Paper 77-2010

Characteristic Examination of the Hospitalized Children with Brain Cancer in 2006: A Population Based Approach

Yubo Gao The University of Iowa Hospitals and Clinics, Iowa City, Iowa

Abstract

The study examined the demographic and hospitalization characteristics of children hospitalized with brain cancer (BC) using 2006 national discharge data from the Healthcare Cost and Utilization Project (HCUP) Kids' Inpatient Database (KID). Of the nearly 3.1 million cases in the database, 8932 (0.285%) were identified through diagnostic coding of BC. There are more boys than girls having BC, and more than half had private insurance as their primary payer. About half of the children were ages between 3 and 12 years, but all ages were represented from age 0 to 20. White children weighed 60 percent. 97 percent BC children stayed in urban hospitals, and 87 percent hospitals had teaching status. All patients had an average length of stay (LOS) 6.89 days. The average number of diagnoses is 5.33, and the average number of procedures is 2.14. Total average charge per child is $50,236. The average number of comorbidities increased as patient got older. There are 4.4% more odds of child death in non-teaching hospitals than in teaching hospitals, and there are 331% more odds of child death in rural hospitals than in urban hospitals.

Introduction

According to the American Cancer Society[1], brain and other nervous system cancers are the second most common cancers (behind Leukemia) in children, making up about 21% of childhood cancers. To our knowledge, no population-based study had been conducted to date that examined the characteristics of children patients who were hospitalized for malignant neoplasm of brain, or brain cancer (BC). The purpose of this study is to report the specific demographic and hospitalization characteristics for children with brain cancers in 2006 as culled from a database of inpatient hospitalization usage by children across the United States. We hypothesized that the death rates of children with BC during hospitalization are lower in an urban or teaching hospital.

Methods

The Healthcare Cost and Utilization Project (HCUP) Kids' Inpatient Database (KID)[2] is the only dataset on hospital use, outcomes, and charges designed to study children's use of hospital services in the United States. The 2006 KID contains approximately 3.1 million pediatric discharges from 3739 community, non-rehabilitation hospitals in 38 states representing all 4 geographic census regions (Northeast, Midwest, West, and South). This KID database includes a sampling of all hospital discharges where the patient was age 20 or less at admission during the year 2006. The sample is weighted by design to be representative of all community hospitals in the American Hospital Association annual survey of hospitals, thus allowing for extrapolation to a national estimation of 7.6 million pediatric hospital discharges. However, all our analyses were performed on the KID sample. Patient demographic variables include age at time of admission, sex, race, and median household income quartiles based on the ZIP code of the family's residence. Hospitalization variables include admission month and source, diagnostic and procedure codes, duration of stay, total charges, expected payer, and discharge disposition. Hospitals included in this database are divided into strata using 6 characteristics: ownership/control, bed size, teaching status, rural/urban location, US region, and hospital type (pediatric vs other). Bed capacity is categorized into small, medium, or large, and varied in specific bed capacity depending on whether the hospital was located in a rural area or was an urban non-teaching or urban teaching hospital.

Data for this study was culled from the 2006 KID database using the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*)[3] codes for brain cancer (Table 1).

Table1. ICD-9 Codes for Malignant Brain Cancer

ICD 9 Codes	General Description
191.0	Cerebrum, except lobes and ventricles
191.1	Frontal lobe
191.2	Temporal lobe
191.3	Parietal lobe
191.4	Occipital lobe
191.5	Ventricles
191.6	Cerebellum NOS

191.7	Brain stem
191.8	Other parts of brain
191.9	Brain, unspecified

The patient was included in this study if the patient had at least one of the above BC code from the 15 available diagnosis codes for each patient in the database. Specifically, we used a data step to pull out our sample from the database Core2006 provided by KID as.

```
data BC;
      set Core2006
      braincancer=0;

      diagnoses=catx(' ',dx1, dx2, dx3, dx4, dx5, dx6, dx7, dx8, dx9, dx10,
                     dx11, dx12, dx13, dx14, dx15);
      procedures=catx(' ',pr1, pr2, pr3, pr4, pr5, pr6, pr7, pr8, pr9, pr10,
                     pr11, pr12, pr13, pr14, pr15);
      if (rxmatch('1910',diagnoses)>0) then braincancer=1;
      if (rxmatch('1911',diagnoses)>0) then braincancer=1;
      if (rxmatch('1912',diagnoses)>0) then braincancer=1;
      if (rxmatch('1913',diagnoses)>0) then braincancer=1;
      if (rxmatch('1914',diagnoses)>0) then braincancer=1;
      if (rxmatch('1915',diagnoses)>0) then braincancer=1;
      if (rxmatch('1916',diagnoses)>0) then braincancer=1;
      if (rxmatch('1917',diagnoses)>0) then braincancer=1;
      if (rxmatch('1918',diagnoses)>0) then braincancer=1;
      if (rxmatch('1919',diagnoses)>0) then braincancer=1;
      *if (rxmatch('1629',diagnoses)>0) then braincancer=1;
      if braincancer=1;
run;
```

The study group was then further divided into 4 age groups based on ages that would likely participate in common activities from a developmental standpoint[4]. The four groups are: (1) infants (age, <1 year), (2) toddlers (age, 1 or 2 years), (3) children (age, 3-12 years), and (4) adolescents (age, 13-20 years). In SAS, the following did that.

```
data BC;
      set BC;
      if age=. then delete;
      if age<1 then agegroup=1;
      else if 1<=age<=2 then agegroup=2;
      else if 3<=age<=12 then agegroup=3;
      else if 13<=age<=20 then agegroup=4;
run;
```

Data is presented in two categories: patient demographic characteristics, and hospitalization-related data. Descriptive statistics including means, standard deviations and percentages are reported, and Chi-Square χ^2 test or unbalanced analysis of variance was used for determining differences between age-group specific means or proportions. For all analyses, a *P* values less than 0.05 was considered statistically significant. Data were analyzed using SAS software (version 9.1.3; SAS Institute, Cary, NC)[5].

The results in Table 3 are obtained from the following SAS codes.

```
proc freq data=BC;
      tables female*agegroup/chisq;
run;
proc freq data=BC;
      tables race*agegroup/chisq;
run;
proc freq data=BC;
      tables pay1*agegroup/chisq;
run;
```

```
proc freq data=BC;
      tables elective*agegroup/chisq;
run;
proc freq data=BC;
      tables hosp_bedsize*agegroup/chisq;
run;
proc freq data=BC;
      tables hosp_location*agegroup/chisq;
run;
proc freq data=BC;
      tables hosp_teach*agegroup/chisq;
run;
proc freq data=BC;
      tables died*agegroup/chisq;
run;
proc freq data=BC;
      tables zipinc_qrtl*agegroup/chisq;
run;
```

While the results in Table 4 are obtained from the following SAS codes.

```
proc freq data=BC;
      tables agegroup*(female race pay1 zipinc_qrtl  elective hosp_bedsize
                       hosp_location hosp_region hosp_teach died);
run;
```

Results

There are a total of 8,932 observations related to BC disease out of 3.1 million records in database KID. Since there were 22 records without age, the final study group consists of 8,910 records. Generally, the number of patients went down as age increased, see Table 2 and Figure 1. Meantime, the correlation between age and the number of patients is -0.8527 (P<0.0001), which further substantiated the trend.

Table 2. Patients Number versus Age

Age	0	1	2	3	4	5	6	7	8	9	10
Count	700	810	942	612	494	452	432	681	429	327	287
Age	11	12	13	14	15	16	17	18	19	20	
Count	271	489	287	283	224	319	225	230	236	180	

Figure 1. Patients Number versus Age

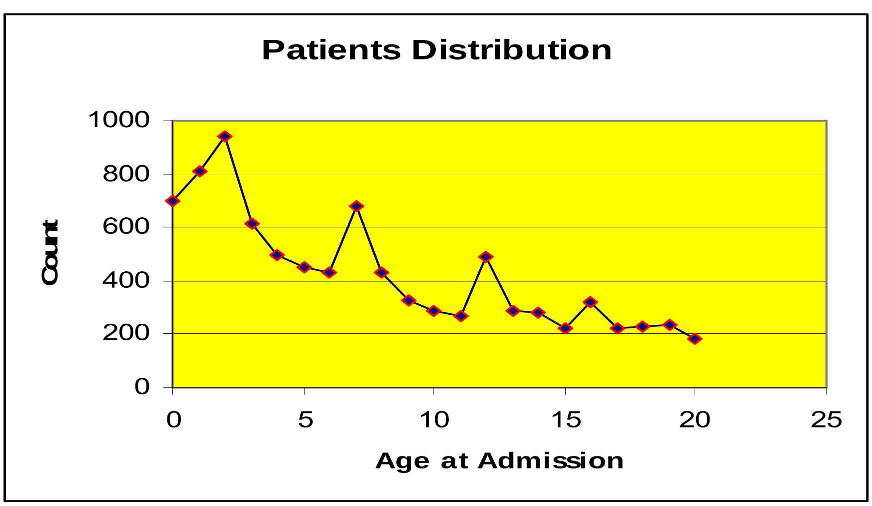

Table 3 showed that more than half (50.21%) occurred in children age between 3 and 12. The overall incidence rate is higher in boys (58%) than in girls (42%), which is consistent with the statistics from American Cancer Society[1] that boys are affected slightly more often than are girls; however, for different age groups, much more cases in infants occurred in girls (61%), whereas there was a male predominance in children older than 1 year old, and Chi-Square χ^2 test showed that the age-group specific gender proportions were statistically different (P=0.0065). The largest single racial group was white for accounting for 60%, and the age-group specific racial proportions were statistically different (P<0.0001). More than half patients had private insurance (54%), and the age-group specific primary payer proportions were statistically different (P<0.0001). The percentages among age-group specific median household income quartile for patient's ZIP Code were statistically different (P<0.0001).

Table 3. Demographic Characteristics of Hospitalized Children with BC

	Age Group				
Parameter	<1 y, No. (%)	1-2 y, No.(%)	3 -12 y, No. (%)	13-20 y, No. (%)	Total (%)
No. patients (total)	700 (7.86)	1752 (19.66)	4474 (50.21)	1984 (22.27)	8910 (100)
Average age (y)	0	1.54	7.05	16.23	7.46
Sex					
Female	270(61)	796(45)	1886(42)	815(41)	3767(42)
Male	430(39)	954(55)	2568(58)	1165(59)	5117(58)
Missing	26				
Race					
White	285(55)	791(57)	2064(60)	963(66)	4103(60)
Black	86(16)	167(12)	372(11)	135(9)	760(11)
Hispanic	104(20)	304(22)	703(20)	234(16)	1345(20)
Asian or Pacific	15(3)	36(3)	119(3)	59(4)	229(3)
Native American	3(1)	8(1)	17(1)	10(1)	38(1)
Other	30(5)	91(5)	156(5)	66(4)	343(5)
Missing	2092				
Primary payer					
Medicare			2(.04)	5(.25)	7(.08)
Medicaid	391(56)	724(41)	1585(35.47)	640(32.31)	3340(37.53)
Private	272(39)	894(51)	2494(55.82)	1159(58.51)	4819(54.15)
Self-pay	10(1)	35(2)	149(3.33)	73(3.69)	267(3)
No charge		5(1)	7(.16)	4(.20)	16(.18)
Other	27(4)	93(5)	231(5.17)	100(5.05)	451(5.06)
Missing	10				
Median household income quartiles for patient's ZIP Code					
1	211(30.8)	510(29.98)	955(21.7)	420(21.56)	2096(24)
2	183(26.72)	418(24.57)	1155(26.25)	424(21.77)	2180(24.96)
3	163(23.8)	379(22.28)	1028(23.36)	517(26.54)	2087(23.9)
4	128(18.69)	394(23.16)	1262(28.68)	587(30.13)	2371(27.15)
Missing	176				

Table 4. Hospitalization Data Characteristics for Children with BC: Categorical Data

	Age Group				
Parameter	<1 y, No. (%)	1-2 y, No.(%)	3 -12 y, No. (%)	13-20 y, No. (%)	Total (%)
Admission type					
Elective	157(23)	603(35)	1750(40)	770(39)	3280(37)
Non elective	538(77)	1141(65)	2679(60)	1195(61)	5553(63)
Missing	77				
Hospital bed size					
Small	116(18)	301(18)	572(13)	239(13)	1228(15)
Medium	171(26)	452(28)	1118(27)	460(24)	2201(26)
Large	371(56)	894(54)	2497(60)	1195(63)	4957(59)
Missing	524				
Hospital location					
Urban	602(91)	1578(96)	4101(98)	1854(98)	8135(97)
Rural	56(9)	69(4)	86(2)	40(2)	251(3)
Missing	524				
Hospital teaching status					
Teaching	508(77)	1362(83)	3781(90)	1671(88)	7322(87)
Non teaching	150(23)	285(17)	406(10)	223(12)	1064(13)

Missing	524				
Hospital region					
Northeast	75(11)	235(13)	674(15)	348(18)	1332(15)
Midwest	218(31)	455(26)	1081(24)	523(26)	2277(26)
South	272(39)	699(40)	1557(35)	533(27)	3061(34)
West	135(19)	363(21)	1162(26)	580(29)	2240(25)
Died (%)	13(1.86)	18(1.03)	70(1.56)	46(2.32)	147(1.65)

Non elective admission accounted for 63% (Table 4) for all patients, also predominating in all four age groups, and the age-group specific admission type proportions were statistically different (P<0.0001). 59% of patients in all chose the hospitals with large bed size, also every age group patients preferring the large bed size hospitals, and the age-group specific hospital bed size proportions were statistically different (P<0.0001). Most hospitals (97%) are located in urban areas, and the age-group specific hospital location proportions were statistically different (P<0.0001). 87% of hospitals in all had teaching status, and the age-group specific hospital teaching status proportions were statistically different (P<0.0001). The death proportions among four age groups were statistically different (P=0.0183).

Next, we did a separate analysis for each continuous hospitalization data, that is, LOS, number of diagnoses, number of procedures, total charges, and number of comorbidities by the following SAS codes.

```
proc means data=BC maxdec=2;
      class agegroup;
      var los ndx npr totchg age ncm;
run;
```

Table 5. LOS (days) Distributions

	Age Group (years)				
Parameters	<1	1-2	3-12	13-20	Total
N	699	1752	4474	1984	8909
Mean	10.6	7.12	6.3	6.7	6.89
StdDev	23.67	14.33	10.07	11.3	12.85
Range	0-326	0-313	0-166	0-148	0-326

The average LOS for all patients is 6.89 days (Table 5) with standard deviation 12.85 days. Intuitively, the average LOS decreased as patient age went up. The infants stayed the longest in hospital. The unbalanced analysis of variance model showed that these age-group specific LOS means were statistically different (P<0.0001).

Table 6. No. Diagnoses Distributions

	Age Group (years)				
Parameters	<1	1-2	3-12	13-20	Total
N	700	1752	4474	1984	8910
Mean	6.25	5.38	5.12	5.44	5.33
StdDev	4.37	3.33	3.23	3.6	3.45
Range	0-29	0-25	0-26	0-25	0-29

The average number of diagnoses for all patients is 5.33 with standard deviation 3.45 (Table 6). The unbalanced analysis of variance model showed that these age-group specific diagnosis means were statistically different (P<0.0001).

Table 7. No. procedures Distributions

	Age Group (years)				
Parameters	<1	1-2	3-12	13-20	Total
N	700	1752	4474	1984	8910

Mean	2.46	2.3	2.07	2.04	2.14
StdDev	3.61	2.82	2.36	2.27	2.56
Range	0-21	0-21	0-21	0-29	0-29

The average number of procedures each patient received for all patients is 2.14 with standard deviation 2.56 (Table 7). Strictly, the average number of procedures patient received decreased as patient age went up. The unbalanced analysis of variance model showed that these age-group specific procedure means were statistically different (P<0.0001).

Table 8. Total Charges ($) Distributions

	Age Group (years)				
Parameters	<1	1-2	3-12	13-20	Total
N	690	1731	4411	1921	8753
Mean	60,097	46,848	46,351	58,669	50,236
StdDev	127,037	86,968	78,274	99,312	89,720
Range	813-954,655	1,028-864,046	562-947,215	716-961,049	562-961,049

The average charge for all patients is $50,236 with standard deviation $89,720 (Table 8). For different age groups, the unbalanced analysis of variance model showed that these age-group specific charge means were significantly different (P<0.0001).

Table 9. No. Comorbidities Distributions (1)

	Age Group (years)				
Parameters	<1	1-2	3-12	13-20	Total
N	700	1752	4474	1984	8910
Mean	0.58	0.83	0.84	0.97	0.85
StdDev	0.85	0.98	0.96	1.06	0.98
Range	0-5	0-5	0-6	0-6	0-6

The study group of patients had a total of 7,574 (=8910*0.85) comorbidities with an average of 0.85 and a range from 0 to 6 (Table 9). In average, patients had more comorbidities as patient age went up. The unbalanced analysis of variance model showed that age-group specific comorbidity means were significantly different (P<0.0001). We further tabulated the number of comorbidities by age group and number of comorbidities. See Table 10. About 46% patients had no comorbidity, only 5 patients had 6 comorbidities, and these 5 patients are in the two older age groups. There are more proportions of infants patients that showed zero comorbidity than any other patient categories, and proportions of patients that showed zero comorbidity decreased as age went up. That probably indicated that older children could easily show more comorbidities than younger patients. Oldest patients had higher proportions of higher number of comorbidities such as 2, 3, 4, and 6 than any other patient categories. Chi square χ^2 test showed these proportions were significantly different (P<0.0001).

Table 10. No. Comorbidities Distributions (2)

No. Comorbidities	Age Group (years)				
	<1, No. (%)	1-2, No. (%)	3 -12, No. (%)	13-20, No. (%)	Total (%)
0	424(60.57)	820(46.8)	2000(44.7)	818(41.23)	4062(45.59)
1	178(25.43)	574(32.76)	1552(34.69)	662(33.37)	2966(33.29)
2	75(10.71)	235(13.41)	634(14.17)	322(16.23)	1266(14.21)
3	17(2.43)	89(5.08)	222(4.96)	124(6.25)	452(5.07)
4	4(0.57)	32(1.83)	54(1.21)	51(2.57)	141(1.58)
5	2(0.29)	2(0.11)	9(0.2)	5(0.25)	18(0.20)
6	0(0)	0(0)	3(0.07)	2(0.1)	5(0.06)
Total	700(100)	1752(100)	4474(100)	1984(100)	8910(100)

Impacts of hospital teaching status and hospital location on death of BC during hospitalization were examined by using univariate logistic regression model with the following results (Table 11). In SAS, we did this using

```
proc logistic data=BC;
      class  hosp_teach;
      model died= hosp_teach;
run;
proc logistic data=BC;
      class  hosp_location ;
      model died= hosp_location;
run;
```

Table 11. Odds Ratio Analyses

Hospital Characteristics	Odds Ratio (95% CI)
Hospital Teaching Status	
Non teaching	1.044 (0.626-1.74)
Teaching	1.00
Hospital Location	
Rural	4.314 (0.601-30.962)
Urban	1.00

That is, there are 4.4% more odds of child death in non-teaching hospitals than in teaching hospitals, and there are 331% more odds of child death in rural hospitals than in urban hospitals.

A linear regression model among total charges, LOS, number of procedures, and age was built to have the equation as total charges=5403*los+6836*npr+1023*age. We could see that total charges will increase $5,403, $6,836, and $1,023 for one more day stay, one more procedure, and one more year old, respectively. These predictors are significant in that their *P* values are less than 0.05. In SAS, we use the codes below.

```
proc glm data=BC;
      model totchg=npr los age/solution;
run;
quit;
```

Discussions

In conclusion, this study has provided the understanding of the demographic and hospitalization characteristics about children with brain cancer. There were more boys than girls for this disease. White children weighed 60 percent. About one half of the children were ages between 3 and 12 years. The number of patients decreased as patient got older. More than half of patients had private insurance as their primary payer. All patients had an average length of stay 6.89 days with average charge $50,236. Each patient had an average 5.33 of diagnoses with an average 2.14 of procedures performed. The patient average number of comorbidities increased as patient got older. Urban or teaching hospitals were associated with lower odds of in-hospital mortality

The current study had several limitations. We used administrative discharge data, and the data set did not have information regarding disease stage or tumor size. Thus, cancer severity is not captured. It is well known that patients who have advanced-stage cancers have poor outcomes. Costs, reported in the KID as billed charges, may differ from the actual amount paid because of discounts, deductibles, copayments, and coinsurances.[6,7,8] The true costs are likely to be underestimated in this study, because non covered charges, professional fees, time lost from school and work by parents and other caregivers, and other societal costs are also not included in these estimates. Another limitation, although these databases allow for extrapolation to national estimates, is the potential for error in the extrapolation with this particular population[7], since the KID does not include all US states and does not sample the regions of the United States equally. Regional variations in the BC incidence are possible.

As is the case with all large data sets, missing data and coding errors can potentially lead to bias in the data. There are 9 out of 38 states who did not report race information in the KID, or 20.8% (=651026/3131324) records (9 states had total 651026 records, while there are 3131324 records in the KID). This is, in itself, makes a potential source of bias, since each such state has different race distribution. In this study, there are 2092 records missing race, accounting 23.4% (=2092/8910). This percentage difference further complicates this issue. According to the 2000 US Census data regarding children younger than 20 years, the racial breakdown is 61% white, 15% black, 17% Hispanic, and 7% other races, compared with 60% white, 11% black, 20% Hispanic, and 9% other races in this study with race information. Thus, it seems that Hispanic children are overrepresented while black children are underrepresented in percentage of children diagnosed with BC.

According to the American Cancer Society[1], brain and other nervous system cancers are the second (after leukemia) most common cancers in children, making up about 21% of childhood cancers, around 3,400 central nervous system tumors are diagnosed each year in children under the age of 20, and about one fourth of these are considered benign tumors. The incidence rate of these cancers has not changed much in recent years[1]. Now in KID database, we found 8,932 records with BC. The discrepancy here lies that the KID database contains discharge-level records, not patient-level records, therefore, due to without uniform patient identifier available that allows a patient-level analysis with the KID[2], individual patients who are hospitalized multiple times in one year may be present in the KID multiple times. Another possible reason is that some patients diagnosed in previous years were present for treatment in 2006.

Meantime, from the visualizations the five variables, LOS, total charges, the number of diagnose, the number of procedures, and the number of comorbidities, are quite right-skewed. There is a need to try to transform these data through some techniques, for example, log. Here, based on the Central Limits Theorem in statistics, we did not do that for easy explanations.

References

1.http://www.cancer.org/docroot/CRI/content/CRI_2_4_1X_What_are_the_types_of_childhood_cancers_7.asp. Accessed on June 30, 2009.
2. Healthcare Cost and Utilization Project (HCUP), *Kids' Inpatient Database* 2006 (KID). Rockville, MD: Agency for Healthcare Research and Quality. Issued June 2008. http://www.hcup-us.ahrq.gov/kidoverview.jsp
3. *International Classification of Diseases, Ninth Revision, Clinical Modification for Hospitals* (ICD-9-CM). Vols.1,2, and 3. 6th ed. Salt Lake City, UT: Ingenix St Anthony Publishing; 2003.
Or look at http://icd9.chrisendres.com/
4. Loder RT and Feinberg JR. Orthopaedic Injuries in Children With Nonaccidental Trauma Demographics and Incidence From the 2000 Kids' Inpatient Database. *J Pediatr Orthop*. 2007; 27:421-426.
5. SAS Institute, Cary, NC. *SAS Version* 9.1.3.
6. Allareddy V, Konety BR, Characteristics of Patients and Predictors of In-Hospital Mortality after Hospitalization for Head and Neck Cancers. *Cancer.* 2006;106:2382-2388.
7. Diehr P, Yanez D, Ash A, Hornbrook M, Lin DY. Methods for analyzing health care utilization and costs. *Annu Rev Public Health*. 1999;20:125-144.
8. Lubicky JP, Feinberg JR. Fractures and Amputations in Children and Adolescents Requiring Hospitalization After Farm Equipment Injuries. *J Pediatr Orthop* 2009;29:435–438

Contact Information (Header 1)

Your comments and questions are valued and encouraged. Contact the author at:

Name: Yubo Gao
Enterprise: The University of Iowa Hospitals and Clinics
Address: 01066 JPP
City, State ZIP: Iowa City, IA 52242
Phone: (319)356-1674
Fax:
E-mail:yubo-gao@uiowa.edu
Web: