Background

The purpose of this study is to compare the length of stay of neonate cases that have been managed by the external contractor ("MANAGED"), to the length of stay of neonate cases that have not been managed by said external contract ("UNMANAGED") and to the length of stay of neonate cases from an external benchmark data ("NPIC"). This study is using the data in which information such as Source of Data ("SOURCE"), Length of Stay ("LOS"), Birth Weight ("WEIGHT"), Case Year ("YEAR"), and APR DRG are the variables. For the purpose of this study, we used two variables to help us compare the three different sources of data – Length of Stay and Birth Weight.

Assumptions

In Exhibit 1, we averaged the length of stay by three other variables – the average length of stay by the case year, the average length of stay by the source of data, and the average length of stay by the APR DRG. The vertical axis in Exhibit 1 splits the data by the APR DRG, while the horizontal axis splits it by the case year for each data source. Since our main purpose of this study is to compare the average length of stay of neonate cases for each source of data, we believe it would be easier if we have the horizontal axis of the table split by the data source, and then for each data source, it is split again for each case year.

In Exhibit 2, since our purpose of this study is to compare the MANAGED cases to those of UNMANAGED and NPIC, we believe it is appropriate to do the comparison by using the average birth weight, the average length of stay, as well as the case count for each of the seven birth weight ranges.

Results

Exhibit 1

The result in Exhibit 1 shows the yearly average length of stay for each APR DRG on each data source. As we can infer from Exhibit 1, there are more NPIC cases compared to the MANAGED and UNMANAGED cases. Moreover, the highest average length of stay is at 115.57 days from an NPIC case with an APR DRG of **589 – Neonate Birth Weight < 500g** in the year 2014, while the lowest is at 7 days from two NPIC cases with APR DRGs of **950 – Extensive Procedure Unrelated to Principal Diagnosis** in the year 2015 and **138 – Bronchiolitis & RSV Pneumonia** in the year 2014.

To further summarize the data, we took the average length of stay of each year for each data source, as we can see in Table 1. The purpose of Table 1 is to compare the yearly average length of stay for each data source – we can see that the average length of stay is higher in 2014 than in 2015 in all data sources, with the average length of stay of 28.36 days in 2014 from the MANAGED cases being the highest, and 21.82 days in 2014 from the UNMANAGED cases being the lowest.

Table 1						
SOURCE	MANAGED		UNMANAGED		NPIC	
YEAR	2014	2015	2014	2015	2014	2015
Case Count by Year, by Source	851	923	894	720	8,995	9,646
Average LOS by Year, by Source	28.36	27.11	22.31	21.82	26.06	25.37

Furthermore, we combine each of the case year averages in each data source to calculate the total average length of stay for each data source in Table 2. From the result in Table 2, we can see that the average length of stay of the MANAGED cases is the highest compared to the other data source, with an average of 27.71 days, and the UNMANAGED cases would have the lowest average length of stay of 22.09 days.

Table 2			
SOURCE	MANAGED	UNMANAGED	NPIC
Case Count by Source	1,774	1,614	18,641
Average LOS by Source	27.71	22.09	25.70

Even though comparing the raw averages in Table 2 is a good way to evaluate how the average length of stay differs in each data source, we took a step further by calculating the Case-Mix Adjusted average length of stay for each of the data sources. Since each data source has varying case counts, one way to have a more accurate comparison is to normalize each APR DRG case count of the UNMANAGED and NPIC cases to the distribution of that of the MANAGED cases *(further explanation is covered in the Methodology section)*.

Since we have the distribution of the MANAGED cases as a reference point, both its raw average LOS and case-mix adjusted average LOS would have the same value, as we can see in Table 3 below. However, since their case count-mix for each APR DRG has changed to that of the MANAGED case count, the UNMANAGED case-mix adjusted average LOS would increase from 22.09 days to 25.48 days, and that of NPIC would increase from 25.70 days to 29.06 days. With these changes, the NPIC cases would have the highest average LOS of 29.06 days, and the UNMANAGED remains to have the lowest average LOS of 25.48 days.

Table 3			
SOURCE	MANAGED	UNMANAGED	NPIC
Raw Average LOS by Source	27.71	22.09	25.70
Case-Mix Adjusted Average LOS by Source	27.71	25.48	29.06

Last but not least, we have included two additional summaries – the average LOS by APR DRG, as shown in Table 4 below, and the average LOS by each year, as shown in Table 5. When combining the case year and the source of data, the highest average length of stay by the APR DRG is at 100.09 days for the APR DRG of **589 – Neonate Birth Weight < 500g**, while the lowest average length of stay is at an average of 7 days for the APR DRG **950 – Extensive Procedure Unrelated to Principal Diagnosis**.

Table 4

APR DRG	Average LOS by APR DRG
004	56.00
021	47.50
044	75.00
058	16.00
115	25.00
120	105.00
121	70.33
132	26.00
137	15.50
138	8.00
143	18.43
167	74.00
169	22.50
173	55.25
197	13.00
200	29.00
221	47.20
246	16.00
248	48.00
254	8.00
283	8.50
421	11.20
425	8.00
484	8.00
588	97.03
589	110.09
591	92.36
593	71.63
602	53.67
603	41.63
607	40.30
608	28.62
609	41.12
611	26.74
612	26.01
613	20.59
614	17.53
621	17.65
622	15.52
623	12.23
625	14.89
626	11.38
630	28.39
631	24.83
633	14.58
634	11.11
636	9.21
639	11.80
640	9.51
663	38.00
690	37.00
710	14.50
722	13.00
724	11.50
850	28.13
861	13.00
863	31.83
950	7.00
952	41.57

Consistent with the data from Table 1, Table 5 shows that the average length of stay in the year 2014 is higher than the average length of stay in the year 2015.

Table 5

YEAR	**Average LOS by Year**
2014	25.93
2015	25.29

Exhibit 2

This exhibit is comparing the MANAGED cases to UNMANAGED and NPIC cases using the Birth Weight variable in seven birth weight ranges, as we can see in Table 6 below. Table 6 summarizes the data that we use to compare, which are the case count, average weight, and average length of stay in each birth weight range for each source of data. As we can see in Table 6, the NPIC case in the birth weight range of "<750 grams" has the most average length of stay of 101 days, while the shortest average length of stay is at 12 days, which is that of the UNMANAGED case in the birth weight range of ">2,499 grams". It also shows that the MANAGED and UNMANAGED cases have the same average weight for each birth weight range.

Table 6									
SOURCE	MANAGED			UNMANAGED			NPIC		
Birth Weight Range (in grams)	Case Count	Avg Weight	Avg LOS	Case Count	Avg Weight	Avg LOS	Case Count	Avg Weight	Avg LOS
<750	57	749	91	37	749	77	512	633	101
750 - 999	116	999	71	60	999	69	930	877	76
1000 - 1249	118	1,249	55	78	1,249	46	1,200	1,129	52
1250 - 1499	162	1,499	40	127	1,499	33	1,569	1,381	37
1500 - 1999	447	1,999	23	487	1,999	21	4,858	1,763	22
2000 - 2499	413	2,499	16	534	2,499	13	4,513	2,229	15
>2499	461	2,500	13	291	2,500	12	5,059	3,167	13
Total Cases	1,774			1,614			18,641		

The next metric we use to compare the three data sources is their average birth weight, as we can see in Table 7 below. When we compare the raw average birth weight of each source of data, the UNMANAGED cases will have the highest average birth weight of 2,113.34 grams, while the MANAGED cases would have the lowest average birth weight of 2,044.50 grams. Like exhibit 1, the birth weight range case count for each data source is distributed differently. Therefore, it would be more accurate if we normalize them and compare their case-mix adjusted averages. By taking the case-mix adjusted average birth weight, the MANAGED and UNMANAGED would have the same average birth weight; this is unsurprising since they have the same average birth weight for each birth weight range. Meanwhile, the NPIC cases would have the highest case-mix adjusted average birth weight of 2,064.98 grams.

Table 7			
SOURCE	MANAGED	UNMANAGED	NPIC
Avg Birth Weight	2,044.50	2,113.34	2,108.58
Case-Mix Adjusted Avg Birth Weight	2,044.50	2,044.50	2,064.98

The last metric we used to compare the three data sources is the average and case-mix adjusted average length of stay. The difference between this case-mix adjusted average and that of Exhibit 1 is that in this exhibit, the case-mix is adjusted by the case count distribution of the birth weight ranges, while that of Exhibit 1 is adjusted by the case count distribution of the APR DRGs, as seen in Table 8. For the case-mix adjusted by birth weight range average LOS, the highest value would be that of MANAGED cases of 27.71 days, while the lowest would be that of UNMANAGED cases of 24.57 days.

Table 8			
SOURCE	**MANAGED**	**UNMANAGED**	**NPIC**
Avg LOS	27.71	22.09	25.70
Case-Mix Adjusted by Birth Weight Range Avg LOS	27.71	24.57	27.26
Case-Mix Adjusted by APR DRG Avg LOS	27.71	25.48	29.06

Finally, one last data that we compared is the case count for each birth weight range. Table 9 shows that the birth weight range of ">2499 grams" has the highest case count with a value of 5,811 cases, while the weight range of "<749 grams" has the lowest value of 606 cases.

Table 9	
Birth Weight Range (in grams)	**Case Count by Range**
<750	606
750 - 999	1,106
1000 - 1249	1,396
1250 - 1499	1,858
1500 - 1999	5,792
2000 - 2499	5,460
>2499	5,811

Data

The data that we were provided with came in with 5 variables and a total of 22,029 cases. As a disclaimer, 3 cases do not have what APR DRG is assigned to them, so we have decided to exclude these 3 cases from our analysis. Below is a Table of the variables that are used, as well as their respective values:

Table 10	
Variable	**Value**
Source of Data:	MANAGED, UNMANAGED, NPIC
Length of Stay:	7 – 200 days
Birth Weight:	6 – 9999 grams
Case Year:	2014, 2015
APR DRG:	004, 021, 044, 058, 115, 120, 121, 132, 137, 138, 143, 167, 169, 173, 197, 200, 221, 246, 248, 254, 283, 421, 425, 484, 588, 589, 591, 593, 602, 603, 607, 608, 609, 611, 612, 613, 614, 621, 622, 623, 625, 626, 630, 631, 633, 634, 636, 639, 640, 663, 690, 710, 722, 724, 850, 861, 863, 950, 952

The APR DRG stands for *All Patient Refined Diagnosis Related Groups*, which in this case is a categorical variable with values ranging from 001 to 956 as shown above in Table 10. For further information regarding APR DRG, please refer to the file *APR-DRGsV20MethodologyOverviewandBibliography.pdf*.

One issue with the data is that the NPIC cases have more types of APR DRG handled compared to the MANAGED and UNMANAGED cases, with both MANAGED and UNMANAGED having the same APR DRGs. On top of that, the case count of the NPIC cases seems to be much higher than the MANAGED and UNMANAGED cases, with 18,641 NPIC cases compared to 1,774 MANAGED cases and 1,614 UNMANAGED cases. However, this issue, along with the different distribution of each APR DRG for each data source, will not be a problem when comparing the data using their Case-Mix Adjusted Average, since we compensated for this discrepancy by normalizing the UNMANAGED and NPIC case distribution using MANAGED cases' distribution as a reference, which is further explained in the Methodology section.

Methodology

The data is directly provided to us; thus, this section only covers how the data is processed, rather than how the data was obtained. The data is mainly processed using conditional arithmetic formulas such as COUNTIF(S), AVERAGEIF(S), arithmetic formulas such as SUMPRODUCT and SUM, as well as data organization formulas such as UNIQUE and SORT. The arithmetic formulas would perform a normal calculation of a certain selected dataset, while the conditional arithmetic formulas would perform them under a provided condition. Meanwhile, the data organization formulas, particularly UNIQUE and SORT, were used to obtain the unique values in a dataset as well as sort them accordingly.

Exhibit 1

The first metric of Exhibit 1, the Average LOS by APR DRG, by Year, by Source, is calculated by taking the average of each LOS for a specific APR DRG in a given year from one of the data sources. For example, the MANAGED case has six APR DRGs of 603 in the year 2014, and each case has a value of 17, 26, 25, 32, 28, and 39; thus, the average of these six APR DRGs in the year 2014 from the MANAGED cases is 27.83 (refer to cell D36 Exhibit 1).

Similarly, the Average LOS by Year, by Source, has the same calculation concept as that of the Average LOS by APR DRG, by Year, by Source, but the average is combined over all APR DRGs. Lastly, the Average LOS by source also has the same concept as the two, more detailed, preceding averages, but for this one, the data is averaged over all APR DRGs and year, for each data source.

The Case-Mix Adjusted (by APR DRG) Average LOS by Source, however, is different from the raw Average LOS by Source. Since the APR DRGs in each data source have different case counts, it would be more accurate when the comparison is done if they are the same; this is called normalizing the case-count distribution by the APR DRG. Since we want to compare the MANAGED cases to the others, we normalize the other data sources' APR DRG case counts to that of MANAGED. This is done by taking the sum of the product between each data source's average LOS and MANAGED's case count for all APR DRG and dividing it by MANAGED's total case count. For example, the Case-Mix Adjusted Average for the UNMANAGED cases is calculated by taking the product between the UNMANAGED's average LOS for a specific APR DRG and MANAGED's case count for the same APR DRG and then taking the sum across all APR DRGs, and lastly dividing it by MANAGED's total case count.

In short, the Case-Mix Adjusted Average LOS by Source would have this equation:

$$Case - Mix\ Adjusted\ Average\ LOS_{\ Data\ Source} = \frac{\sum_{APR\ DRG}(Average\ LOS_{Data\ Source,APR\ DRG} \times Case\ Count_{MANAGED,APR\ DRG})}{Case\ Count_{MANAGED}}$$

Exhibit 2

The method of calculating the Average LOS in Exhibits 2 is very similar to that of Exhibit 1, including the method of calculating the Case-Mix Adjusted Averages. For example, in Exhibit 1, the average LOS of the MANAGED case in the birth weight range of 750 – 999 grams in Exhibit 2 is calculated by taking the average LOS for all birth weights inside this range.

Similar to Exhibit 1, instead of using each data source's case count by birth weight range and the total data source case count, the Case-Mix Adjusted Average LOS for the UNMANAGED and NPIC in Exhibit 2 is calculated by normalizing their case count for each of the birth weight range to that of the MANAGED case. Lastly, the Case-Mix Adjusted Average Birth weight has also the same calculation method as the Case-Mix Adjusted Average LOS in Exhibit 2; the only difference is that the Average Birth Weight (BW) of each birth weight range replaces the Average LOS of each birth weight range in the equation.

Below are the equations for both the Case-Mix Adjusted Average LOS and Birth Weight by Source for Exhibit 2, respectively:

$$Case - Mix\ Adjusted\ Average\ LOS_{Data\ Source} = \frac{\sum_{Birth\ Weight\ Range}(Average\ LOS_{Data\ Source,Birth\ Weight\ Range} \times Case\ Count_{MANAGED,Birth\ Weight\ Range})}{Case\ Count_{MANAGED}}$$

$$Case - Mix\ Adjusted\ Average\ BW_{Data\ Source} = \frac{\sum_{Birth\ Weight\ Range}(Average\ BW_{Data\ Source,Birth\ Weight\ Range} \times Case\ Count_{MANAGED,Birth\ Weight\ Range})}{Case\ Count_{MANAGED}}$$

Conclusion

By looking at the raw Average LOS of the neonate cases in each data source, we can conclude that the MANAGED cases have the most average length of stay, with 27.71 days on average, compared to the UNMANAGED cases' average of 22.09 days and NPIC case's average of 25.70 days. Similarly, if we compare their Case-Mix Adjusted by Birth Weight Range Average LOS, the MANAGED cases still have the highest Case-Mix Adjusted Average LOS of 27.71 days, while the NPIC cases place 2nd with an average of 27.26 days, and lastly the UNMANAGED case for having the smallest Case-Mix Adjusted Average LOS of 24.57 days. However, the Case-Mix Adjusted by APR DRG Average LOS of the UNMANAGED and NPIC changes to 29.06 days and 25.48 days, respectively. Since the MANAGED cases have the same average of 27.71 days, it becomes the second highest now, as the NPIC cases would have the highest average of 29.06 days, while the UNMANAGED cases would still have the lowest average of 25.48 days.

This concludes that the MANAGED cases would have the highest raw Average Length of Stay and Case-Mix Adjusted by Birth Weight Range Average Length of Stay compared to those of UNMANAGED and NPIC cases, but the MANAGED cases would be in the middle if we compare it with the other data source cases according to their respective Case-Mix Adjusted by APR DRG Average Length of Stay.