

# Interpretation of the data quality scores applied to REF submissions

10 May 2006

Prepared by REF Analyst  
Enquiries: [m.govuzela@medicalschemes.com](mailto:m.govuzela@medicalschemes.com)  
Prepared for the Council of Medical Schemes



COUNCIL FOR MEDICAL SCHEMES

## Table of Contents

<b>Executive summary .....</b>	<b>iii</b>
<b>1. Introduction .....</b>	<b>1</b>
1.1 Purpose of Quality Scores .....	1
1.2 Automated data quality scores developed .....	1
1.3 Testing of method for bias .....	1
<b>2. Publication of scheme specific information on the website.....</b>	<b>3</b>
<b>3. Automated data quality score .....</b>	<b>4</b>
3.1. Total REF = Total SR .....	4
3.3. Age 85+ reasonable score.....	5
3.4. Age under one reasonable .....	7
3.5 NON cases submitted.....	11
3.6 HIV cases submitted.....	13
3.7 Total CDL count.....	15
3.8 Multiple CDL count .....	17
3.9 Total Score .....	19

## Executive summary

Medical schemes provide consolidated information on the age and chronic diseases profile to the Council for Medical Schemes in the form of quarterly submissions to the Risk Equalisation Fund (REF). The format for the data for the shadow period is known as the REF Grids. These provide for the number of beneficiaries by gender and in five-year age bands. The number of beneficiaries is shown separately for each of the chronic diseases in the PMBs; those with multiple conditions; those with none of the chronic conditions; the number of maternity cases and the number being treated with anti-retroviral treatment for HIV/AIDS.

The automated data quality score was developed to quantify/measure deviations from expected industry profiles and Statutory Returns data. The expected prevalence of chronic diseases published with the REF Contribution Table is used as a benchmark. This replaces the manual data quality index used in Q1 2005. The following 8 items are checked for each scheme in an automated manner:

1. Degree of correlation between REF submissions and the Statutory Returns data.
2. Changes in number of beneficiaries in the 85+ age category.
3. Changes in number of beneficiaries in the below 1 year age category.
4. Number of maternity cases submitted
5. The number of beneficiaries without any chronic diseases submitted.
6. The number of HIV cases submitted.
7. Chronic diseases in line with known prevalence and age-profile information.
8. Multiple chronic diseases in relation to the expected for this age profile.

Based on the outcome of these evaluations, each scheme is assigned a score in accordance with the quality of the data. The data quality scores will range from 0 to 10, with zero reflecting good quality data and 10 reflecting data with many problems. The automated data quality scores were evaluated against the manual evaluation of data submissions. In additions, these cores were tested for bias against scheme size (number of beneficiaries) and scheme risk (net payments to REF). No relationship between the data scores and the scheme's risk or the size to the schemes could be demonstrated.

The purpose of this report is to demonstrate the progress made in improving the quality of data and to identify areas that are still lagging behind. Schemes will also be able to see how they fare in relation to other schemes. Scheme ranks will also be published to assist in this regard. Schemes are also to ascertain that the reported deviations are valid

## Observations

Table 1: Median scores by quarter per test

Test	Median Score		
	Quarter 1	Quarter 2	Quarter 3
Degree of correlation between REF submissions and the Statutory Returns data.	1.02	0.90	0.74
Changes in number of beneficiaries in the 85+ age category.	3.10	2.50	2.10
Changes in number of beneficiaries in the below 1 year age category.	3.25	3.71	3.19
Number of maternity cases submitted	1.42	1.36	1.28
The number of beneficiaries without any chronic diseases submitted.	2.75	3.34	3.03
The number of HIV cases submitted.	5.08	5.09	5.04
Chronic diseases in line with known prevalence and age-profile information.	6.36	6.44	6.49
Multiple chronic diseases in relation to the expected for this age profile	5.69	5.83	6.34
Total Score	3.48	3.58	3.41

The summary of the results of the data evaluations are presented in the table above. There has been some improvement in the quality of data in the following categories: “Degree of correlation between REF submissions and the Statutory Returns data”; “Changes in number of beneficiaries in the 85+ age category”; “Changes in number of beneficiaries in the below 1 year age category” and “Number of maternity cases submitted”. No major changes were observed in the quality of data for the number of beneficiaries without any chronic diseases submitted. The data quality scores have remained high HIV, chronic and multiple chronic disease cases submitted. The median total score has slightly improved from 3.48 in Quarter 1 to 3.41 in Quarter 3.

Analysis performed on Quarter 4 data will be published later. The results of this analysis show a major improvement in the quality of submitted by medical schemes.

# 1. Introduction

## 1.1 Purpose of Quality Scores

The purpose of automated quality scores is to quantify deviation of reported data from expected industry profiles and Statutory Returns data. The scores range from zero to 10, zero indicating perfect data, and 10 indicating data with many problems. The automated scoring system replaces the previously reported on fault index. The automated data quality scores will be used to detect irregular data, gaming or fraud before effecting financial transfers.

## 1.2 Automated data quality scores developed

The scores were checked manually for each scheme to ensure that they correctly describe scheme's data quality. The scores were found to be useful and consistent with the quality of reported data. Submissions with fewer data problems were correctly assigned low scores and high scores were given to submissions with many data problems

## 1.3 Testing of method for bias

Correlation analysis was performed to test for bias for each month. No significant correlation was found between automated scores and scheme risk (net transfer and net payment). The automated scores were also shown not to be biased towards large schemes. The  $R^2$  values are reported in the table below.

Table 1:  $R^2$  values for correlation analysis

Month	Scheme Risk vs. Automated Score		Size vs. Automated Score
	Net payment to REF	Net payment to REF pbpm	
Jan	0.0930	0.0049	0.0041
Feb	0.0002	0.0006	0.0028
Mar	0.0004	0.0002	0.0028
Apr	0.0221	0.0134	0.0102
May	0.0210	0.0000	0.0091
Jun	0.0068	0.0001	0.0133
Jul	0.0018	0.0005	0.0020
Aug	0.0002	0.0273	0.0024
Sep	0.0006	0.0207	0.0008

The correlation analysis graphs for the month of September are shown below. All the three correlation graphs show that there is no significant correlation between the developed scores or number of beneficiaries in a scheme and the risk experienced by schemes.

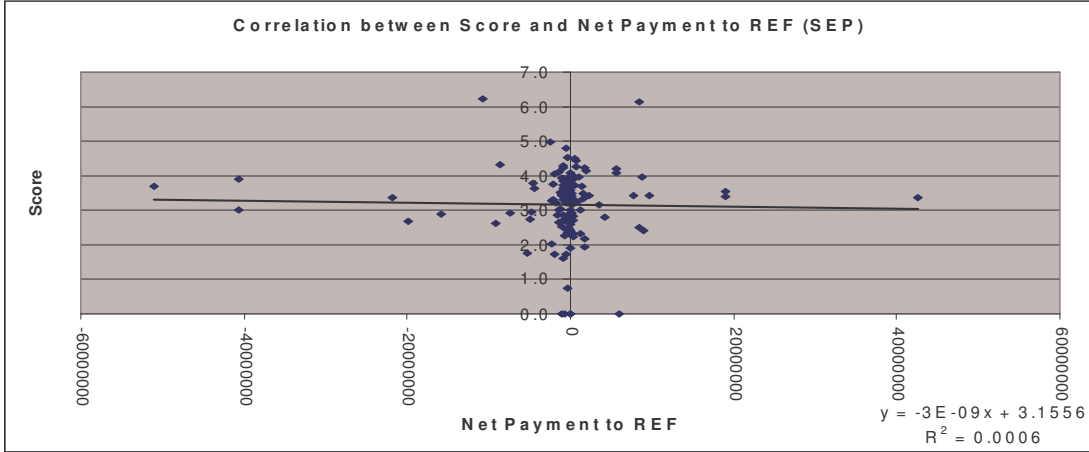


Figure 1: Correlation between automated score and Net Payment to REF

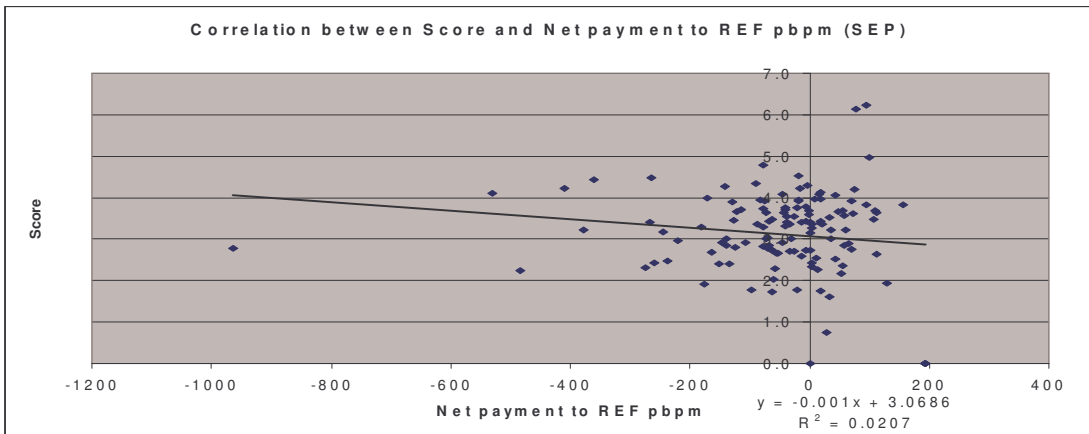


Figure 2: Correlation between automated score and Net Payment to REF pbpm

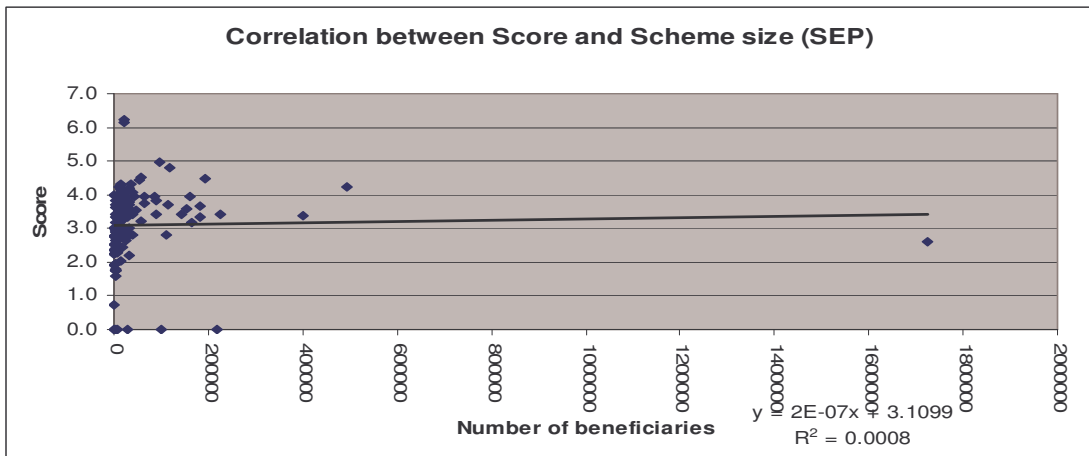


Figure 3: Correlation between automated score and Scheme Size

## 2. Publication of scheme specific information on the website

The scheme specific automated data quality scores will be published on the council website. Schemes are encouraged to ascertain that the reported deviations are valid. The scores will be accompanied by “Scheme Rank”. The rank gives a scheme’s position in relation to other schemes with regards to data quality. The scheme with the lowest value (1) has the best data quality scores.

The benchmark used in the calculation of these scores is available on the council website ([www.medicalschemes.com](http://www.medicalschemes.com), under REF Reports). The Assumed REF Grid Count (REF contribution Table 2006 vFinal) gives the expected number of cases for each CDL per thousand beneficiaries per month. For maternity, the annual rate is given.

The screenshot shows an Excel spreadsheet titled "REF Contribution Table 2006 vFinal". The active cell is D9, containing the value 0.069. The spreadsheet is divided into sections:

- Industry Assumptions:**
  - Cell D1: REF Contribution Table 2006
  - Cell D2: in the cell for the period, per 1,000 exposed
  - Cell H2: Base Period
  - Cell I2: HIV/AIDS expected in 2006, all other diseases from REF Grids Q3 2005
- Chronic Disease Rates:**

	AST	BCE	BMD	CHF	CMY	COP	CRF	CSD	DBI	DM1	DM2	DY
3	5.624	0.022	0.000	0.076	0.043	0.087	0.076	0.000	0.011	0.043	0.000	
4	23.412	0.004	0.004	0.081	0.016	0.069	0.049	0.004	0.004	0.275	0.004	
5	32.657	0.015	0.053	0.077	0.022	0.071	0.087	0.003	0.012	0.724	0.040	
6	26.368	0.024	0.124	0.050	0.009	0.050	0.094	0.015	0.018	1.090	0.094	
7	17.831	0.015	0.453	0.077	0.034	0.028	0.120	0.083	0.015	1.443	0.231	
8	11.769	0.021	0.911	0.076	0.055	0.055	0.227	0.181	0.029	1.957	0.411	
9	10.405	0.007	1.110	0.100	0.093	0.081	0.285	0.374	0.022	2.264	1.077	
10	12.943	0.012	1.213	0.204	0.166	0.125	0.338	0.408	0.020	2.697	2.207	
11	14.649	0.021	1.375	0.505	0.248	0.230	0.433	0.415	0.024	3.472	4.853	

Figure 4: REF Contribution Table 2006 vFinal

### 3. Automated data quality score

#### 3.1. Total REF = Total SR

##### 3.1.1. Definition

This data quality test measures the degree of correlation between scheme totals in REF submissions and the scheme totals in Statutory Returns data. A percentage of REF submission to SR submission is calculated to estimate the level of correlation. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the REF and SR datasets are equal, the score will be zero. The score gets worse as the difference between the totals in the two data sets increases. The maximum score possible is 10. This score is can be calculated fro each month's data in a quarterly submission.

#### 3.2. Average "Total REF = Total SR" score by month

The degree of correlation between REF submissions and Statutory Returns has remained constant with slight improvements in the second and third quarters. The average score in September is less than one. This translates to an average deviation of about one per cent.

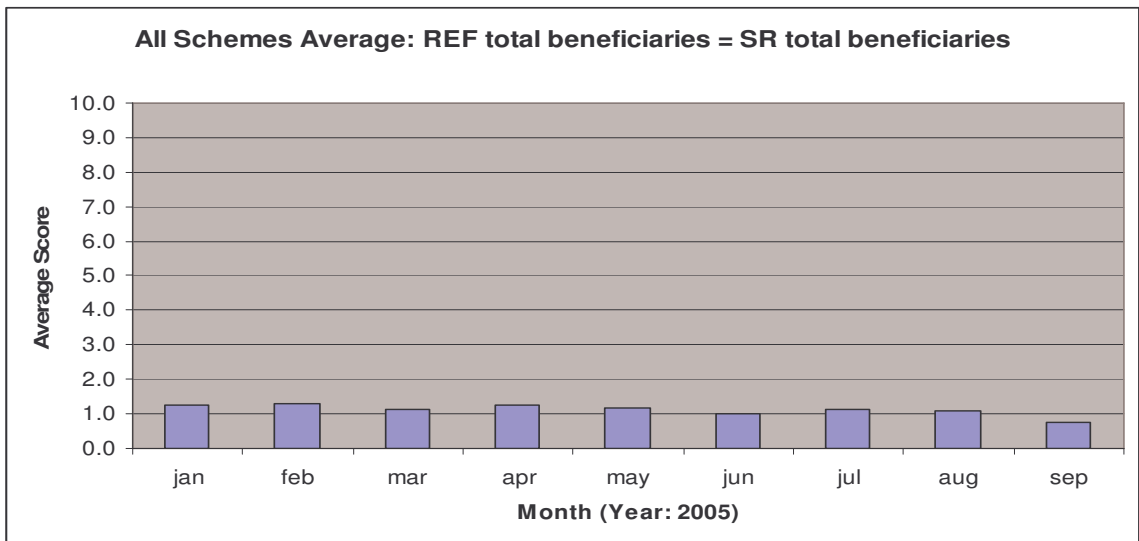


Figure 5: Average score for the degree of correlation between REF submissions and the Statutory Returns data.

### 3.2.1. Distribution by quarter

Most schemes have scores below 2 (about 2% deviation) for the degree of correlation between REF submissions and the Statutory Returns data. The median score 1.02, 0.9 and 0.74 for quarter 1, 2 and 3 respectively. This shows some improvement in the quality of data in quarter 3.

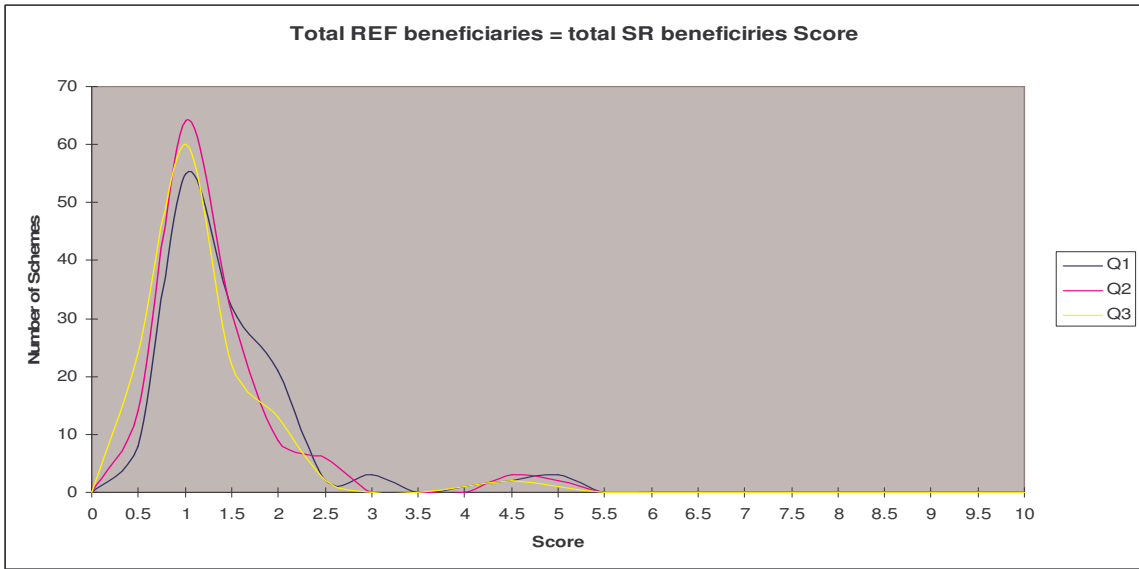


Figure 6: The distribution of scores for the degree of correlation between REF submissions and the Statutory Returns data.

### 3.3. Age 85+ reasonable score

#### 3.3.1. Definition

This score estimates changes in the number of beneficiaries in the 85+ age category. It is expected that the ratio of beneficiaries in the 85+ age category to the total number of beneficiaries in a scheme should be the same in both the REF and SR submissions. A percentage of the two ratios is calculated to estimate the level of correlation. The score is obtained by finding a cube root of the percentage difference between the resulting and the expected percentage (100%). If the ratios in both the REF and SR datasets are the same, the score will be zero. The score gets worse as the difference between the ratios in the two data sets increases. The maximum score possible is 10. This score is can only be calculated from each quarter's data in a quarterly submission.

### 3.3.2. Average “Age 85+ reasonable” score by month

The degree of correlation between number of beneficiaries in the 85plus age band for REF submissions and Statutory Returns has improved markedly from quarter 1 to quarter 3. Number of beneficiaries in the 85plus age band as reported in the last month of the quarter in the Statutory Returns was used to compute scores for the months in which this data is not given. The average score in September is 2.1. This translates to an average deviation of about 2%.

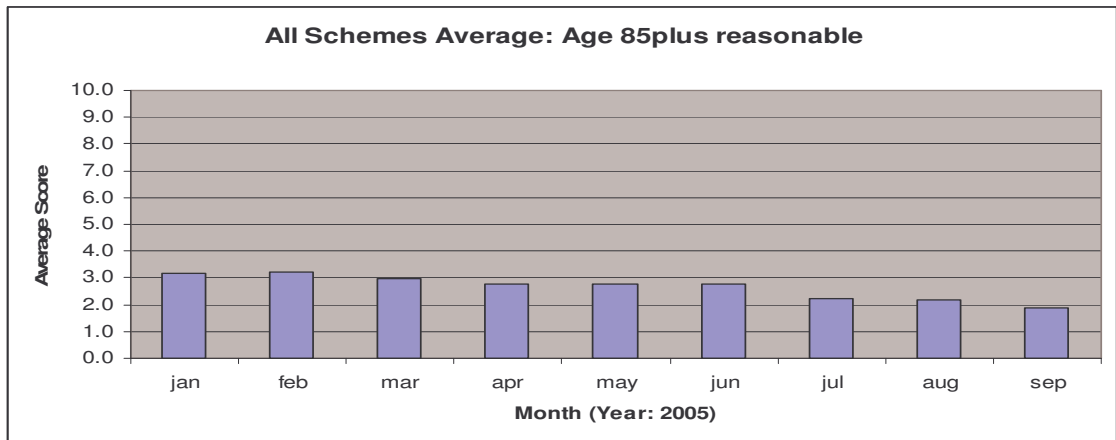


Figure 7: Average score from the changes in number of beneficiaries in the 85+ age category.

### 3.3.3. Distribution by Quarter

The median score for the data reported for the 85plus age band has improved from 3.1 in Quarter 1 to 2.1 in Quarter 3.

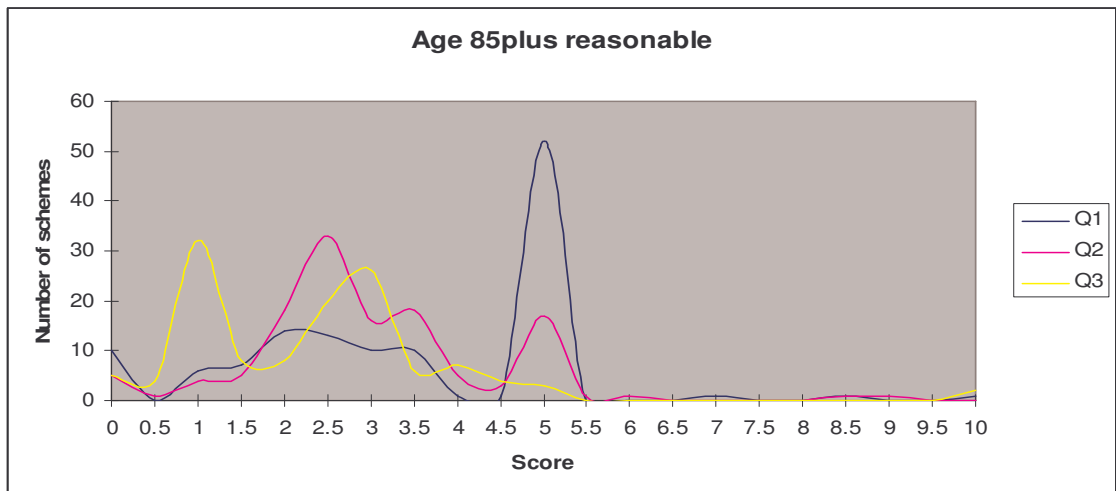


Figure 8: Distribution of scores for the changes in number of beneficiaries in the 85+ age category.

### 3.4. Age under one reasonable

#### 3.4.1. Definition

This score estimates changes in the number of beneficiaries in the below 1 year age category. It is expected that the ratio of beneficiaries in the below 1 year age category to the total number of beneficiaries in a scheme should be the same in both the REF and SR submissions. A percentage of the two ratios is calculated to estimate the level of correlation. The score is obtained by finding a cube root of the percentage difference between the resulting and the expected percentage (100%). If the ratios in both the REF and SR datasets are the same, the score will be zero. The score gets worse as the difference between the ratios in the two data sets increases. The maximum score possible is 10. This score is can only be calculated fro each quarter's data in a quarterly submission.

#### 3.4.2. Average "Age Under one reasonable" score by month

The degree of correlation between number of beneficiaries in the below 1 age band for REF submissions and Statutory Returns has improved markedly from quarter 1 to quarter 3. Number of beneficiaries in the below 1 age band as reported in the last month of the quarter in the Statutory Returns was used to compute scores for the months in which this data is not given. The average score in September is 2.4.

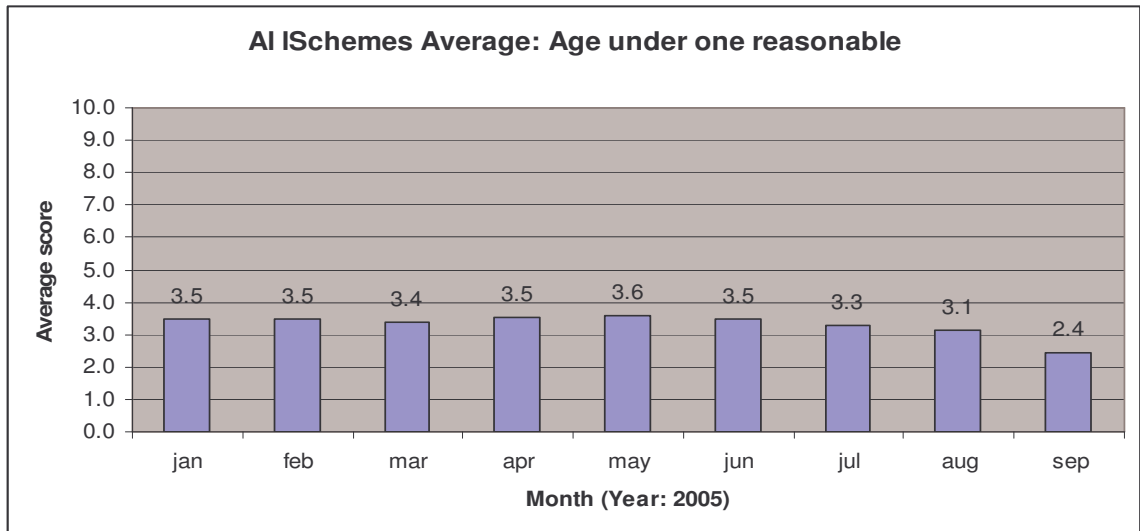


Figure 9: Average score for the changes in number of beneficiaries in the below 1 year age category

### 3.4.3. Distribution by Quarter

The graph below shows that there has been some improvement on the data submitted in the below 1 age category. Some schemes have scored badly in Quarter 3 compared to Quarter 1 and Quarter 2. The median score is 3.3, 3.7 and 3.2 in Quarters 1, 2 and 3 respectively.



Figure 10: Distribution of scores for the changes in number of beneficiaries in the below 1 year age category

### 3.5. Maternity cases score

#### 3.5.1. Definition

The maternity cases score measures the reasonableness of maternity cases submitted against the expected industry profile for the scheme. The final score is obtained by finding an average of scores in a) and b) below:

- a) A percentage of total reported to the total expected maternity cases is calculated. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the reported and expected totals are equal, the score will be zero.
- b) If the reported data for each age bands is within the expected range, a score of zero is given. If not, the score is calculated by finding the cube root of the sum of percentage deviations from upper or lower limits (assumed deviation from expected values) of reported data for each age band. A deviation of zero given if the expected and reported cases are less than ten in an age band. A score which is greater than zero indicates that the reported cases are less than the lower limit or higher than the upper limit.

#### 3.5.2. Average “Maternity cases” scores by month

The number of maternity cases was high in the first quarter. The quality of maternity data has improved in the second and third quarter. The average score in the first quarter is 2 and 1.5 in the second and third quarter.

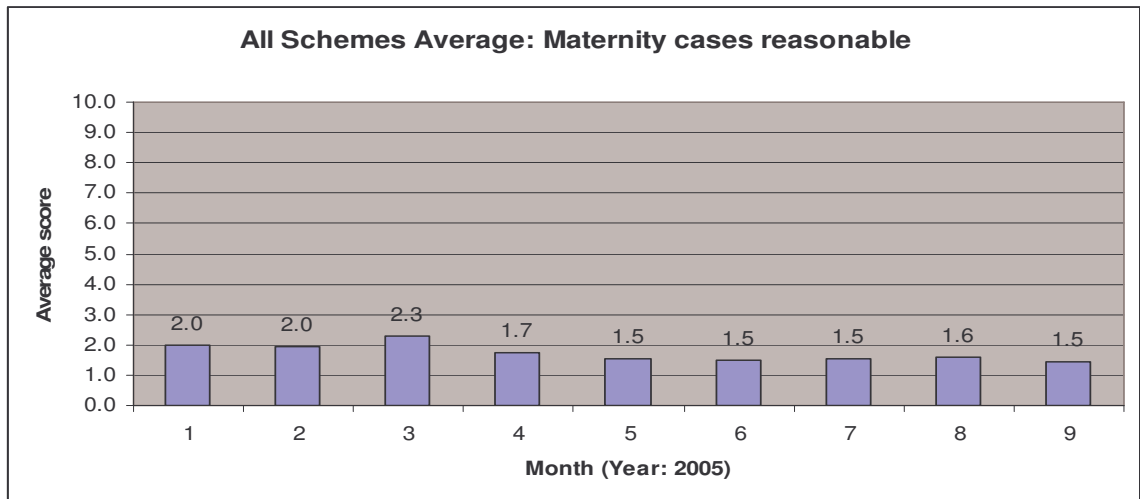


Figure 11: Average score for the number of maternity cases submitted

### 3.5.3. Distribution by Quarter

The median scores for the three quarters are 1.4, 1.4 and 1.3 for quarters 1, 2 and 3 respectively. There has been a slight improvement in the quality of maternity data from quarters 1 to 3.

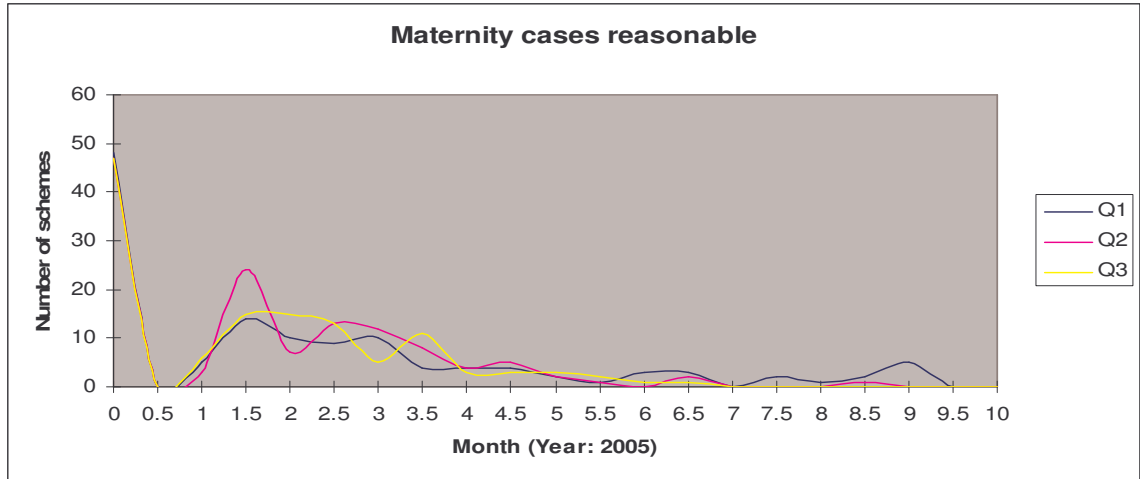


Figure 12: Distribution of scores for the number of maternity cases submitted

### 3.5 NON cases submitted

#### 3.5.1 Definition

The NON cases score measures the reasonableness of number of beneficiaries without any chronic diseases submitted against the expected industry profile for the scheme. The final score is obtained by finding an average of scores in a) and b) below:

- c) A percentage of total reported to the total expected NON cases is calculated. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the reported and expected totals are equal, the score will be zero.
- d) If the reported data for each age bands is within the expected range, a score of zero is given. If not, the score is calculated by finding the cube root of the sum of percentage deviations from upper or lower limits (assumed deviation from expected values) of reported data for each age band. A deviation of zero given if the expected and reported cases are less than ten in an age band. A score which is greater than zero indicates that the reported cases are less than the lower limit or higher than the upper limit.

#### 3.5.2 Average “NON cases submitted” scores by month

The average data quality score for the number of NON cases submitted has been stable form quarters 1 to 3 with an average of 2.7, 3.2 and 3.0 in quarters 1, 2 and 3, respectively.

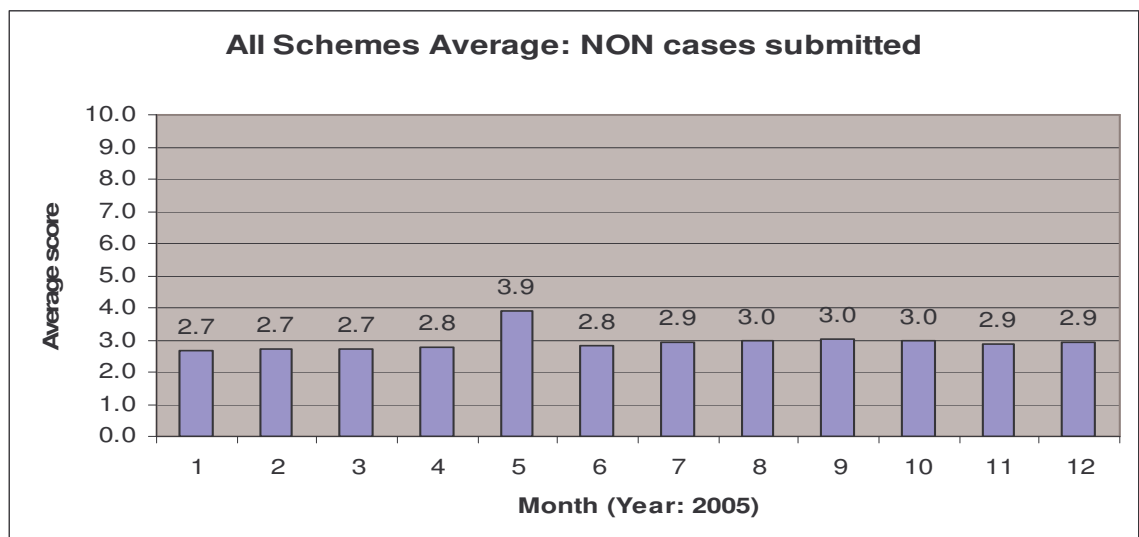


Figure 13: Average score for the number of beneficiaries without any chronic diseases submitted

### 3.5.3 Distribution by quarter

The distribution of scores for the number of beneficiaries without any chronic diseases submitted have followed a similar pattern in quarters 1, 2 and 3 without any obvious improvement. The median score is 2.7, 3.3 and 3.0 for quarters 1, 2 and 3, respectively.

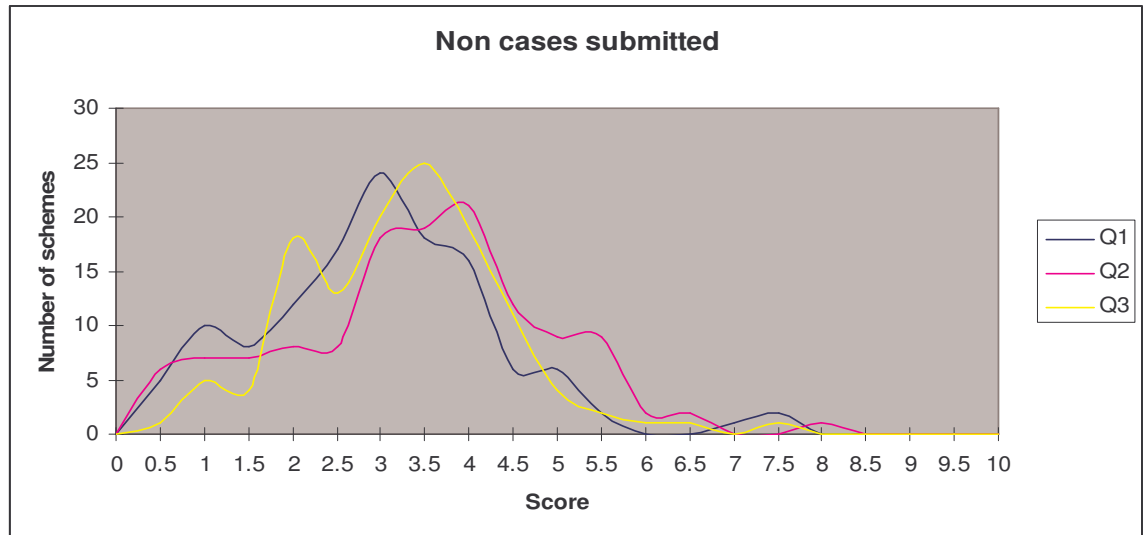


Figure 14: Distribution of scores for the number of beneficiaries without any chronic diseases submitted

### 3.6 HIV cases submitted

#### 3.6.1 Definition

The HIV cases score measures the reasonableness of HIV cases submitted against the expected industry profile for the scheme. The final score is obtained by finding an average of scores in a) and b) below:

- e) A percentage of total reported to the total expected HIV cases is calculated. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the reported and expected totals are equal, the score will be zero.
- f) If the reported data for each age bands is within the expected range, a score of zero is given. If not, the score is calculated by finding the cube root of the sum of percentage deviations from upper or lower limits (assumed deviation from expected values) of reported data for each age band. A deviation of zero given if the expected and reported cases are less than ten in an age band. A score which is greater than zero indicates that the reported cases are less than the lower limit or higher than the upper limit.

#### 3.6.2 Average by month

The average score for the number of HIV cases submitted has remained relatively high and stable over the three quarters. The mean scores are 4.5, 4.6 and 4.4 for quarters 1, 2 and 3 respectively.

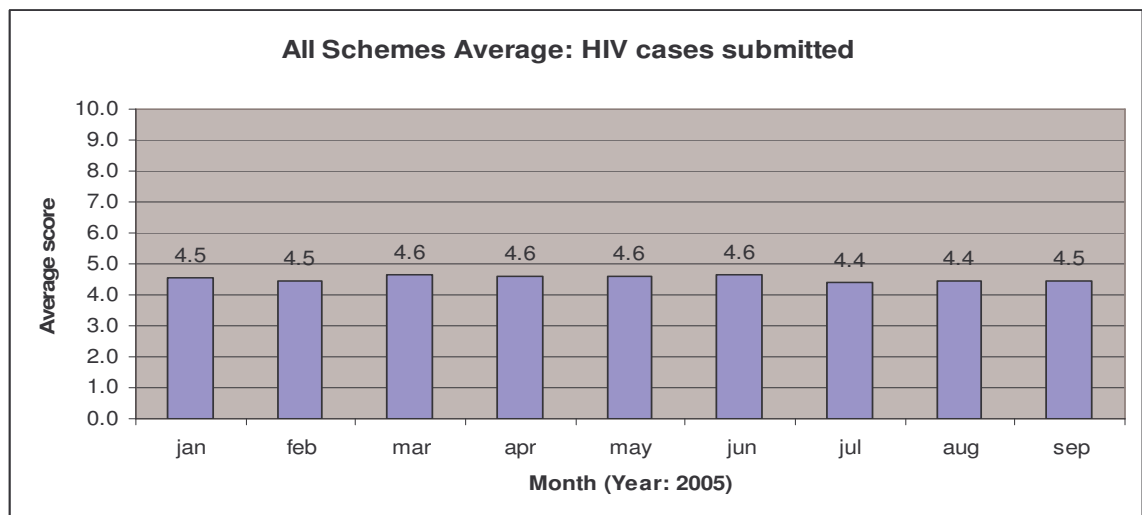


Figure 15: Average score for the number of HIV cases submitted.

### 3.6.3 Distribution by Quarter

As reported in the previous graph, the graph below shows that the quality of HIV data has not improved over the three quarters. The median value is 5 for the each of the quarters.

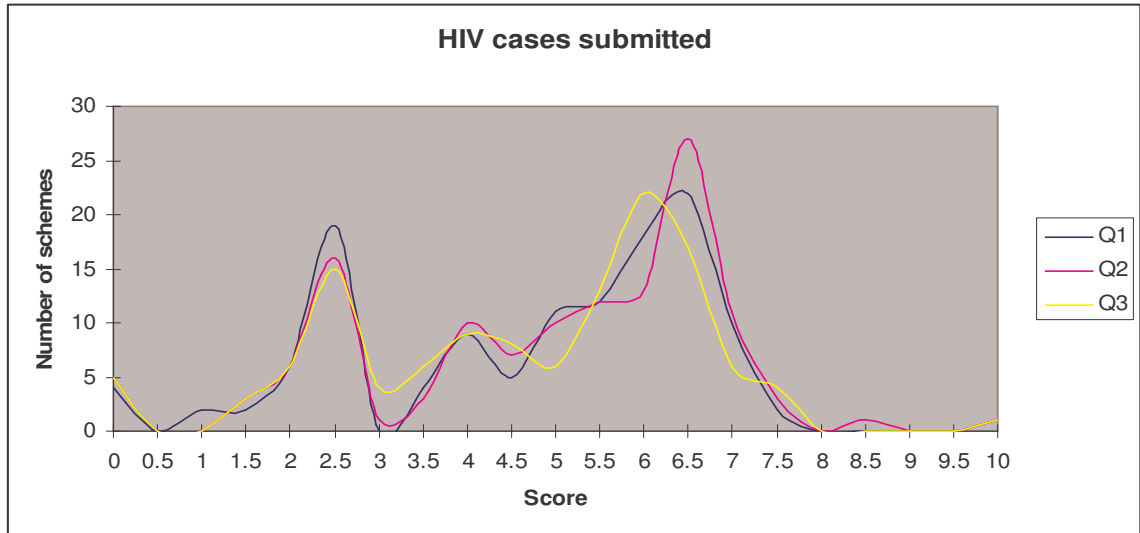


Figure 16: The distribution of scores for the number of HIV cases submitted.

### 3.7 Total CDL count

#### 3.7.1 Definition

The Total CDL cases score measures the reasonableness of the total chronic diseases against the known prevalence and age-profile information. The final score is obtained by finding an average of scores in a) and b) below:

- g) A percentage of total reported to the total expected CDL cases is calculated. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the reported and expected totals are equal, the score will be zero.
- h) If the reported data for each age bands for each CDL is within the expected range, a score of zero is given. If not, the score is calculated by finding the cube root of the sum of percentage deviations from upper or lower limits (assumed deviation from expected values) of reported data for each age band. A deviation of zero given if the expected and reported cases are less than ten in an age band. A score which is greater than zero indicates that the reported cases are less than the lower limit or higher than the upper limit.

#### 3.7.2 Average by month

The average score for the number of CDL cases submitted has remained relatively high and stable over the three quarters. The mean scores are 6.2, 6.3 and 6.3 for quarters 1, 2 and 3 respectively.

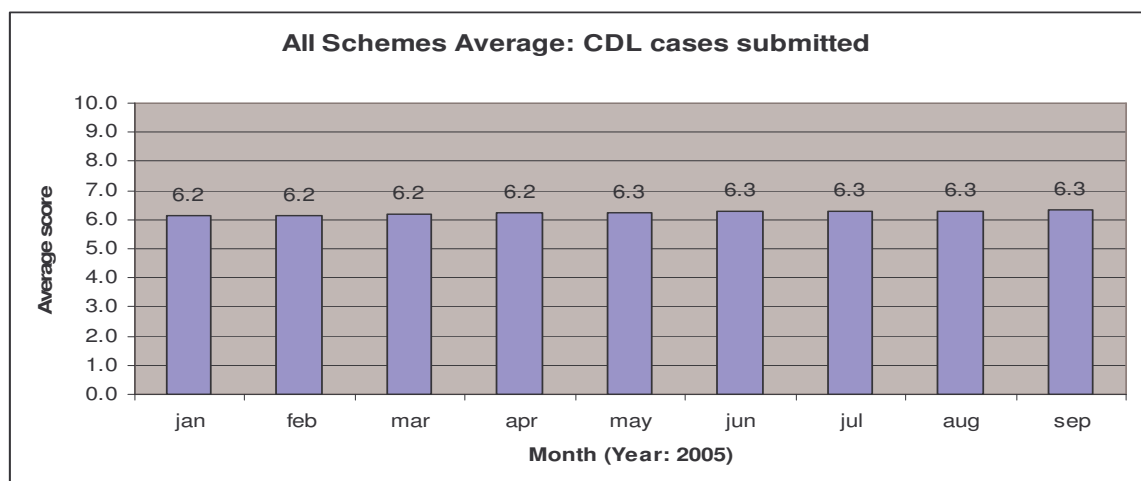


Figure 17: Average score for the total number of CDL cases submitted

### 3.7.3 Distribution by quarter

The quality of CDL cases data submitted has remained unchanged over the three quarters. The median scores are 6.4, 6.4 and 6.5 in quarters 1, 2 and 3 respectively.

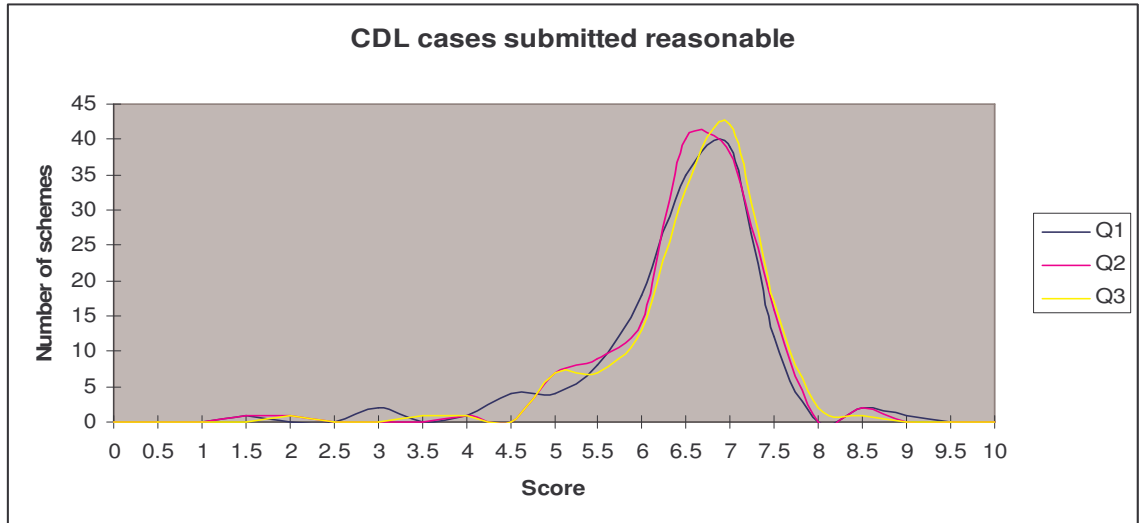


Figure 18: Distribution of scores for the total number of CDL cases submitted

### 3.8 Multiple CDL count

#### 3.8.1 Definition

The Multiple CDL count score measures the reasonableness of the total multiple chronic diseases against the known prevalence and age-profile information. The final score is obtained by finding an average of scores in a) and b) below:

- i) A percentage of total reported to the total expected multiple CDL cases is calculated. The score is obtained by finding a cube root of the percentage point difference between the resulting and the expected percentage (100%). If totals in both the reported and expected totals are equal, the score will be zero.
- j) If the reported data for each age bands for each multiple CDL is within the expected range, a score of zero is given. If not, the score is calculated by finding the cube root of the sum of percentage deviations from upper or lower limits (assumed deviation from expected values) of reported data for each age band. A deviation of zero given if the expected and reported cases are less than ten in an age band. A score which is greater than zero indicates that the reported cases are less than the lower limit or higher than the upper limit.

#### 3.8.2 Average “Multiple CDL count” scores by month

The average score for the number of multiple CDL cases submitted has remained relatively high and stable over the three quarters. The mean scores are 5.4, 5.6 and 5.8 for quarters 1, 2 and 3 respectively.

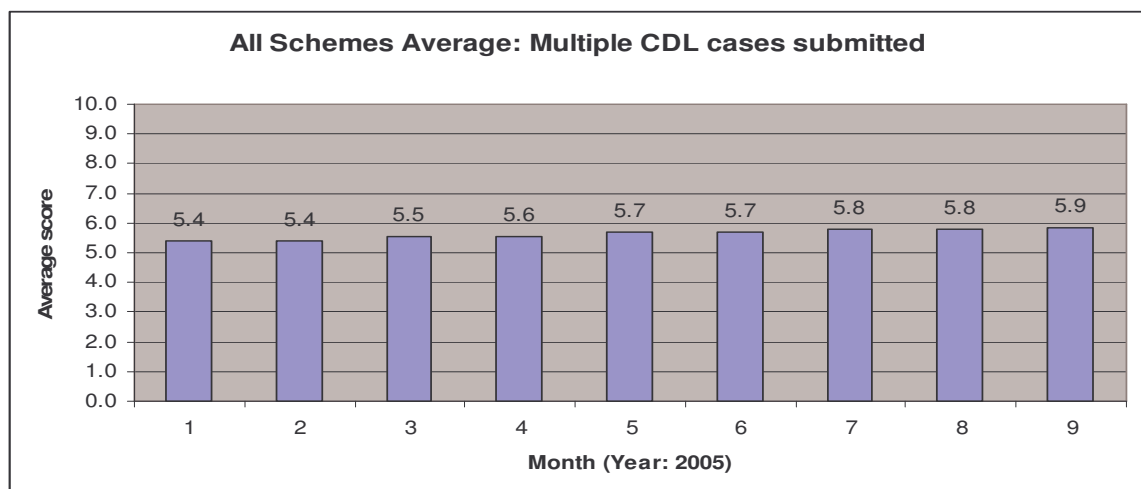


Figure 19: Average scores for the total number of multiple CDL cases submitted

### 3.8.3 Distribution by Quarter

The quality of multiple CDL cases data submitted has worsened over the three quarters. The median scores are 5.6, 5.8 and 6.3 in quarters 1, 2 and 3 respectively.

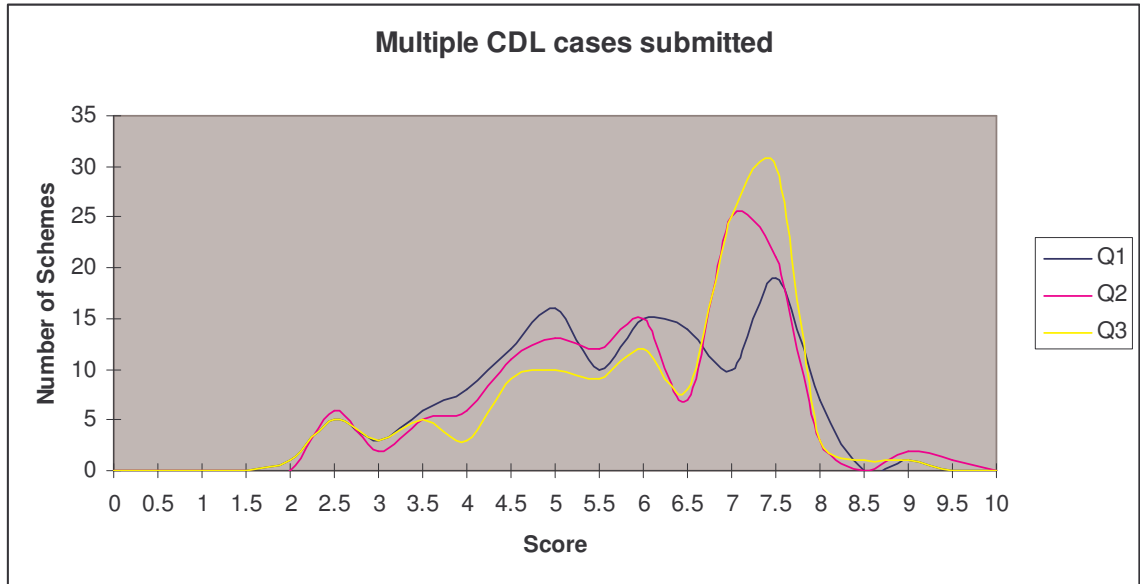


Figure 21 Distribution of scores for the total number of multiple CDL cases submitted

### 3.9 Total Score

#### 3.9.1 Definition

The Total Score is a simple arithmetic average of all computed scores. The score ranges from zero to 10. A score of zero indicates a perfect data set, and a score of 10 indicates a data set with many problems.

#### 3.9.2 Average by month

The average total score has slightly improved from 3.6 in quarter 1 to 3.4 in quarter 3.

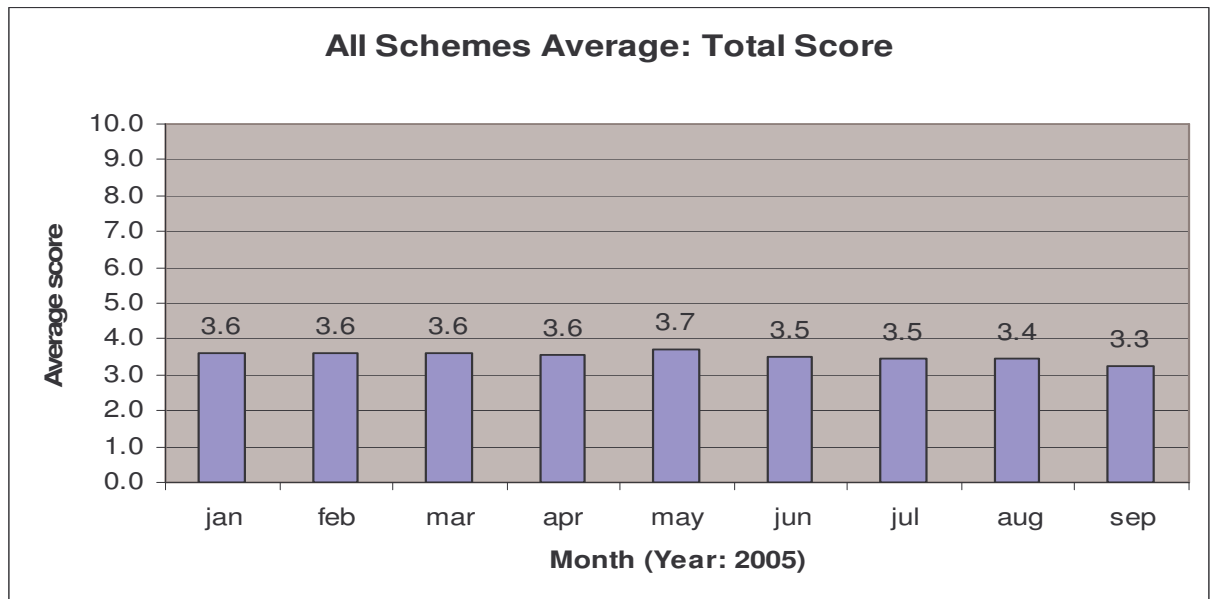


Figure 22: Average score for all data submitted

### 3.9.3 Distribution by Quarter

The median values for the total score are 3.5, 3.6 and 3.4 for quarters 1, 2 and 3 respectively.

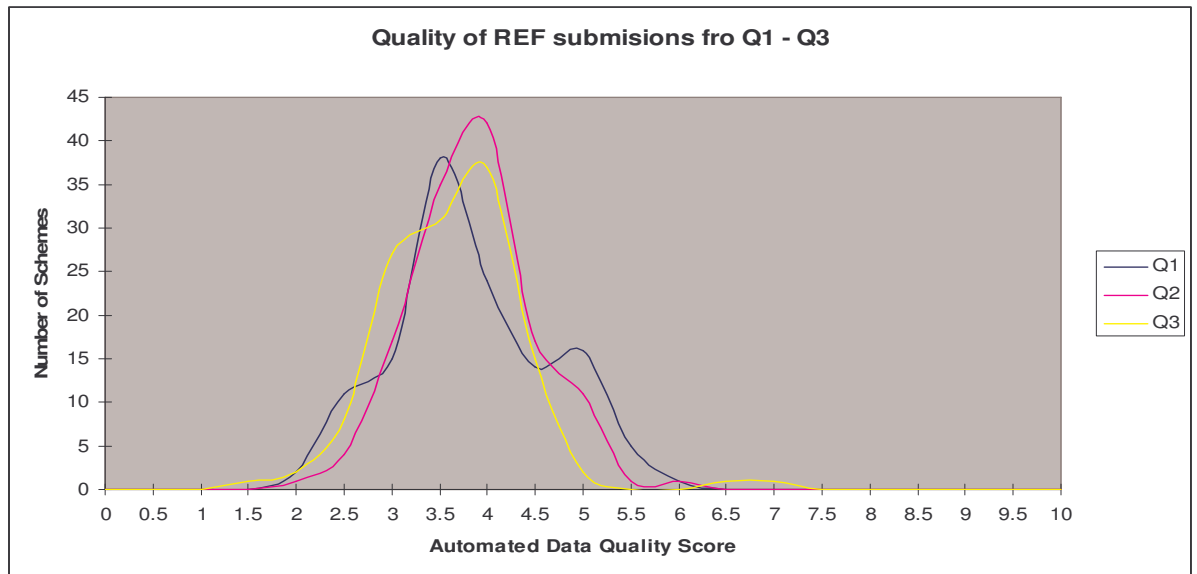


Figure 22: Distribution of scores Total Score for all data submitted