

**Human Resources Development
Canada**

CPP/QPP Analysis of Billing Discrepancies

September 30, 1999

TABLE OF CONTENTS

	<u>Page</u>
1 BACKGROUND	2
2 OUR APPROACH	3
3 METHODOLOGY AND FINDINGS OF THE CPP/QPP STUDY	4
3.1 OVERVIEW	4
3.2 ESTIMATE OF TOTAL AND PROJECTED DISCREPANCIES	5
3.3 AREAS OF DISCREPANCIES	5
3.3.1 Dual payment arrears in CSC system for retirement and disability benefit	5
3.3.2 Withholding rate for QPP	6
3.3.3 Dual payments to children in the CSC system	7
3.3.4 Addition error on report Y0142S (now Y0143)	7
3.3.5 Percentage error for QPP in the CSC system	7
4 ANALYSIS OF DATA	8
4.1 POPULATION CHANGES	8
4.2 BILLING TRENDS	9
5 SUMMARY OF ANALYSIS	15
5.1 ANALYTICAL ASSUMPTIONS	15
5.2 ALTERNATIVES FOR IMPROVING SAMPLE SELECTION	15
5.2.1 Confirm Existence of Time Trend	15
5.2.2 Confirm Existence of Seasonality in the Data	16
5.3 KEY CONSIDERATIONS	16

1 Background

Human Resources Development Canada (HRDC) has requested a review of the methodology and samples used by Business Delivery Systems/Canada Pension Plan (BDS/CPP) Earnings – National Information and Benefits Services (NIBS), to calculate the Canada Pension Plan billing discrepancies to the Quebec Pension Plan (QPP).

Over a period of approximately twelve years there has been billing discrepancies by the CPP to the QPP for benefits. The main reason for these discrepancies relates to the incorrect calculations made by the computer applications for these benefits. The Systems and Technology Management Directorate (STMD) have identified five such discrepancies and corrective action has been taken to resolve them.

The five problems identified by the STMD are as follows:

- Dual payment arrears in CSC system for retirement and disability benefit
- Withholding rate for QPP
- Dual payments to children in the CSC system
- Addition error on report Y0142S (now Y0143)
- Percentage error for QPP in the CSC system

HRDC maintains electronic records of the CPP/QPP monthly transactions for three months, therefore, only hard copies are available for the period of twelve years. Each month has approximately 2700 pages of transactions with 50 entries on each page.

The STMD has prepared a summary report of the five types of discrepancies noted above using the data for a period of twenty months, as the electronic disks were held over for this exercise. The period was from February 1996 to September 1997. Based on the error rate from data available for this period of twenty months, STMD has estimated the potential discrepancy amount incurred over the twelve-year period from 1986 to 1998.

Accordingly, HRDC has requested a consultant to review the methodology and samples of the five types of discrepancies identified.

2 Our Approach

In order to critically evaluate the methodology used by HRDC to develop the estimate of billing discrepancies incurred over the 12-year period in question, it was necessary to distil the available information. The components of this approach included the following:

- Meet with client group;
- Review documentation;
- Document assumptions, procedures and calculations;
- Analyze outputs;
- Analyze factors that may have an impact on analysis (e.g., population); and,
- Develop recommendations for further action.

Client meetings were conducted to discuss and review the methodology, including the procedures used to develop the estimates of billing discrepancies. During these discussions, we noted the procedures used to develop the estimates. Our review of the documentation included the document entitled “The Dual Billing Problems: An Estimate”, prepared by BDS. This document summarized the cause of each of the five discrepancies, providing a calculation of the total discrepancies accrued in the twenty months analyzed, and a projection of the total, overall discrepancy for the twelve-year period from 1986 to 1998.

Analysis of outputs included a verification of the calculations themselves by duplicating the calculations for specific months for each type of discrepancy. Using data provided by HRDC, we calculated selected monthly discrepancies to confirm data used in the error estimate. We also conducted an analysis of external factors that may have an impact on the estimate, such as changes in population over time.

Upon completion of the analysis, we prepared our findings and comments based on the quantity and quality of the data that was available for our review in the limited time frame available for this study.

3 Methodology and Findings of the CPP/QPP Study

3.1 Overview

Based on our preliminary review of the CPP/QPP billing discrepancy issue, the data tested by HRDC does not provide a statistical basis (e.g., margin of error, statement of statistical precision such as statistical significance) or a confidence level associated with the total estimated discrepancy as the data used in the calculation is not sample based. Measures of statistical precision are based on a number of parameters, including but not limited to the stipulation that data are selected using some type of sampling procedure, which implies a degree of random selection. Accordingly, the sample selected and results were prepared on a computational and common sense basis by HRDC.

For this analysis, the data selected by HRDC was over a 20-month period and did not fully represent the entire time period that the discrepancies are known to exist (1986 to 1998) because not all data points from this time period were analyzed. Therefore, this becomes a computational exercise rather than a statistical exercise. Accordingly, the focus of the analysis was on computational issues.

In this section of our report, we note the results of our methodology review, and comment on the estimates of each of the specific issues. For each discrepancy we note the specific methodology that was used to calculate the estimate and the analysis of the calculation.

In reviewing the procedures used to calculate the estimates, we noted that HRDC accounted for every discrepancy that occurred in each month of the analysis. For each discrepancy type, HRDC used physical records to calculate the total discrepancy for each month from all available data. Accordingly, a key assumption is that these discrepancies were exhaustive of all discrepancies in each month. If any data were missing from this, then these estimates and the analysis of them would not be accurate.

A key finding from the analysis is that the data used for this project display a strong time trend. In the analysis that follows, each of the five discrepancies is discussed separately, followed by the data graphs in Section 4 that captures and extrapolates the data provided by HRDC.

3.2 Estimate of Total and Projected Discrepancies

HRDC calculates the total amount resulting from these billing discrepancies over the 12-year period from 1986 to 1998 as \$4,936,667. After a preliminary review of the computation of the figures, the calculations have been adjusted to \$4,640,048, a difference of \$296,619. Data in Exhibit 1, below, show the total and projected discrepancies for each of the five issues as identified by HRDC. The following is an analysis of each separate billing discrepancy.

Exhibit 1: Total and Projected Amount for Each Billing Discrepancy

Issue	Dual Billing Issues	HRDC		Study Results		Differences	
		Discrepancy in sample (14-20 mths)	Estimated Discrepancy for 1986/98	Discrepancy in sample (14-20 mths)	Estimated Discrepancy for 1986/98	Discrepancy in sample (14-20 mths)	Estimated Discrepancy for 1986/98
1	Net Arrears	\$ 315,477	\$ 2,449,324	\$ 194,438	\$ 2,152,705	\$ 121,039	\$ 296,619
2	RRQ Withholds	\$ 503,914	\$ 3,023,482	\$ 503,914	\$ 3,023,482	\$ -	\$ -
	RRQ Withholds, Excluding DSB Benefits	\$ 361,312	\$ 722,625	\$ 361,312	\$ 722,625	\$ -	\$ -
3	Child Benefits	\$ (97,453)	\$ (1,258,764)	\$ (97,453)	\$ (1,258,764)	\$ -	\$ -
4	Addition Arrears	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
5	RRQ % > 100%	\$ -	\$ -	\$ -	\$ -	\$ -	\$ -
	Total Billing Error	\$ 1,083,250	\$ 4,936,667	\$ 962,211	\$ 4,640,048	\$ 121,039	\$ 296,619

3.3 Areas of Discrepancies

3.3.1 Dual payment arrears in CSC system for retirement and disability benefit

HRDC discovered that the CSC system switches the CPP and RRQ contribution percentages. For example, if the CPP value shows 75% and the RRQ value shows 25%, the system switches them and RRQ is billed 75%, while CPP is billed 25%. In cases where the original CPP value is higher than the RRQ value, RRQ is overbilled. In cases where the original CPP value is lower than the RRQ value, RRQ is underbilled.

Based on an analysis of 20-months of data (February 1996 to September 1997), the total net billing discrepancy due to this error was calculated to be \$315,477. Pro-rated over a one-year period, the annual discrepancy was estimated by HRDC to be \$189,625 by using the annualized average overbilling and underbilling amounts. Using this annual discrepancy estimate, the total discrepancy for the period in question (1986 through November 1998) would be \$2,449,324.

Our analysis of this data was based on the 14-months of data (July 1996 to September 1997, excluding January 1997), reported in the documentation provided by HRDC, as the data for the 20-months was not available for our study. Based on these 14-months, the total net billing discrepancy was calculated to be \$194,438, which when annualized becomes \$166,661. Using this annual error estimate, the total discrepancy for the period in question would be \$2,152,705.

Basing the calculation on 14-months instead of 20-months results in a difference in the calculation of the total error of \$296,619, a decrease of 12.1% from the estimate based on 20-months. This difference suggests that a high degree of variability in the CSC system data switching, therefore, it is necessary to use more data points to substantiate the calculation of this billing discrepancy.

3.3.2 Withholding rate for QPP

The RRQ billing process was entered onto the CSC system in December 1989. HRDC discovered that when the system calculates the amount from which the billing percentage must be taken, the RRQ withhold was not deducted automatically. Hence this has resulted in a billing discrepancy with the RRQ.

The calculation of the impact of these errors determined that the total resulting discrepancy, covering eight years from December 1989 to December 1997, was a net projected billing discrepancy of \$3,746,107. This total is derived from two calculations, one for a six-year and one from a two-year period.

For the first, the total discrepancy from 14-months of data was calculated as \$587,899, which when annualized becomes \$503,914. Since DSB benefits were manually adjusted beginning in 1996, this annual discrepancy was projected for the six years that did not include the manual adjustment (December 1989 through December 1995). The resulting total projected discrepancy was \$3,023,482.

For the second, HRDC's calculation of the discrepancy associated for the two-year time period in which DSB benefits are manually adjusted was \$722,625, based on \$361,312 per year for two years.

We noted these same values in the analysis of the data provided to us.

3.3.3 Dual payments to children in the CSC system

The CSC system has a problem with several Dual benefits paid to children (ORP and DCC). These benefits are either not billed to RRQ or are billed incorrectly. HRDC notes that in all cases the RRQ percentage present on CICS 1B2 is different from the one appearing on the Y0142 report. This problem has existed since the CSC system was created in 1986.

Based on 14-months of available data from July 1996 to September 1997, HRDC calculated the error as \$113,695. The annualized discrepancy becomes \$97,453, which when projected over the historical duration results in a total discrepancy of \$1,258,764. We noted these same values in the analysis of the data provided to us.

3.3.4 Addition error on report Y0142S (now Y0143)

HRDC discovered discrepancies in the total billings shown in the different columns of the Y0142S report. Essentially the problem was that the CSC did not calculate sub-totals correctly. The estimated total discrepancy resulting from this problem was \$337,000. However, in a co-ordinated effort between HRDC and RRQ to discover the cause of this problem, it was determined that no actual billing discrepancy had occurred because RRQ used calculations based on the totals, not the sub-totals. Therefore, HRDC has concluded that there is no discrepancy to be resolved on this issue.

This billing discrepancy study has excluded any review of the CSC system, including any specific inputs or computer codes, therefore, we are not able to comment on this analysis.

3.3.5 Percentage error for QPP in the CSC system

In reviewing the Y0142R register for the different dual billing problems, HRDC discovered that it was possible for RRQ to have a percentage higher than 100%. The investigation of this problem indicated that the CSC system did not have any edits that would restrict the user from entering values greater than 100%. Two discrepancies were found, which resulted in corrective action. Billing adjustments totalling \$285,339 were made to RRQ. HRDC indicated that all existing discrepancies of this nature have been corrected, and RRQ credited accordingly.

To correct this problem, edits were programmed to restrict the value of this percentage to 99.99%.

4 Analysis of Data

The following section firstly analyzes the changes in population for Canada and Quebec over the time period of analysis and, secondly analyzes the overall trend in billing during the months analyzed by HRDC.

4.1 Population Changes

The following table analyzes the changes in population for Canada and Quebec over the time period of analysis. From 1986 to 1998, the total Quebec population as a proportion of the total population of Canada decreased from 25.81% to 24.21%. Also the Quebec population has increased at a lower rate than the total Canadian population, as noted in the table below.

Exhibit 2: Comparison of Population Changes – 1986 to 1998 *

		All ages	65 +
1986	Canada	25,309,330	2,697,580
	Quebec	6,532,460	650,635
	Quebec % of Total	25.81%	24.12%
1991	Canada	27,296,855	3,169,970
	Quebec	6,895,960	770,920
	Quebec % of Total	25.26%	24.32%
	Canada 86-91 Percentage Increase	7.85%	17.51%
	Quebec 86-91 Percentage Increase	5.56%	18.49%
1996	Canada	28,846,760	3,527,845
	Quebec	7,138,795	860,705
	Quebec % of Total	24.75%	24.40%
	Canada 91-96 Percentage Increase	5.68%	11.29%
	Quebec 91-96 Percentage Increase	3.52%	11.65%
1998	Canada	30,301,185	3,737,072
	Quebec	7,334,502	911,523
	Quebec % of Total	24.21%	24.39%
	Canada 96-98 Percentage Increase	5.04%	5.93%
	Quebec 96-98 Percentage Increase	2.74%	5.90%

* Source: Statistics Canada

Traditionally, Canadians begin claiming pension benefits between the ages of 60 and 65. A potential impact on billing estimates, therefore, is the pension-eligible population. Analysis of population trends between 1986 and 1998 shows that the Quebec population aged 65+ as a proportion of the Canadian population aged 65+ increased from 24.12% to 24.39%.

Another component of population change is provincial migration. Statistics Canada reports that, over time, there has been a net decrease in Quebec’s population due to inter-provincial migration, and it is concentrated in the working age population (18 to 65). In combination, changes in population may have an important impact on this dual billing issue, which requires a more in depth analysis.

4.2 Billing Trends

In the series of charts that follow, data illustrate that there is a clear time trend in the total billing data. For example, in Exhibit 3, below, the three billing discrepancies for which corrections still need to be made demonstrate trends over the 14-month period. Slight increases over time can be observed in the Arrears Underbilled and the DCC discrepancies, while decreases can be observed in the remaining discrepancies, although these decreases are more noticeable in the RRQ Withhold factor.

Visual inspection of these data suggests that, projected backward, each discrepancy may be significantly different at the time of its first occurrence (see Exhibit 4 for Projection of Errors to Point of Occurrence). One of the key assumptions of this is that the conditions that exist during the months these data were analyzed remain constant over the entire historical period.

Exhibit 3: Comparison of Errors Over 14-Month Study Period

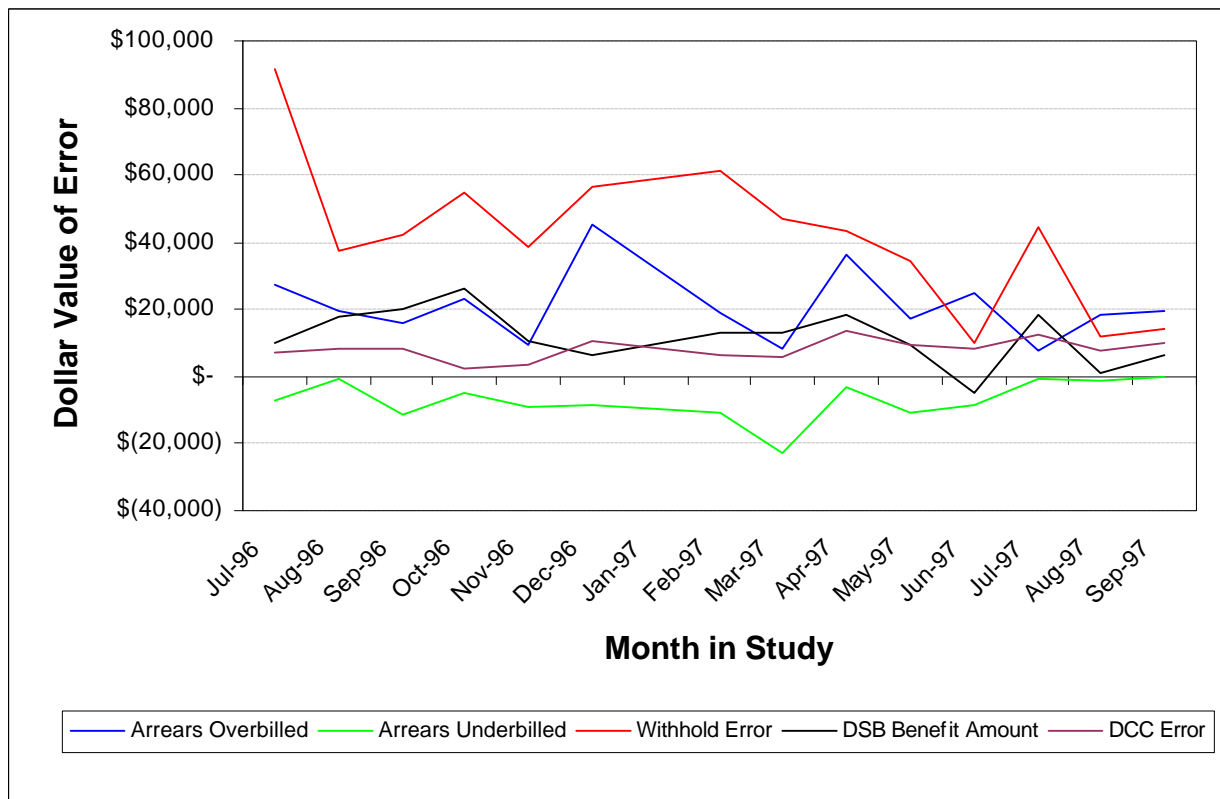
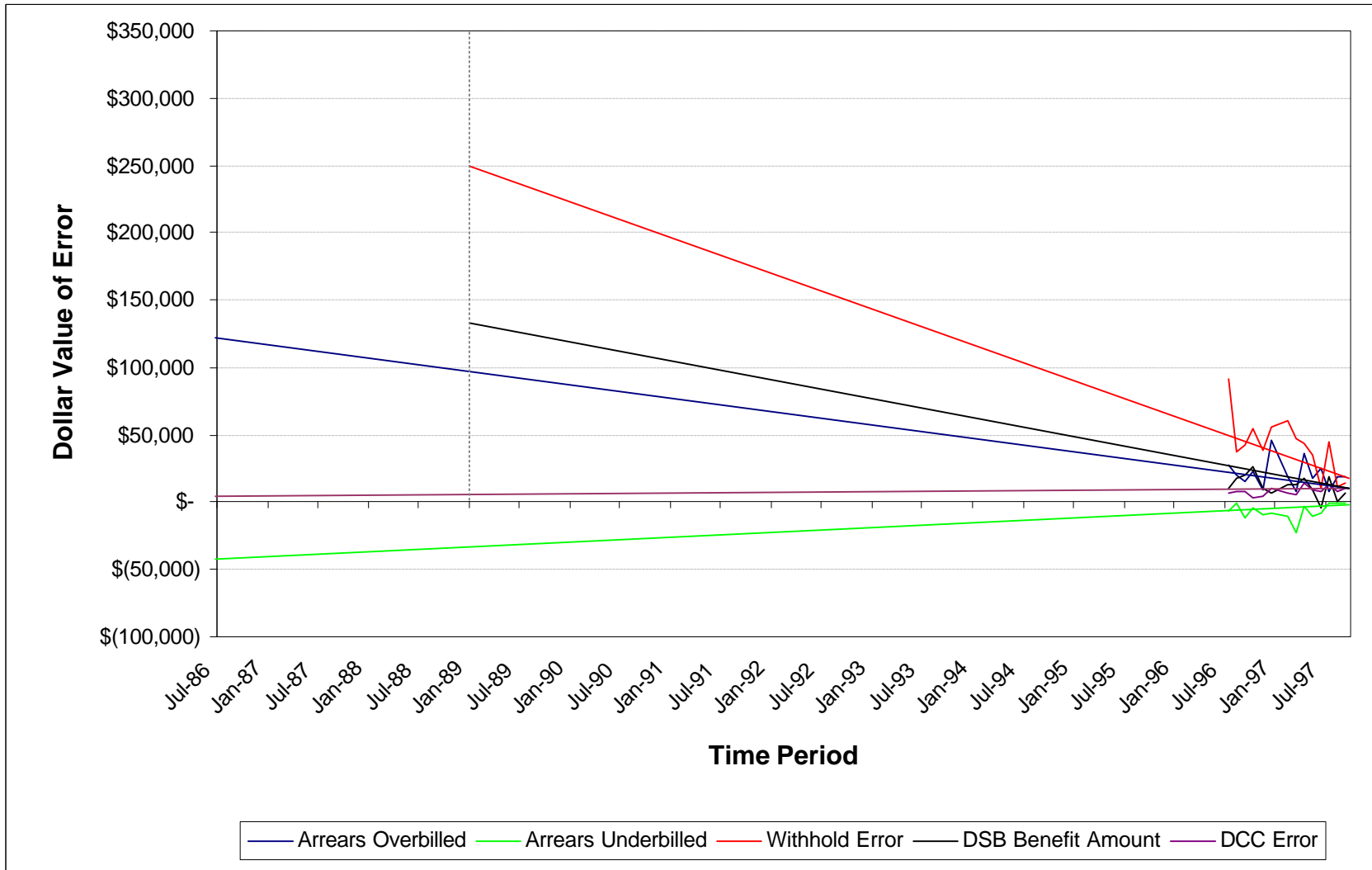
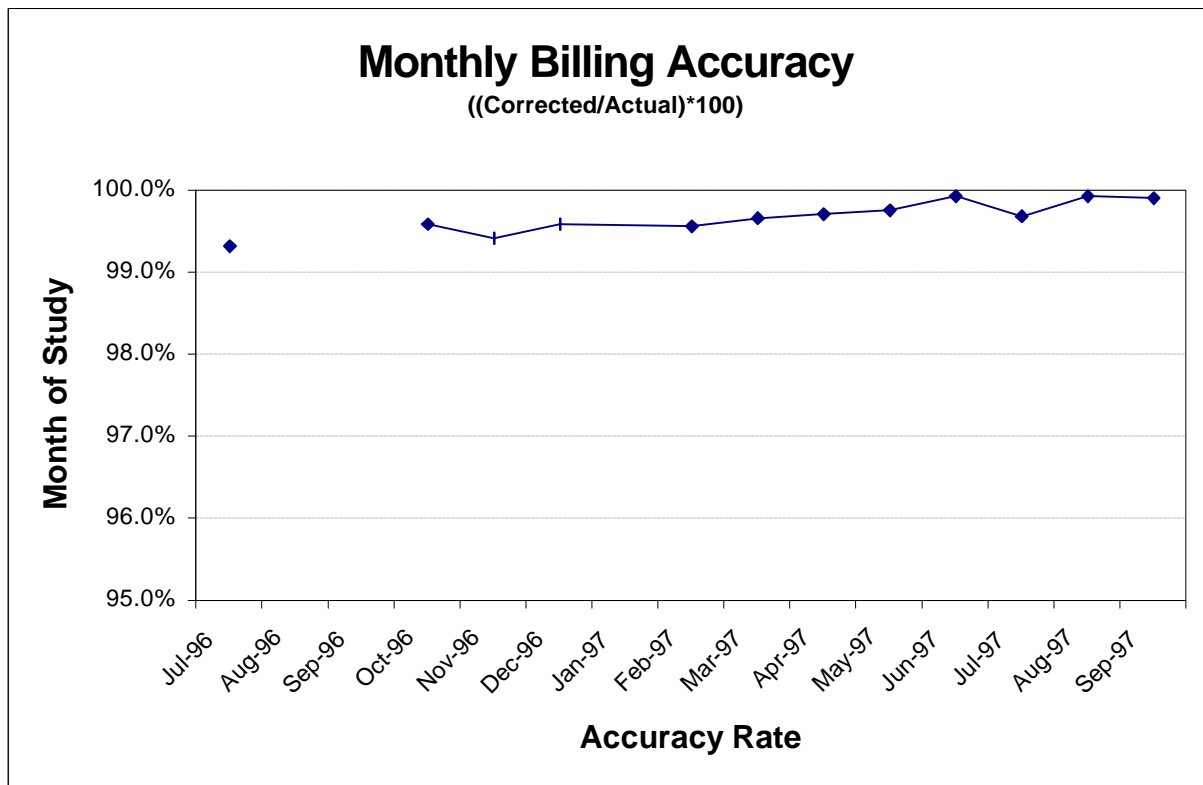


Exhibit 4: Backward Projection of Errors to Point of Occurrence



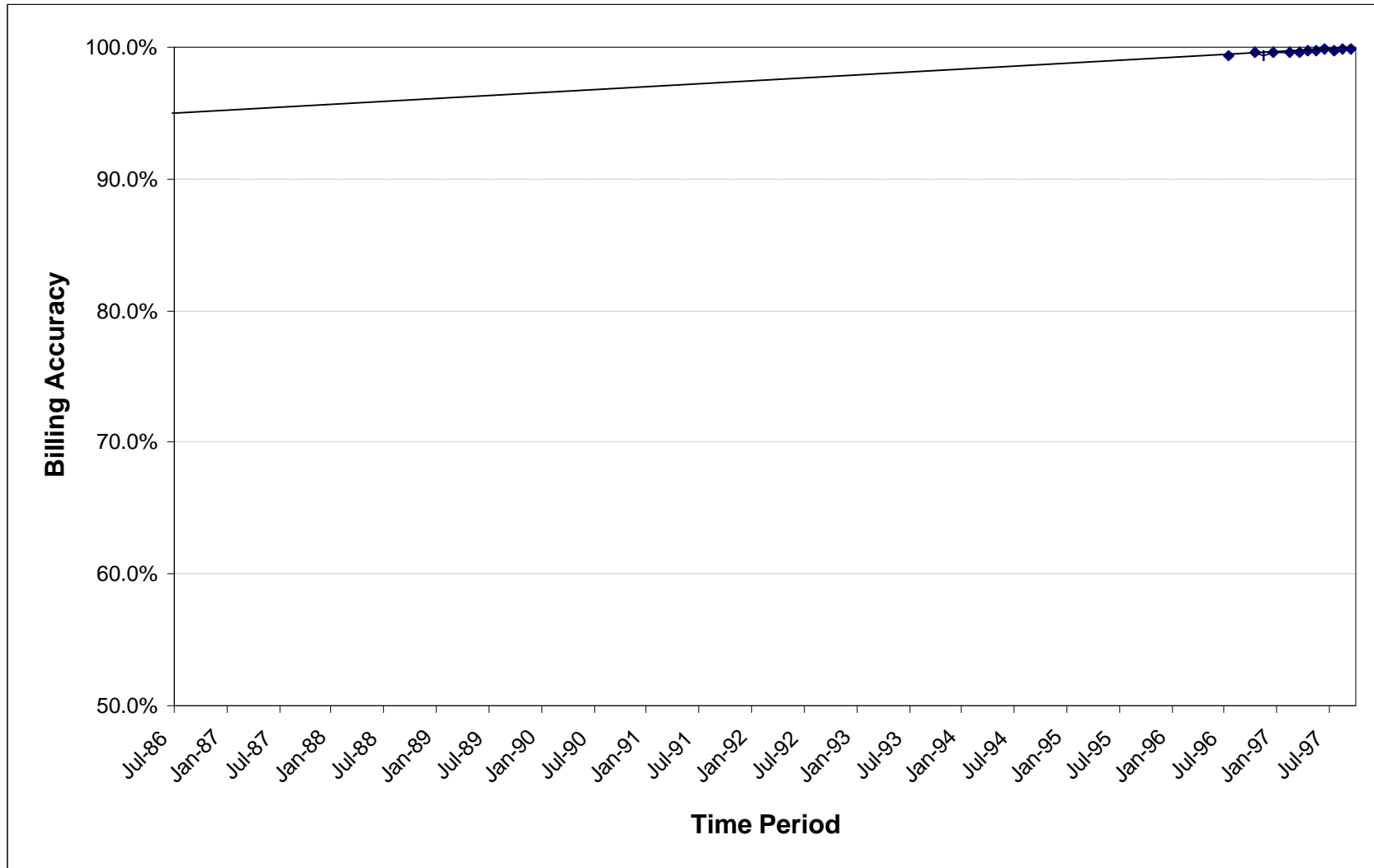
We also examined the total billings each month, both actual and corrected. This analysis also indicates a strong time trend. Exhibit 5, below, illustrates the increase in the accuracy of total billings to RRQ each month. The billing accuracy is determined by the corrected billing as a percentage of the actual billing.

Exhibit 5: Accuracy of Total RRQ Billing Over 14-Month Study Period



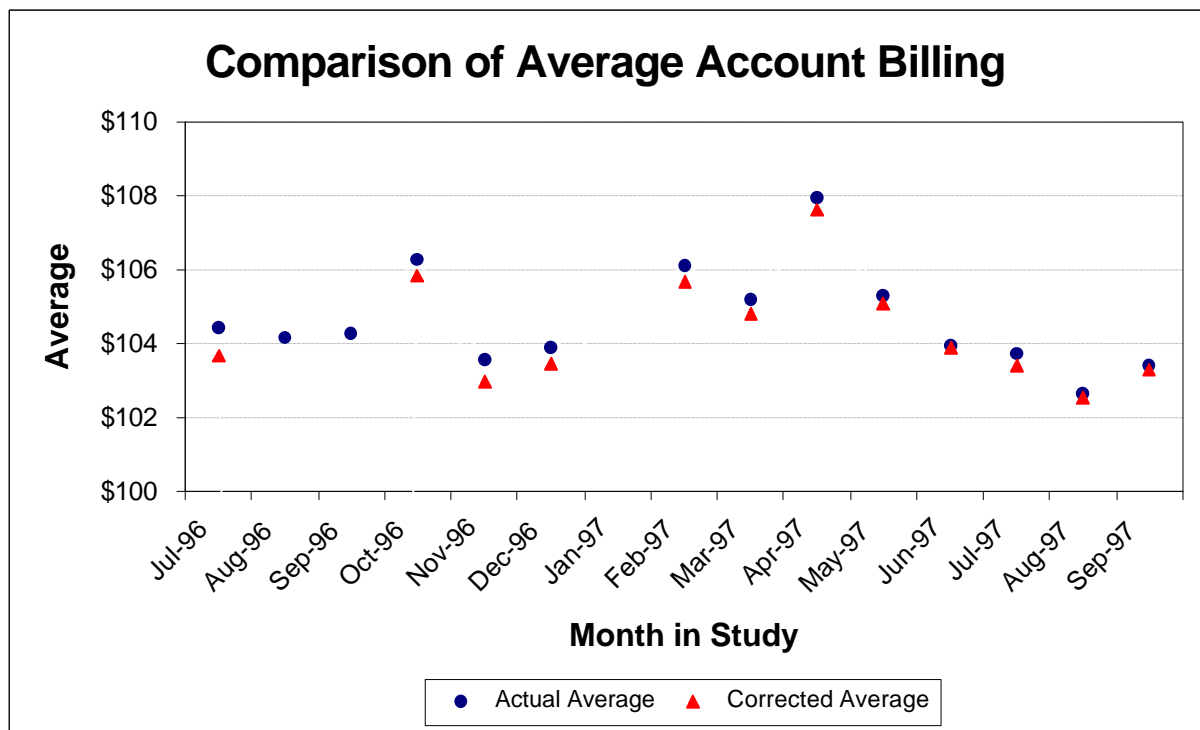
Visual inspection for this component also suggests that a backward extrapolation would result in a decrease in the accuracy of corrections over time. The projected accuracy in 1986 would be approximately 95.2%, compared to an average of 99.7% accuracy during the 14-months from which data were extrapolated (see Exhibit 6, on the following page for the “Backward Projection in Monthly Billing Accuracy to 1986”). Again, more time would be required to conduct a more detailed analysis of this issue if the sample size was expanded.

Exhibit 6: Backward Projection in Monthly Billing Accuracy to 1986



Lastly, we examined the average monthly account billing, actual and corrected. This calculation is a measure of the actual and corrected billings as a proportion of the total accounts for each month of the analysis. Data in Exhibit 7, below, show that the difference between the actual average monthly billing and the corrected average monthly billing has decreased over the 14 months from which data were extrapolated. This suggests that differences between actual and corrections back to 1986 may be different from what has occurred from July 1996 to September 1997.

Exhibit 7: Actual vs. Corrected Average Monthly Account Billing



In summary, the time trend indicates two potential outcomes. First, if the linearity in these trends holds over time, then the total correction associated with these discrepancies may be larger than the estimate calculated by HRDC. However, it is not possible to determine the actual amount without further analysis.

Second, the 14-months of data used for this current analysis contains multiple occurrences of summer months (July, August, September) and only a single occurrence of all other months, and excludes January altogether. It is possible that seasonality exists in these discrepancies that has not been discovered here. For example, due to workloads and available resources it is possible that a pattern may exist in processing of records. Staff may be unavailable to process data during the latter months of the fiscal year (December through March), resulting in more records being processed during summer months (June through August) when workloads tend to be somewhat less than normal. This is one potential explanation for differences in discrepancies between the different times of the year.

Again, without further analysis, it is not possible to determine the potential impact of seasonality in the discrepancies themselves.

5 Summary of Analysis

The above study was to provide HRDC with an analysis of the methodology and samples used for the projected estimate of the historical billing discrepancies by CPP to RRQ. Our review and analysis has shown that the billing discrepancies estimated by HRDC have limitations based on the sample methodology. Analysis also suggests that a time series effect likely exists in these data, which warrants further analysis. Accordingly, our comments focus on future analysis of these billing discrepancies. However, before presenting our comments, it is important to state two assumptions that have underscored our analysis and findings.

5.1 Analytical Assumptions

The first of the assumptions is that a time trend appears to be evident in the overall data. Total billings and accuracy of corrected billings relative to actual billings demonstrates a linear trend for the months examined. If this linear trend is constant for the entire period that the discrepancies have existed, then amounts will be greater back in time. This would suggest that the estimate developed by HRDC is biased because it is based on more current data, which tends to be less prone to discrepancies.

The second assumption is that some seasonality exists in the existence of errors from month to month. Analysis indicates that there is a relationship between data from summer months compared to those from the rest of the year. The clarity of this pattern varies by billing discrepancy, but it does appear that this factor requires further investigation.

5.2 Alternatives for Improving Sample Selection

5.2.1 Confirm Existence of Time Trend

The time trend identified in the above analysis is based on a period of 14-months. Since many of these discrepancies have existed since 1986, at least 144-months may be affected. Therefore, it is necessary to confirm the existence of the time trend back in time because conditions may exist prior to July 1996 (i.e. the start date of the sample period) that cause the time trend to disappear. For example, there may be economical and/or legislative changes which may have an impact on these issues. However, it is necessary to confirm the linearity of this trend over time

Confirmation of the time trend can be accomplished by analyzing consecutive months of data (periods of 12-months) from multiple points in the time period. At a minimum, we recommend that the first twelve months from 1986 should be used to make a comparison to the trend observed in 1996 to 1997. If the trend can be confirmed based on the beginning and end points, then the discrepancy can be calculated with greater accuracy. Alternatively, multiple 12-month periods could be selected from the 12-year period to examine a more robust series of data. Any such analysis will involve a great deal of staff resources due to the large number of transactions in a month/year.

5.2.2 Confirm Existence of Seasonality in the Data

Similar to the time trend, it is important to confirm the existence, and then the impact of seasonality in these errors. The total estimated discrepancy produced by HRDC has not been based on a balanced selection of months or seasonal periods. Therefore, it cannot be assumed that seasonal shifts in claims or other factors do not have an impact on the calculation of the billing discrepancies.

Confirmation of seasonality can be accomplished by collecting a sample of months from the 12-year period. This sample would need to be systematically conducted, to ensure coverage of months and years. It would be necessary to collect at least 48 data points for this analysis because without consecutive data points as in 5.2.1 above, the effect of seasonality will not be easy to determine. This could be accomplished by selecting four consecutive 12-month periods, or by selecting a sample of four months from each of the twelve years to be analyzed. Note that the checking of the samples will involve a great deal of staff resources to perform the collection of data.

5.3 Key Considerations

Based on this CPP/QPP study of billing discrepancies, more analysis is required. The issues that need to be addressed more fully, include, but are not limited to the following:

- More data elements need to be added to this analysis. In calculating errors in Arrears over 14 and 20-month periods, the difference was 12% (section 3.3.1). This is a large enough difference to suggest that more data elements should be used in the analysis.
- This is a computational issue, not a sampling issue. Due to monthly variability in discrepancies, both positive and negative, a higher degree of accuracy is needed to clearly state the exact nature of the total discrepancy.
- The cost/benefit issue has to be considered as any increase in sample size will involve a great deal of time and effort, in the way of staffing resources to manually perform the sample checks.

Our comments are based on the preliminary results of our analysis during the period from September 16-23, 1999, and based on the two key assumptions noted in 5.1 above. Also, any additional research and sample checks would be affected by the availability of comprehensive historical data dating back to 1986 for those issues that occurred at that point in time. If such data is not available, then it may not be possible to conduct the necessary confirmation analyses that would give a better sample base.