Brandeis UNIVERSITY

Advanced Topics in Healthcare Data Analytics and Data Mining

Case Study: Hospital Monopoly in the US

Erin Cao

Bo Chen

Qingyue Su

Peihan Tian

Jiujun Zhang

Kaihang Zhao

04/01/2020

I. Introduction

In recent years, it is noticed that hospitals in rural towns across the U.S. are in crisis. Some reports show that 121 rural hospitals have closed since 2010. And, the National Rural Health Association says more than one-third of all rural hospitals in the U.S. are at serious risk of shutting down, which is a very serious phenomenon we are now facing. Since larger hospitals have more capital to put into the research on some tough diseases, it is much easier for these hospitals to attract those patients with some specific problems. Besides, hospitals offering better service and equipment will attract more high-income patients, which increases sharply the revenue of hospitals, which reinforces this attraction. On the contrary, hospitals lacking money will get hurt badly on their profit to maintain the further research job, and finally lead them to getting closed. That's also why we did this research.

As the preparation for making the Origin-Destination matrices to model the referral patterns among hospitals, we first calculate the market share for each hospital so that we can pick the hospitals with highest market share here. And to test whether in more complex procedures among the Major Diagnostic Categories MDCs the larger hospitals attract more of the well-paid Commercial insurance patients and fewer government paid Medicare beneficiaries. To begin with, we first take a look of the whole dataset.

	hnum2	ATYPE	asour	intage	TXTZIP	sex	dstat	PPAY	CHRGS	DX1	...	AFLAG	UNIQ	ADMID_QTR	DISCD_QTR	CHRGS_HCIA	SCUD
0	11	3	1	14	050	2	6	1	67375.25	Z5189	...	1	254	4	1	67375.25	
1	11	3	3	14	050	1	2	1	22886	Z5189	...	0	265	4	1	22886	
2	11	3	1	14	057	2	2	1	120544.05	Z4789	...	0	271	4	1	120544.05	
3	11	3	1	14	051	2	6	1	32325.8	S72001D	...	0	285	4	1	32325.8	
4	11	3	1	12	050	2	6	7	29244.1	C3490	...	0	305	4	1	29244.1	
5	11	3	1	12	050	1	6	1	124611.4	Z5189	...	0	310	4	1	124611.4	
6	11	3	4	13	050	2	6	1	71292	S7221XD	...	0	313	4	1	71292	
7	11	3	20	13	037	2	8	1	22413.4	Z5189	...	0	334	4	1	22413.4	
8	11	3	1	14	037	2	6	1	106723	Z5189	...	0	339	4	1	106723	
9	11	3	1	14	050	1	6	1	77200.65	Z4789	...	0	349	4	1	77200.65	
10	11	3	1	14	037	2	6	1	12460	S42025D	...	0	360	4	1	12460	
11	11	3	1	11	037	2	5	7	8471	Z4782	...	0	361	4	1	8471	
12	11	3	1	14	050	1	6	1	53412.1	Z4789	...	0	373	4	1	53412.1	

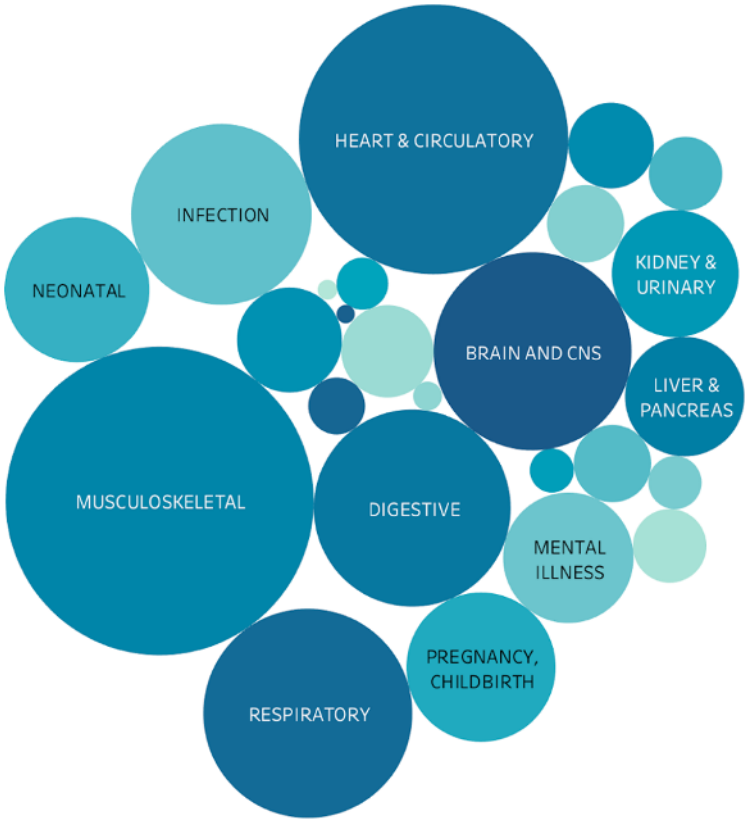
TableI-1 Glimpse of Whole Data

Since we want to focus on whether monopoly exists in the hospital, we select some relevant variables instead of processing data altogether. The description of those key variables is shown as below. To make it simple, we use major diagnostic category as our key variable to represent the diseases patients have without considering specific diagnosis (DX).

Variables	Representative meaning	Description
hnum2	Hospital Number	The unique number of every hospital
PPAY	Principal Payment Source	Medicard, Medicare, Self-paid, etc.
CHRGS	Total Charges	Patient paid for the treatment
MDC	Major Diagnostic Category	The unique number of every kind of disease

Table I-2 Description of Key Variables

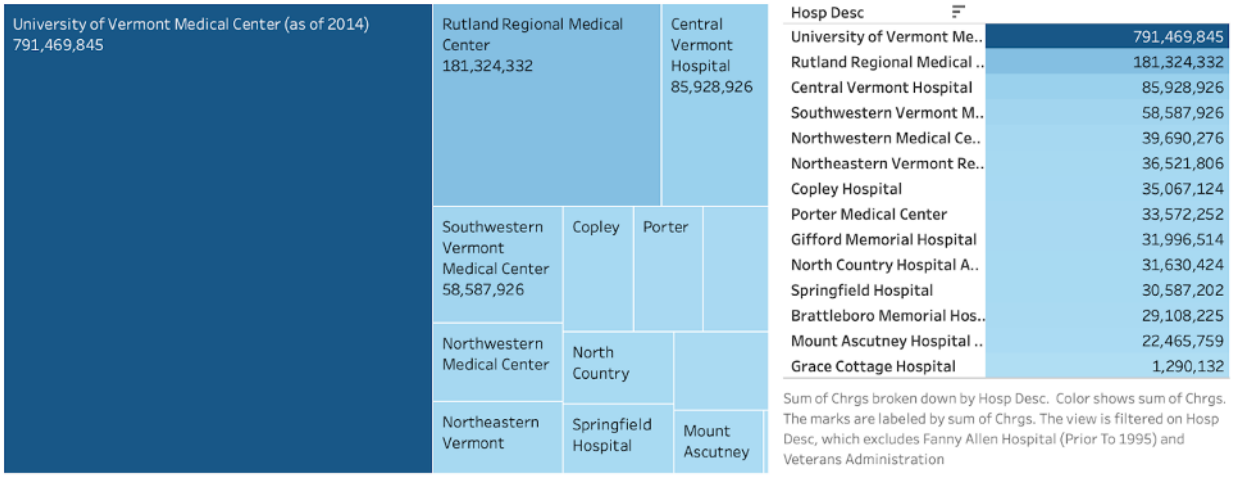
In order to know whether there are only a few dominant players among all the hospitals to have the ability to offer the treatment to some specific tough diseases with highest charges, we have to calculate the cost of different major diagnostic categories first, and we get the outcome as follows. From this graph, we can tell that the top 5 expensive major diagnostic categories are respectively musculoskeletal diseases, heart & circulatory diseases, respiratory problems, brain & CNS diseases, and digestive problems. The table on the right hand of this graph reveals the exact charges of these diseases respectively which is calculated from all the inpatient data in 2016.



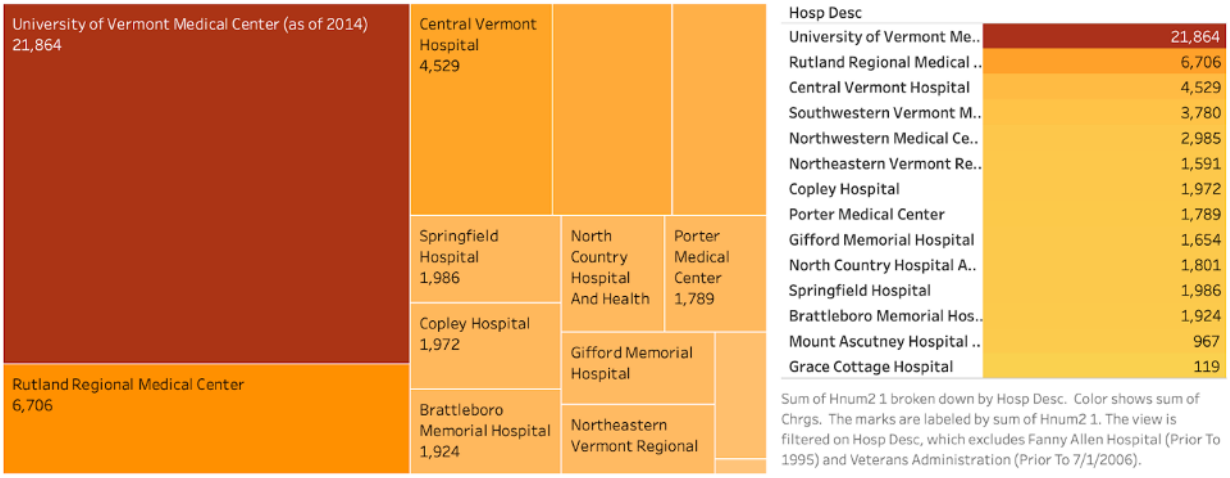
Mdc Cat Name	
MUSCULOSKELETAL	281,979,521
HEART & CIRCULATORY	215,457,944
RESPIRATORY	129,805,945
BRAIN AND CNS	116,309,419
DIGESTIVE	115,140,770
INFECTION	96,758,023
PREGNANCY, CHILDBIRTH..	66,142,373
NEONATAL	62,361,387
MENTAL ILLNESS	50,509,303
KIDNEY & URINARY	47,912,298
LIVER & PANCREAS	42,350,487
ENDOCRINE	32,695,998
ALL OTHER	25,220,452
SKIN AND BREAST	21,829,117
INJURY, TOXIC EFFECTS	18,146,970
LYMPHATIC	17,772,243
TRAUMA	16,115,936
SPLEEN & BLOOD	16,037,768
EAR, NOSE & THROAT	9,636,701
SUBSTANCE ABUSE	8,420,401
FEMALE REPRODUCTIVE	8,015,874
MALE REPRODUCTIVE	5,790,921
BURNS	2,490,797
HIV	1,167,002
EYE	1,010,990

GraphI-1 MDCs and Charges of MDCs

Afterward, we take it from another side, which is to check this problem from the hospital part. Since the calculation method of market share of hospitals does not have the only interpretation, we tried to figure out the market share distribution of these hospitals in two ways, one is calculated by the number of patients, and another is calculated by the number of money patients paid in this hospital. After calculating, it is discovered that there is no specific difference between these two methods, which can be told from the outcome below.



GraphI-2 Market share calculated by the costs of patients

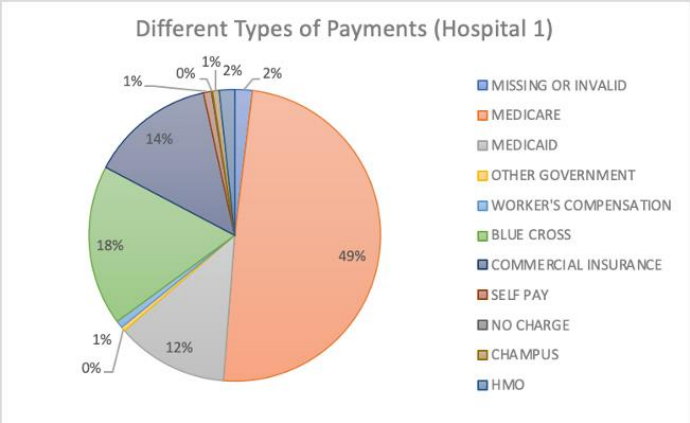


GraphI-3 Market share calculated by the number of patients

Furthermore, we chose the top three hospitals with highest market share, which are respectively University of Vermont Medical Center, Rutland Regional Medical Center, and Central Vermont Hospital, to explore what kind of patients takes up the most percentage of the total patients set. To be more understandable, we just want to figure out what types of payment accounts for the most the revenue in these giant hospital with the lion share of the whole market.

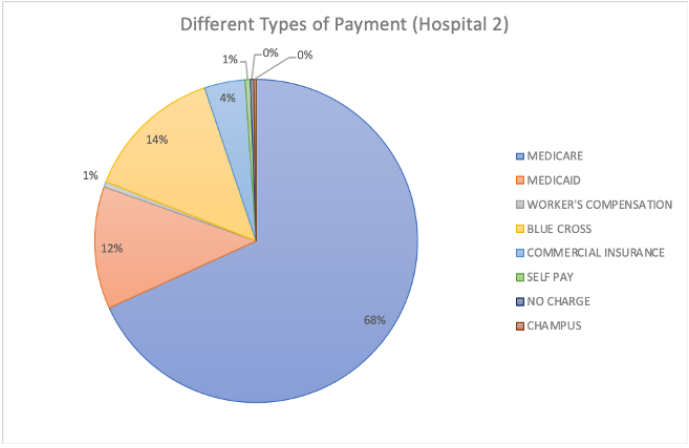
In the revenue construction of University of Vermont Medical Center, patient with government-aid insurance are the largest part of the whole patient (61%), which is the combination of patients with Medicare and those with Medicaid, followed by some patients with commercial insurance

of large insurance companies (32%), which is the combination of patients with commercial insurance of Blue Cross and those of commercial insurance.



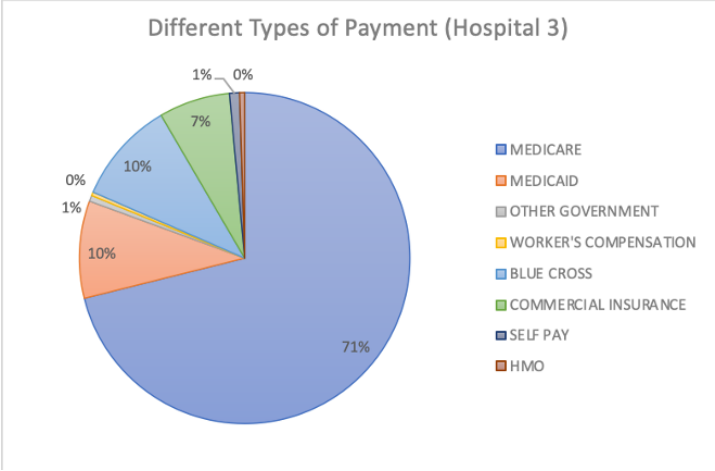
GraphI-5 Different types of payment resources - University of Vermont Medical Center

The situation is kind of similar in the revenue construction of Rutland Regional Medical Center. Patient with Medicare have the absolute predominance among all patients (68%), which means that elderly people with health insurance is the main patients of Rutland Regional Medical Center. Besides, the poor people with Medicaid and the people with Blue Cross are the other two main part contributing to the whole patient of this medical center.



Graph I-6 Different types of payment resources - Rutland Regional Medical Center

This phenomenon is much stronger in the Central Vermont Hospital, whose patient with Medicare takes up 71% among all patients, which means that elderly people with health insurance is the main patients of Rutland Regional Medical Center, and it is much higher than the 68% in Rutland Regional Medical Center, and higher than 49% in University of Vermont Medical Center. So, though the government-aid insurance is the main source of payment, the role commercial insurance plays is also a very important part to help with the hospital have a larger market share to turn to monopoly.



Graph I-7 Different types of payment resources - Central Vermont Hospital

II Exploration on hospital monopoly

Examples: Heart & Circulatory and Musculoskeletal System and Connective Tissue

I/ Heart & Circulatory Medicare

We tested to see if monopolistic powers existed amongst the regions in Vermont by looking at the movement of diagnosis and care in the form of an origin-destination matrix. In this case, we are focusing on Heart and Circulatory (MDC 5) and Medicare payers (PPAY=1) as a proxy to specialty care.

The data used for this specific analysis had 3831 observations.

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_ Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**99. 23%**	0. 35%	0. 00%	0. 07%	0. 35%
RR2--Barre	32. 99%	**64. 02%**	0. 69%	2. 30%	0. 00%
RR3--Newport/St. J	9. 80%	2. 88%	**86. 74%**	0. 29%	0. 29%
RR4--Randolph	5. 98%	1. 84%	0. 23%	**88. 97%**	2. 99%
RR5--Rutland	13. 05%	0. 00%	0. 13%	0. 67%	**86. 15%**
Hospital Market Share %	**48. 60%**	**15. 14%**	**8. 07%**	**10. 81%**	**17. 38%**

O-D Matrix 1

Furthermore, the table below shows an Origin-Destination Matrix of MDC 5. We found that when looking at the care associated with the heart and circulatory diagnosis, it was oftentimes executed locally, instead of taking care outside of the local region (Local execution rate=99.23%,64.02%,86.74%,88.97%, and 86.15%).

Moreover, we found that Burlington, home of the state's largest hospital- University of Vermont Medical Center holds the lion share of the market at 48.6%; compared to the next largest, RR5- Rutland, at

17.38%. This suggests that of all heart and circulatory diagnosis, nearly half (48.6%) of the care was rendered in Burlington.

Additionally, we tested to see if monopolistic powers existed amongst the regions in Vermont by looking at the movement of diagnosis and care in the form of an origin-destination matrix. In this case, we are focusing on Heart and Circulatory (MDC 5) and Commercial payers (PPAY=6&7) as a proxy to specialty care.

The data used for this specific analysis had 1018 observations.

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**99.23%**	0.57%	0.00%	0.00%	0.19%
RR2--Barre	59.36%	**40.18%**	0.00%	0.46%	0.00%
RR3--Newport/St. J	15.69%	3.92%	**80.39%**	0.00%	0.00%
RR4--Randolph	15.63%	4.69%	0.00%	**75.00%**	4.69%
RR5--Rutland	37.89%	0.00%	0.00%	2.48%	**59.63%**
Hospital Market Share %	**71.48%**	**9.44%**	**4.03%**	**5.21%**	**9.83%**

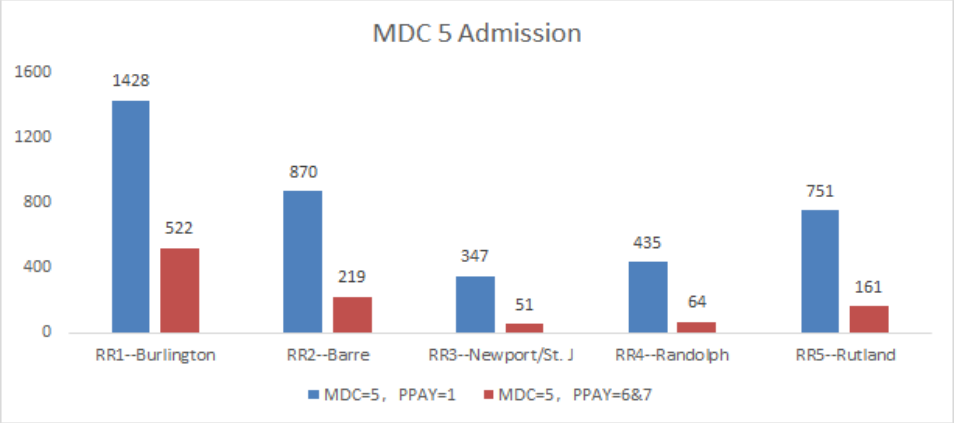
O-D Matrix 2

Furthermore, the table below shows an Origin-Destination Matrix of MDC 5. We found that when looking at the care associated with the heart and circulatory diagnosis, it was oftentimes executed locally, instead of taking care outside of the local region except for the case of Barre which only 40.18% people took care locally but other local execution rates are exceptionally high: 99.23%,80.39%, 75%, and 59.63%.

Moreover, we found that Burlington, home of the state's largest hospital- University of Vermont Medical Center holds the amazingly largest share of the market at 71.48%; compared to the next largest, RR5- Rutland, at 9.83%. This suggests that of all heart and circulatory diagnosis, nearly half

(71.48%) of the care was rendered in Burlington. It seems that there is an obvious monopolistic power existing in Burlington.

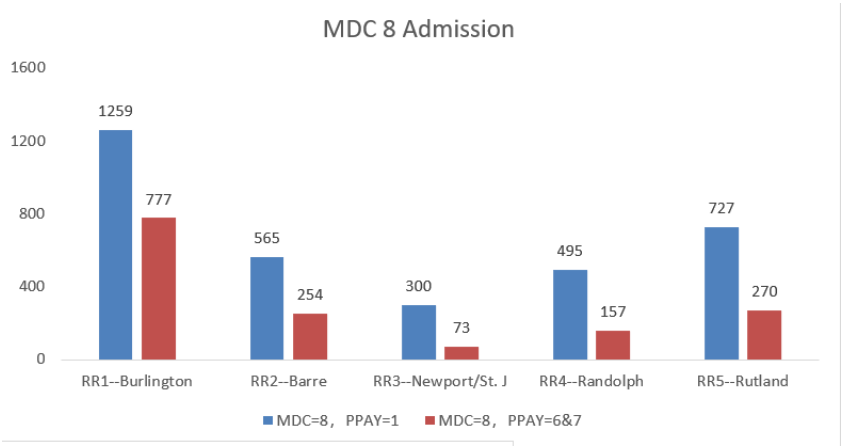
We also plotted the distribution of MDC=5. It indicates that Medicare cases overwhelm commercial cases quantitatively but the trend for both cases is similar: Burlington is the major power in the regions in Vermont.



Graph II-1

II/ Musculoskeletal System and Connective Tissue

Assuming MDC 8, we observe if monopolistic powers existed amongst selected regions in Vermont in the form of the origin-destination matrix. MDC 8 refers to the Musculoskeletal System and Connective Tissue, which covers various topics such as skeleton, muscles, connective system, joints and so on. In this example, by filtering MDC=8, we got 4,877 observations and divided it into two parts, 3,346 Medicare patients and 1,531 commercial patients. The distribution of MDC 8 is shown below.



Graph II-2

In general, no matter whether they were Medicare patients or commercial patients, our matrix remained majorly concentrated in the diagonal elements with few leakages to other areas, which could demonstrate that most patients utilized localized and well-organized healthcare services. It was not necessary for patients to travel outside for better services. Moreover, we found that home of the state's largest hospital-University of Vermont Medical Center, was located at Burlington, and it empirically proved that most leakages traveled there for diagnosis. 24.02% of Patients in Barre and 12.33% of those in Newport/St_Jns were more likely to run outside their home to Burlington.

Commercial patients afford their private health insurance by their employers or themselves. They don't have to worry about the shortage of insurance compensation on healthcare. In contrast, Medicare is a U.S. federal government health insurance program that subsidizes the poor or homeless to satisfy their fundamental need for healthcare. Therefore, more commercial patients ran to Burlington for better services. (Example: Region Barren: 24.02%>13.98%, Region Newport/St_Jns: 12.33%>7.67%)

To research on the potential monopoly market share distribution of MDC 8, we discovered that Burlington holds the majority for Medicare patients at 38.02%; followed by Rutland (22.50%), Barre (17.54%), Randolph (14.14%), and Newport/St_Jns (7.80%). Meanwhile, Burlington holds the majority for Medicare patients at 49.77%; followed by Rutland (18.88%), Barre (17.31%), Randolph (10.25%), and Newport/St_Jns (3.79%).

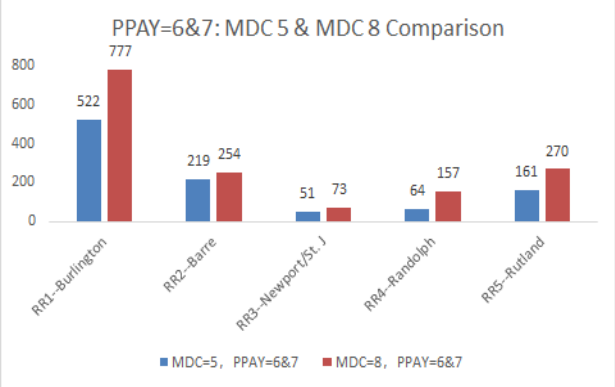
MDC=8	PPAY=6&7	Commercial			
RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**86.74%**	10.42%	0.00%	0.13%	2.70%
RR2--Barre	24.02%	**66.54%**	1.57%	7.48%	0.39%
RR3--Newport/St. J	12.33%	12.33%	**71.23%**	2.74%	1.37%
RR4--Randolph	1.91%	2.55%	1.27%	**78.34%**	15.92%
RR5--Rutland	5.56%	0.74%	0.00%	4.44%	**89.26%**
Hospital Market Share %	**49.77%**	**17.31%**	**3.79%**	**10.25%**	**18.88%**

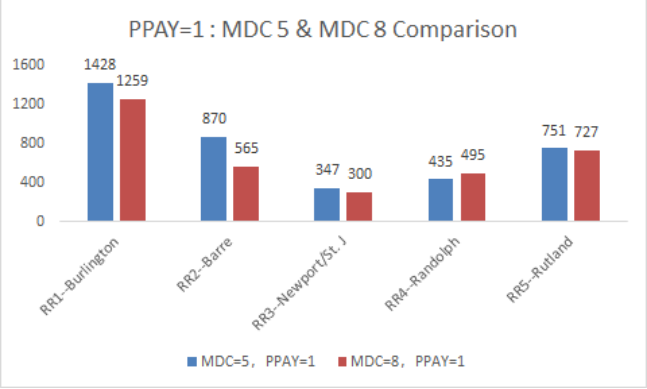
MDC=8	PPAY=1	Medicare			
RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**89.20%**	8.34%	0.16%	0.24%	2.07%
RR2--Barre	13.98%	**79.65%**	1.59%	4.07%	0.71%
RR3--Newport/St. J	7.67%	6.67%	**82.33%**	3.33%	0.00%
RR4--Randolph	2.02%	2.02%	0.61%	**84.85%**	10.51%
RR5--Rutland	5.09%	0.28%	0.00%	2.34%	**92.30%**
Hospital Market Share %	**38.02%**	**17.54%**	**7.80%**	**14.14%**	**22.50%**

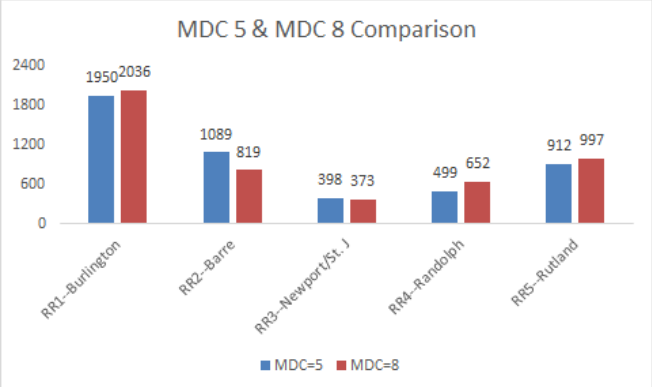
O-D Matrix 3&4

III/ Comparison between MDC 5 and MDC 8

We also did a comparison between Heart & Circulatory and musculoskeletal cases based on the following bar plots. It can be inferred that there are more commercial musculoskeletal cases than commercial Heart & Circulatory cases and Medicare musculoskeletal cases are very close to Medicare Heart & Circulatory cases. The overall musculoskeletal cases are also very close to Medicare Heart & Circulatory cases. More significantly, for all the cases, we can see the same trend: Burlington always takes control of the regions of Vermont and there should be monopolistic power existing in this area.







Graph II-3&4&5

III. Further exploration

Based on our analysis above, we can see that the market share difference among hospitals in the regions of Vermont in 2 disease categories. To further dig the major DRG in MDC categories which would cause the monopolistic practices, we set the sample size as higher than 180, then we define monopoly as the high market share. By analyzing data, we choose 0.5 as the baseline , which means if market share is higher than 0.5, it would be regarded as a high monopoly DRG, if market share is lower than 0.5, it would be regarded as a low monopoly DRG. Then we get 2 pairs of DRG: commercial pair and medicare pair.

	DRG	Monopoly	MDC	DESC
Commercial Pair	774	high	5	Vaginal delivery w complicating diagnoses
	885	low	14	Psychoses
Medicare Pair	949	high	19	Aftercare w CC/MCC
	312	low	23	Syncope & collapse

Table III-1

To compare the distribution among different hospitals, we made Origin-Destination Matrix as below:

Commercial pair

DRG 774 Commerical

O-D Matrix with Counts

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland	Total Admissions
RR1--Burlington	**193**	3	-	-	-	**196**
RR2--Barre	27	**23**	1	7		**58**
RR3--Newport/St. J	-	-	**9**	-		**9**
RR4--Randolph	-	2	1	**23**	-	**26**
RR5--Rutland	10			4	**19**	**33**
Destination Hosp SubTotal:	**230**	**28**	**11**	**34**	**19**	**322**

Percentages

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**98%**	2%	0%	0%	0%
RR2--Barre	47%	**40%**	2%	12%	0%
RR3--Newport/St. J	0%	0%	**100%**	0%	0%
RR4--Randolph	0%	8%	4%	**88%**	0%
RR5--Rutland	30%	0%	0%	12%	**58%**
Hospital Market Share %	**71%**	**9%**	**3%**	**11%**	**6%**

DRG 885 Commerical

O-D Matrix with Counts

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland	Total Admissions
RR1--Burlington	**299**	59	-		12	**370**
RR2--Barre	17	**71**	3	9	1	**101**
RR3--Newport/St. J	2	6	**32**	-	-	**40**
RR4--Randolph	1	3	1	**50**	18	**73**
RR5--Rutland	6	1		4	**118**	**129**
Destination Hosp SubTotal:	**325**	**140**	**36**	**63**	**149**	**713**

Percentages

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**81%**	16%	0%	0%	3%
RR2--Barre	17%	**70%**	3%	9%	1%
RR3--Newport/St. J	5%	15%	**80%**	0%	0%
RR4--Randolph	1%	4%	1%	**68%**	25%
RR5--Rutland	5%	1%	0%	3%	**91%**
Hospital Market Share %	**46%**	**20%**	**5%**	**9%**	**21%**

Graph III-1

Brandeis UNIVERSITY

Medicare Pair

DRG 949 Medicare

O-D Matrix with Counts

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland	Total Admissions
RR1--Burlington	**23**	-	-		-	**23**
RR2--Barre	7	**4**	-	4	-	**15**
RR3--Newport/St. J	3	-	**1**	4	-	**8**
RR4--Randolph		-	-	**149**	-	**149**
RR5--Rutland	1			2	**-**	**3**
Destination Hosp SubTotal:	34	4	1	159	-	198

Percentages

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**100%**	0%	0%	0%	0%
RR2--Barre	47%	**27%**	0%	27%	0%
RR3--Newport/St. J	38%	0%	**13%**	50%	0%
RR4--Randolph	0%	0%	0%	**100%**	0%
RR5--Rutland	33%	0%	0%	67%	**0%**
Hospital Market Share %	**17%**	**2%**	**1%**	**80%**	**0%**

DRG 312 Medicare

O-D Matrix with Counts

RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland	Total Admissions
RR1--Burlington	**47**	1	-	1	-	**49**
RR2--Barre	5	**35**	-	2	-	**42**
RR3--Newport/St. J	-	1	**17**	-	-	**18**
RR4--Randolph	-	1	-	**36**	1	**38**
RR5--Rutland	1			-	**41**	**42**
Destination Hosp SubTotal:	53	38	17	39	42	189

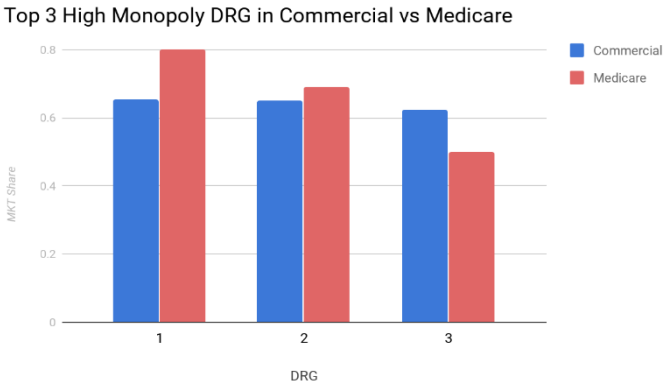
Percentages

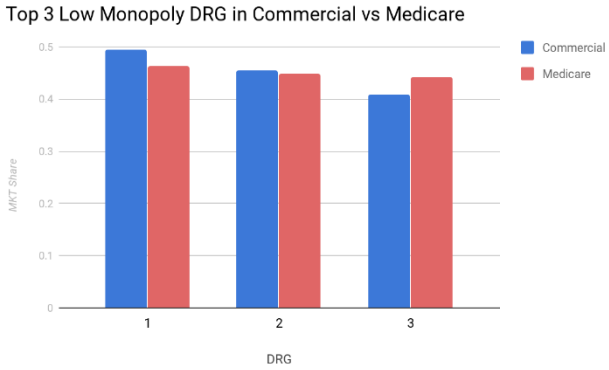
RRName	RR1--Burlington	RR2--Barre	RR3--Newport/St_Jns	RR4--Randolph	RR5--Rutland
RR1--Burlington	**96%**	2%	0%	2%	0%
RR2--Barre	12%	**83%**	0%	5%	0%
RR3--Newport/St. J	0%	6%	**94%**	0%	0%
RR4--Randolph	0%	3%	0%	**95%**	3%
RR5--Rutland	2%	0%	0%	0%	**98%**
Hospital Market Share %	**28%**	**20%**	**9%**	**21%**	**22%**

Graph III-2

As we can see from the commercial pair matrix, Burlington, home of the state's largest hospital, holds the largest share of the market at 71%; However, the second largest Randolph only takes 11%. From the Medicare pair matrix, we can see that Randolph holds the largest share of the market at 80%, the second largest Burlington takes 17%. Thus, we can conclude that Burlington has more advantage in the commercial area, while Randolph has more resources in the Medicare area.

On the other hand, through the description of 4 DRGs, we found that both 774 and 949 are high-end diseases, which means hospitals need to invest a lot of resources in these DRGs, including more professional doctors, expensive facilities, and high technology. In this case, small hospitals do not have enough budget on that. That is why these high-end diseases are monopolized by largest hospitals. Conversely, the low-end DRG 885 and 312, which would be satisfied by hospitals' basic resources, shows low monopoly. Thus, we can conclude that high or low end DRG is a factor of monopoly.
To further compare the difference between the behavior of Commercial patients and Medicare patients, we list top 3 market share of DRGs as below:





Graph III-3&4

The chart shows that in high monopoly areas, the monopoly difference between commercials is not higher than between Medicare, the reason behind that maybe the top 3 DRGs were monopolized by one hospital. However, in low monopoly areas, the monopoly difference between commercials is higher than between Medicare. Considering patients behavior, For commercial patients, they paid medical bills with their private health insurance. They would not care much about the high expenditure because of the high reimbursement of the insurance. However, Medicare patients are those who have insurance provided by the U.S. federal government health insurance program, which have the lowest standard of reimbursement, thus these patients would worry about the high expenditures.

IV. Summary

Based on our analysis, we can get 5 conclusions:
1/ Elderly people with health insurance are the main patients of Rutland Regional Medical Center. Besides, the poor people with Medicare and the people with Blue Cross are the other two main parts contributing to the whole patient of this medical center.

2/ Although government-aid insurance is the main source of payment, the role commercial insurance plays is also a very important part to help hospitals have a larger market share to turn to monopoly.

3/ For all the cases, we can see the same trend: Burlington always takes control of the regions of Vermont and there should be monopolistic power existing in this area.

4/ High or low end DRG is a factor of monopoly. High-end DRG are always monopolized by largest hospitals.

5/ Considering patients behavior, for commercial patients, they paid medical bills with their private health insurance. However, Medicare patients are those who have insurance provided by the U.S. federal government health insurance program, which have the lowest standard of reimbursement, thus these patients would worry about the high expenditures.