

Insurance Claims Data & Analytics

**Group 1 | Huijue Ye | Jiaying Yan | Jie Tang | Ying Tung Lau | Zhaoyi Wang | Zoey Yang**



November 18, 2019

Brandeis university

**Patient vignettes**

**Patient with UNIQ 40436**

The patient (Uniq ID = 40436) is a 70-74 years old lady with urgent heart and circulatory problem and was transferred from a hotel. She stayed in University of Vermont Medical Center for 1 day and was discharged home afterwards. Her charge by the hospital was $70K and her insurance was covered by Medicare. She was diagnosed with Perc cardiovasc proc w drug-eluting stent w MCC or 4+ vessels/stents. She received major services including cardiology (costs $23275.36), implants (costs $10146.48), cardiac catheter lab (costs $8387.29), CT scan for body (costs $4112.29) and clinical diagnostic (costs $2226.59).

**Patient with UNIQ 507033**

The patient (Uniq ID = 507033) is a 25-29 years old pregnant lady from non-health care facility point of origin. She stayed in Northwestern Medical Center for 1 day and was discharged to another faculty afterwards. Her charge by the hospital was $3.2K and her insurance was covered by BLUE CROSS. She was diagnosed with Vaginal delivery w/o complicating diagnoses. She received major services including labor (costs $1273.76), room and board (semi-private 2 beds) (costs $1002.13), clinical diagnostic (costs $341), sterile supplies (costs $334.33) and pharmacy: IV solutions (costs $92.1).

**Patient with UNIQ 859382**

This patient (Uniq ID = 859382) is from emergency department. He is a 30-34 years old man with emergent injury or toxic effect from non-health care facility point of origin. He stayed in Rutland Regional Medical Center for 1 day but he was against to medical advice afterwards. His charge by the hospital was $13K and he paid by himself. He was diagnosed with Poisoning & toxic effects of drugs w MCC. He received major services including intensive care (costs $4450), clinical diagnostic (costs $3339.63), respiratory services (costs $2575.58), emergency room (costs $1227) and pharmacy (costs $923.99).

**Patient with UNIQ 1585831**

| **REVCODE** | **REVCHRGS** | **REVUNITS** |
| --- | --- | --- |
| 200 intensive care | 4450 | 1 |
| 250 Pharmacy | 2024.81 | 1505 |
| 258 Pharmacy: IV solutions | 182.57 | 10 |
| 259 Pharmacy: Other | 248.43 | 1 |
| 270 Medical/Surgical Supplies | 51.77 | 1 |
| 300 Laboratory - Clinical Diagnostic | 2392.11 | 23 |
| 320 Radiology - Diagnostic | 656 | 2 |
| 351 CT Scan: Head | 1648 | 1 |
| 410 Respiratory Services | 3723.1 | 3 |
| 450 Emergency Room | 1227 | 1 |
| 730 EKG/ECG | 490 | 2 |

**Table 1** Service received by patient with UNIQ 1585831

The patient (with UNIQ 1585831) was served in Rutland Regional Medical Center. This patient was admitted from NON-HEALTH CARE FACILITY POINT OF ORIGIN which includes patients coming from home, a physician’s office, or workplace. The patient was admitted by emergency, meaning that the patient required immediate medical intervention as a result of severe, life threatening or potentially disabling conditions. Generally, the patient was admitted through the emergency room. The patient’s age is between 40-44 years old, and gender of this patient is female. She was under coverage of Medicaid, and she was diagnosed as Poisoning & toxic effects of drugs w MCC, which belongs to category of Injury and Toxic effects. The length of stay in hospital was just 1 day. The discharge status for this patient is DIED. Therefore, we can infer that this patient has died in hospital the second day she came to stay in hospital. The total charges for this patient were 17093.79 dollars. From the revenue code in **Table 1**, we can clearly see that intensive care, respiratory services, clinical diagnostic and pharmacy cost the most, which charged the patient respectively 4450, 3723.1, 2393.11 and 2024.81 dollars. Therefore, we can infer this patient was sent to emergency room from other places, doctors have rapidly discussed the solution after CT scan and gave the patient respiratory services and intensive care. However, all rescue measures proved ineffectual, the patient died the next day.

**Patient with UNIQ 200760**

|  |  |  |
| --- | --- | --- |
| **REVCODE** | **REVCHRGS** | **REVUNITS** |
| 120 Room & Board (Semi-Private 2 beds) | 6768 | 4 |
| 250 Pharmacy | 1387.36 | 321 |
| 272 Medical/Surgical Supplies: Sterile supplies | 409.96 | 12 |
| 278 Medical/Surgical Supplies: Other implants | 10696.77 | 14 |
| 300 Laboratory - Clinical Diagnostic | 610.08 | 17 |
| 320 Radiology - Diagnostic | 833.28 | 2 |
| 360 Operating Room Services | 14055.8 | 220 |
| 370 Anesthesia | 2590.07 | 227 |
| 420 Physical Therapy | 138.63 | 1 |
| 424 Physical Therapy: Evaluation/re-evaluation | 350.34 | 1 |
| 450 Emergency Room | 2582.43 | 4 |

**Table 2** Service received by patient with UNIQ 1585831

The patient (with UNIQ 200760) was served in University of Vermont Medical Center. He was admitted from NON-HEALTH CARE FACILITY POINT OF ORIGIN which includes patients coming from home, a physician’s office, or workplace. This patient’s admission type is emergency, meaning that the patient required immediate medical intervention as a result of severe, life threatening or potentially disabling conditions. Generally, the patient was admitted through the emergency room. The age of patient is between 18-24 years old, and this patient is a female. She is covered under COMMERCIAL INSURANCE. The patient was diagnosed as Lower Extrem & Humer Proc Except Hip, Foot, Femur W/O Cc/Mcc, which belongs to Musculoskeletal. The total length of her stay in hospital is 4 days, during diagnosis she was charged 49533.15 dollars in total. After he was discharged from hospital, she has used Home Health service to maintain her status. It probably because she was not absolutely cured in hospital, and she has much work to do (since she is under commercial insurance coverage). Therefore, she chose to use home health service instead. From **Table 2**, we can clearly see that the two most costly diagnoses are medical/surgical supplies which have cost 10696.77 dollars and services in operating room which were charged 14055.8 dollars in total. Other than that, Room & Board costs her 6768 dollars, since she lived in semi-private 2 beds during these 4 days.

**Patient with UNIQ 3692**

|  |  |  |
| --- | --- | --- |
| **REVCODE** | **REVCHRGS** | **REVUNITS** |
| 124 Psychiatric | 106662 | 58 |
| 250 Pharmacy | 5608.26 | 843 |
| 300 Laboratory - Clinical Diagnostic | 1716.92 | 31 |
| 320 Radiology - Diagnostic | 1873.08 | 4 |
| 450 Emergency Room | 1983.03 | 1 |
| 730 EKG/ECG | 52 | 1 |

**Table 3** Service received by patient with UNIQ 1585831

The patient (with UNIQ 3692) was served in University of Vermont Medical Center. He was admitted from NON-HEALTH CARE FACILITY POINT OF ORIGIN which includes patients coming from home, a physician’s office, or workplace. This patient’s admission type was urgent, meaning that this patient is required immediate attention for the care and treatment of a physical or mental disorder. Generally, the patient needed to be admitted to the first available and accommodation. The patient’s age is between 18-24, and his address is in Milton. He was covered under BLUE CROSS insurance, and he was diagnosed as Psychoses, which belongs to mental illness. The total length of his stay in hospital is 58 days, during which he was charged 117895.29 dollars in total. After he was discharged from hospital, this patient has been taken care of his family at home. It probably means that his mental disorder is not too damaged to be mandated to stay in hospital. From **Table 3**, we can clearly see the big picture of patient diagnose procedure by his revenue code. He was sent to psychiatric diagnose 58 times, which charges him 106662 dollars. He has bought 843 units of drugs in pharmacy. Also, he has done clinical diagnose 31 times, charging him 1716.92. Finally, Emergency room and Electrocardiography charged him 1983.03 dollars and 53 dollars respectively. From the above information, we can infer this patient has suffered mental disorder for a long time, and even was sent to emergency room once. We need to keep eyes on this patient.

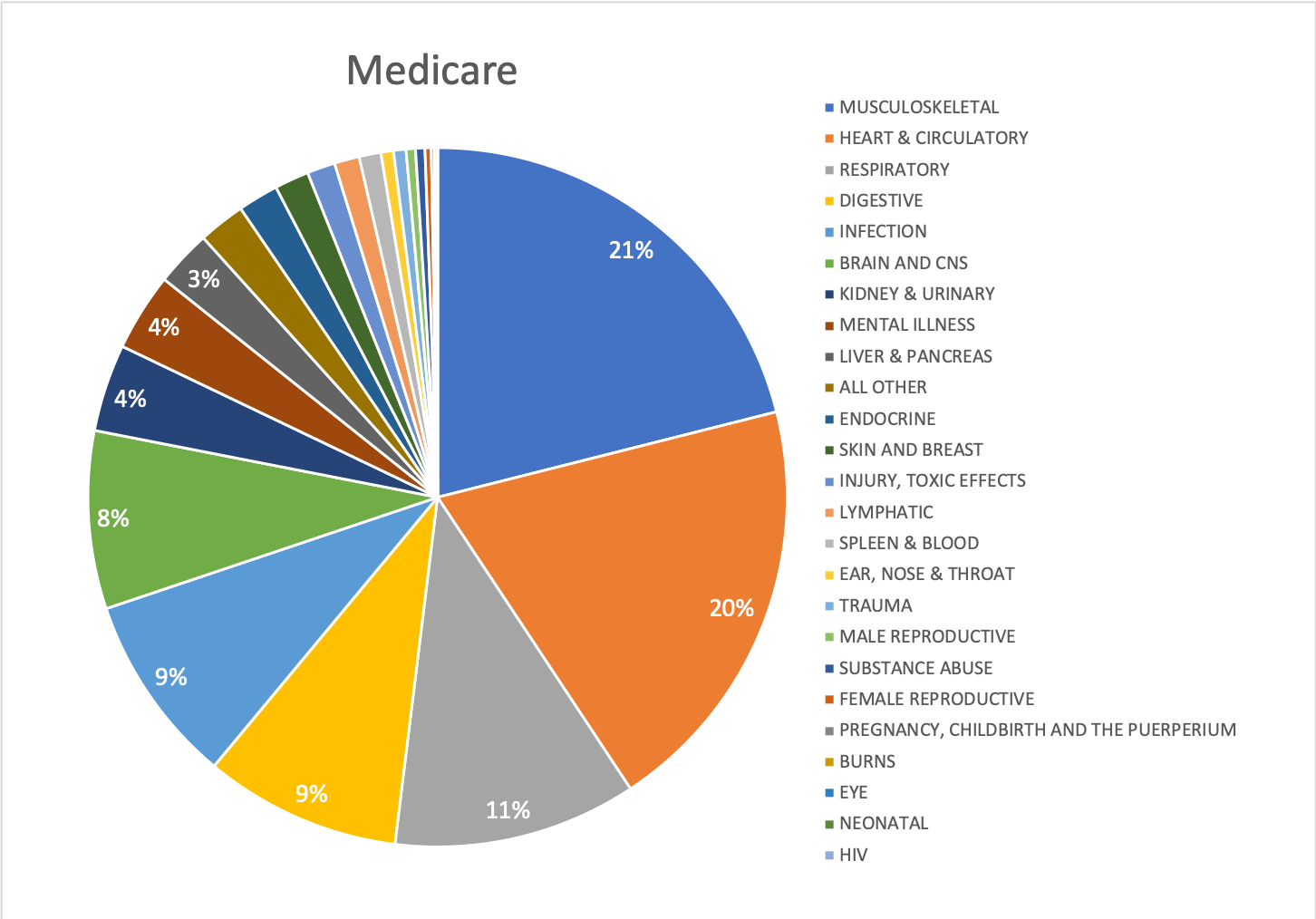
**Patient with UNIQ 690326**

The patient (Uniq ID = 690326 ) is a 40-44 years old lady with skin and breast problem from non-health care facility point of origin. She stayed in University of Vermont Medical Center for 3 days and was discharged to another faculty afterwards. Her charge by the hospital was $43K and she paid by herself. She was diagnosed with Other skin, subcut tiss & breast proc w/o CC/MCC. She received major services pulmonary function (costs $25619.89), room and board (semi-private 2 beds) (costs $5076), anesthesia (costs $4609.64), CT Scan for body (costs $3273.01) and pharmacy (costs $1351.25).

**Service and Cost Profile of Major Insurances**

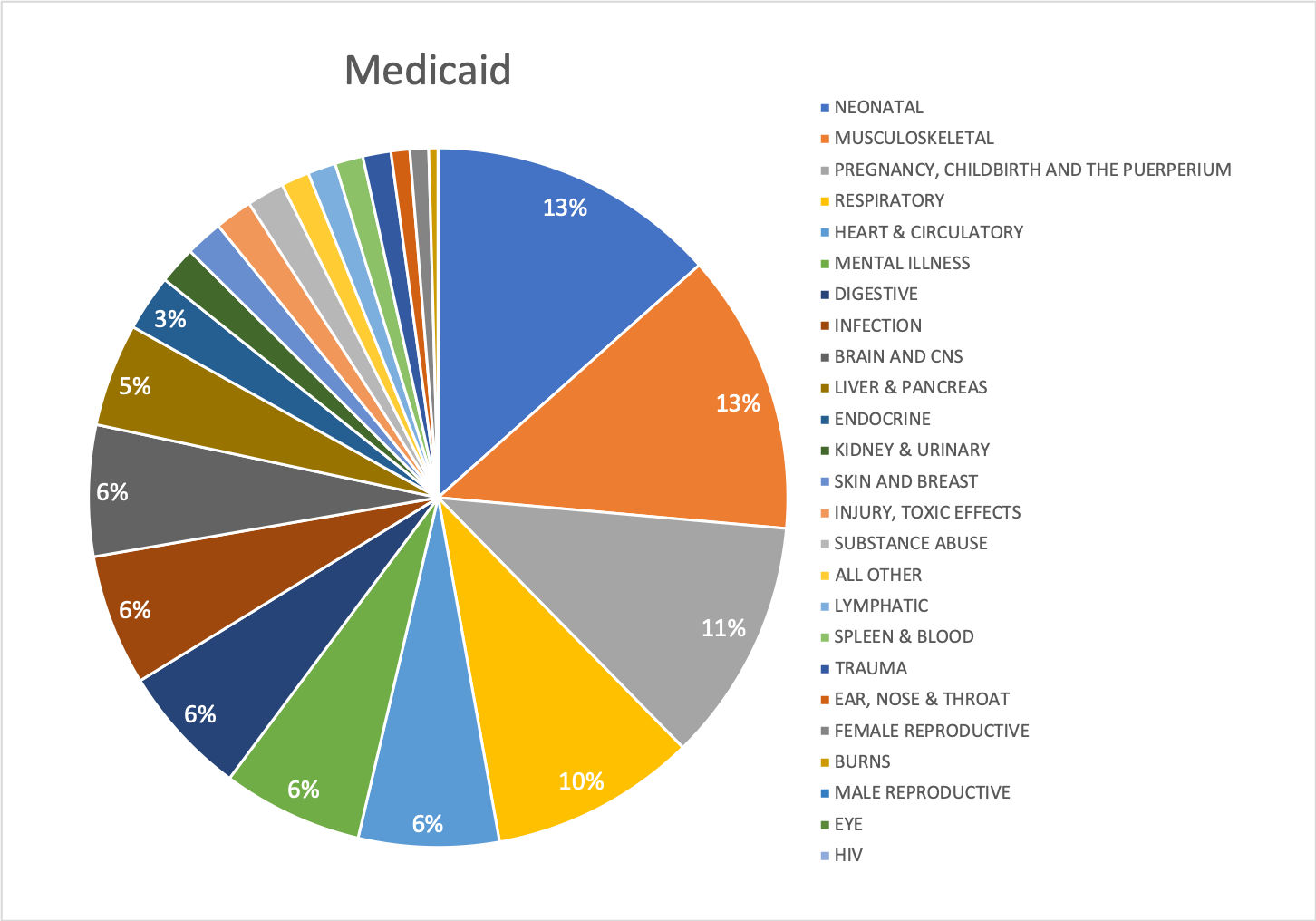
| **MDC Name** | **Medicare** | **Medicaid** | **Commercial Payers** |
| --- | --- | --- | --- |
| BRAIN AND CNS | 57 | 14 | 38 |
| EYE | 0 | 0 | 0 |
| EAR, NOSE & THROAT | 4 | 2 | 3 |
| RESPIRATORY | 78 | 22 | 24 |
| HEART & CIRCULATORY | 136 | 15 | 54 |
| DIGESTIVE | 63 | 14 | 32 |
| LIVER & PANCREAS | 18 | 11 | 12 |
| MUSCULOSKELETAL | 146 | 30 | 88 |
| SKIN AND BREAST | 11 | 4 | 6 |
| ENDOCRINE | 13 | 6 | 9 |
| KIDNEY & URINARY | 28 | 4 | 8 |
| MALE REPRODUCTIVE | 3 | 0 | 2 |
| FEMALE REPRODUCTIVE | 2 | 2 | 4 |
| PREGNANCY, CHILDBIRTH AND THE PUERPERIUM | 1 | 26 | 34 |
| NEONATAL | 0 | 31 | 28 |
| SPLEEN & BLOOD | 7 | 3 | 5 |
| LYMPHATIC | 8 | 3 | 6 |
| INFECTION | 61 | 14 | 18 |
| MENTAL ILLNESS | 25 | 15 | 10 |
| SUBSTANCE ABUSE | 3 | 4 | 1 |
| INJURY, TOXIC EFFECTS | 9 | 4 | 4 |
| BURNS | 1 | 1 | 0 |
| ALL OTHER | 15 | 3 | 6 |
| TRAUMA | 4 | 3 | 8 |
| HIV | 0 | 0 | 0 |
| BRAIN AND CNS | 0.08 | 0.06 | 0.10 |
| EYE | 0.00 | 0.00 | 0.00 |
| EAR, NOSE & THROAT | 0.01 | 0.01 | 0.01 |
| RESPIRATORY | 0.11 | 0.10 | 0.06 |
| HEART & CIRCULATORY | 0.20 | 0.06 | 0.14 |
| DIGESTIVE | 0.09 | 0.06 | 0.08 |
| LIVER & PANCREAS | 0.03 | 0.05 | 0.03 |
| MUSCULOSKELETAL | 0.21 | 0.13 | 0.22 |
| SKIN AND BREAST | 0.02 | 0.02 | 0.02 |
| ENDOCRINE | 0.02 | 0.03 | 0.02 |
| KIDNEY & URINARY | 0.04 | 0.02 | 0.02 |
| MALE REPRODUCTIVE | 0.00 | 0.00 | 0.01 |
| FEMALE REPRODUCTIVE | 0.00 | 0.01 | 0.01 |
| PREGNANCY, CHILDBIRTH AND THE PUERPERIUM | 0.00 | 0.11 | 0.09 |
| NEONATAL | 0.00 | 0.13 | 0.07 |
| SPLEEN & BLOOD | 0.01 | 0.01 | 0.01 |
| LYMPHATIC | 0.01 | 0.01 | 0.02 |
| INFECTION | 0.09 | 0.06 | 0.05 |
| MENTAL ILLNESS | 0.04 | 0.06 | 0.03 |
| SUBSTANCE ABUSE | 0.00 | 0.02 | 0.00 |
| INJURY, TOXIC EFFECTS | 0.01 | 0.02 | 0.01 |
| BURNS | 0.00 | 0.00 | 0.00 |
| ALL OTHER | 0.02 | 0.01 | 0.02 |
| TRAUMA | 0.01 | 0.01 | 0.02 |
| HIV | 0.00 | 0.00 | 0.00 |

**Table 4** The sum of the dollar value (million dollars) of the charges for the MDCs for Medicare, Medicaid and Commercial Payers



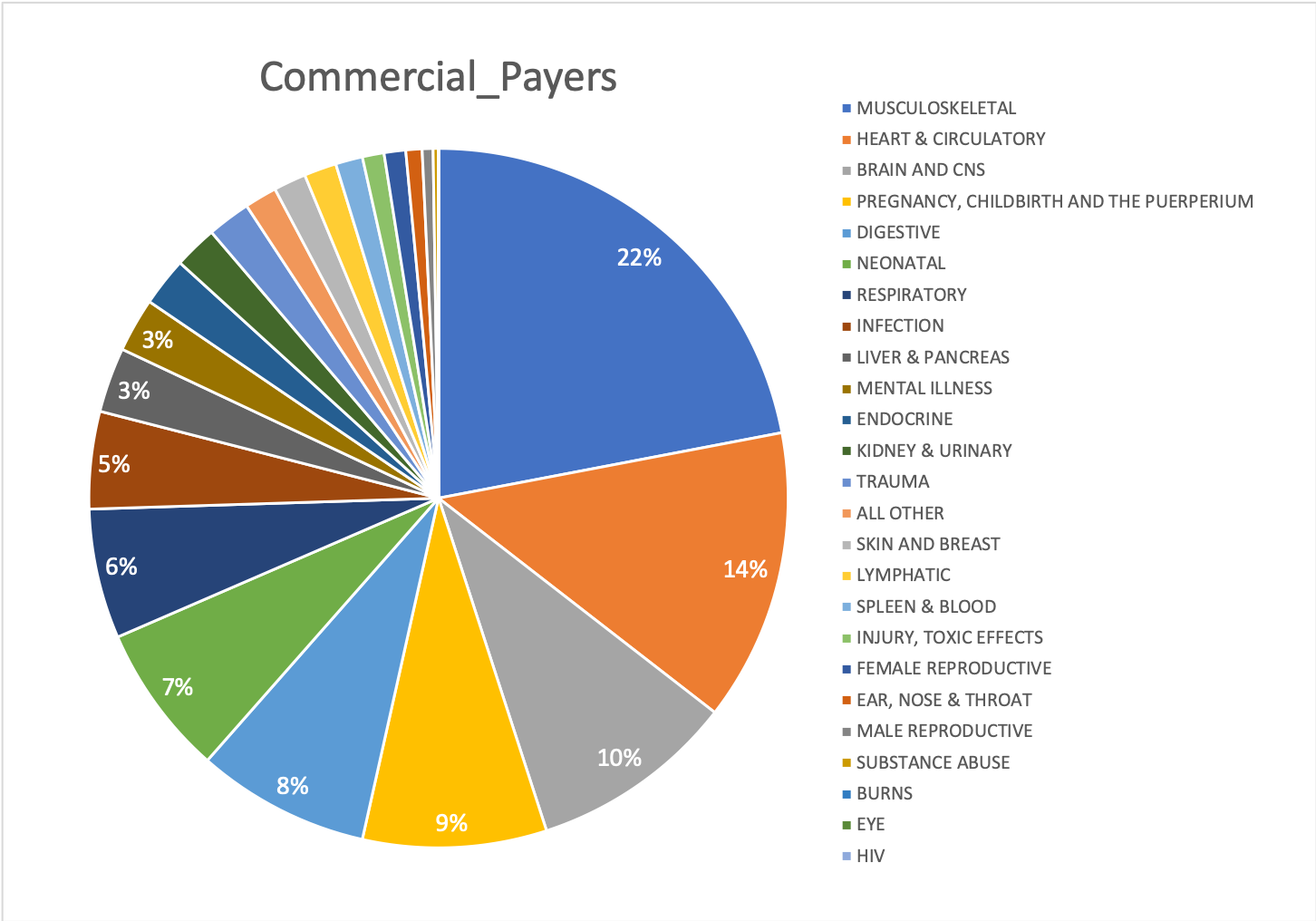
**Chart 1** The inpatient services portfolio for Medicare

From the pie chart of Medicare (**Chart 1**), we are observing that the top-5 MDCs are Musculoskeletal, Heart & Circulatory, Respiratory, Digestive and Infection, and these five MDCs even take up 70% of Medicare’s charges. Because Medicare mainly provides health insurance for citizens aged 65 or older, and these diseases are more common among the elderly. This reason can also be justified by the distribution of patients’ age group (**Chart 4**). According to the chart, we can find that more than 50% of the patients covered by Medicare are 75 or older (age group 14), and about 80% of the patients are aged 65 or older (age group 12-14). Both female and male patients show the similar situation in **Chart 5**.



**Chart 2** The inpatient services portfolio for Medicaid

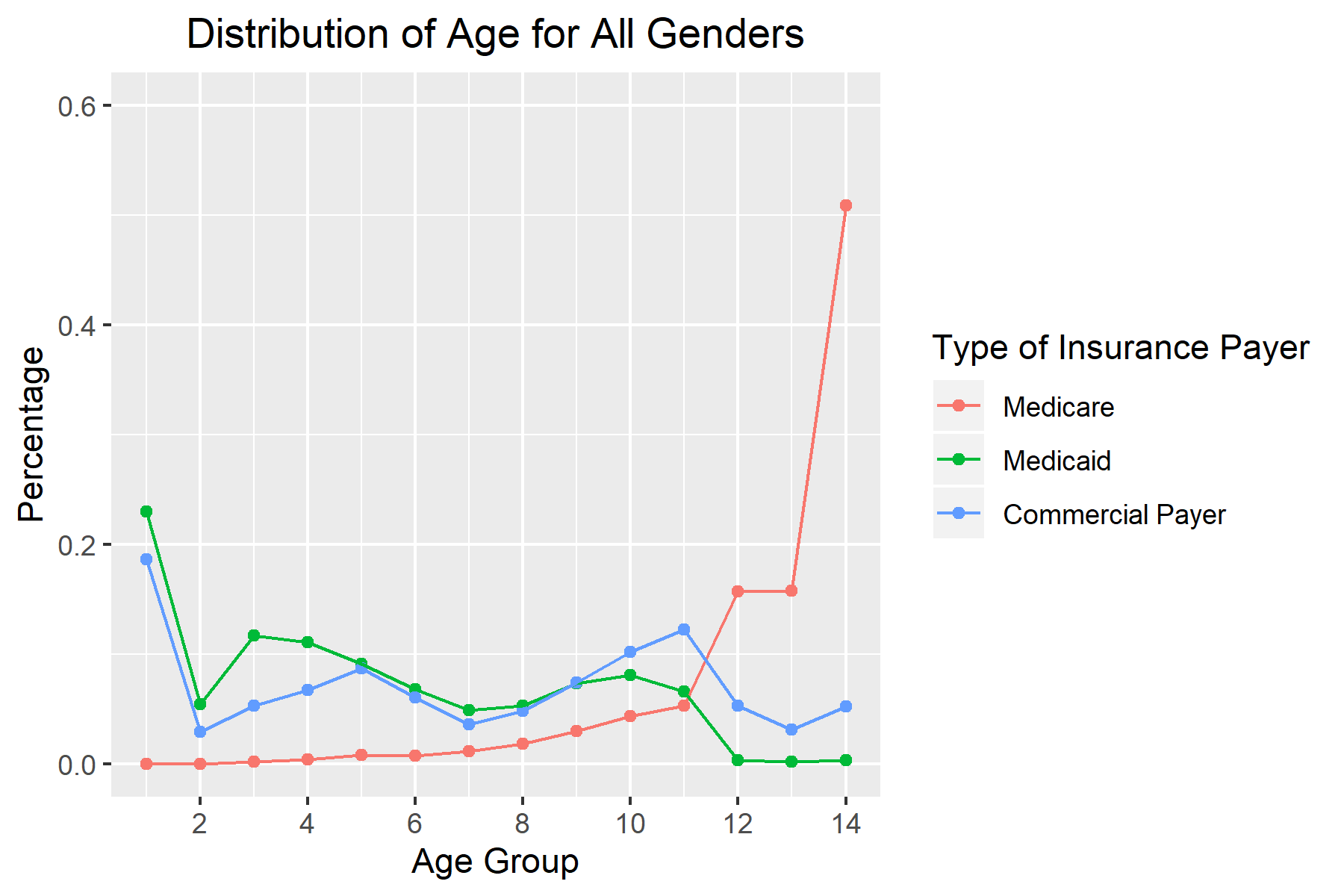
However, for Medicaid and Commercial Payers, the situation is quite different. From the pie chart of Medicaid, we are observing that the top5 MDCs are Neonatal, Musculoskeletal, Pregnancy, Childbirth and the Puerperium, Respiratory and Heart & Circulatory (Mental Illness, Digestive, Infection and Brain an CNs all take up 6% of Medicaid’s cost). As a government insurance program for persons of all ages whose income and resources are insufficient to pay for health care, infants under 1 years old accounts for more than 20% of those covered by Medicaid, as **Chart 4** shows. In addition, the percentage shows an increase after children are over the age of 18. This may because that, in the United States, children tend to pay for themselves (such as student loans) when they turned 18. They may not have enough money to pay for health care.



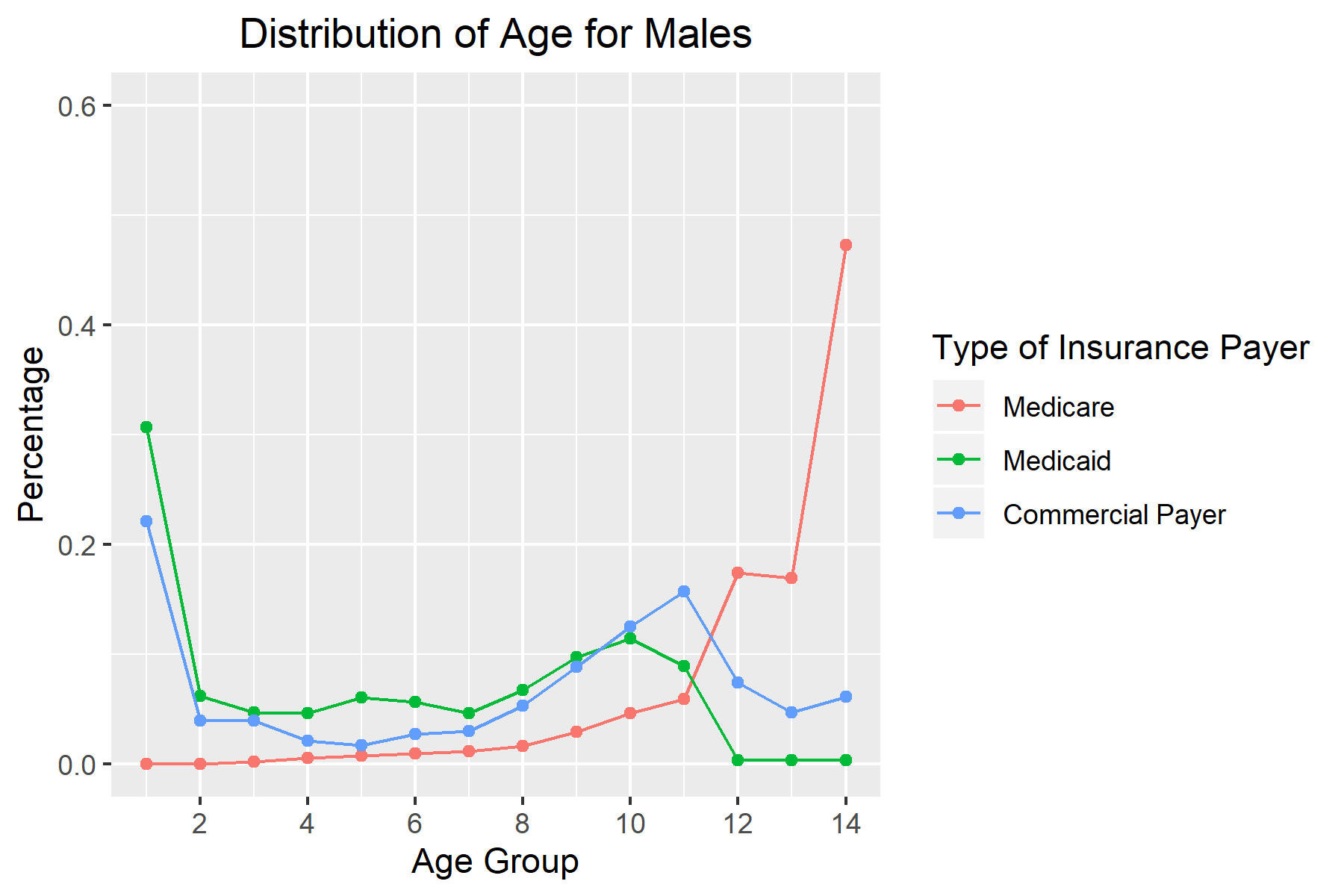
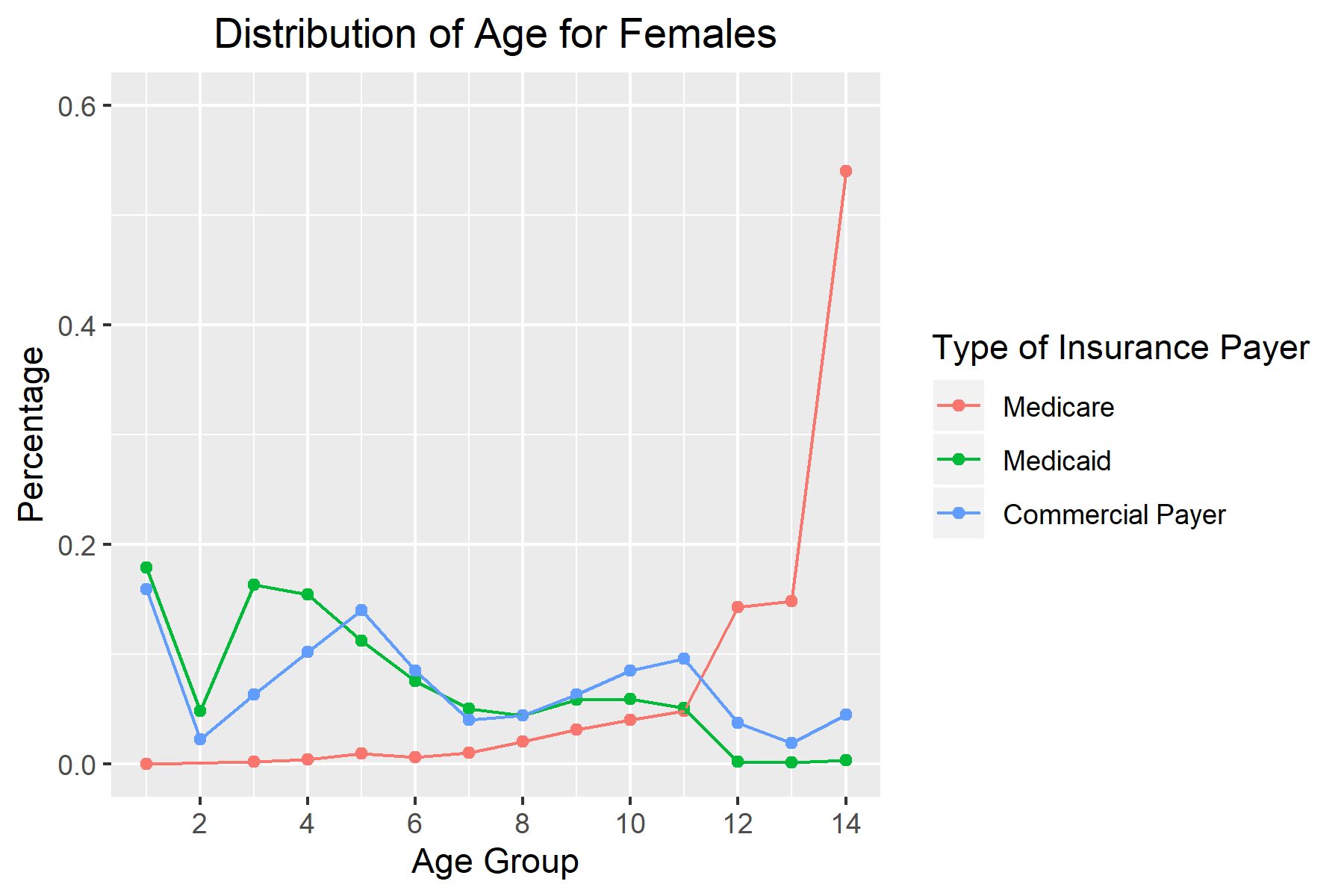
**Chart 3** The inpatient services portfolio for Commercial Payers

From the pie chart of Commercial Payers (**Chart 3**), the top-5 MDCs are Musculoskeletal, Heart & Circulatory, Brain and CNs, Pregnancy, Childbirth and the Puerperium, and Digestive.

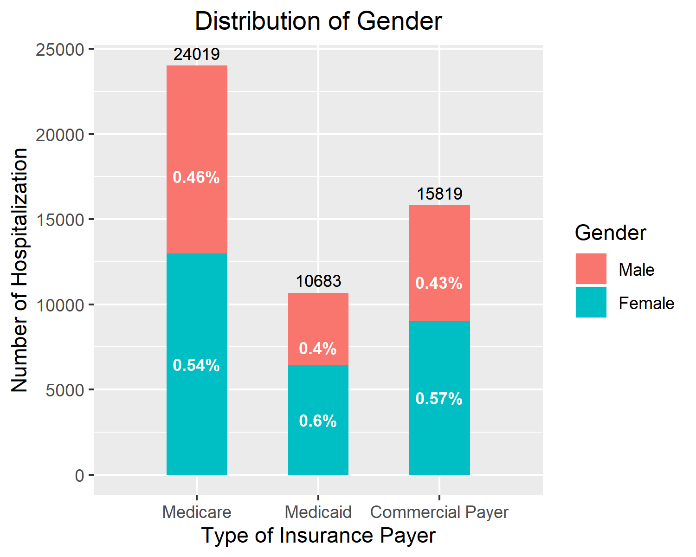
Musculoskeletal and Heart & Circulatory seems to be a very common disease cross genders and ages. **Chart 4** shows very small difference between different age groups in the percentage of patients covered by Medicaid, except that patients under 1 years old account for nearly 20%, relatively higher than other ages. However, it’s worth noting that, as **Chart 5** shows, there are some difference between females and males. Among those covered by Commercial Payers and aged from 1 to 34 (age group 2-5), the percentage is increasing when the age group getting higher for females, contrary to the situation for males. And among those covered by Commercial Payers and aged from 30 to 44 (age group 5-7), the percentage is decreasing when the age group getting higher for females, also contrary to the situation for males.



**Chart 4** The percentage of each age group for all genders covered by Medicare, Medicaid, and Commercial Payer



**Chart 5** The percentage of each age group for females and males covered by Medicare, Medicaid, and Commercial Payer



**Chart 6** The gender composition of Medicare, Medicaid, and Commercial Payer

**Chart 6** shows the gender composition of three types of insurance payers. As we found that Childbirth are among the top-5 MDCs for both Medicaid and Commercial Payer, the reason behind could be the higher proportion (about 60%) of females covered by these two insurance payers. Mothers tend to pay for their children.

**Study of the drug abuse related cases**

Using the ED claims data, we examine the enormity of the health crisis related to illicit drugs and prescription opioids use/abuse/overdose. The findings are as follows:

* The number of ED visits diagnosed as drug user/abuser is **2151**.
* There is no evidence supporting the gender bias myth that men are more likely to have drug use/abuse problem than women (**1009** male drug users/abusers vs. **1141** female drug users/abusers from our data).
* The dollar amount spent on the identified drug abuse related patients is 30741220. Of each of the total payments for major insurance payers, the share of Medicare’s expenditure on drug abuse related cases is **2.88%**, the share of Medicaid’s is **2.78%**, and the share of Commercial Payers’ is **1.78%.**
* **156** patients have been brought to ED for diagnosis related to synthetic narcotics or amphetamines, accounting for **0.06** **percent** of all of the patients brought to ED.
* The 3 zip code regions with highest numbers of drug use/abuse cases are **054**, **057**, and **05701**.
* The 10 most common diagnoses of drug use/abuse are:

|  |  |  |
| --- | --- | --- |
| **NO.** | **CODE** | **COUNT** |
| 1 | T401X1A | 258 |
| 2 | T402X5A | 256 |
| 3 | T424X2A | 123 |
| 4 | T424X5A | 114 |
| 5 | T40605A | 112 |
| 6 | T43222A | 82 |
| 7 | T402X1A | 81 |
| 8 | T424X1A | 80 |
| 9 | T426X5A | 75 |
| 10 | T426X2A | 74 |

**Table 5** The 10 most common diagnoses of drug use/abuse

**Appendix**

1. **# Load the Data ###**
2. setwd("C:/Users/ytlau/Desktop/Brandeis/193HS-256F-1 Healthcare Data Analytics and Data Mining/HW3")
3. library(data.table)
4. ed = fread("VTED16.TXT")
5. inp = fread("VTINP16\_upd.TXT")
6. outp = fread("VTOUTP16.TXT")
7. revcode = fread("VTREVCODE16.TXT")
8. **# Question 2 ####**
9. ## Subset three major insurance payers
10. inp\_insurance = inp[inp$PPAY == 1 | inp$PPAY == 2 | inp$PPAY == 6 | inp$PPAY == 7,]
11. inp\_insurance$PPAY = ifelse(inp\_insurance$PPAY==6|inp\_insurance$PPAY==7, 3, inp\_insurance$PPAY)
12. ## Create cross tabulation MDC X PPAY
13. tabulation = xtabs(CHRGS ~ MDC + PPAY, inp\_insurance)
14. rownames(tabulation) = c('BRAIN AND CNS', 'EYE', "EAR, NOSE & THROAT", "RESPIRATORY", "HEART & CIRCULATORY", "DIGESTIVE","LIVER & PANCREAS", "MUSCULOSKELETAL", "SKIN AND BREAST", "ENDOCRINE", "KIDNEY & URINARY", "MALE REPRODUCTIVE", "FEMALE REPRODUCTIVE", "PREGNANCY, CHILDBIRTH AND THE PUERPERIUM","NEONATAL", "SPLEEN & BLOOD", "LYMPHATIC", "INFECTION", "MENTAL ILLNESS", "SUBSTANCE ABUSE", "INJURY, TOXIC EFFECTS", "BURNS", "ALL OTHER","TRAUMA", "HIV")
15. colnames(tabulation) = c("Medicare", "Medicaid", "Commercial\_Payers")
16. ## Transform matrix to data.table
17. write.csv(tabulation, file = "tabulation.csv")
18. table\_tabulation = fread("tabulation.csv")
19. ## Number Formatting
20. table\_tabulation = table\_tabulation[,.(MDC\_name = V1, Medicare = round(Medicare/1000000), Medicaid = round(Medicaid/1000000), Commercial\_Payers = round(Commercial\_Payers/1000000))]
21. ## Calculate the shares
22. table\_tabulation$Medicare\_share = round(table\_tabulation$Medicare/sum(table\_tabulation$Medicare),4)
23. table\_tabulation$Medicaid\_share = round(table\_tabulation$Medicaid/sum(table\_tabulation$Medicaid),4)
24. table\_tabulation$Commercial\_Payers\_share = round(table\_tabulation$Commercial\_Payers/sum(table\_tabulation$Commercial\_Payers),4)
25. write.csv(table\_tabulation, file = "tabulation.csv")
26. ## Plot the distribution of age group
27. library(ggplot2)
28. age\_m = inp\_insurance[sex==1]
29. age\_f = inp\_insurance[sex==2]
30. age\_count = inp\_insurance[,.N, by=.(intage,PPAY)][,.(intage,N,sum\_by\_PPAY = sum(N)),by=PPAY][order(PPAY,intage)]
31. age\_count$pct=round(age\_count$N/age\_count$sum\_by\_PPAY,3)
32. age\_count\_m = age\_m[,.N, by=.(intage,PPAY)][,.(intage,N,sum\_by\_PPAY = sum(N)),by=PPAY][order(PPAY,intage)]
33. age\_count\_m$pct=round(age\_count\_m$N/age\_count\_m$sum\_by\_PPAY,3)
34. age\_count\_f = age\_f[,.N, by=.(intage,PPAY)][,.(intage,N,sum\_by\_PPAY = sum(N)),by=PPAY][order(PPAY,intage)]
35. age\_count\_f$pct=round(age\_count\_f$N/age\_count\_f$sum\_by\_PPAY,3)
36. ## Distribution of age for three major insurance payers for both genders
37. (p\_age = ggplot(age\_count, aes(x=intage, y=pct, color=factor(PPAY)))+ geom\_point()+ geom\_line()+ labs(x = 'Age Group', y = 'Percentage', title = 'Distribution of Age for All Genders',color = 'Type of Insurance Payer')+ scale\_x\_continuous(breaks=seq(0,14,2))+ ylim(0,0.6)+ scale\_colour\_discrete(labels = c('Medicare','Medicaid','Commercial Payer'))+ theme(plot.title = element\_text(hjust=0.5)))
38. ## Distribution of age for three major insurance payers for females
39. (p\_age\_f = ggplot(age\_count\_f, aes(x=intage, y=pct, color=factor(PPAY)))+ geom\_point()+ geom\_line()+ labs(x = 'Age Group', y = 'Percentage', title = 'Distribution of Age for Females',color = 'Type of Insurance Payer')+ scale\_x\_continuous(breaks=seq(0,14,2))+ ylim(0,0.6)+ scale\_colour\_discrete(labels = c('Medicare','Medicaid','Commercial Payer'))+ theme(plot.title = element\_text(hjust=0.5)))
40. ## Distribution of age for three major insurance payers for males
41. (p\_age\_m = ggplot(age\_count\_m, aes(x=intage, y=pct, color=factor(PPAY)))+ geom\_point()+ geom\_line()+ labs(x = 'Age Group', y = 'Percentage', title = 'Distribution of Age for Males',color = 'Type of Insurance Payer')+ scale\_x\_continuous(breaks=seq(0,14,2))+ ylim(0,0.6)+ scale\_colour\_discrete(labels = c('Medicare','Medicaid','Commercial Payer'))+ theme(plot.title = element\_text(hjust=0.5)))
42. ## Plot gender difference for three major insurance payers
43. library(plyr)
44. sex\_count = inp\_insurance[, .N, by = .(PPAY,sex)][,.(sex=factor(sex),N,sum=sum(N)),by=PPAY][order(PPAY,-sex)]
45. sex\_count$pct = sex\_count[, N/sum]
46. sex\_count = ddply(ddply(sex\_count, 'PPAY', transform, label\_pct\_pre = ifelse(sex==2,N/2,N)), 'PPAY', transform,label\_pct=cumsum(label\_pct\_pre))
47. ## Distribution of gender for three major insurance payers
48. (p\_sex = ggplot(sex\_count, aes(x=PPAY, y=N, fill=sex))+ geom\_bar(stat='identity',width=0.5)+ labs(x = 'Type of Insurance Payer', y = 'Number of Hospitalization', title = 'Distribution of Gender',fill = 'Gender')+ scale\_x\_discrete(limits=c('Medicare','Medicaid','Commercial Payer'))+ scale\_fill\_discrete(labels = c('Male','Female'))+ geom\_text(aes(y=label\_pct,label=paste(round(pct,2),'%',sep='')),size=3,color='white',fontface='bold')+ geom\_text(aes(y=sum,label=sum,vjust=-0.5),size=3)+ theme(plot.title = element\_text(hjust=0.5)))
49. **# Question 3 ####**
50. ## Create a new column for subsetting the data
51. ed$newcol <- paste(ed$DX1, ed$DX2, ed$DX3, ed$DX4, ed$DX5,ed$DX6, ed$DX7, ed$DX8, ed$DX9, ed$DX10, ed$DX11, ed$DX12, ed$DX13, ed$DX14, ed$DX15, ed$DX16, ed$DX17, $DX18, ed$DX19, ed$DX20, sep='-')
52. ## Subset patients with at least one drug abuse related ICD-10 code
53. ed\_drug <- ed[grepl("T40", ed[["newcol"]]) | grepl("T41", ed[["newcol"]]) | grepl("T42", ed[["newcol"]]) | grepl("T43", ed[["newcol"]]), ]
54. ed\_drug$newcol = NULL
55. ## How many ED visits exactly have been diagnosed as drug user/abuser?
56. nrow(ed\_drug)
57. ## Can you check if your data supports this gender bias myth?
58. ed\_drug[,.(count=.N), by = .(sex)]
59. ## Can you find the exact dollar amount for your identified patients in this question?
60. sum(ed\_drug$CHRGS)
61. ## Of the three insurances in Question 2, what was share of each of the total payments?
62. insurance\_payment = ed[,.(sum\_payment=sum(CHRGS, na.rm=TRUE)), by = .(PPAY)]
63. pay\_medicare = insurance\_payment[insurance\_payment$PPAY==1, "sum\_payment"]
64. pay\_medicaid = insurance\_payment[insurance\_payment$PPAY==2, "sum\_payment"]
65. pay\_commercial = insurance\_payment[insurance\_payment$PPAY==6, "sum\_payment"] + insurance\_payment[insurance\_payment$PPAY==7, "sum\_payment"]
66. insurance\_payment\_drug = ed\_drug[,.(sum\_payment=sum(CHRGS, na.rm=TRUE)), by = .(PPAY)]
67. pay\_medicare\_drug = insurance\_payment\_drug[insurance\_payment\_drug$PPAY==1, "sum\_payment"]
68. pay\_medicaid\_drug = insurance\_payment\_drug[insurance\_payment\_drug$PPAY==2, "sum\_payment"]
69. pay\_commercial\_drug = insurance\_payment\_drug[insurance\_payment\_drug$PPAY==6, "sum\_payment"] + insurance\_payment\_drug[insurance\_payment\_drug$PPAY==7, "sum\_payment"]
70. pay\_medicare\_drug/pay\_medicare
71. pay\_medicaid\_drug/pay\_medicaid
72. pay\_commercial\_drug/pay\_commercial
73. ## Use the ICD-10 codes of T404xxx and T4362xx to identify only a small sample of such patients.
74. ## How many of patients have been brought to ED for diagnosis related to synthetic narcotics or amphetamines?
75. ed\_narcotics\_amphetamines <- ed[grepl("T404", ed[["newcol"]]) | grepl("T4362", ed[["newcol"]])]
76. nrow(ed\_narcotics\_amphetamines)
77. nrow(ed)
78. ## Name the 3 zip code regions with the highest numbers of drug use/abuse cases.
79. ed\_drug[,.(count=.N), by = .(TXTZIP)][order(-count)]
80. ## What are the 10 most common diagnoses of drug use/abuse?
81. ed\_new = as.data.table(gather(ed\_drug, key = "DX", value = "Code", `DX1`, `DX2`, `DX3`, `DX4`, `DX5`, `DX6`, `DX7`, `DX8`, `DX9`, `DX10`, `DX11`, `DX12`, `DX13`, `DX14`, `DX15`, `DX16`, `DX17`, `DX18`, `DX19`, `DX20`))
82. ed\_new\_2 = ed\_new[grepl("T40", ed\_new[["Code"]]) | grepl("T41", ed\_new[["Code"]]) | grepl("T42", ed\_new[["Code"]]) | grepl("T43", ed\_new[["Code"]])]
83. ed\_diagnose = ed\_new\_2[,.(count=.N), by = .(Code)][order(-count)]
84. ed\_diagnose[1:10,]