Healthcare Data Analytics for Parkinson's Disease Patients: A Study of Hospital Cost and Utilization in the United States

Sunanda Mukherjee, MS[1], Huanmei Wu, Phd[1, 2], Josette Jones, Phd[1]
[1]School of Informatics and Computing, IUPUI, Indianapolis, IN, USA;
[2]Department of Computer and Information Technology, IUPUI, Indianapolis, IN, USA;

Abstract

Parkinson's Disease (PD), a prevalent problem, especially for the aged populations, is a progressive but non-fatal nervous system disorder. PD patients have special motor as well as non-motor symptoms over time. There are several limitations in the study of PD such as unavailability of data, proper diagnosis and treatment methods. These limitations significantly reduce the quality of PD patient life quality, either directly or indirectly. PD also imposes great financial burdens to PD patients and their family. This project aims to analyze the most common reasons for PD patient hospitalization, review complications that occur during inpatient stays, and measure the costs associated with PD patient characteristics. Using the HCUP NIS data, comprehensive data analysis has been performed. The results are customized visualized using Tableau and other software systems. The preliminary findings sheds light into how to improve the life quality of PD patients.

Keywords

Parkinson's disease, Data Analytics, Hospitalization utilization and cost, Tableau, HCUP NIS database

1. Introduction

Parkinson's disease (PD) is a slowly progressive, chronic but non-fatal neurological movement disorder or disorder of the brain. It is the second most common neurodegenerative disease in US. It is reported that there is an average of 0.3% of PD prevalence throughout the entire population as estimated by the individual studies based on the projection of economic burden of PD[1]. Previous study also showed there was about approximately 2% and 5% increase rate in the number of patients identified with PD over age 65 and over 85, respectively. In addition, the overall prevalence and global economic burden of PD in World's top 10 densely populated nations has been investigated and it is estimated that the PD occurrence rate will be double by 2030.[2]

Although the actual cause of PD remains unclear, it has been seen that the symptoms of PD is associated with the deficiency of Dopamine, a chemical in the brain. There is no explicit tests to diagnose PD. Thus, the disease is diagnosed only based on the patient's report and the combination of exhibited symptoms. Confirmed PD is typically treated with the combinations of medications and exercises. Medicines can improve the symptoms to some extent, but the care and management is very challenging and costly for the patients and their healthcare providers.

The major symptoms of PD disease is characterized by a collection of motor as well as non-motor symptoms, which is the results from dopaminergic neuron degeneration[1, 3]. The major PD motor symptoms is clustered into an acronym called TRAP: Tremor, Rigidity, Akinesia or Bradykinesia and Postural instability. The non-motor symptoms include mood disorders, sleep disorders, depression, anxiety and decreased motivation etc. These symptoms have a huge influence to the potential adverse effects on health-related life quality of PD patients. The severity of such symptoms indicates the progression of the disease, which resulting in a higher costs of medical treatments for hospitalization and care coordination.

The major objectives of this study are to perform data analytics, to highlight the most common reasons for PD patient hospitalization, to determine hospitalization costs, and to measure the co-morbidities associated with this disease. The long term goal of this study is to lessen the symptoms of PD and hence, improve the quality of life of the patients by lowering the hospitalization costs and inpatient hospital stays.

2. Materials and Methods

2.1 Data Source:

A retrospective analysis was performed using the Healthcare Cost and Utilization Project's (HCUP) Nationwide Inpatient Sample (NIS)[4] data for the year 2012. The Agency for Healthcare Research and Quality (AHRQ) sponsors

the HCUP databases that are specifically designed to determine and identify patterns in hospital utilization and cost across the United States. The HCUP-NIS database is the largest inpatient database available in the US, which represents a sample of non-federal US community hospitals. Just in 2012, there are 1,049 hospitals from 46 states in US participated in NIS projects. More than 7 million records of hospital stays were reported by these hospitals. It is estimated that the sampling includes over 95 percent of discharges from the hospitals participated in NIS. The estimated samples were weighted in order to minimize the margin of error and to reflect all 50 states across the US. The large sample size of the database enabled us to analyze rare conditions, uncommon treatments and special patient populations[4].

To protect the privacy of individual patients, physicians, and hospitals, the state and hospital identifiers are de-identified. There are many clinical and non-clinical hospitalization data elements recorded in the HCUP NIS database[4]. Sample non-clinical information are patient demographic information, payment related information, hospital characteristics (such as region, census division, and location), and total charges. Sample clinical related information include principal and secondary diagnosis and procedure (includes both ICD-9 and CCS codes), discharge status and Length of stay.

2.2. Variables of Interest:

At first, based on the ICD-9-CM diagnosis codes, the individuals who were primarily diagnosed as PD patients at the time of admission had been identified. In HCUP databases, more than 14,000 ICD-9-CM diagnosis codes and 3,900 procedure codes had been mentioned which were further classified and clustered into lesser number of clinically appropriate categories by the AHRQ's Clinical Classification Software (CCS). This enabled the database to capture a large population of relatively similar conditions into a single group by making the information more useful for performing statistical analyses and developing reports.[5] PD patients were identified using CCS diagnosis code as 79.

One of the dependent or outcome variable of interest was the total hospitalization cost resulting from the entire hospital stay. The total hospital charges counted all the expenses for services provided to the patients excluding physician or professional fees.[6]

For the analysis, among the predictor variables, the primary and secondary disease conditions and procedures and length of the hospital stays were important. The preliminary diagnoses had been identified at the time of admission whereas the comorbidities or the secondary disease conditions were recorded throughout the entire hospital stay. The similar manner has been followed for the determination of the primary and secondary procedures. The length of stay was the time period from the admission to the discharge of latest hospital stay of each patient.

2.3. Approaches:

Retrospective analysis was performed over the HCUP-NIS database focusing on the determination of the hospital utilization cost for PD patients. Descriptive statistics were used to summarize the results. The mean and standard deviation was used to explain the continuous variables. On the other hand, the categorical variables were shown in percentage values.

Statistical approaches have been used to determine the correlation of clinical features including diagnosis, current medical conditions and costs. For the Univariate analysis, the two major characteristics of the variable analyzed were the distribution and the central tendency.

Data were analyzed using R language[7] and Tableau software. The Tableau software provides a great user interface and visualization tools for the analytical results.

3. Results & Findings

Extensive analysis were performed for the 2439 PD patient hospitalization records. The preliminary results are presented below.

3.1 Demographics:

Age Distribution: The patients were distributed according to their age group as shown in the (Figure 1). The result shows that a total of more than 99% of patients were in the range of 50 to 90 age group where approximately 26%, 34% and 25% patients were in 61-70, 71-80, and 81-90 age groups, respectively. Although, in earlier studies,[8,9,10,11] there is no significant association has been found between the aging process and the acceleration of the disease, still it is a debatable fact whether the aging influences the progression of the disease or not.

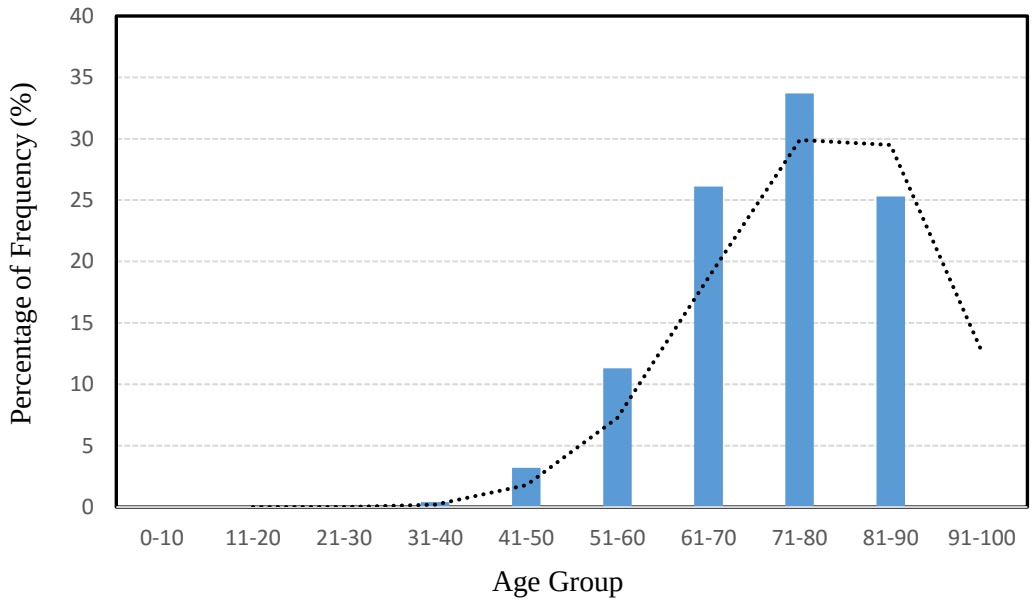

Figure 1. The distribution of PD patients by the age group.

Patients across the census division: The data is geographically divided into 9 divisions from 4 regions. Figure 2 below shows all the states in each of the 4 US Census Bureau Regions and 9 divisions. The states highlighted in grey color do not participate in HCUP.

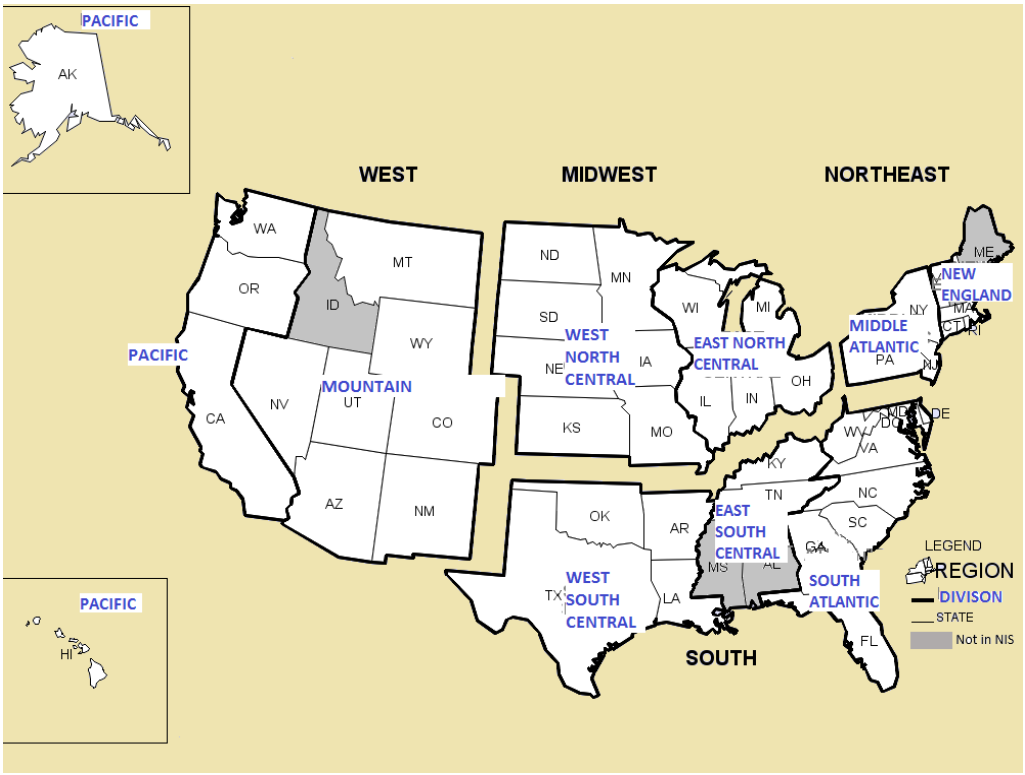

Figure 2. The map depicting the 9 census divisions and 4 regions in US.

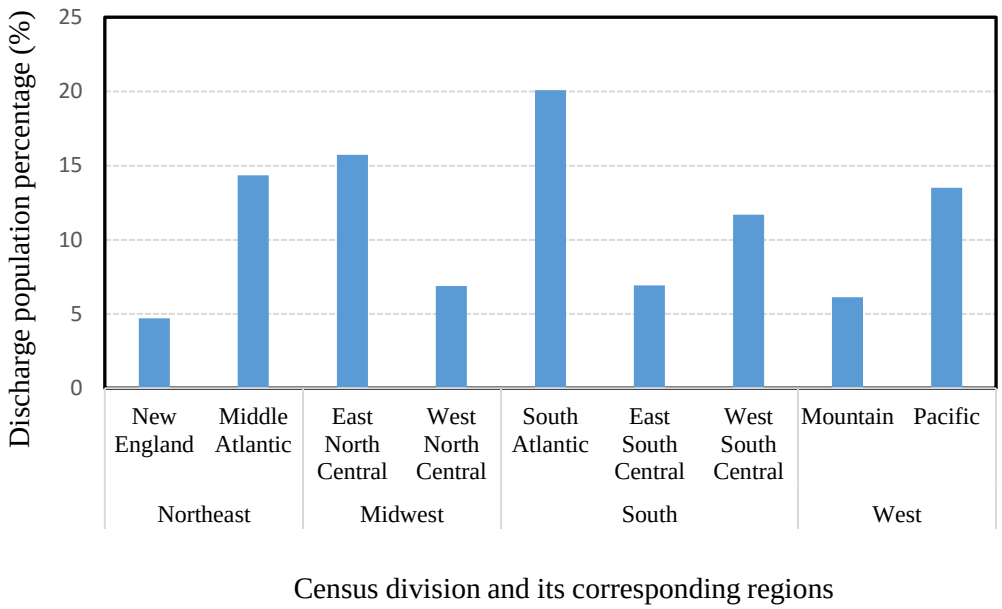

Figure 3. The percentage of PD patients across the census division and the region in US.

For the four geographical divisions, the Middle Atlantic division in the Northeast Region contributes to the highest number of PD patient sample (20% of total 2439 patients) whereas the New England division from the same region contributes to the lowest number of PD patient sample (5% of total 2439 patients) as described by the (Figure 2). Next to the Middle Atlantic division, the South Atlantic division from South region, East North Central division from Midwest region and Pacific division from the West region contributes to the more number of inpatient samples.

3.2. Diagnosis Related Group (DRGs):

DRG is a system of classification that has been developed with the intent to identify the presence of complications or comorbidities and to replace "cost based" reimbursement that had been used up to that point by the hospitals. In contrast, according to the literature by Muller et al.[12] published in 2010, a significant decrease in the reimbursement of treatment costs has been noticed in spite of the presence of increased operation risk and medical conditions that in turn results in an increased economic burden for the PD patients.

The DRGs, identified for the PD patients, included craniotomy and endovascular intracranial procedures, spinal procedures, extracranial procedures, peripheral/ cranial nervous system procedures and degenerative nervous system disorders. The DRGs also indicated the severity of the disease as it has identified the patients with or without minor or major complications and co-morbidities.

The top 6 DRGs having the percentage of the total number of PD patients in each group is shown in (Figure 4). The description of these DRGs are -

- *Degenerative Nervous System Disorders W/O MCC (57),*
- *Craniotomy & Endovascular Intracranial Procedures W/O CC/MCC (27),*
- *Degenerative Nervous System Disorders W MCC (56),*
- *Cranio W Major Dev Impl/Acute Complex CNS PDX W/O MCC (24),*
- *Craniotomy & Endovascular Intracranial Procedures W CC (26)*
- *Periph/Cranial Nerve & Other Nerv Syst Proc W/O CC/MCC (42)*

where W CC, W MCC, W/O MCC and W/O CC/MCC corresponds to With Complications or Co-morbidities, With Major Complications or Co-morbidities, Without Major Complications or Co-morbidities, Without Complications or Co-morbidities/ Major Complications or Co-morbidities respectively.

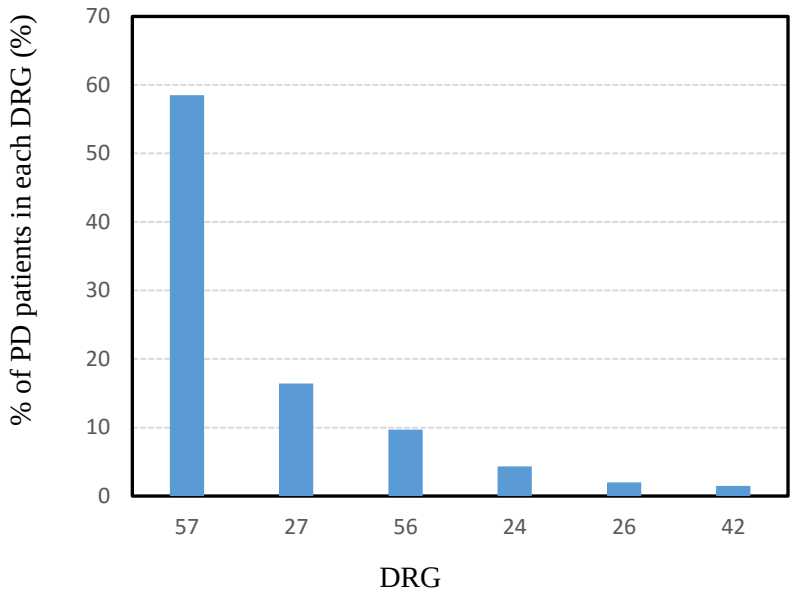

Figure 4. The percentage of PD patients in top 6 disease related groups (DRGs)

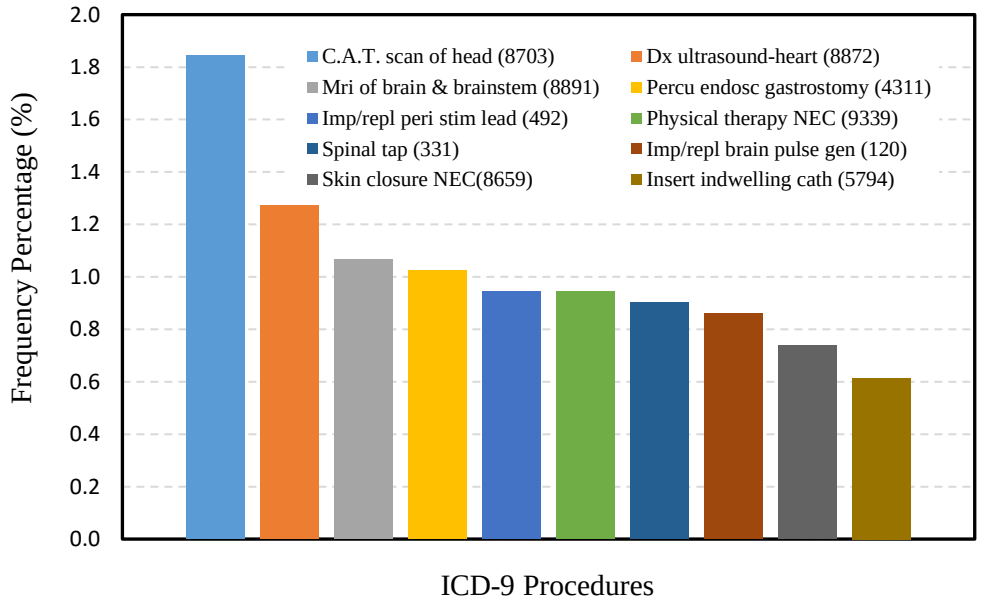

Figure 5. Frequently used 10 principal procedures for the PD patients

3.3. Frequently Used Procedures:

The study has identified total top 11 primary procedures that has been frequently used for the PD patients. The Figure 5 shows 10 primary procedures having percentage frequency in between 0.5% to 2% and excludes the procedure with ICD-9 code as 293 having the highest frequency of 20%. The Table 1 listed the top 10 primary procedures based on

their average cost per day. Considering both the Figure 5 and Table 1, the study found it interesting that the procedures having higher average cost per day has been frequently used for the PD patients. This analysis has also listed 3 interesting procedures based on their cost and usage frequency:

1) *The procedure Imp/repl brain stim lead (ICD code 293):* It having the highest frequency, is used for 20% of the total patients. The description of the procedure is known as Implantation or replacement of intracranial neurostimulator lead(s). This is applied to the implantation, insertion, placement, or replacement of intracranial brain pacemaker [neuropacemaker], depth electrodes, foramen ovale electrodes and intracranial electrostimulator.
2) *Implantation or replacement of peripheral neurostimulator lead(s) (ICD code 492):* This procedure having the highest average cost per day is also used frequently for the patients.
3) *The cranial implantation or replacement of neurostimulator pulse generator (ICD code 120):* It is a procedure which is also found in the top 10 list, and is followed along with the above procedure (ICD code 293).

Table 1. The top 10 procedure codes, description of procedure, average cost/day and standard deviation

Procedure 1	Description	Average cost/day	Standard Dev
492	Imp/repl peri stim lead	64636.29	30449.76
8695	Ins/re pls gn no rechrg	57862.33	23096.03
293	Imp/repl brain stim lead	50798.37	27336.41
120	Imp/repl brain pulse gen	39664.58	2474.76
8841	Contr cerebr arteriogram	19949.33	17334.49
8659	Skin closure NEC	13138.78	10156.98
8919	Video/radio eeg monitor	9873.53	4955.67
9339	Physical therapy NEC	9572.88	660.61
8944	Cardiac stress test NEC	9269.11	3877.17
4525	Clos large bowel biopsy	8925.88	7322.97

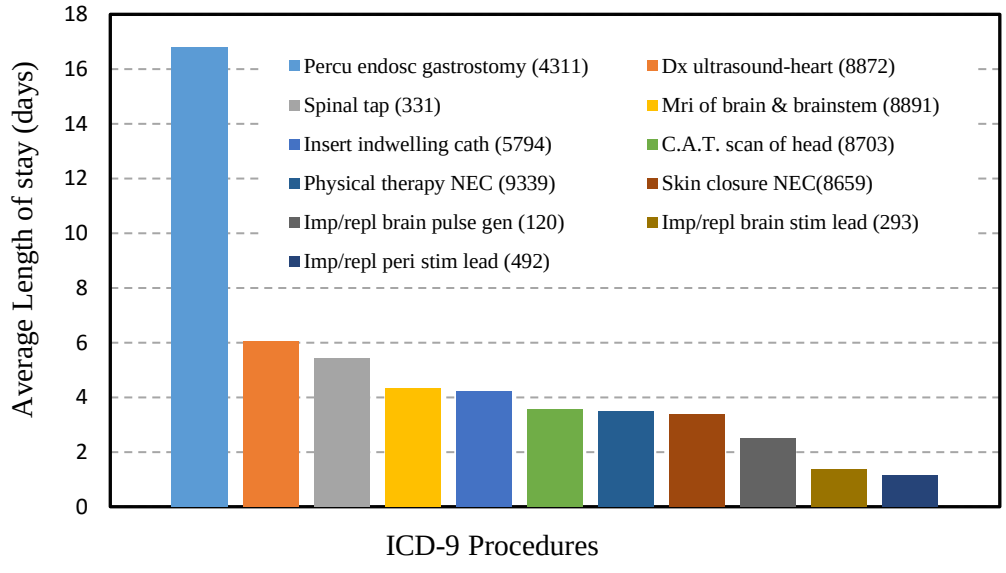

Figure 6. The length of hospital stays for PD patients

3.4 Length of Patient Stay:

The average length of stay for the top 10 primary procedures was approximately 4.8 days, with a range of 1-16 days. Qualitatively, the PD patients treated with the procedure called percutaneous endoscopic gastronomy had the longest stays whereas the patients treated with insertion of catheter had the shortest stays as described by the (Figure 6).

3.5. Average cost per procedure:

The study calculated the average cost for each of the primary procedures. The below diagram shows the plots for top 15 primary procedures and each boxplot represented one primary procedure as shown in (Figure 7). The plots were sorted based on the mean value of cost per day in an ascending order.

The average cost for all the top 10 procedure was approximately $28400 (range ~ 9000-65000) with the standard deviation as approximately $12800 (range ~ 660-30450). The below table (Table 1) shows the average cost and standard deviation for the top 10 primary procedures.

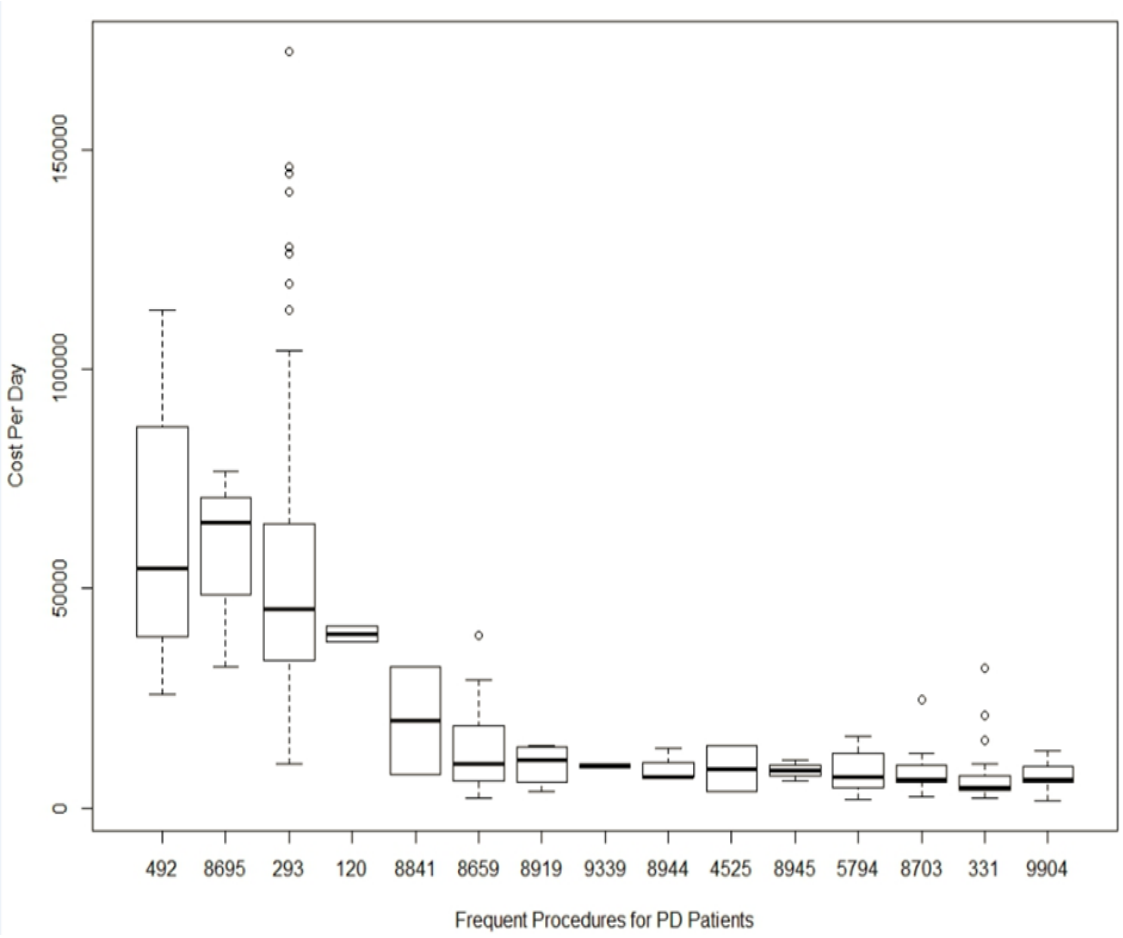

Figure 7. The average hospitalization cost for top 15 frequent procedures

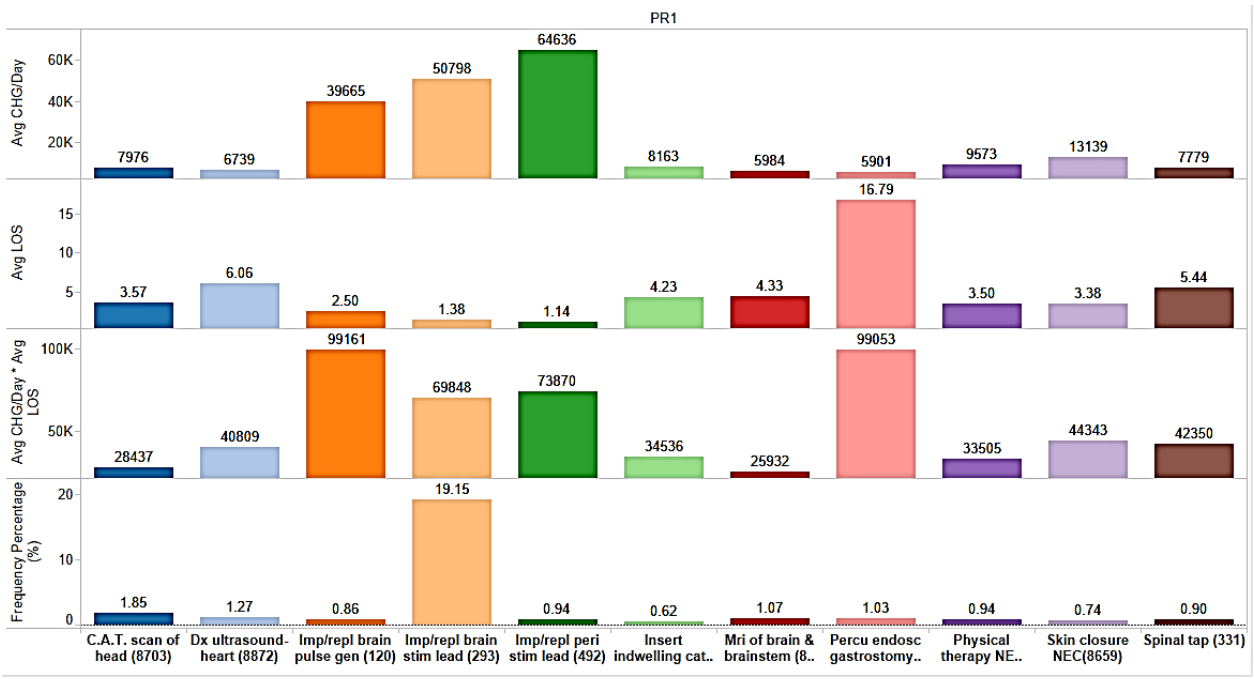

Figure 8. The comparison of average hospitalization cost, average hospital stays and frequent procedures

When the study compared the procedures by taking average cost, average length of stay and frequency into consideration, it has identified the below procedures for which the average cost, length of stay and the frequency was higher as described by the (Figure 8) -

- Imp/repl brain pulse gen (120)
- Imp/repl brain stim lead (293)
- Imp/repl peri stim lead (492)
- Percu endosc gastrostomy (4311)

4. Discussion:

This analysis of population-based hospital discharge data from patients with PD reveals variations in hospital costs by patient demographic characteristics. Average hospital cost was significantly higher for hospitalizations. Several important observations were generated. A major component of the overall economic burden of healthcare is the initial cost. The study identified the procedures, having higher cost and length of stay, was related to the Deep Brain Stimulation (DBS) which is an important treatment option for PD patients experiencing dyskinesia and motor complications. Literature related to DBS showed that the most common reasons for prolonged hospitalization after DBS are mental status changes and cerebral hemorrhage.[13,14]

Although the study illustrated the large economic burden of PD in terms of medical costs, still it couldn't able to demonstrate the true burden of the disease. The presence of comorbidities and its treatment procedures had been significantly reduced the quality of life for people with PD and their family members. It has been further projected that the pressure of such chronic conditions will substantially increase in next few years with the increase of size of the aged populations.[15,16] Such projections demonstrate the need for the innovation of old techniques and identification of new treatment procedures including vigorous exercises in order to prevent or delay the onset of the disease; hence, increasing the quality of life for the PD patients.

The present study had several limitations common to administrative databases. It was susceptible to errors for categorizing patients and calculating costs. The hospital charges analysis does not constitute the hospital-specific

differences in charge practices. There may be variation in cost-to-charge ratio (CCR) by patient population and by individual hospital departments that could have resulted in errors in calculating costs. Thus, it is difficult to develop comprehensive framework of current payment models focusing on cost structure on PD patients because up-to-the minute charge data are not available, and charge practices may be lost in the heterogeneity of hospital type.

However, based on these data, a predictive model of cost can be built in future so that the model can be additionally utilized in the cost management debate and the creation of data-driven policies. This can fuel further studies in the field and provide elements for the design of prospective investigations.

5. Conclusion

This study has successfully identified the most common reasons for hospitalization in PD, determined costs associated with nationwide hospitalizations of individuals with PD and measured the co-morbidities associated with the disease using the nationwide discharge sample data across the US. Future work, incorporating the development of a prediction model of hospitalization utilization cost can fuel the further studies in the field of healthcare.

References

1. Kowal SL, Dall TM, Chakrabarti R, Storm MV, Jain A. The current and projected economic burden of Parkinson's disease in the United States. Mov Disord. 2013;28(3):311-8.
2. Dorsey ER, Constantinescu R, Thompson JP, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology. 2007;68(5):384-6.
3. J Jankovic. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry 2008;79:368-376
4. HCUP Nationwide Inpatient Sample (NIS). Healthcare Cost and Utilization Project(HCUP). 2006–2011. Agency for healthcare research and quality R, MD. http://www.hcup-us.ahrq.gov/nisoverview.jsp
5. HCUP Clinical Classifications Software for ICD-9-CM (M). Healthcare Cost and Utilization Project (HCUP). 2006–2011. Agency for healthcare research and quality R, MD. http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp
6. Simeone RM, Oster ME, Cassell CH, Armour BS, Gray DT, Honein MA. Pediatric Inpatient Hospital Resource Use for Congenital Heart Defects. Clinical and molecular teratology. 2014;100(12):934-943.
7. The R Project for Statistical Computing. https://www.r-project.org/
8. Critchley M. The neurology of old age. Lancet.1931;1:1221-1230.
9. Agid YBlin, JBonnet AM et al. Does aging contribute to aggravation of Parkinson's disease? Raven Press. 1989;115-123.
10. Blin, JDubois, BBonnet, AMVidailhet, MBrandabur, MAgid Y. Does ageing aggravate parkinsonian disability? J Neurol Neurosurg Psychiatry.1991;54:780-782.
11. Wolters, ECCalne DB. Is Parkinson's disease related to aging? Raven Press. 1989;125-132.
12. M Müller, G Hemicker, U Jüptner, et al. Inadequate reimbursement for surgical patients with Parkinson's disease according to diagnosis-related grouping. ZENTRALBLATT FÜR CHIRURGIE. 2010;136(2):168-71
13. Chou KL, Zamudio J, Schmidt P, et al. Hospitalization in Parkinson disease: A survey of National Parkinson Foundation Centers. Parkinsonism & related disorders. 2011;17(6):440-445.
14. Aminoff MJ, Christine CW, Friedman JH, et al. Management of the hospitalized patient with Parkinson's disease: Current state of the field and need for guidelines. Parkinsonism & related disorders. 2011;17(3):139-145.
15. Missios S, Bekelis K. Drivers of hospitalization cost after craniotomy for tumor resection: creation and validation of a predictive model. BMC Health Services Research. 2015;15:85.
16. Nicole C. R. McLaughlin, Irene Piryatinsky, Gary Epstein-Lubow, Louis Marino, and Joseph H. Friedman. Neuropsychiatric Symptoms in an Inpatient Parkinson's Disease Sample. Parkinson's disease 2014; Article ID 420240, 5 pages.