



Statistical Analysis of Complaints in Relation to Ethnicity

Results (Draft)

April 2002

CONFIDENTIAL



Working for a safer London

CONTENTS

1. INTRODUCTION	1
2. METHODOLOGY	1
3. RESULTS	2
3.1 'Public Complaints (Recorded) Complainants'	2
3.2 'Public Complaints (Recorded) Officers'	2
3.3 'Stop and Search <u>Only</u> (Recorded)'	3
3.4 'Racial Discriminatory Behaviour Allegations <u>Only</u> (Recorded)'	4
3.5 'Internal Investigations (Recorded) Officers'	4
3.6 'Suspended Officers'	4
3.7 'Written Warnings'	5
3.8 'Informal Discipline as a Result of Investigation	5
(Public Complaint - <u>Excluding</u> Written Warnings)'	5
3.9 'Informal Discipline as a Result of Investigation'	6
(Internal Investigations - <u>Excluding</u> Written Warnings).....	6
3.10 'Misconduct/Discipline Hearings (Public complaints)'	6
3.11 'Misconduct/Discipline Hearings (Internal Investigations)'	6
3.12 'Dismissals or Required Resignation as a Result of Misconduct/Discipline Hearings (Public complaints)'	6
3.13 'Dismissals or Required to Resign Sanctions as a Result of Misconduct/Discipline Hearings (Internal Investigations)'	7
3.14 'Commissioner's Appeal/AC Review (Public complaints)'	7
3.15 'Commissioner's Appeal/AC Review (Internal Investigations)'	7
3.16 'Secretary of State/Police Appeals Tribunal (Public complaints)'	7
3.17 'Secretary of State/Police Appeals Tribunal	7
(Internal Investigations)'	7
3.18 'Grievances'	7
3.19 'Employment Tribunals'	8
4. SUITABILITY OF THE STATISTICAL METHODS USED IN THE TRAINING GUIDE.....	9
APPENDIX A GRAPHICAL RESULTS OF CHI-SQUARE ANALYSIS	
Public Complaints (Recorded) Complainants	A1
Stop and Search <u>Only</u> (Recorded)	A2
Racial Discriminatory Behaviour Allegations <u>Only</u> (Recorded)	A3
Employment Tribunals	A4

1. Introduction

DPS DCC8 (2) have a number of groups of data relating to complaints, such as number of complaints made, number of complaints resulting from Stops and Searches etc. This information is available according to ethnicity of officers.

DPS DCC8 (2) tasked Consultancy Group to carry out some statistical analysis on this data to ascertain whether a relationship existed between the number of complaints recorded and the ethnic categories.

Consultancy Group was also asked by DPS DCC8 (2) to read the National Police Training guide 'Performance Management Statistics', and to comment on the suitability of a statistical method, described in the guide, when comparing recorded complaint levels among the ethnic groups used by them.

2. Methodology

Consultancy Group chose chi-squared tests to carry out the analysis because it is appropriate to use on counts of data (eg: the number of complaints recorded). Chi-squared tests were used to determine whether a relationship existed between the number of recorded complaints and the ethnic groups identified.

Chi-squared calculates an 'expected' value for each count. ie: if a count was the result of chance, this is the value that would be expected. If the actual numbers are sufficiently different from this, then the result will be statistically significant.

If a result was found to be significant this would mean that a relationship existed between the number of recorded complaints and the ethnic groups. If a result was not significant then the numbers could be a result of chance and no interpretation of a set of data such as this should be made.

Note: a significant result suggests that a relationship is present, it can not offer possible causes of that relationship and conclusions should be avoided.

In order for a chi-squared test to be valid, a calculation should not contain any 'expected' value less than 1 or more than 20% of values less than 5. Because some of the numbers that DCC8 (2) collected were very small, this meant that a number of tests that Consultancy Group performed were invalid. In an attempt to validate the tests, three years' data received from DCC8 (2) were aggregated in order to produce sufficiently high 'expected' values. In some cases 'other ethnicity' was removed from the three years' aggregated data, as the expected values were too small to be used. The tables were also collapsed, to form 2x2 tables containing just 'white' and 'visibly ethnic' categories. The validation criterion for a 2x2 table with an overall total greater than 40, is that none of the 'expected' values should be less than 1.

In order to perform chi-squared analysis both the number of complaints recorded and the population within each ethnic group was required. For instance, in the category 'Public Complaints (Recorded) Officers', the required information was: the number of complaints recorded within a given year according to ethnic group, as well as the total number of officers employed during the same time period according to ethnic group. Difficulties were encountered in attempting to obtain data for the relevant populations. These are explained in the results section of this report.

3. Results

Three years' data was supplied by DCC8 (2): 1999/2000, 2000/2001 and 2001/2002. The results of the chi-squared tests are outlined in this section, with tables for those tests where it is considered appropriate to interpret the results.

3.1 'Public Complaints (Recorded) Complainants'

The required population for this test was 'the number and ethnicity of people who come into contact with the Metropolitan police'. The population used for this test was the estimated resident population of London in 2000, as this was the only data available. However, it was acknowledged that this data source was unsatisfactory, because use of this data is based on the assumption that contact with the police is proportional to the resident population, which was felt unlikely to be the case. For this reason, extreme caution needs to be applied when considering the results of the test, as they could be an incorrect representation of the real situation.

The result of this test was both valid and significant. This means that the data used showed that a relationship existed between the number of public complaints and the ethnic group. A far higher number of complaints recorded were from black members of the public than would be expected if the numbers were a result of chance. Table 1 compares the actual complaints recorded with those expected using chi-square analysis. A graph displaying the actual number of complaints recorded during 2000/01, compared to the number expected, is shown at Appendix A1.

Table 1. The actual number of public complaints recorded during 2000/01, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Other	Total
No Complaint	5874838	591134	573163	326241	7365376
Expected	5874024	591819	573230	326303	
Complaint	1836	952	326	209	3323
Expected	2650	267	259	147	
Total	5876674	592086	573489	326450	7368699

3.2 'Public Complaints (Recorded) Officers'

The population used for this test was the number of officers employed by the MPS during 2001/02 according to ethnic group. It should be noted that police officer populations are a measure of all officers and not just those that come into contact with the public. There may be different proportions of officers in non-operational posts, and this has the potential to offer a result that is not consistent with the true picture.

The result of this test was both valid and significant. This means that the data used showed that a relationship existed between the number of public complaints and the ethnic group. A higher number of complaints were recorded of both black and Asian officers than would be expected if the numbers were a result of chance. Table 2 compares the actual complaints recorded to those expected using chi-square analysis.

Table 2. The actual number of public complaints of officers recorded during 2000/01, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Other	Total
No Complaint	20759	360	253	427	21799
Expected	20750	388	266	395	
Complaint	4721	116	74	58	4969
Expected	4730	88	61	90	
Total	25480	476	327	485	26768

3.3 'Stop and Search Only (Recorded)'

DCC8 (2) were only able to supply ten months data for 2001/02. The final data was extrapolated using mean averages. As this was recognised as being crude, this test was run using both 2000/2001 and 2001/2002 data.

The population used was the number of people stopped and searched during the relevant time period according to ethnicity. Note that for both these years, 2% of the stops and searches recorded 'ethnicity unknown'. It is unlikely that this 2% is proportional to the population that has been recorded.

The results for the two years were both valid and significant. This means that the data used showed that a relationship existed between the number of complaints made, as a result of stops and searches, and the ethnic group.

A lower number of complaints were received from both white and Asian complainants than expected, and a higher number of complaints recorded were from both black and other ethnic complainants than would be expected if the numbers were a result of chance. Table 3 compares the actual complaints recorded as a result of stop & searches, during 2000/01 (2001/02 contained extrapolated data), compared to those expected using chi-square analysis. A graph displaying the actual number of complaints recorded during 2000/01, compared to the number expected, is shown at Appendix A2.

Table 3. The actual number of public complaints resulting from Stop & Searches recorded during 2000/01, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Other	Total
No Complaint	93214	51222	16681	2845	163962
Expected	93175	51267	16670	2849	
Complaint	109	127	16	9	261
Expected	148	82	27	5	
Total	93323	51349	16697	2854	164223

3.4 'Racial Discriminatory Behaviour Allegations Only (Recorded)'

The required population for this test, and that used, was as outlined in 3.1, with the same data problems. As mentioned in 3.1, extreme caution needs to be applied when considering the results of the test, as they could be an incorrect representation of the real situation.

The result of this test was both valid and significant. This means that the data used showed that a relationship existed between the number of public complaints and the ethnic group. A far higher number of allegations of discriminatory behaviour recorded were from black members of the public than would be expected if the numbers were a result of chance. Table 4 compares the actual complaints recorded with those expected using chi-square analysis. A graph of the actual number of complaints recorded during 2000/01, compared to the number expected, is shown at Appendix A3.

Table 4. The actual number of allegations of discriminatory behaviour recorded during 2000/01, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Other	Total
No Allegation	5876630	591928	573433	326405	7368525
Expected	5876535	592072	573475	326442	
Allegation	44	158	56	45	174
Expected	139	14	14	8	
Total	5876674	592086	573489	326450	7368699

3.5 'Internal Investigations (Recorded) Officers'

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.6 'Suspended Officers'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result from this test was valid and also found to be significant. This means that the data used showed that a relationship existed between the number of suspended officers and the ethnic group. A slightly lower number of white officers were suspended than would be expected by chance, the reverse was found to be true of visibly ethnic officers. However, caution should be applied when considering these results as the numbers involved are small. Table 5 compares the actual number of suspended officers with those expected using chi-square analysis.

Table 5. The actual number of suspended officers during 2000/01, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Visibly Ethnic	Total
Not Suspended	25462	1284	26746
Expected	25459	1287	
Suspended	18	4	22
Expected	21	1	
Total	25480	1288	26768

3.7 'Written Warnings'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.8 'Informal Discipline as a Result of Investigation (Public Complaint - Excluding Written Warnings)'

In order to validate the test, three years' data was used with 'Other ethnicity' removed. The result from this test was valid and also found to be significant. This means that the data used showed that a relationship existed between the number of informal discipline (public complaints) and the ethnic group. A higher number of informal discipline proceedings were recorded of black officers and lower number of were recorded of Asian officers than would be expected if the numbers were a result of chance. However, this result should be viewed with caution, as using a collapsed table with just 'white' and 'visibly ethnic' categories with 2001/02 data produced a result that was valid but not significant. Table 6 shows the significant result and compares the actual number of officers who received informal discipline over 3 years, with those expected using chi-square analysis.

Table 6. The actual number of officers who received informal discipline over 3 years, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Total
Not Disciplined	74542	1362	763	76667
Expected	74536	1370	761	
Disciplined	313	14	1	328
Expected	319	6	3	
Total	74855	1376	764	76995

3.9 'Informal Discipline as a Result of Investigation' (Internal Investigations - Excluding Written Warnings)

In order to validate the test, three years' data was used with 'Other ethnicity' removed. The result from this test was valid and also found to be significant. This means that the data used showed that a relationship existed between the number of informal discipline (internal investigations) and the ethnic group. A higher number of informal discipline proceedings were recorded of both black officers and Asian officers than would be expected if the numbers were a result of chance. A slightly higher number would be expected of white officers. Table 7 compares the actual number of officers subject to informal discipline as a result of internal investigation with those expected using chi-square analysis. With this set of data, the collapsed table also showed both a valid and a significant result.

Table 7. The actual number of officers subject to informal discipline as a result of internal investigation over 3 years, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Total
Not Disciplined	74569	1356	756	76681
Expected	74550	1370	761	
Disciplined	286	20	8	314
Expected	305	6	3	
Total	74855	1376	764	76995

3.10 'Misconduct/Discipline Hearings (Public complaints)'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.11 'Misconduct/Discipline Hearings (Internal Investigations)'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.12 'Dismissals or Required Resignation as a Result of Misconduct/Discipline Hearings (Public complaints)'

Results for all tests undertaken with this data were not valid. No further interpretation of these results should be undertaken.

3.13 'Dismissals or Required to Resign Sanctions as a Result of Misconduct/Discipline Hearings (Internal Investigations)'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.14 'Commissioner's Appeal/AC Review (Public complaints)'

Results for all tests undertaken with this data were not valid. No further interpretation of these results should be undertaken.

3.15 'Commissioner's Appeal/AC Review (Internal Investigations)'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result of this test was valid but not significant, meaning there was no association found between the number of internal investigations and the ethnic groups. Any differences could be the result of chance and no further interpretation of these results should be undertaken.

3.16 'Secretary of State/Police Appeals Tribunal (Public complaints)'

Results for all tests undertaken with this data were not valid. No further interpretation of these results should be undertaken.

3.17 'Secretary of State/Police Appeals Tribunal (Internal Investigations)'

Results for all tests undertaken with this data were not valid. No further interpretation of these results should be undertaken.

3.18 'Grievances'

In order to validate the test, three years' data was used. The result from this test was valid and also found to be significant. This means that the data used showed that a relationship existed between the number of grievances and the ethnic group. A higher number of grievances were recorded from black, Asian and Other ethnicity officers than would be expected if the numbers were a result of chance. A lower number was recorded than would be expected of white officers. Table 8 compares the actual number of recorded grievances with those expected using chi-square analysis. With this set of data, the collapsed table also showed both a valid and a significant result.

Table 8. The actual number of grievances over 3 years, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Black	Asian	Other	Total
No Grievance	74507	1356	749	1147	77759
Expected	74458	1369	760	1173	
Grievance	348	20	15	32	415
Expected	397	7	4	6	
Total	74855	1376	764	1179	78174

3.19 'Employment Tribunals'

In order to validate the test, it was necessary to collapse the table to contain just 'white' and 'visibly ethnic' categories.

The result from this test was valid and also found to be significant. This means that the data used showed that a relationship existed between the number of employment tribunals and the ethnic group. A considerably lower number of white officers went to employment tribunals than would be expected by chance, the reverse was found to be true of visibly ethnic officers. Table 9 compares the actual number of employment tribunals with those expected using chi-square analysis. A graph of the actual number of employment tribunals during 2001/02, compared to the number expected, is shown at Appendix A4.

Table 9. The actual number of employment tribunals, by ethnicity, compared to those expected using chi-square analysis

Ethnicity	White	Visibly Ethnic	Total
No Tribunal	25423	1241	26664
Expected	25381	1283	
Tribunal	57	47	104
Expected	99	5	
Total	25480	1288	26768

4. Suitability of the Statistical Methods Used in the Training Guide

In order to carry out the statistical procedures outlined in the 'Performance Management Statistics' guide, the data being used must be both quantitative and from, what is referred to in statistical terms, a normal distribution. If the data were converted to percentages of complaints then it would be quantitative, but would have the associated problems with small counts of data.

There are various ways to carry out tests for a normal distribution, but the simplest is to check both the mean and the median of a set of data. If they are similar values then the data is likely to be from a normal distribution. Apart from having a very small number of data which makes the result spurious, the mean of 'percentage number of public complaints (officers) is 25, whilst the median is 1.9. Quite clearly these data are not from a normal distribution. The procedure outlined in the 'Performance Management Statistics' guide would not, therefore, be appropriate to use for this set of data.

The procedure could be used, however, to compare the rate of complaints across boroughs or other police services for each ethnicity group, to see whether the same pattern of complaint rates were recorded across the groups (providing the test for normality was satisfied). It is not possible, however, to compare the ethnicity groups directly using this method (eg: it cannot be used to compare the rate of complaints against white officers with the rate of complaints against black officers).