formulary, a list of approved medications that the hospital will fund for patients. The clinical guidelines mandate metformin and insulin, however there is flexibility for other diabetes medications, which is to be decided.

Each additional medication that is carried by the pharmacy adds to the hospital overhead costs, and thus unnecessary medications are to be avoided. However, each physician at the hospital is likely to have their preferred diabetes treatment protocols; those whose preferred medications are not on the list are likely to question why they have not been included.

You and your team of external consultants have been hired by the director of pharmacy to help make a few data-driven recommendations. She provides you with a dataset containing 10 years of hospital data (1998 to 2008) from US hospitals that includes information such as anonymized patient demographics, medications they were taking in hospital, and whether they were readmitted to the hospital. Data legend is provided in Appendix A.

Your task is to analyze the data (using tools of your choice) and make a recommendation for 2 drugs, from different drug classes (see Appendix B) to be listed on the formulary. You may assume that the drugs all cost relatively the same. The primary consideration should be given to readmissions, a large driver of costs to the hospital. The director would like to list medications that lead to the lowest number of readmissions when controlling for the baseline illness (for example, patients who have multiple health conditions are naturally more likely to be readmitted). In addition to the main recommendation, consultants should also present 2 alternative medications for consideration. The audience for the presentation includes the director of pharmacy, the head of endocrinology, and the CEO.

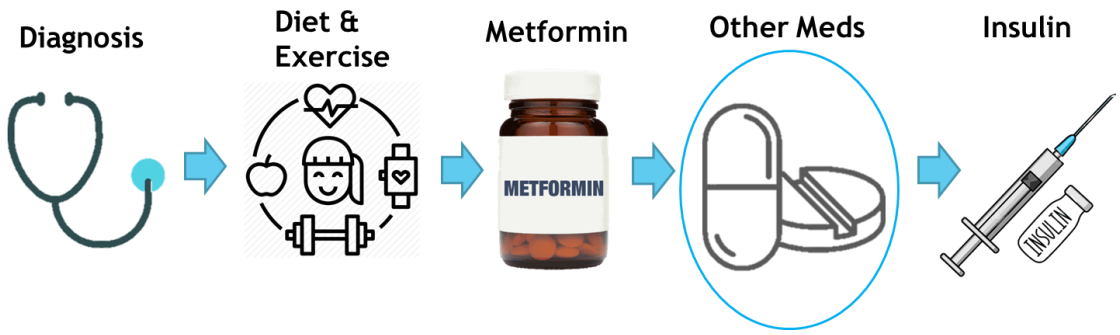
Background: Diabetes

Diabetes is one of the most prevalent serious diseases in North America. In 2008 (at the time of hospital's planned opening), the US diabetes population was estimated at roughly 18 million. Long-term complications of diabetes include heart disease, stroke, kidney failure, limb amputations, and death. Poorly managed diabetes can lead to readmissions to the hospital.

A typical patient will be initially diagnosed with early-stage diabetes by their family doctor and prescribed a regimen of diet and exercise. If that fails to stop the progression of the

This case was originally prepared by Scott Shi for the 2019 RBACxHMA SAS IQVIA Case Challenge; it has been slightly modified by Dmitry Krass for use in the MMA Datathon. This case was developed solely as the basis for this competition and is not intended to serve as an endorsement, source of primary data, or illustration of effective or ineffective management.

disease, the patient will be prescribed Metformin – a very effective drug that will be sufficient for many patients. If however, the disease continues to progress, other medications will be employed – these are listed on Table B below. Finally, a patient might progress to insulin injections. A typical patient journey is presented below.



Since the hospital must carry both Metformin and Insulin, the focus of your analysis will be on the "other medications" group. These medications fall into five different classes:

- Alpha-glucosidase inhibitors
- DPP-4 Inhibitors
- Glitazones
- Meglitinides
- Sulfonylureas

The particular medications belonging to each class, the primary effects of each class, and some additional notes can be found on Table B below.

The data for this case was extracted from the Health Facts database, a US national data warehouse that collects comprehensive clinical records across hospitals throughout the United States. Data for all hospital visits (called "encounters") between 1999 and 2008 was extracted. Additional filters were applied to ensure:

- It was an inpatient encounter (a hospital admission)
- It was a "diabetic" encounter, that is, one during which any kind of diabetes was entered to the system as a diagnosis.
- The length of stay was at least 1 day and at most 14 days.
- Laboratory tests were performed during the encounter
- Medications were administered during the encounter.

This resulted in a dataset consisting of 101,766 encounters. The data dictionary is given on Table A below. Several variables are of particular importance:

- **A1C result**: this test is used to measure what percentage of patient's hemoglobin — a protein in red blood cells that carries oxygen — is coated with sugar (glycated). Generally, the lower the reading the better. Readings of < 7% are considered normal, readings in the 7%-8% range are considered elevated. A reading of >8% places the patient at an elevated risk for diabetes complications

and indicates the disease is not being effectively managed. Unfortunately, this test is not always administered during an encounter.

- **Medication columns**: there are 22 columns, one for each medication or a combination of a particular medication and metformin. Values indicate whether this medication was prescribed and, if so, whether the dosage was increased, decreased, or kept steady.
- **Readmission**: this column indicates whether the patient was readmitted within 30 days of the last hospital admission ("<30"), more than 30 days after the last readmission on record (">30") or there were no records of previous admissions in the data ("No"). Note that readmissions within the last 30 days are tracked particularly closely and are often penalized by government oversight and insurance agencies; readmissions that occur more than 30 days since the previous admission are considered less serious.

Data structure (i.e., first five lines of data) can be found in *Pharmacy Case Data Structure.csv* . Full data can be found in *Pharmacy Case Full Data.csv .*

Deliverables:

- A 10-20 slide presentation (PPT) with appendix (no limit). While the Appendix is not limited in size, the relevance of every page in the Appendix and how it supports the main body of your presentation should be very clear to the audience (reader).
- Submit via Quercus website .

Appendix A: Data Legend

Feature Name	Type	Description of Values	% Missing
encounter_id	Numeric	Unique identifier of an encounter	0
patient_nbr	Numeric	Unique identifier of a patient	0
race	Nominal	Values: Caucasian, Asian, African American, Hispanic, and other	2
gender	Nominal	Values: male, female, and unknown/invalid	0
age	Nominal	Grouped in 10-year intervals: [0,10), {10,20),…[90,100)	0
time_in_hospital	Numeric	Integer number of days between admission and discharge	0
num_lab_procedures	Numeric	Number of lab tests performed during the encounter	0
num_procedures	Numeric	Number of procedures (other than lab tests) performed during the encounter	0
num_medications	Numeric	Number of distinct generic names administered during the encounter	0
number_outpatient	Numeric	Number of outpatient visits of the patient in the year preceding the encounter	0
number_emergency	Numeric	Number of emergency visits of the patient in the year preceding the encounter	0
number_inpatient	Numeric	Number of inpatient visits of the patient in the year preceding the encounter	0
number_diagnoses	Numeric	Number of diagnoses entered into the system as a result of the encounter	0
A1Cresult	Nominal	Indicates the range of the result or if the test was not taken. Values: “>8” if the result was greater than 8%, “>7” if the result was greater than 7% but less than 8%, “normal” if the result was less than 7%, and “none” if not measured.	0
Medications (22 columns)	Nominal	Separate column for each medication (some columns correspond to medication + metformin). The value indicates whether the drug was prescribed and whether there was a change in the dosage: “up” if the dosage was increased, “down” if the dosage was decreased, “steady” if the dosage did not change, and “no” if the drug was not prescribed	0
readmitted	Nominal	Days to inpatient readmission. Values: “<30” if the patient was readmitted in less than 30 days, “>30” if the patient was readmitted in more than 30 days, and “No” for no record of readmission.	0

Appendix B: Medication Classes

Name	Class	Primary Effect of This Class	Notes
Metformin	Biguanide	N/A	
Repaglinide	Meglitinides	Increases Insulin Release	
Nateglinide	Meglitinides	Increases Insulin Release	
Chlorpropamide	sulfonylurea (1st generation)	Increases Insulin Release	
Glimepiride	sulfonylurea (3rd generation)	Increases Insulin Release	
Acetohexamide	sulfonylurea (1st generation)	Increases Insulin Release	Rarely used
Glipizide	sulfonylurea (2nd generation)	Increases Insulin Release	
Glyburide	sulfonylurea (2nd generation)	Increases Insulin Release	
Tolbutamide	sulfonylurea (1st generation)	Increases Insulin Release	
Pioglitazone	Glitazones	Reduces Resistance to Insulin	
Rosiglitazone	Glitazones	Reduces Resistance to Insulin	
Acarbose	Alpha-glucosidase inhibitors	Reduces Glucose Absorption	
Miglitol	Alpha-glucosidase inhibitors	Reduces Glucose Absorption	
Troglitazone	Glitazones	Reduces Resistance to Insulin	Withdrawn from market in 2000
Tolazamide	sulfonylurea (1st generation)	Increases Insulin Release	
Sitagliptin	DPP-4 Inhibitor	Maintains Signal for Insulin Release	Newest medication; recently approved
Insulin	Insulin	N/A	

Note: All generations of sulfonylureas are one class of medications.