Hospital Readmission & Medicare Cost Analysis

Daniel Kattoula

PROJECT OVERVIEW

Research Question:

- Is there a relationship between hospital readmission rates and Medicare spending? What hospital characteristics influence this relationship?

Why This Matters:

- Hospital readmissions cost Medicare billions annually
- CMS penalizes hospitals with high readmission rates
- Understanding cost-readmission relationships helps identify improvement opportunities
- Can inform healthcare policy and hospital management decisions

Project Goals:

- Quantify the relationship between readmissions and costs
- Identify which hospital types perform better/worse
- Provide actionable insights for targeting interventions

DATA SOURCES

Data Sources (CMS.gov):

1. **Hospital Readmissions Reduction Program** (18,510 records) https://data.cms.gov/provider-data/dataset/9n3s-kdb3
   a. 6 conditions: Heart attack, heart failure, pneumonia, COPD, hip/knee surgery, CABG
   b. Excess Readmission Ratio (1.0 = national average)

2. **Medicare Spending Per Patient** (4,627 hospitals) https://data.cms.gov/provider-data/dataset/rrqw-56er
   a. Spending per beneficiary (1.0 = national average)

3. **Hospital General Information** (5,421 hospitals) https://data.cms.gov/provider-data/dataset/xubh-q36u
   a. Ownership type, size, emergency services, quality ratings

METHODOLOGY

- Merged 3 datasets on Hospital Facility ID
- Aggregated 6 condition-level readmission rates to hospital-level averages
- Created performance categories based on cost/readmission combinations
- Applied correlation analysis and multiple regression models

Final Analysis Dataset: 2,778 U.S. hospitals (2020-2023 data)

Data Cleaning Challenges

Major Challenges Solved:

1. Format Inconsistencies

- Hospital IDs had leading zeros in one dataset (010001 vs 10001)
- Solution: Converted to integer format

2. Text in Numeric Fields

- "N/A", "Not Available", "Too Few to Report" in numeric columns
- Solution: Converted to NA (missing values) and handled appropriately

3. Multiple Rows Per Hospital

- Each hospital had 6 rows (one per medical condition)
- Solution: Aggregated to single hospital-level average readmission ratio

4. Missing Data

- 307 hospitals removed due to incomplete key variables
- 324 hospitals missing quality ratings (retained but noted)

Key Decision: Averaged readmission ratios across all 6 conditions to create a composite hospital performance metric

Descriptive Statistics

Dataset Overview:

- 2,778 hospitals analyzed across all 50 states
- Analysis period: July 2020 - June 2023

Metric	Mean	Range
Readmission Ratio	0.999	0.58 - 1.64
Medicare Spending	0.993	0.51 - 1.51
Total Discharges	816	0 - 9,965
Total Readmissions	137	0 - 1,896

Hospital Performance Distribution:

- 884 hospitals (32%) - Low Cost, Low Readmit (Best)
- 639 hospitals (23%) - High Cost, High Readmit (Worst)
- 747 hospitals (27%) - Low Cost, High Readmit (Need readmission focus)
- 508 hospitals (18%) - High Cost, Low Readmit (Need cost focus)

Top States by Hospital Count: California (249), Texas (244), Florida (162)

Correlation Analysis

Statistical Test: Pearson Correlation

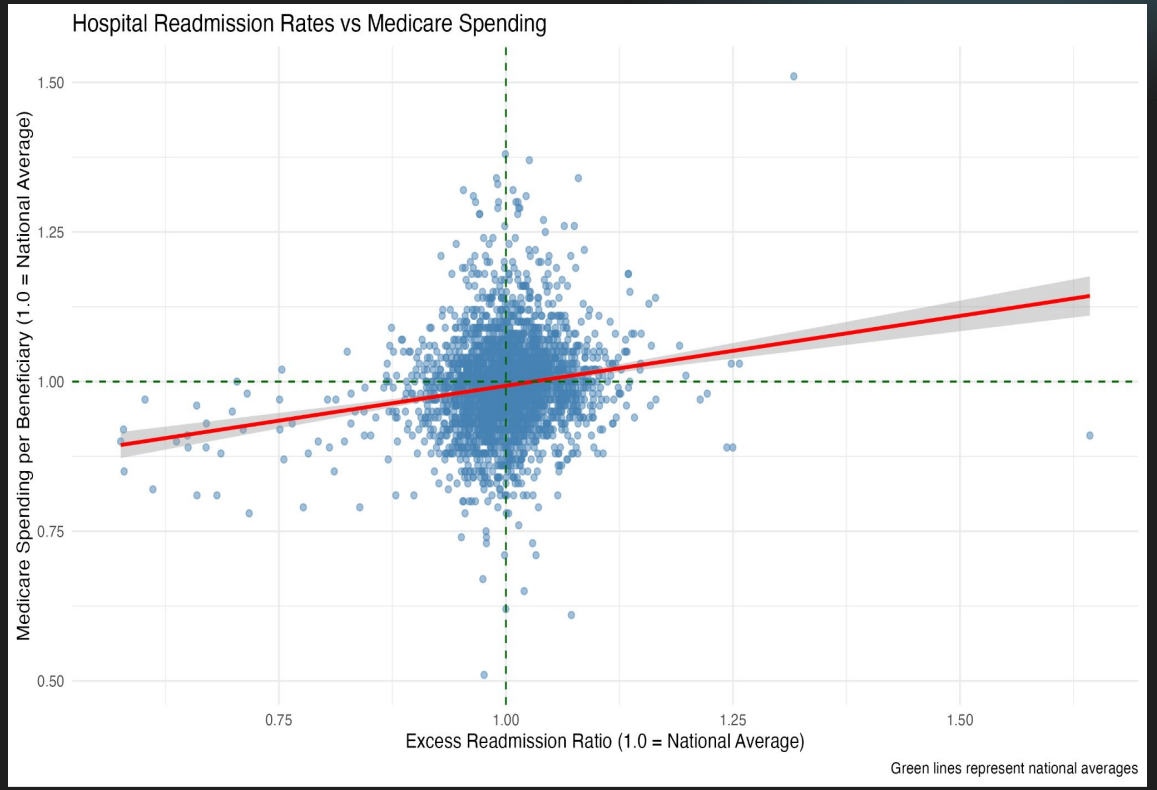
- Correlation coefficient (r) = 0.169
- P-value < 0.001 (highly statistically significant)
- 95% Confidence Interval: [0.132, 0.205]

Interpretation:

- Positive correlation: Higher readmission rates ARE associated with higher Medicare spending
- The relationship is statistically significant but moderate in strength
- This is NOT due to random chance (p < 0.001 means < 0.1% probability of chance)

Visual Insights:

- Red trend line shows positive slope
- Green dashed lines mark national averages (1.0)
- Most hospitals cluster around the national benchmark
- Notable outliers exist in both directions



Regression Models

Model 1: Simple Linear Regression

- Medicare Cost = 0.760 + 0.233 × Readmission Ratio
- R^2 = 0.028 (2.8% of cost variation explained)
- Coefficient = 0.233 ($p < 0.001$)
- Interpretation: For every 0.10 increase in readmission ratio, costs increase by 2.3%

Model 2: Multiple Regression (Better Model)

- Medicare Cost = f(Readmission Ratio, Hospital Size, Ownership, Emergency Services)
- R^2 = 0.078 (7.8% of cost variation explained - almost 3× better!)
- Readmission coefficient = 0.203 (still significant after controlling for other factors)

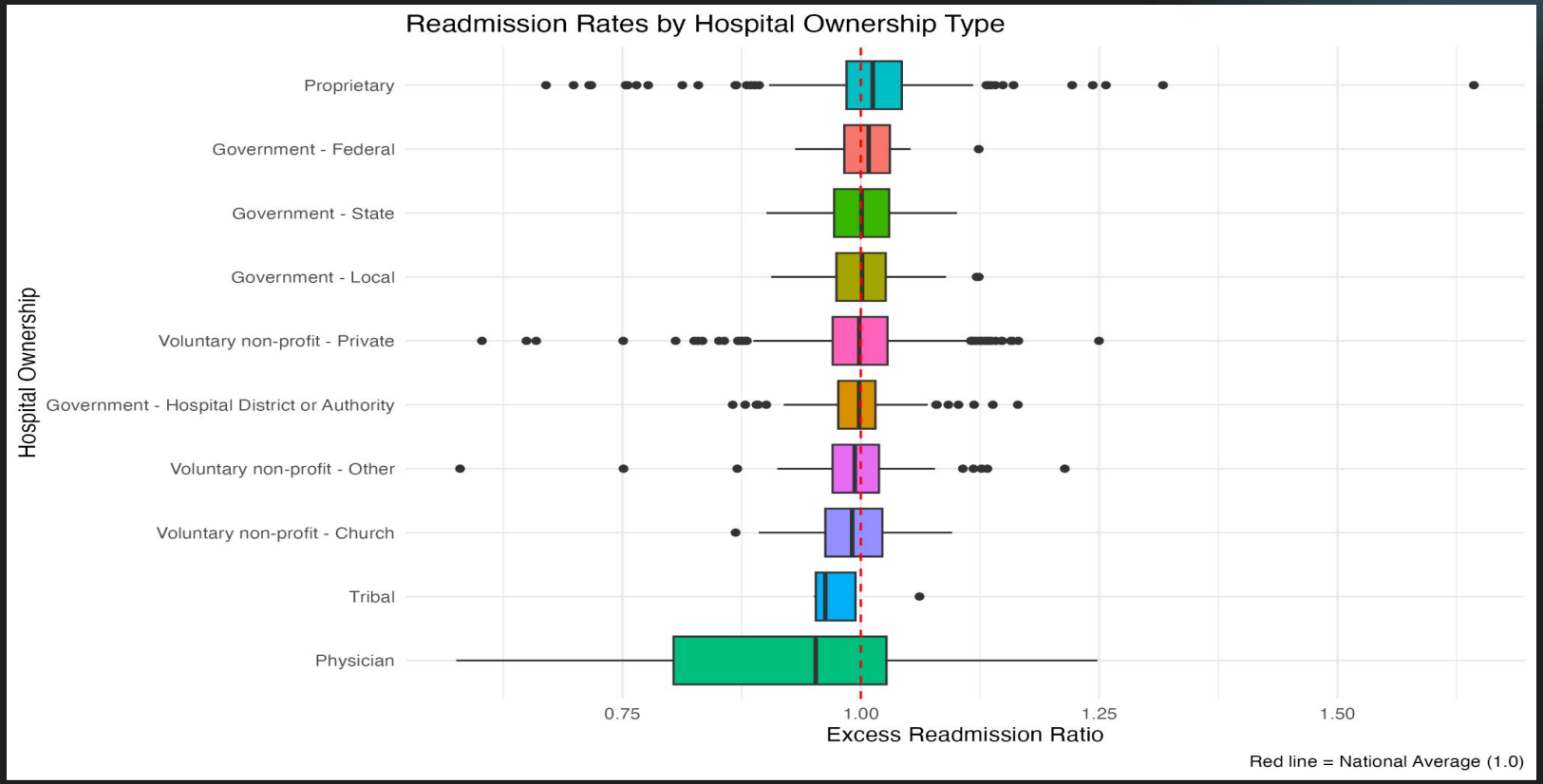
Other Significant Predictors:

- Total Discharges ($p < 0.001$): Larger hospitals have slightly higher costs
- Proprietary Ownership ($p < 0.001$): For-profit hospitals 14.7% more expensive
- Emergency Services ($p = 0.932$): NOT significant

Why Low R^2?

Healthcare costs are complex with many unmeasured factors (patient demographics, disease severity, social determinants). Our analysis shows readmissions DO matter significantly, but they're one piece of a bigger puzzle.

Hospital Ownership Matters



Best Performers

Ownership Type	Avg Readmission	Avg Cost	Count
Physician -owned	0.909	0.957	53
Tribal	0.984	0.888	4
Voluntary non-profit - church	0.993	0.982	202

Worst Performers

Ownership Type	Avg Readmission	Avg Cost	Count
Proprietary (for-profit)	1.013	1.018	543
Government - Federal	1.006	0.869	15
Government - State	1.003	0.983	38

Key Insights:

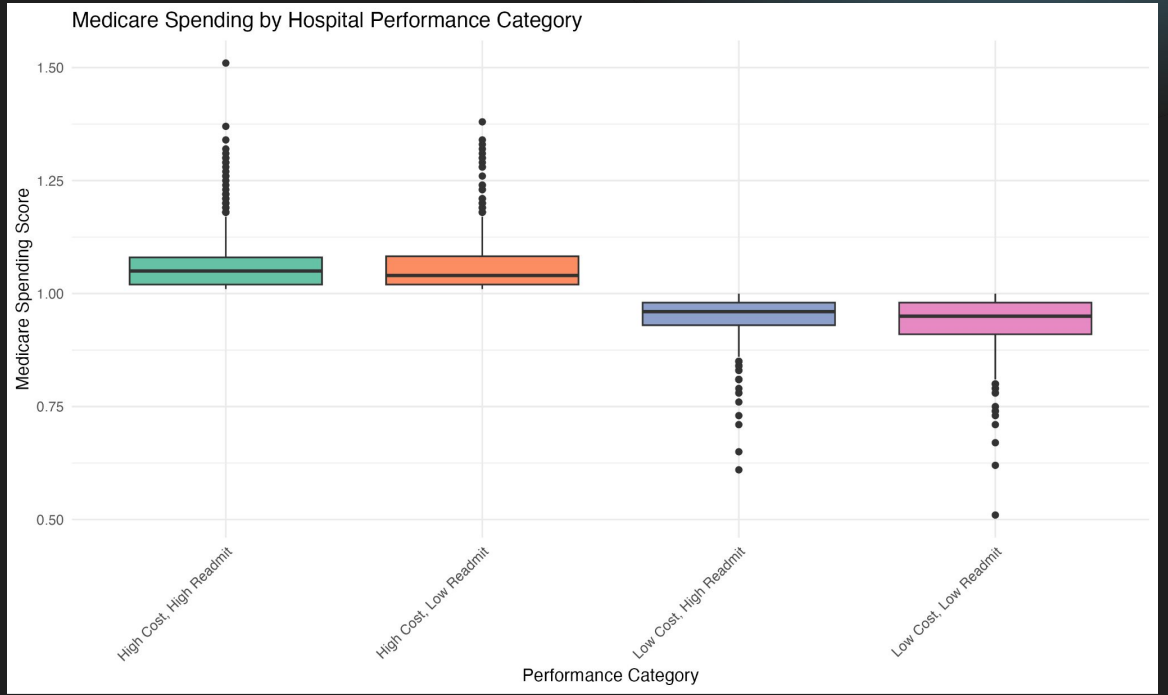
- Physician-owned hospitals outperform by ~11% on readmissions
- For-profit hospitals have both highest readmissions AND costs
- Smaller, independent facilities may provide more personalized care
- Organizational structure and incentives significantly impact outcomes

Performance Categories

Distribution Across 4 Performance Categories:

1. **High Cost, High Readmit (639 hospitals, 23%)**
   - Median spending: 1.07
   - These are the hospitals most in need of intervention
2. **High Cost, Low Readmit (508 hospitals, 18%)**
   - Good quality but expensive operations
   - May serve sicker populations or be in high-cost regions
3. **Low Cost, High Readmit (747 hospitals, 27%)**
   - Efficient but need quality improvement
   - Potential quick wins by adopting best practices
4. **Low Cost, Low Readmit (884 hospitals, 32%)**
   - Best performers - models for others to follow
   - Median spending: 0.94, Good quality outcomes

Visual Insight: The boxplot shows clear separation between high and low cost categories, with significant overlap in the high readmission groups suggesting varied cost management approaches.



Conclusions and Recommendations

Main Conclusions:

1. **Readmissions and costs are positively correlated (r = 0.169, p < 0.001)**
   - Statistically significant relationship confirmed
   - Effect persists even after controlling for hospital characteristics
2. **Hospital ownership significantly impacts both metrics**
   - Physician-owned facilities outperform for-profit hospitals by 11%
   - Organizational incentives matter for patient outcomes
3. **23% of hospitals (639) struggle with both high costs AND readmissions**
   - Clear target population for focused interventions
4. **7.8% of cost variation explained by our model**
   - Other factors (patient demographics, severity, social determinants) also critical

Recommendations:

1. For Medicare/CMS:
   - Target the 639 high-cost, high-readmission hospitals with intervention programs
   - Study and disseminate best practices from top-performing physician-owned hospitals
   - Consider ownership structure in quality incentive programs
2. For Hospital Administrators:
   - Benchmark against physician-owned and church-affiliated hospitals
   - Investigate why for-profit model correlates with worse outcomes
   - Focus on the specific conditions driving readmissions
3. For Future Research:
   - Add patient-level demographics and clinical severity measures
   - Analyze specific conditions separately rather than averaging
   - Conduct time-series analysis to establish causality
   - Investigate social determinants of health impact

Thank you!

danielkattoula@gmail.com
https://github.com/danielkattoula