Joan & Irwin Jacobs
Center for Health Innovation
AT UC SAN DIEGO HEALTH

DATA SCIENCE ALLIANCE

Surgeon Preference Cards Optimization: Project Report

August 1, 2024

Table of Contents

Abstract

The Data Science Alliance (DSA) collaborated with The Joan & Irwin Jacobs Center for Health Innovation (JCHI) at UC San Diego (UCSD) to develop a proof-of-concept (POC) predictive model for generating the most effective surgery cards, using data science methods to optimize surgery preference cards for surgeons, surgical techs, and OR nurses. This project was fostered in efforts to enhance patient safety and satisfaction, promote standardized care delivery practices, and reduce operational inefficiencies across various surgery procedures.

Project Background

Surgical operations are intricate procedures that demand precise coordination of personnel, equipment, and time. Any inefficiency in these areas can result in delays, increased costs, and, more importantly, adverse impacts on patient outcomes and satisfaction. A major source of inefficiency in surgical operations is the creation and maintenance of surgery preference cards.

Surgery preference cards contain all information needed to safely perform a case, including details on room preparation, patient preparation, positioning, instruments, and supplies. However, the current manual system often fails to accurately capture and update individual surgeon preferences, leading to potential errors or omissions in surgical preparations. This, in turn, can cause operation delays, resource wastage, and contribute to surgeon burnout and compromised patient care.

At the Urology department at UCSD Health, for example, only 32% of preference cards have been used more than twice in 2023, and there is no standard review process in place to evaluate effective card use. Surgeons are often unaware of how many cards they own and what procedures may be linked to them. With over 9,000 preference cards linked to just UC San Diego Health, the scale of this problem becomes apparent.

The inefficiencies and errors in the surgical preparation process, including issues with surgery preference cards, can lead to adverse outcomes ranging from minor complications to more severe incidents:

- **Surgical Delays:** Missing or incorrect items on a preference card can lead to extended anesthesia time for the patient, increasing the risk of complications.
- **Wrong Site Surgery:** Inadequate preparation or miscommunication, potentially exacerbated by errors in preference cards, can contribute to wrong-site surgeries.

- **Infections:** The use of incorrect instruments or materials, not specified or improperly sterilized due to oversight in the preference card setup, can increase the risk of post-operative infections.
- **Increased Stress and Errors:** The added pressure and confusion caused by discrepancies in preference cards can increase stress levels among surgical staff, leading to a higher likelihood of errors during the procedure.

To address these challenges, a data science approach could be employed to develop a model that accurately predicts and updates surgeon preference cards based on past procedures and patient data. This model would aim to ensure all necessary equipment and supplies are correctly prepared and available for each surgery performed. The result would be a decrease in operational delays and resource wastage, along with an increase in overall patient safety.

This project aligns with the goals of DSA to foster a Responsible Data Science (RDS) ecosystem by helping mission-driven organizations find an RDS approach to solving problems. By leveraging data science to improve surgical efficiency and patient outcomes, this project demonstrates the potential of RDS to make a significant impact in healthcare.

Project Scope

This project aims to utilize a data science model to accurately optimize surgical preference cards in the UCSD health system, to improve patient outcomes, patient experience, and operations

Goals

1. Enhance Patient Safety and Satisfaction: Ensure optimal surgical tools and setups are used to reduce the risk of complications, minimize the likelihood of rescheduled surgeries, and improve overall patient experience.
2. Reduce Operational Inefficiencies: Streamline the surgical preparation process to minimize delays and unnecessary resource utilization.
3. Promote Standardization in Care Delivery: Identify and diminish practice variation among surgeons to foster a more consistent and efficient delivery of care.
4. Continual Learning and Improvement: Utilize ongoing case data to refine and update preference cards, adapting to evolving surgical practices and individual preferences.

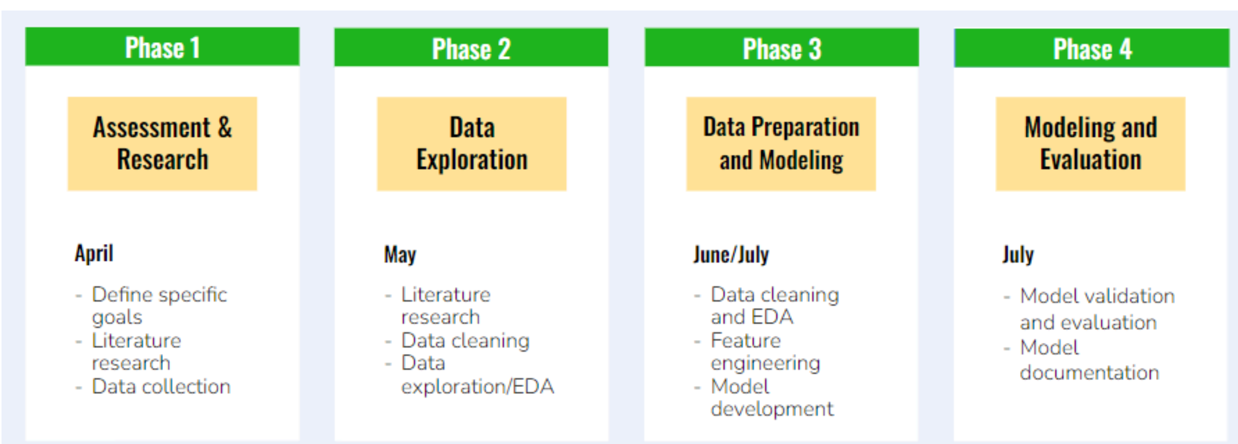
Objectives

1. Reduction in OR Delays: Track improvements in surgical start times and reductions in unexpected delays.
2. Patient Safety Incidents: Monitor changes in the frequency of tool-related safety incidents.
3. Surgeon and Patient Satisfaction: Assess feedback from surgeons and patients regarding their experiences and outcomes.
4. Waste Reduction: Measure decreases in the usage of unnecessary tools and materials.

Methods

Project Plan

A project plan was developed through communications between DSA and JCHI. A summary of the project plan that was executed by DSA can be seen below. Holding consistent meetings throughout each step created a dynamic approach to the project, in which all parties communicated effectively and applied their expertise to achieve the project goal.



End-User Partnership

As the literature suggests, model design/development/implementation should be performed with end-users to increase model acceptance. DSA worked with a team of UC San Diego Healthcare physicians and staff on a weekly basis to ensure that we correctly developed the model's purpose, inputs/outputs, and goals.

Technology

The model was developed by a team of 5 data scientists and data science interns using Amazon Web Services. The data inputted in the model was pulled from the UC San Diego Healthcare system and the algorithm was built in Python notebooks. The model aims to minimize waste by optimizing the quantities of surgical supplies for procedures based on historical usage data.

Regulatory & Legal

To comply with HIPAA Privacy regulations and protect patient privacy, our data is processed into a Limited Data Set (LDS). **PHI data remains within secure hosted environments** (Epic: Clarity, Epic: Caboodle, AWS: UCSD Cloud). PHI does not leave UCSD environments. UCSD infrastructure is under Health Information Security team responsibility. All users must use Health Active Directory (AD) accounts to access Epic of AWS hosted environments. Non-contracted third-party access is not permissible. Epic conducts annual audits of their hosted systems, including Clarity/Caboodle and provides the report to UCSD.

Methodology

Our approach builds upon the innovative work of Scheinker et al. (2021), who developed a sophisticated model to optimize surgical preference cards. This model addresses several key challenges in surgical supply management, particularly the high dimensionality of preference card data and the sparsity of procedure-specific data points.

- **Dimensionality:** Preference cards often list hundreds of items, creating a high-dimensional outcome space relative to the number of procedures performed by each surgeon. This imbalance poses challenges for traditional supervised learning models.
- **Data Sparsity:** Pooling data across providers or extending the timeframe to increase data points introduces complexities due to interprovider differences and temporal changes in supply usage patterns.
- **Item Definition:** The model considers both items listed on the preference card and those historically used by the surgeon in the associated procedure, ensuring a comprehensive approach to supply management.

Model Components and Algorithm

Data Inputs

The algorithm requires three key inputs:

- The quantity of each item to be opened and provisioned as listed on the current preference card
- Historical quantity of each item used in each procedure
- The date and time of the beginning of each procedure

Time-Series Regression Model

For each item associated with a surgeon-specific preference card, a linear time-series ordinary least squares regression model is fitted. The model uses data points from every instance of the surgeon performing the procedure, with:

- **Dependent Variable:** Quantity of the item used in that instance of the procedure
- **Independent Variable:** Time (measured in days from patient in-room) between that instance of the procedure and the procedure in which the item was first used during the previous year
- **Estimated Quantity:** The value of the linear function corresponding to the end of the 1-year period

Target Quantity Calculation

- **To be Opened:** Defined as the maximum of zero and the estimated quantity from the regression.
- **To be Provisioned:** Calculated as the difference between the 90th percentile of historical usage and the target quantity to be opened. If negative, it's set to zero.

Relative Likelihood of Inaccuracy

Computed as the absolute difference between the current quantity and the target quantity. Items are ranked in decreasing order of this likelihood.

Model Intuition and Handling of Usage Patterns

- **Discontinued Items:** For items used early in the year but discontinued later, the regression will fit a line with a negative slope, resulting in an estimate close to zero or negative. This corresponds to a target quantity of 0 to be opened.
- **Newly Introduced Items:** For items introduced later in the year, the regression starts from the first use, producing an estimate based on the linear trend of usage over that time.
- **Consistently Used Items:** For items used throughout the year, the regression provides a stable estimate based on the linear trend of usage over the entire period.
- **Provisioning Logic:** The use of the 90th percentile for provisioning ensures adequate supply for most cases while minimizing excessive stock. For example, if an item typically uses between 0 and 4 units per surgery, with a time-series estimate of 3 and a 90th percentile of 4, the target quantity to be provisioned would be 1 (the difference between the 90th percentile and the time-series estimate).

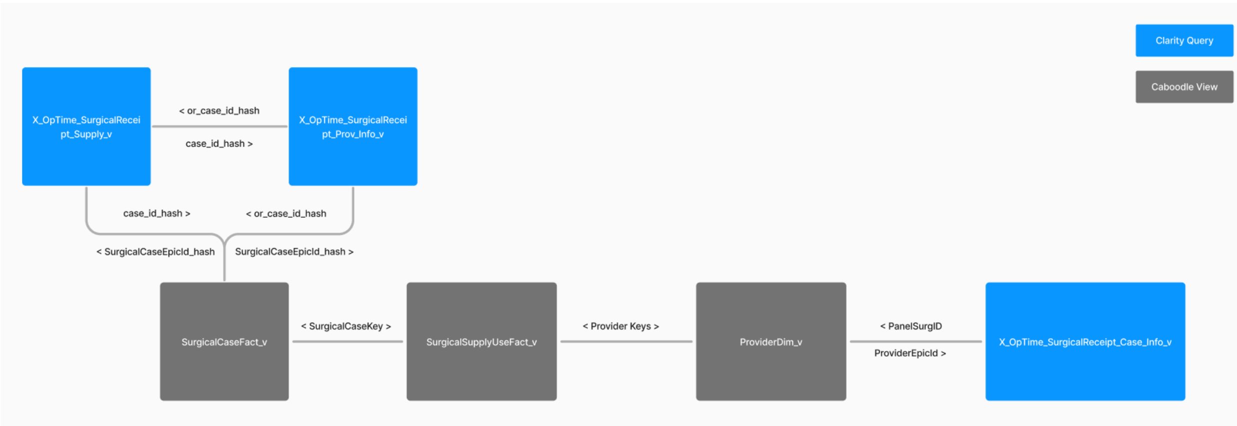
Data Preparation and Model Implementation

Data Sources and Extraction

The data used in this model is classified as ePHI (electronic Protected Health Information) and is collected by UC San Diego Health. It is stored in the AWS system environment and contains surgery data from 2019-2024. The data is stored in the following tables:

1. **ProviderDim_v**: Detailed information about providers/doctors
2. **X_OpTime_SurgicalReceipt_Prov_Info_v**: Detailed information about cases and providers/doctors
3. **SurgicalCaseFact_v**: Detailed information about cases
4. **X_OpTime_SurgicalReceipt_Case_Info_v**: Detailed information about cases
5. **SurgicalSupplyUseFact_v**: Detailed information about supply linked to cases
6. **X_OpTime_SurgicalReceipt_Supply_v**: Detailed information about supply linked to cases

The figure below shows how the tables are linked.



The data extraction process involves running an SQL query that joins multiple tables in Amazon Athena to create the main dataset, with a focus on Urology, Oncology, and Colorectal data (see Appendix A1 for the SQL query). Here's a breakdown of the data extraction process:

1. **Data Sources**: The query pulls data from four main tables:
   - x_optime_surgicalreceipt_prov_info_v
   - x_optime_surgicalreceipt_supply_v
   - surgicalcasefact_v
   - surgicalsupplyusefact_v
2. **Filtering Criteria**:
   - Only elective cases (caseclass = 'F) Elective')
   - Only completed cases (schedstatus = 'Completed')
   - Only cases with one panel (Panel ≤ 1)
   - Specific to Urology, Oncology, and Colorectal (claritysubservice = 'Urology', 'Surgical Oncology', 'Surgery: Colon and Rectal')
3. **Data Joining Process**: a. Create a base table of Urology provider information. b. Join with supply information to get preference card IDs. c. Join with surgical case information to get surgeon and procedure details. d. Join with supply use information to get details on items used.
4. **Key Variables Extracted**:
   - or_case_id_hash: Unique identifier for each case
   - doneorpnm: Procedure name
   - primaryprocedurename: Primary procedure name
   - panel: Panel number
   - inroom: Timestamp for when the patient entered the room
   - pref_card_id: Preference card identifier

    - primarysurgeondurablekey: Unique identifier for the primary surgeon
    - primarysurgeonname: Name of the primary surgeon
    - surgicalsupplydurablekey: Unique identifier for each surgical supply
    - surgicalsupplyname: Name of the surgical supply
    - unitcost: Cost per unit of the supply
    - numberopen: Number of items opened
    - numberprn: Number of items brought in but not opened unless needed
    - numberused: Number of items actually used
    - proc_pc_id: A combined identifier of procedure and preference card (created in the query)
5. **Data Aggregation**: The query creates a proc_pc_id by concatenating the procedure name (doneorpnm) and the preference card ID (pref_card_id).
6. **Data Ordering**: The final dataset is ordered by proc_pc_id and inroom timestamp.

After the SQL query is executed, the resulting dataset is stored in S3 buckets with specific URLs for each specialty (Urology, Oncology, and Colorectal). This comprehensive data extraction process ensures that all necessary information for the surgical preference card optimization is captured, including details about the procedures, surgeons, supplies used, and their associated costs.

Data Loading and Initial Processing

Once the query is run in Athena and the S3 bucket URL is retrieved, the next step is to load the data into a Pandas dataframe within an Amazon SageMaker notebook. Only preference cards with IDs starting with 'M' are kept and the data is filtered to include only records from 2019 to 2023.

Key Variables and Feature Engineering

The following features are used in the analysis:

- proc_pc_id: Unique identifier for each procedure-preference card pair
- doneorpnm: Procedure name
- pref_card_id: Preference card identifier
- surgicalsupplydurablekey: Unique identifier for each surgical supply item
- surgicalsupplyname: Name of the surgical supply item
- numberopen: Number of items opened before the start of the case
- numberprn: Number of items brought into the room but not opened unless needed
- numberused: Number of items actually used during the case
- inroom: Timestamp indicating when the patient entered the room

- unitcost: Cost per unit of the supply item

Additional features are engineered during the analysis:

- time_from_start: Number of days since the first use of an item in the dataset
- TargetOpen: Estimated number of items to be opened
- TargetProvisioned: Estimated number of items to be provisioned
- RelativeLikelihood: Measure of the likelihood of inaccuracy in item quantity

Analysis Functions

Several functions are defined to perform the analysis:

1. `fit_least_squares()`: Fits a linear regression model to the usage data.
2. `estimate_usage()`: Estimates future usage based on the regression model.
3. `calculate_target_quantities()`: Calculates target quantities for opening and provisioning.
4. `calculate_relative_likelihood()`: Calculates the likelihood of inaccuracy in item quantity.
5. `analyze_preference_card()`: Performs the complete analysis for a given preference card.

Analysis Process

The analysis is conducted separately for each specialty (Urology, Colorectal, and Oncology) and follows these steps:

1. For each unique proc_pc_id (procedure-preference card pair): a. Filter the data for the current proc_pc_id. b. Analyze data from the most recent year. c. For each unique surgical supply item:
    - Fit a linear regression model to the usage data.
    - Estimate future usage.
    - Calculate target quantities for opening and provisioning.
    - Calculate the relative likelihood of inaccuracy.

   d. Combine results for all items into a single DataFrame. e. Sort items by relative likelihood of inaccuracy. f. Save results to a CSV file named after the proc_pc_id.

2. The analysis is repeated at three levels: a. For each Procedure-Preference Card pair (as described above). b. For each Preference Card separately. c. For each Procedure separately.

Output Generation

The final results are CSV files of optimized preference cards at the three levels mentioned above for the three surgery specialties. They include both the quantities listed in the current preference cards as well as the quantities estimated by the model. Each CSV file contains the following columns:

- surgicalsupplydurablekey
- supplyname
- numberopen (current)
- numberprn (current)
- numberopen_estimated
- numberprn_estimated
- unitcost
- relative_likelihood

The files are named after the proc_pc_id, the pref_card_id, or the doneorpnm with non-alphanumeric characters replaced by underscores for safety. The figure below is an example of the model's output.

	surgicalsupplydurable...	supplyname	numberopen	numberprn	numberopen_estimated	numberprn_estimated	unitcost	relative_likelihood
1	14326	BAG DRAINAGE NE...	0.0	0.0	10	0	8.29	8.5
2	49244	GLOVE SURGICAL B...	0.0	0.0	7	0	0.84	5.666666666666667
3	47992	URETEROSCOPE LI...	0.0	1.0	6	0	1500.0	5.5
4	273443	ACCESS SHEATH FL...	0.0	1.0	5	0	123.2	4.5
5	4672	GLOVE SURGEON B...	1.0	0.0	1	2	0.84	1.744186046511628
6	54471	GLOVE BIOGEL PI U...	0.0	0.0	2	1	1.43	1.5714285714285714
7	54468	GLOVE BIOGEL PI U...	0.0	0.0	1	2	1.43	1.53125
8	38626	GLOVE BIOGEL PI I...	0.0	0.0	3	0	0.69	1.4
9	30780	BALLOON DILATATI...	0.0	1.0	2	0	237.7	1.3333333333333333
10	38625	GLOVE BIOGEL PI I...	0.0	0.0	1	2	0.69	1.3076923076923077
11	52299	SOLUTION IRR BAG ...	0.0	0.0	0	4	16.05	1.25
12	204112	FIBER LASER MOSE...	0.0	0.0	1	2	487.58	1.25
13	50723	GOWN SIRUS XLG B...	1.0	0.0	1	1	3.04	1.0116279069767442
14	273467	GUIDEWIRE SENSO...	0.0	0.0	1	1	41.62	1.0
15	320000	STENT ASCERTA FI...	0.0	0.0	1	1	75.08	0.9166666666666666
16	54473	GLOVE BIOGEL PI U...	0.0	0.0	2	0	1.43	0.9
17	42654	HOLDER FOLEY CA...	0.0	1.0	0	1	2.91	0.8833333333333333
18	32061	CATHETER URETHR...	0.0	1.0	0	1	66.24	0.88
19	13437	BAG URINE DRAIN 2...	0.0	1.0	0	1	2.57	0.8589743589743589
20	30691	GOWN SURGICAL U...	1.0	0.0	1	1	2.33	0.8421052631578947
21	54858	GOWN SURGICAL U...	0.0	0.0	0	2	2.65	0.8
22	33834	GUIDEWIRE SENSO...	1.0	1.0	1	1	41.3	0.7191011235955056
23	38629	GLOVE BIOGEL PI I...	0.0	0.0	1	1	0.69	0.7142857142857143
24	945	CONTAINER PRECIS...	0.0	1.0	0	1	0.51	0.7023809523809523
25	277565	STENT URETERAL T...	0.0	0.0	1	1	199.0	0.7

This comprehensive analysis provides a detailed view of surgical supply usage patterns and recommendations for optimizing preference cards across multiple specialties and at various levels of granularity.

Validation and Evaluation

Evaluation Metric: Waste Cost

We use waste cost as the primary evaluation metric of the results of the model because it directly aligns with the core objectives of this project:

1. **Addressing Inefficiencies**: As highlighted in the project background, a major source of inefficiency in surgical operations is the inaccurate creation and maintenance of surgery preference cards. This leads to over-preparation of supplies, resulting in waste.
2. **Financial Impact**: Waste in surgical supplies translates directly to unnecessary costs for the hospital. By focusing on waste cost, we can quantify the financial impact of our model in terms that are immediately relevant to hospital administrators and decision-makers.
3. **Patient Care**: While not directly measured, reducing waste can indirectly improve patient care. Resources saved through better supply management can be reallocated to other areas of patient care or used to keep healthcare costs down.
4. **Environmental Considerations**: Reducing waste in surgical supplies also has positive environmental implications, aligning with broader sustainability goals in healthcare.
5. **Operational Efficiency**: Waste cost serves as a proxy for operational efficiency. Lower waste indicates more accurate preference cards and smoother surgical operations, potentially reducing delays and improving overall efficiency.

By using waste cost as our primary metric, we can demonstrate the model's effectiveness in addressing the key challenges identified in the project background, namely the inefficiencies in surgical supply management and their associated costs.

Data Splitting

The validation and evaluation process begins with a strategic split of the available data:

- Training data: Historical data from 2019 to 2023
- Evaluation data: Data from 2024

This split allows the model to learn from a substantial period of past usage patterns while providing a true test of its predictive capabilities on unseen data.

Waste Cost Calculation

Central to the evaluation is the calculation of waste costs:

1. Actual waste: Based on 2024 data
   - Calculated as: (total_actual_usage - numberused) * unitcost
2. Estimated waste: Based on model predictions
   - Calculated as: (total_estimated_usage - numberused) * unitcost

These quantities are converted to costs using the unit prices of each item, providing a tangible measure of the model's potential financial impact.

Performance Metrics

A couple of key metrics are used to evaluate the model's performance:

1. Total waste cost savings: actual_waste_cost - estimated_waste_cost
2. Percentage of waste cost savings: (waste_cost_savings / actual_waste_cost) * 100

These metrics are calculated at multiple levels:

- Individual procedure-preference card pairs
- Individual procedure
- Overall hospital department

Detailed Analysis

The evaluation includes a detailed examination of the model's performance:

1. Top-performing procedures: Identifies procedure-preference card pairs and procedures with the highest cost savings.
2. Procedures with increased waste costs: Investigates instances where the model may lead to increased waste.

Sensitivity Analysis

A sensitivity analysis is conducted by evaluating the model's performance across different relative likelihood thresholds (0 to 1, in 0.1 increments). This analysis helps in understanding the model's stability.

Results

The evaluation was conducted for three specialties (Urology, Oncology, and Colorectal) and both individual procedures and procedure-preference card ID pairs. Below are the results for each category:

1. Urology

Procedure-level Analysis

- Overall waste cost savings: $251,880.08 (19.2% reduction)
- Top 10 procedures with highest waste cost savings:
    1. URETEROSCOPY, STONE: $76,219.1 (39.3%)
    2. DA VINCI XI- NEPHRECTOMY, PARTIAL, ROBOT-ASSISTED : $64,451.9 (49.3%)
    3. NEPHRECTOMY: $40,197.3 (95.3%)
    4. NEPHROLITHOTOMY, PERCUTANEOUS, OR PERCUTANEOUS NEPHROLITHOTRIPSY: $33,855.9 (10.5%)
    5. ENUCLEATION, PROSTATE, USING LASER: $15,587.3 (73.0%)
    6. DA VINCI XI- NEPHRECTOMY, RADICAL, ROBOT-ASSISTED: $15,462.7 (35.3%)
    7. ABLATION, PROSTATE, USING IRREVERSIBLE ELECTROPORATION: $11,180.2 (87.6%)
    8. NEPHROLITHOTRIPSY, MINI PERCUTANEOUS: $10,742.6 (40.3%)
    9. URETEROSCOPY, NON-STONE: $9,991.94 (31.8%)
    10. DA VINCI XI- NEPHROURETERECTOMY, LAPAROSCOPIC, ROBOT-ASSISTED: $9,310.69 (65.5%)
- Bottom 10 procedures with lowest waste cost savings:
    1. IMPLANTATION, PERIPHERAL NERVE STIMULATOR: -$41,949.1 (-1,723.6%)
    2. INSERTION, ARTIFICIAL SPHINCTER, URINARY: -$20,195.3 (-1,055.17%)
    3. INSERTION, NEUROSTIMULATOR, SACRAL : -$19,772.2 (-568.3%)
    4. DA VINCI XI- PROSTATECTOMY, RADICAL, ROBOT-ASSISTED: -$12,523.9 (-53.6%)
    5. PROSTATECTOMY, RADICAL, ROBOT-ASSISTED, USING DA VINCI XI, WITH PELVIC LYMPHADENECTOMY : -$6,024.96 (-25.8%)
    6. CRYOABLATION, PROSTATE, TRANSPERINEAL: -$5,652.9 (-1,390.8%)
    7. HYSTERECTOMY, VAGINAL, WITH COMBINED ANTEROPOSTERIOR COLPORRHAPHY AND UTEROSACRAL LIGAMENT VAULT SUSPENSION: -$5,360.16 (-60.5%)

8. CYSTOSCOPY, WITH URETERAL STENT INSERTION : -$4,261.01 (-71.8%)
9. IMPLANTATION, URETHRAL LIFT DEVICE, TRANSPROSTATIC, CYSTOSCOPIC (UROLIFT): -$2,161.6 (-21,151.3%)
10. DA VINCI XI- PROSTATECTOMY, SIMPLE, ROBOT-ASSISTED: -$1,875.76 (-15.9%)

Procedure-Preference Card Pair Analysis

- Overall waste cost savings: $384,268 (36.8% reduction)
- Top 10 procedure-preference card pairs with highest waste cost savings:
    1. DA VINCI XI- NEPHRECTOMY, PARTIAL, ROBOT-ASSISTED_M-2109: $52,877.4 (53.7%)
    2. NEPHRECTOMY_M--19687--5022000: $42,458.7 (94.8%)
    3. URETEROSCOPY, STONE_M-11294: $38,588.4 (85.4%)
    4. NEPHROLITHOTOMY, PERCUTANEOUS, OR PERCUTANEOUS NEPHROLITHOTRIPSY_M-3720: $28,596.7 (17.7%)
    5. URETEROSCOPY, STONE_M-8916: $24,718.7 (63.1%)
    6. DA VINCI XI- NEPHRECTOMY, RADICAL, ROBOT-ASSISTED_M-6515: $17,498.9 (64.6%)
    7. URETEROSCOPY, STONE_M-14609: $12,981.7 (32.9%)
    8. ENUCLEATION, PROSTATE, USING LASER_M-17565: $12,728.4 (97.5%)
    9. DA VINCI XI- NEPHRECTOMY, PARTIAL, ROBOT-ASSISTED_M-6516: $11,464.8 (58.3%)
    10. ABLATION, PROSTATE, USING IRREVERSIBLE ELECTROPORATION_M-19973: $11,180.2 (87.6%)
- Bottom 10 procedure-preference card pairs with lowest waste cost savings:
    1. INSERTION, ARTIFICIAL SPHINCTER, URINARY_M-3206: -$19,810.7 (-9082%)
    2. NEPHROLITHOTOMY, PERCUTANEOUS, OR PERCUTANEOUS NEPHROLITHOTRIPSY_M-15372 : -$4,076.93 (-5.8%)
    3. HYSTERECTOMY, VAGINAL, WITH COMBINED ANTEROPOSTERIOR COLPORRHAPHY AND UTEROSACRAL LIGAMENT VAULT SUSPENSION_M-14497: -$2,653.38 (-60.1%)
    4. CYSTOSCOPY, WITH TRANSURETHRAL NEEDLE ABLATION OF PROSTATE_M-5024: -$1,605.69 (-inf) *[There was no waste for this pair in 2024.]*
    5. RESECTION, PROSTATE, TRANSURETHRAL_M--10216--52612 : -$1177.09 (-74.8%)
    6. DISSECTION, LYMPH NODE, RETROPERITONEAL, OPEN_M-16875 : -$1015.02 (-150.6%)
    7. RESECTION, BLADDER TUMOR, TRANSURETHRAL_M--25985--52500: -$836.65 (-26.3%)
    8. VAGINOPLASTY, ROBOT-ASSISTED, USING DA VINCI SP_M-17588: -$500.57 (-184.42%)
    9. PROSTATECTOMY, RADICAL, ROBOT-ASSISTED, USING DA VINCI XI, WITH PELVIC LYMPHADENECTOMY_M-19141: -$471 (-16.7%)
    10. DA VINCI XI- PROSTATECTOMY, RADICAL, ROBOT-ASSISTED_M-2396: -$459.5 (-2.9%)

2. Oncology

Procedure-level Analysis

- Overall waste cost savings: $395,498.81 (42.6% reduction)
- Top 10 procedures with highest waste cost savings:
    1. GASTRECTOMY, SUBTOTAL, LAPAROSCOPIC: $41,509.6 (77.5%)
    2. LAPAROSCOPY, DIAGNOSTIC: $33,434.5 (45.7%)
    3. GASTRECTOMY, LAPAROSCOPIC: $21,833.3 (45.8%)
    4. LUMPECTOMY OR PARTIAL MASTECTOMY, WITH SENTINEL NODE BIOPSY: $21,476.7 (32.3%)
    5. OPEN WHIPPLE WITH DIAGNOSTIC LAPAROSCOPY: $19,458 (56.4%)
    6. HIPEC - CHEMOTHERAPY, INTRAPERITONEAL, HYPERTHERMIC: $19,456.7 (42%)
    7. ABLATION, NEOPLASM, PANCREAS, USING IRREVERSIBLE ELECTROPORATION: $18,869.2 (83.8%)
    8. LAPAROTOMY, EXPLORATORY: $17,711.5 (32.9%)
    9. PANCREATECTOMY, LAPAROSCOPIC: $15,450.6 (76.6%)
    10. CHOLECYSTECTOMY, LAPAROSCOPIC: $14,512.9 (37.9%)
- Bottom 10 procedures with lowest waste cost savings:
    1. BIOPSY, LYMPH NODE, SENTINEL, AXILLARY: -$5,506.02 (-77.9%)
    2. AUGMENTATION, BREAST: -$4,308.19 (-70.2%)
    3. INSERTION, TISSUE EXPANDER, BREAST: -$2,762.24 (-479.2%)
    4. REVISION, RECONSTRUCTION, BREAST: -$2,389.91 (-57.9%)
    5. MASTOPEXY,UNILATERAL: -$593.28 (-165.2%)
    6. HYSTERECTOMY, TOTAL, ABDOMINAL, WITH BILATERAL SALPINGO-OOPHORECTOMY: -$463.55 (-257.1%)
    7. BIOPSY, BREAST: -$318.13 (-7.3%)
    8. DA VINCI XI - EXCISION, CYST, OVARY, LAPAROSCOPIC: -$287.02 (-23.5%)
    9. DA VINCI XI- HYSTERECTOMY, ABDOMINAL, ROBOT-ASSISTED, LAPAROSCOPIC: -$252.62 (-36.6%)
    10. APPENDECTOMY, LAPAROSCOPIC: -$240.34 (-6.4%)

Procedure-Preference Card Pair Analysis

- Overall waste cost savings: $213,304.34 (55.2% reduction)
- Top 10 procedure-preference card pairs with highest waste cost savings:
    1. GASTRECTOMY, SUBTOTAL, LAPAROSCOPIC_M-1101: $21,810.8 (59.9%)

2. ABLATION, NEOPLASM, PANCREAS, USING IRREVERSIBLE ELECTROPORATION_M-3173: $19,319.8 (84.3%)
3. DA VINCI XI- HYSTERECTOMY, ABDOMINAL, ROBOT-ASSISTED, LAPAROSCOPIC, WITH REMOVAL OF TUBES/OVARIES, WITH PELVIC SENTINEL LYMPH NODE MAPPING_M-13614: $14,639.2 (34.2%)
4. HIPEC - CHEMOTHERAPY, INTRAPERITONEAL, HYPERTHERMIC_M-3787: $14,638 (62.1%)
5. LAPAROSCOPY, DIAGNOSTIC, WITH EXPLORATORY LAPAROTOMY AND LIVER RESECTION_M-19872: $11,353.6 (61.2%)
6. LAPAROSCOPY, DIAGNOSTIC_M-3787: $10,879.9 (91.8%)
7. CHOLECYSTECTOMY, LAPAROSCOPIC_M-11689: $8,322.64 (52.9%)
8. LAPAROSCOPY, DIAGNOSTIC_M-12682: $7,329.55 (37.6%)
9. THYROIDECTOMY_M-11605: $7,001.91 (54.3%)
10. PARATHYROIDECTOMY_M-11603: $6,077.64 (34.5%)

- Bottom 10 procedure-preference card pairs with lowest waste cost savings:
  1. HYSTERECTOMY, TOTAL, ABDOMINAL, WITH BILATERAL SALPINGO-OOPHORECTOMY_M--18664--5815000: -$463.55 (-257.1%)
  2. GASTRECTOMY, SUBTOTAL_M-106: -$171.34 (-12.9%)
  3. APPENDECTOMY, LAPAROSCOPIC_M-3781: -$166.29 (-6.3%)
  4. CONIZATION, CERVIX_M--18664--57520522: -$30 (-62.8%)
  5. CHOLECYSTECTOMY, LAPAROSCOPIC_M-11448: -$18.45 (-7.4%)
  6. VULVECTOMY, RADICAL_M-15872: -$0.68 (-1.3%)
  7. EXAM UNDER ANESTHESIA, PELVIS_M-15851: $9.8 (10.8%)
  8. LEEP PROCEDURE_M-17752: $16.1 (45.6%)
  9. BIOPSY, LYMPH NODE, AXILLARY_M-12679 : $37.5 (36.9%)
  10. VULVECTOMY, RADICAL_M-15874: $38.4 (40.6%)

3. Colorectal

Procedure-level Analysis

- Overall waste cost savings: $491,846.79 (69.7% reduction)
- Top 10 procedures with highest waste cost savings:
  1. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC: $133,956 (68.9%)
  2. COMPLEX ILEOSTOMY, TAKE-DOWN: $37,634 (81.2%)
  3. DA VINCI XI - COLECTOMY, LEFT OR SIGMOID, LAPAROSCOPIC: $34,477.9 (74.1%)

4. EXAM UNDER ANESTHESIA, ANORECTAL: $34,368.5 (76.8%)
5. DA VINCI XI- RESECTION, ABDOMINOPERINEAL, WITH COLOSTOMY: $27,244.1 (69.6%)
6. DA VINCI XI - COLECTOMY, RIGHT, ROBOT-ASSISTED: $23,148.7 (56.8%)
7. COLONOSCOPY: $20,718.3 (68.5%)
8. DA VINCI XI - TAMIS ROBOTIC: $20,532.4 (71.6%)
9. HEMORRHOIDECTOMY: $17,087.3 (84.0%)
10. CLOSURE, COLOSTOMY: $15,135.6 (79.9%)

- Bottom 10 procedures with lowest waste cost savings:
    1. EXCISION, HIDRADENITIS: -$445.03 (-282.5%)
    2. CYSTOSCOPY: -$427.54 (-4,060.2%)
    3. COLECTOMY, TOTAL, ROBOT-ASSISTED, LAPAROSCOPIC, USING DA VINCI XI: -$321.79 (-1309.7%)
    4. EXCISION OF PRESACRAL TUMOR: -$64.06 (-128.8%)
    5. FISTULECTOMY, ANAL: -$11.54 (-190.1%)
    6. SURGERY, TRANSANAL APPROACH, MINIMALLY INVASIVE: -$2.08 (-2.5%)
    7. COLOSTOMY TAKEDOWN WITH COLORECTAL ANASTOMOSIS: $15.94 (0.78%)
    8. REPAIR, HERNIA, PARASTOMAL: $61.64 (30.6%)
    9. REPAIR, HERNIA, INCISIONAL: $106.21 (86.7%)
    10. ANOSCOPY, WITH BIOPSY: $170.63 (70%)

Procedure-Preference Card Pair Analysis

- Overall waste cost savings: $454,875.05 (76.3% reduction)
- Top 10 procedure-preference card pairs with highest waste cost savings:
    1. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-2100: $52,639.6 (76.6%)
    2. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-2099: $42,304 (73.6%)
    3. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-5240: $40,134.5 (70.0%)
    4. COMPLEX ILEOSTOMY, TAKE-DOWN_M-1889: $27,004.9 (98.0%)
    5. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-2098: $22,520.5 (70.8%)
    6. DA VINCI XI - TAMIS ROBOTIC_M-3112: $17,320.9 (85.2%)
    7. DA VINCI XI - COLECTOMY, LEFT OR SIGMOID, LAPAROSCOPIC_M-20062: $17,102 (76.3%)
    8. COMPLEX ILEOSTOMY, TAKE-DOWN_M-1533: $5,913.1 (97.4%)

9. DA VINCI XI- RESECTION, ABDOMINOPERINEAL, WITH COLOSTOMY_M-5032: $15,408 (85.8%)
10. CLOSURE, COLOSTOMY_M-1550: $10,662 (98.0%)

- Bottom 10 procedure-preference card pairs with lowest waste cost savings:
    1. DA VINCI XI - COLECTOMY, LEFT OR SIGMOID, LAPAROSCOPIC_M-12937: -$8,423.71 (-637.8%)
    2. COMPLEX ILEOSTOMY, TAKE-DOWN_M-5935: -$3,935.31 (-265.12%)
    3. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-12934: -$1,479.49 (-962.1%)
    4. DA VINCI XI- RECTOPEXY, ROBOT-ASSISTED_M-3185: -$1,012.72 (-242.1%)
    5. EXCISION, HIDRADENITIS_M-10341: -$296.79 (-242.4%)
    6. DA VINCI XI - COLECTOMY, RIGHT, ROBOT-ASSISTED_M-8669: -$289.52 (-49.1%)
    7. CLOSURE, COLOSTOMY_M-13150: -$254 (-1899.8%)
    8. DA VINCI XI- LOW ANTERIOR RESECTION, ROBOT-ASSISTED, LAPAROSCOPIC_M-14123: -$163.68 (-110.3%)
    9. LAPAROTOMY, EXPLORATORY_M-6828: -$149.92 (-5.9%)
    10. EXCISION, PILONIDAL CYST_M-14200: -$47.36 (-68.9%)

These results demonstrate significant potential for waste reduction and cost savings across all three specialties, with the procedure-level analysis generally showing higher savings than the procedure-preference card pair analysis. The findings also highlight specific procedures and preference cards that may require further investigation or refinement to improve their performance.

Limitations and Future Work

While the current iteration of the POC model for surgical supply optimization demonstrates promising results in reducing waste and costs, there are several avenues for potential enhancement and future work:

1. **Intelligent Recommendation System**: Develop a recommendation system that works in tandem with the predictive model. This system could leverage machine learning algorithms like collaborative filtering or content-based filtering to analyze patterns across different surgeons, procedures, and hospitals, offering increasingly personalized and context-aware suggestions over time.
2. **Expanded Data Sources**: Incorporate a wider range of data sources to bolster the model's predictive power. This could include:

   - Patient-specific factors such as comorbidities or BMI for more personalized and accurate supply predictions.
   - Time-based trends or seasonality in surgical supply usage.
   - Qualitative feedback from surgeons and operating room staff to refine the model.
3. **Continuous Learning and Adaptation**: Implement a feedback loop system where post-surgical reports on actual supply usage are quickly fed back into the model. This would allow for continuous learning and adaptation, improving the model's accuracy over time.
4. **Comprehensive Evaluation Metrics**: Expand the model to consider not just cost savings from reduced waste, but also potential impacts on patient outcomes, surgical efficiency, and overall healthcare delivery. This more comprehensive evaluation would provide better insights into the model's benefits and guide further refinements.
5. **Investigation of Increased Waste Cases**: Conduct further investigation into procedures where the model led to increased waste costs to understand the underlying factors and improve predictions for these cases.
6. **Scalability and Generalization**: Test the model's performance across different surgical departments, hospitals, and healthcare systems to ensure its scalability and generalizability.
7. **Integration with Existing Systems**: Explore ways to seamlessly integrate the model with existing hospital inventory management and electronic health record systems for real-time optimization.
8. **User Interface Development**: Design and implement a user-friendly interface for surgeons and hospital staff to interact with the model's predictions and provide feedback.
9. **Long-term Impact Study**: Design and conduct a long-term study to assess the model's impact on not just immediate cost savings, but also on long-term operational efficiency, staff satisfaction, and patient outcomes.

These enhancements and future directions aim to transform the current POC model into a more robust, adaptable, and comprehensive tool for surgical supply optimization. By addressing these areas, the model could potentially have an even greater impact on reducing healthcare costs, improving operational efficiency, and ultimately enhancing patient care.

References

David Scheinker, Matt Hollingsworth, Anna Brody, Carey Phelps, William Bryant, Francesca Pei, Kristin Petersen, Alekhya Reddy, James Wall, The design and evaluation of a novel algorithm for automated preference card optimization, *Journal of the American Medical Informatics Association*, Volume 28, Issue 6, June 2021, Pages 1088–1097, https://doi.org/10.1093/jamia/ocaa275

Appendix

A1. SQL Query

```
WITH urology_prov_info AS (
  SELECT
    DISTINCT or_case_id_hash,
    doneorpnm,
    inroom,
    panel
  FROM
    x_optime_surgicalreceipt_prov_info_v
  WHERE
    panel <= 1
    AND caseclass = 'F) Elective'
    AND schedstatus = 'Completed'
    AND claritysubservice = ‘Urology’
),
urology_supply AS (
  SELECT
    DISTINCT case_id_hash,
    pref_card_id
  FROM
    x_optime_surgicalreceipt_supply_v
  WHERE
    case_id_hash IN (SELECT or_case_id_hash FROM urology_prov_info)
),
urology_surgicalcase AS (
  SELECT
    DISTINCT surgicalcaseepicid_hash,
    surgicalcasekey,
    primarysurgeondurablekey,
    primaryprocedurename,
    primarysurgeonname
  FROM
    surgicalcasefact_v
```

```
  WHERE
    surgicalcaseepicid_hash IN (SELECT or_case_id_hash FROM urology_prov_info)
),
urology_supplyuse AS (
  SELECT
    surgicalcasekey,
    surgicalsupplydurablekey,
    surgicalsupplyepicid,
    surgicalsupplyname,
    unitcost,
    numberopen,
    numberprn,
    numberused
  FROM
    surgicalsupplyusefact_v
  WHERE
    surgicalcasekey IN (SELECT surgicalcasekey FROM urology_surgicalcase)
),
joined_tables AS (
  SELECT
    u.or_case_id_hash,
    u.doneorpnm,
    u.inroom,
    u.panel,
    s.pref_card_id,
    c.surgicalcasekey,
    c.primarysurgeondurablekey,
    c.primaryprocedurename,
    c.primarysurgeonname,
    su.surgicalsupplydurablekey,
    su.surgicalsupplyepicid,
    su.surgicalsupplyname,
    su.unitcost,
    su.numberopen,
    su.numberprn,
    su.numberused,
    CONCAT(u.doneorpnm, '_', s.pref_card_id) AS proc_pc_id
```

```
  FROM
    urology_prov_info u
  JOIN
    urology_supply s ON u.or_case_id_hash = s.case_id_hash
  JOIN
    urology_surgicalcase c ON u.or_case_id_hash = c.surgicalcaseepicid_hash
  JOIN
    urology_supplyuse su ON c.surgicalcasekey = su.surgicalcasekey
)
SELECT
  DISTINCT or_case_id_hash,
  doneorpnm,
  primaryprocedurename,
  panel,
  inroom,
  pref_card_id,
  primarysurgeondurablekey,
  primarysurgeonname,
  surgicalsupplydurablekey,
  surgicalsupplyname,
  unitcost,
  numberopen,
  numberprn,
  numberused,
  proc_pc_id
FROM
  joined_tables
ORDER BY
  proc_pc_id,
  inroom;
```

A2. Model Source Code

There are six Python notebooks attached to this project report. Each file contains a specific aspect of the model implementation and evaluation.

- **2.0-amj-scheinker.ipynb:** This notebook runs the model for specific Procedures and Procedure-Preference Card ID pairs. It also evaluates the reduction in quantity wasted and creates a table comparing the actual quantities in a Preference Card with the estimated quantities.

- **3.0-amj-scheinker-cost-pc.ipynb:** This notebook evaluates the savings in waste cost when using the optimal preference card for each Procedure-Preference Card ID pair. It also performs a sensitivity analysis of the Relative Likelihood threshold.
- **3.0-amj-scheinker-cost-proc.ipynb:** This notebook evaluates the savings in waste cost when using the optimal preference card for each Procedure. It also performs a sensitivity analysis of the Relative Likelihood threshold.
- **4.0-amj-scheinker-pcs.ipynb:** This notebook generates the optimal preference card for each Preference Card ID.
- **4.0-amj-scheinker-procs.ipynb:** This notebook generates the optimal preference card for each Procedure.
- **4.0-amj-scheinker-proc-pc.ipynb:** This notebook generates the optimal preference card for each Procedure-Preference Card ID pair.