UNIVERSITY OF STIRLING

ANN ELIZABETH ROBERTS

HEALTH CARE PERFORMANCE MANAGEMENT – INSIGHTS FROM APPLICATIONS OF DATA ENVELOPMENT ANALYSIS

Submitted for the Degree of Doctor of Philosophy

29th January 2001

CONTENTS

		Page
Acknowledgements		I
Abstract		II
Detailed Table of Contents		III
List of Figures		
List of Tables		
Chapter One	Introduction: Motivation, Key Themes and Overview.	1
Chapter Two	The Measurement of Efficiency and Performance in the Public Sector.	18
Chapter Three	'Squaring Two Circles' – The Changing Organisational Structure of the NHS and its Impact on, and Relation to, Efficiency and Performance Assessment.	57
Chapter Four	The Measurement of Efficiency and Performance in the National Health Service.	105
Chapter Five	The Data Envelopment Analysis Approach to Efficiency Measurement.	155
Chapter Six	Applying DEA to Health Care Organisations: Information from the Literature and Health Care Managers.	243
Chapter Seven	Analysing Health Care Organisations Using Data Envelopment Analysis (STAGE ONE).	279
Chapter Eight	Introducing Restrictions to the Factor Weights in the Data Envelopment Analysis Methodology (STAGE TWO).	357
Chapter Nine	Evaluating the Application Procedure for Data Envelopment Analysis in the Assessment of Hospital Efficiency.	419
Chapter Ten	Conclusions and Recommendations for Further Study.	477
Appendices		493

ACKNOWLEDGEMENTS

Over the last five years, there were many occasions when I thought that finishing my thesis was an impossibility. But here it is, just about on time and fulfilling all the necessary requirements. In the end, the struggle has been worthwhile. After all, 'happiness may be health to the body but hardship is what strengthens the mind' (profound thoughts from South Africa's tight-head prop, Cobus Visagie, The Daily Telegraph, 28/10/99).

I would like to thank my supervisor, Professor Rob Ball, for all his help and encouragement during the last four years. And for all those chats about cricket – England would surely regain the Ashes if we were in charge!

I would also like to express my gratitude to the many other people who have been involved in the course of this research, particularly Richard Thomas, my second supervisor and the numerous others who provided information and additional material.

To my fellow PhD-ers in Stirling, thanks for your friendship, support and advice: Cathy, for putting up with the mess in the office; Caroline, for introducing me to the delights of the www.hamsterdance.com; Peter, for computing help and all those 'herby teas' and to Paul, for lending me your mini-encyclopaedias on a regular basis!

Many thanks are also due to my friends beyond Stirling, for offering encouragement and providing a variety of distractions, which kept me going when it seemed as though things would never be finished: Nicola, for sharing that trip to New York and encouraging my traveller's spirit; Sam (and BT), for those long phone conversations and to the Clitheroe Crew (Stephen, Paul, Neil, Sandra and Jonathan), for being around when I came home, to share a drink or a game of football!

To my new colleagues at MHA and friends from St. James', Clerkenwell, thank you for your help and encouragement through the final stages of the PhD and my move to the 'big city'!

I would like to say thanks to my family, who have always supported me in whatever crazy adventures I have undertaken. Jenni and Nick, 'ta very much' for letting me stay when I needed a break. Last, but probably not least, 'merci beaucoup' to my Mum and Dad for everything you have done: for your help, encouragement and for actually being prepared to read the thesis from beginning to end. You deserve a medal!

And any remaining spelting nistakes or typ0s are dedicated to Winnie the Pooh, 'because my spelling is wobbly. It's good spelling but it wobbles and the letters get in the wrong places' (Winnie the Pooh, by A. A. Milne, published by Methuen, 1926).

ABSTRACT

The comprehensive measurement of efficiency and performance in the Health Service in the UK has become one of the most important managerial developments of recent years. The reasons for this development were examined, particularly in relation to the difficulties involved with performance assessment in such a context. The most widely utilised techniques were evaluated from the perspective of the Health Care Manager and a number of serious limitations were identified.

In response to these limitations, the technique of Data Envelopment Analysis was evaluated as an alternative. It has been proposed as an appropriate and useful tool for the assessment of efficiency, although the literature on DEA showed limited practical application to public sector services in the UK. The many facets of the technique were investigated and literature on its application to hospital data was reviewed.

A two-stage application procedure for the DEA technique was developed in response to this evaluation, to be used in the measurement hospital efficiency. The procedure was based on a deep theoretical understanding of the DEA methodology. The most important elements of the process were related to selection of the initial sample, the identification of the variables to be included in the DEA model and the definition of the weight restrictions to be incorporated. Input from Health Care Managers was used to guide the application and data from a sample of acute hospitals in Scotland was utilised in the analysis. The application procedure showed how the practicalities of the DEA technique could be enhanced, in particular through the inclusion of weight restrictions. This led to the development of efficiency strategies for the inefficient hospitals, which could be related to the policy objectives or managerial structure of the hospitals in the sample.

It was concluded that there were many potential benefits of the DEA approach to efficiency assessment and the two-stage application procedure defined here, which could be seen to fulfil many of the requirements of the Health Care Manager. It was determined that combining theoretical and practical issues can enhance the applicability of the DEA methodology.

DETAILED TABLE OF CONTENTS

Chapter One: Introduction: Motivation, Key Themes and Overview.

1.1	**INTRODUCTION**	**2**
1.2	**MOTIVATION**	**3**
1.3	**RESEARCH OBJECTIVES**	**5**
1.3.1	To Determine the Reasons for the Importance of Performance Assessment in Health Services in the UK	**6**
1.3.2	To Identify the Methods Used for Performance Assessment in the NHS	**6**
1.3.3	To Investigate Alternative Methods for Performance Assessment	**6**
1.3.4	To Develop New Approaches for Performance Assessment	**7**
1.4	**RESEARCH METHODOLOGY**	**7**
1.4.1	Reviewing the Literature	**7**
1.4.2	Including the Views and Perspectives of Individuals	**8**
1.4.3	Applying the DEA Methodology	**9**
1.5	**THESIS OVERVIEW**	**9**
1.5.1	Background Information	**10**
1.5.2	Investigation of the DEA Technique	**12**
1.5.3	Application of the DEA Technique	**13**
1.5.4	Evaluation	**14**
1.5.5	Summary	**15**
1.6	**SUMMARY**	**15**
1.7	**REFERENCES**	**17**

Chapter Two: The Measurement of Efficiency and Performance in the Public Sector

2.1	**INTRODUCTION**	**19**
2.2	**THE PUBLIC SECTOR**	**20**
2.2.1	Physical Inputs	**22**
2.2.2	Physical Outputs	**22**
2.3	**DEFINING EFFICIENCY**	**23**
2.3.1	Types of Efficiency	**24**
2.3.2	Technical Efficiency	**26**
2.3.3	Allocative Efficiency	**27**
2.4	**ACHIEVING EFFICIENCY**	**28**
2.5	**EFFICIENCY AND HEALTH SERVICES**	**29**
2.5.1	Profile of Efficiency	**29**
2.5.2	Response to Efficiency	**31**
2.5.3	Efficient Practice	**32**
2.5.4	Difficulties with Efficiency Measurement in Health Services	**33**
2.6	**EFFICIENCY IN A BROADER CONTEXT**	**35**
2.6.1	Efficiency, Effectiveness and Economy	**35**
2.6.2	Equity and Equality	**36**
2.6.3	Quality of Care	**38**
2.7	**METHODS FOR MEASURING EFFICIENCY AND PERFORMANCE**	**41**
2.7.1	Ratio Analysis	**41**
2.6.2	Econometric Regression Analysis	**44**
2.6.3	Clinical Audit	**46**
2.8	**THE DEA APPROACH TO EFFICIENCY MEASUREMENT**	**48**
2.9	**CONCLUSIONS**	**53**
2.10	**REFERENCES**	**54**

Chapter Three: 'Squaring Two Circles' – The Changing Organisational Structure of the NHS and its Impact on, and Relation to, Efficiency and Performance Assessment.

3.1	**INTRODUCTION**	**58**
3.2	**THE ORIGINS OF THE NHS AND ITS EARLY DEVELOPMENT (1948 – 1974)**	**60**
3.3	**REDEVELOPING THE NHS (1974 – 1982)**	**67**
3.4	**A NEW DIRECTION FOR THE NHS (1982 – 1990)**	**73**
3.5	**CHANGING THE FUNDAMENTAL STRUCTURE OF THE NHS (1991 – 1997)**	**77**
3.6	**THE 'NEW NHS' UNDER NEW LABOUR (1997 AND BEYOND)**	**86**
	3.6.1 Performance Assessment for the 'New NHS'	**91**
3.7	**THE NHS – ISSUES OF FINANCE AND ECONOMICS**	**93**
	3.7.1 The Financial Position of the NHS in 1998	**97**
3.8	**THE FUTURE OF THE NHS**	**98**
3.9	**REFERENCES**	**101**

Chapter Four: The Measurement of Efficiency and Performance in the National Health Service.

4.1	**INTRODUCTION**	**106**
4.2	**PERFORMANCE AND EFFICIENCY MEASUREMENT**	**108**
	4.2.1 Data Availability and Accuracy	**108**
	4.2.2 The Politics of Performance Evaluation	**109**
	4.2.3 The Nature of the Health Service	**109**
	4.2.4 Implications of Environmental Differences in Comparative Assessments	**110**
4.3	**THE NHS PERFORMANCE INDICATORS**	**110**
	4.3.1 Introduction and Development of Performance Indicators	**111**
	4.3.2 Use of Performance Indicators	**115**
	4.3.3 Users of Performance Indicators	**117**
	4.3.4 Interpretation of Performance Indicators	**118**
	4.3.5 Limitations and Criticisms of Performance Indicators	**120**
	4.3.6 Performance Indicators and Outcomes	**125**
	4.3.7 Performance Indicators and Efficiency	**127**
	4.3.8 The Future Development of Performance Indicators	**128**
	4.3.9 Evaluation of Performance Indicators	**132**
4.4	**CLINICAL AUDIT**	**133**
	4.4.1 Definition of Clinical Audit	**133**
	4.4.2 The History and Development of Audit	**134**
	4.4.3 Audit in Practice	**138**
	4.4.3.1 Participation in Audit	**139**
	4.4.3.2 Scope of Audit	**140**
	4.4.3.3 Initiation of Audit	**141**
	4.4.3.4 Reasons for Audit	**141**
	4.4.3.5 Information from Audit	**142**
	4.4.3.6 Impact of Audit	**142**
	4.4.4 Evaluation of Audit	**143**
	4.4.5 The Future of Clinical Audit	**146**
4.5	**ADDITIONAL APPROACHES TO PERFORMANCE ASSESSMENT**	**148**
	4.5.1 QALYS	**148**
	4.5.2 Regression Techniques	**149**
4.6	**SUMMARY**	**150**
4.7	**REFENCES**	**151**
4.8	**ADDITIONAL SOURCES**	**154**

Chapter Five: The Data Envelopment Analysis Approach to Efficiency Measurement.

5.1 INTRODUCTION 156
5.2 THE DEVELOPMENT OF EFFICIENT FRONTIERS METHODOLOGY 158
5.3 APPLICABILITY OF THE DEA METHODOLOGY FOR THE EVALUATION OF HEALTH SERVICES 160
5.4 DEFINITION OF THE DEA METHODOLOGY 163
5.4.1 The Ratio Form of the DEA Model 168
5.4.2 Formulation of the DEA Model as a Linear Programme 170
5.4.3 Interpretation of the Factor Weights 173
5.4.4 Virtual Inputs and Outputs 174
5.4.5 The Output from a DEA Model 176
5.4.6 The Reference Set 178
5.5 EXTENDING THE DEA METHODOLOGY 180
5.5.1 Variable Selection in DEA Models 180
5.5.2 The Nature of the Variables 189
5.5.2.1 Manpower Levels (Inputs) 189
5.5.2.2 Resources (Inputs) 190
5.5.2.3 Environmental Factors (Inputs) 191
5.5.2.4 Activity Levels (Outputs) 193
5.5.2.5 Quality Measures (Outputs) 194
5.6 THE DEVELOPMENT OF ALTERNATIVE DEA MODELS 195
5.6.1 The BCC Model 195
5.6.2 Additive Models 199
5.6.3 Multiplicative Models 201
5.6.4 Evaluation of the Different Model Types 201
5.7 WEIGHT RESTRICTION ISSUES 202
5.7.1 Choosing the Methods for Restricting the Factor Weights 207
5.7.1.1 Minimum Contribution 208
5.7.1.2 Maximum Contribution 209
5.7.1.3 Range of Values 209
5.7.1.4 Grouped Restrictions 209
5.7.1.5 Specifying Inequalities 210
5.7.2 Illustrating the Impact of Weight Restrictions 211
5.7.3 Weight Restriction Methodology in Practice 214
5.8 OTHER EXTENSIONS TO THE DEA METHODOLOGY 215
5.8.1 Different Types of Efficiency in DEA 216
5.8.2 The Inclusion of Categorical Data as Variables 217
5.8.3 Non-Discretionary Variables 217
5.8.4 Longitudinal and Window Analysis 220
5.8.5 Cross-Efficiency in DEA 222
5.8.5.1 Definition of Cross-Efficiency 223
5.8.5.2 Application of Cross Efficiency 223
5.8.6 Recent Developments 225
5.8.6.1 Super Efficiency 225
5.8.6.2 The Free Disposal Hull (FDH) Method 227
5.9 ANALYSIS OF THE DEA RESULTS USING STATISTICAL METHODS 228
5.9.1 Correlation Coefficients 230
5.9.2 The Mann-Witney Test 231
5.9.3 Two-Tailed Z-test for Proportions 232
5.10 SUMMARY: PERSPECTIVES ON DATA ENVELOPMENT ANALYSIS 233
5.10.1 Efficiency Measurement 233
5.10.2 Identification of Inefficiencies 233
5.10.3 Measuring External Influences 234
5.10.4 Inclusiveness 234
5.10.5 Encompassing Local Policies and Innovations 234
5.11 REFERENCES 236

Chapter Six: Applying DEA to Health Care Organisations: Information from the Literature and Health Care Managers.

6.1 INTRODUCTION 244
6.2 INFORMATION FROM THE LITERATURE 244
6.2.1 Ehreth (1994) 245
6.2.2 Ozcan, Luke and Haksever (1992) 248
6.2.3 Sherman (1984) 252
6.2.4 Grosskopf and Valdmanis (1987) 254
6.2.5 Byrnes and Valdmanis (1994) 258
6.2.6 Hollingsworth and Parkin (1995) and Parkin and Hollingsworth (1996) 262
6.2.7 Other Sources 266
6.2.8 Evaluation of the DEA Applications in the Literature 267
6.2.8.1 Modelling Strategy 267
6.2.8.2 DEA Model Formulation 268
6.2.8.3 Variable Selection 268
6.2.8.4 Including Individual Perspectives 268
6.2.8.5 Evaluating and Presenting Results 269
6.3 EVALUATION OF DATA ENVELOPMENT ANALYSIS FROM THE PERSPECTIVE OF THE HEALTH CARE MANAGER 269
6.3.1 Selection of Samples and Variables 271
6.3.2 Providing Useful Information 272
6.3.3 Avoiding complicated Theory 272
6.3.4 Presenting a True Reflection of Performance 272
6.3.5 Relevant at a Local Level 273
6.4 SUMMARY 273
6.5 DEA HEALTH SERVICES BIBLIOGRAPHY 274
6.5.1 Health Care Summary 274
6.5.2 Regional/District Health Authorities of Health Districts 274
6.5.3 Hospitals 274
6.5.4 Hospital Departments 275
6.5.5 Individuals (physicians and/or managers) 275
6.5.6 General Practice/Primary Care 276
6.5.7 Nursing Homes 276
6.5.8 Pharmacies 276
6.5.9 Health Care Programmes 276
6.6 REFERENCES 277
6.7 ADDITIONAL SOURCES 278

Chapter Seven: Analysing Health Care Organisations Using Data Envelopment Analysis (STAGE ONE).

7.1 INTRODUCTION 280
7.2 THE DEA MODELLING PROCESS 282
7.3 DEFINITION OF AN APPRORIATE SAMPLE 283
7.4 FORMULATION OF THE DEA MODEL 287
7.5 SELCECTING VARIABLES TO BE USED IN THE DEA MODEL 290
7.5.1 The Opinion of Experts 295
7.5.2 Previous Use in the Literature 296
7.5.3 Evaluation of the Data 299
7.5.4 Statistical Manipulation 303
7.5.5 A Heuristic Approach 310
7.6 ANALYSING THE OUTPUT FROM DEA MODELS 311
7.6.1 Using 'Discharges' or 'Days' to Measure Inpatient Activity 321
7.6.2 Removing Highly Correlated Factors 324
7.6.3 Investigating Alternative Measures of Output 331
7.6.4 The Impact of Variable Aggregation 337
7.6.4.1 Aggregating the Cost Factors (Inputs) 339
7.6.4.2 Aggregating Inpatient Discharges (Outputs) 341

7.6.5 Evaluation of the Models (Variable Set) 348
7.6.6 Evaluation of the Models (Sample Definition) 351
7.7 **SUMMARY OF THE FIRST STAGE OF THE APPLICATION PROCESS** **354**
7.8 **REFERENCES** **355**
7.9 **ADDITIONAL SOURCES** **356**

Chapter Eight: Introducing Restrictions to the Factor Weights in the Data Envelopment Analysis Methodology (STAGE TWO).

8.1 **INTRODUCTION** **358**
8.2 **ANALYSIS OF THE RESULTS FROM MODEL 8.1** **359**
8.3 **ANALYSIS OF THE FACTOR WEIGHTS FOR MODEL 8.1** **362**
8.3.1 Widely Varying Factor Weights 363
8.3.2 Non-Contribution of Factors 367
8.3.3 Interpretation of the Factor Weights 370
8.3.4 Inappropriate Relationships between Inputs and Outputs 371
8.4 **FURTHER JUSTIFICATION FOR THE INTRODUCTION OF WEIGHT RESTRICTIONS** **373**
8.4.1 The Number of Units Classified as Efficient 373
8.4.2 Discrimination Against the All-Rounder 375
8.4.3 Expert Opinion 376
8.5 **SPECIFICATION OF WEIGHT RESTRICTIONS** **376**
8.6 **SELECTION OF WEIGHT RESTRICTIONS** **378**
8.6.1 Restrictions Attached to the Outputs 379
8.6.1.1 Equal Contributions from All Outputs 379
8.6.1.2 Minimum Contributions from All Outputs 380
8.6.1.3 Externally Defining the Important Outputs and their Contribution 381
8.6.2 Restrictions Attached to the Inputs 383
8.6.3 Definition of the Weight Restriction Scenarios 384
8.6.3.1 Scenario One (All Round Efficiency) 385
8.6.3.2 Scenario Two 385
8.6.3.3 Scenario Three 386
8.6.3.4 Scenario Four 387
8.6.3.5 Scenario Five 387
8.7 **APPLICATION OF WEIGHT RESTRICTIONS** **388**
8.7.1 General Analysis of the Alternative Weights Scenarios 389
8.7.2 Analysis of the Individual DMUs 398
8.7.2.1 The Efficient DMUs 399
8.7.2.2 The Inefficient DMUs 404
8.7.3 Analysis of the Factor Weights 412
8.8 **SUMMARY OF THE SECOND STAGE OF THE APPLICATION PROCESS** **415**
8.9 **REFERENCES** **417**
8.10 **ADDITIONAL SOURCES** **418**

Chapter Nine: Evaluating the Application Procedure for Data Envelopment Analysis in the Assessment of Hospital Efficiency.

9.1 **INTRODUCTION** **420**
9.2 **THE MODELLING STRATEGY** **422**
9.3 **SELECTION OF SAMPLES** **425**
9.3.1 Small Sample Analysis Using DEA 436
9.3.2 Small Samples of Hospitals 438
9.3.2.1 Geographical Region 439
9.3.2.2 Hospital Size 441
9.3.2.3 Functional Classification 442
9.3.3 Comparisons by Patient Type (Inpatients, Outpatients etc.) 445
9.3.4 Comparisons by Patient Category (Ophthalmology, Medical, Gynaecology) 446
9.3.5 Comparisons by Patient Category and Type (Gynaecology Inpatients, Ophthalmology Day Cases) 446

9.4 **SELECTION OF VARIABLES** 448
9.4.1 The Inclusion of a Case-Mix Index for Inpatient Activity 449
9.4.2 Alternative Measures of Input 451
9.4.3 The Influence of the Environment 451
9.4.4 The Importance of Measuring Service Quality 453
9.5 **PROVIDING USEFUL INFORMATION** 455
9.5.1 DEA and the Measurement of Technical Efficiency 456
9.5.2 Relating DEA to Other Techniques 458
9.5.3 The Relevance of the Output from DEA Models 463
9.6 **AVOIDING COMPLICATED THEORY** 468
9.7 **PRESENTING A TRUE REFLECTION OF PERFORMANCE** 470
9.8 **RELEVANT AT A LOCAL LEVEL** 471
9.9 **SUMMARY** 473
9.10 **REFERENCES** 475
9.11 **ADDITIONAL SOURCES** 476

Chapter Ten: Conclusions and Recommendations for Further Study.

10.1 **INTRODUCTION** 478
10.2 **SUMMARY OF FINDINGS** 478
10.3 **THE RESEARCH OBJECTIVES** 484
10.3.1 To Determine the Reasons for the Importance of Performance Assessment in Health Services in the UK 484
10.3.2 To Identify the Methods Used for Performance Assessment in the NHS 484
10.3.3 To Investigate Alternative Methods for Performance Assessment 485
10.3.4 To Develop New Approaches for Performance Assessment 485
10.4 **LIMITATIONS OF THE RESEARCH** 486
10.5 **RECOMMENDATIONS FOR FURTHER STUDY** 488

Appendix One: Key Definitions

1A **GENERAL DESCRIPTION OF FUNCTIONAL CLASSIFICATION** 493
1B **SCOTTISH HEALTH BOARD CODES** 494
1C **DEFINITIONS OF KEY TERMS** 495
1D **DEFINITIONS OF TYPES OF CARE AND THE COMPOSITION OF SPECIALTY GROUPS** 498

Appendix Two: Data and Analysis used in Chapter Five

2A **HOSPITALS IN THE SAMPLE (GYNAECOLOGY INPATIENT SERVICES)** 501
2B **DATA FOR THE ANALYSIS IN CHAPTER FIVE** 502
2C **DEA RESULTS FOR MODEL G3** 503
2D **DEA RESULTS FOR MODEL G8** 504

Appendix Three: Data and Analysis used in Chapter Seven

3A **ACUTE HOSPITALS IN SCOTLAND** 506
3B **DATA USED FOR THE ANALYSIS IN CHAPTER SEVEN** 508
3C **RESULTS FROM MODEL 7.1** 516
3D **RESULTS FROM MODEL 7.11** 518
3E **DMUS EXCLUDED FROM THE ANALYSIS IN CHAPTER EIGHT** 520

Appendix Four: Data and Analysis used in Chapter Eight

4A **HOSPITALS IN THE REFINED SAMPLE (MODEL 8.1)** 522
4B **DATA USED IN THE ANALYSIS IN CHAPTER EIGHT** 523
4C **RESULTS FROM MODEL 8.1** 524
4D **RESULTS FROM WEIGHT RESTRICTION SCENARIOS** 525

LIST OF FIGURES

Figure 1.1: Overview of Thesis Chapters 10

Figure 2.1: Model of Public Sector Performance 21
Figure 2.2: The Relationship between Economy, Efficiency and Effectiveness 36
Figure 2.3: The Efficiency Ratio 41

Figure 3.1: Structure of the NHS 1948-74 (England and Wales) 62
Figure 3.2: The structure of the NHS 1974-82 (England) 69
Figure 3.3: Structure of the NHS 1982 - 90 (England) 74
Figure 3.4: The Contracting Process 80
Figure 3.5: Structure of the NHS 1991 (England) 81
Figure 3.6: Structure of the NHS 1994 83
Figure 3.7: UK NHS Expenditure 1949 – 1995 94
Figure 3.8: Total Health Expenditure Per Person 1992 95
Figure 3.9: Health Care Expenditure as a Percentage of GDP 96

Figure 4.1: Interests in Medical Audit 138

Figure 5.1: Farrellian Definition of Efficiency 159
Figure 5.2: Locating The Efficiency Frontier for Model G1 165
Figure 5.3: The Ratio Form of the DEA Efficiency Equation 168
Figure 5.4: The Efficiency of a DMU #1 169
Figure 5.5: The Output-Oriented CCR Model (Dual) (Output Maximisation) 171
Figure 5.6: The Input-Oriented CCR Model (Dual) (Input Minimisation) 171
Figure 5.7: Locating the Efficiency Frontier with Varying Returns to Scale 196
Figure 5.8: Relating Inefficient Units to the Efficiency Frontier using the BCC Model (Input Orientation) 196
Figure 5.9: The Input-Oriented BCC Model (Dual) 197
Figure 5.10: Relating Inefficient Units to the Frontier using the Additive Model 200
Figure 5.11: DEA Model for Uncontrollable Environmental Input Factors 219
Figure 5.12: The Shape of the Efficiency Frontier in the FDH Model 227
Figure 5.13: Distribution of the Efficiency Scores 229
Figure 5.14: The Spearman Rank Correlation Coefficient 230
Figure 5.15: The Mann-Witney U-Test 231
Figure 5.16: Two-tailed Z-test for Proportions 232

Figure 7.1: An Application of Data Envelopment Analysis to Health Care Data 282
Figure 7.2: Structure of the NHS in Scotland in 1998 284
Figure 7.3: The CCR Model with Input Minimisation Orientation 290
Figure 7.4: The Regression Equation for Total Costs 309
Figure 7.5: The Distribution of Efficiency Scores for Models 7.1, 7.3 and 7.4 326
Figure 7.6: Distribution of Efficiency Scores for Models 7.3, 7.5 and 7.6 333
Figure 7.7: Distribution of Efficiency Scores for Models 7.7, 7.9 and 7.10 342

Figure 8.1: Distribution of the Efficiency Scores for Model 8.1 and the Scenarios 394
Figure 8.2: Input Reduction Targets for DMU *4 (Costs in £000s) 408

Figure 9.1: Comparison of Output Patterns According to Functional Classification 430
Figure 9.2: Comparison of Input Patterns According to Functional Classification 430
Figure 9.3: Cost Per Patient Day (Gynaecology Inpatients) 460
Figure 9.4: Cost Per Patient Discharge (Gynaecology Inpatients) 461

LIST OF TABLES

Table 2.1: Strengths and Weaknesses of Ratio Analysis 43
Table 2.2: Strengths and Weaknesses of Regression Analysis 46
Table 2.3: Strengths and Weaknesses of Data Envelopment Analysis 52

Table 4.1: National Performance Framework in the New NHS 131
Table 4.2: Perspectives on the Scope of Audit from BHRV 140

Table 5.1: The Results from Model G1 172
Table 5.2: Virtual Inputs for Model G2 175
Table 5.3: Impact of Variable Changes on Efficiency Scores 181
Table 5.4: Variation in Efficiency Scores 181
Table 5.5: The Impact of Varying Returns to Scale on Overall Efficiency 198
Table 5.6: Nature of Scale Variance 199
Table 5.7: Analysis of Factors Contributing to Efficiency Scores 205
Table 5.8: The Impact of Weight Restrictions on Overall Efficiency Scores 211
Table 5.9: Analysis of the Virtual Inputs in Relation to the Weight Restrictions and Efficiency Scores for DMUs #2, #4 and #10. 213
Table 5.10: Cross-Efficiency Matrix for a Simplified Example 223
Table 5.11: Variations in the Weights Applied 224
Table 5.12: Super Efficiency Scores 226

Table 7.1: Hospital Trusts in Scotland 285
Table 7.2: Potential Variables 293
Table 7.3: Zero Values in the Output Factors 300
Table 7.4: Non-zero Data in the Output Factors 300
Table 7.5: Pairwise Correlation Coefficients for Output Factors 304
Table 7.6: Regression of Output Factors 308
Table 7.7: Definitions and General Results for the DEA Models under Investigation 312
Table 7.8: Analysis of Efficiency Scores by Health Board 315
Table 7.9: Efficiency Score by Functional Classification 316
Table 7.10: Efficiency Scores and Hospital Size 317
Table 7.11: Distribution of Dominant Outputs 319
Table 7.12: Analysis of the Least Efficient DMUs 320
Table 7.13: Changes in Efficiency Scores 322
Table 7.14: Evaluation of Changes in Efficiency Scores 323
Table 7.15: Comparisons Between Models 7.7 and 7.9 343
Table 7.16: Distribution of the Factor Weights for Model 7.9 344
Table 7.17: Comparisons Between Models 7.9 and 7.10 346
Table 7.18: Distribution of the Factor Weights for Model 7.10 347
Table 7.19: Evaluation of the Models 348
Table 7.20: Factors to be Used in Stage Two of the Application Process 351

Table 8.1: Summary Description of Model 8.1 360
Table 8.2: Newly Efficient DMUs 361
Table 8.3: Analysis of the Virtual Weights by Mean, Standard Deviation, Maximum, Minimum, Count and Dominance 363
Table 8.4: Virtual Factor Weights for Selected DMUs 364
Table 8.5: Analysis of the Virtual Weights 365
Table 8.6: Factor Weight Percentages for an Inefficient DMU and its Peer Group 366
Table 8.7: Contribution of Factors 368
Table 8.8: Distribution of Factors Used to Calculate Efficiency Scores 368
Table 8.9: Differences in Average Cost per Case by Functional Class 382
Table 8.10: Weight Restriction Scenarios 387
Table 8.11: Impact of Weight Restrictions 389
Table 8.12: Correlation Statistics for Efficiency Rankings in Model 8.1 and the Weightings Scenarios 391
Table 8.13: The Five Largest Differences in Efficiency Scores between Model 8.1 and Each Scenario 392

Table 8.14: Average Differences In Efficiency Scores 393
Table 8.15: The Impact of Weight Restrictions on Efficient and Inefficient DMUs 396
Table 8.16: Inputs and Outputs for the Core Group of Efficient Units 399
Table 8.17: Dominant Output Factors for the Core Group of Efficient Units 401
Table 8.18: The Factor Weight Patterns for DMU *13 402
Table 8.19: Efficiency Scores for a Selection of Inefficient DMUs 405
Table 8.20: Targets for Improving Efficiency for a Selection of Inefficient DMUs 406
Table 8.21: References Sets for a Selection of Inefficient DMUs 410
Table 8.22: Virtual Weights for a Selection of Inefficient DMUs in Model 8.1 413
Table 8.23: Efficiency Scores for DMUs Robust against Weight Restrictions 414

Table 9.1: The Two-Stage Application Procedure for Data Envelopment Analysis 423
Table 9.2: Minimum, Maximum and Average Values for Factors used in Model 8.1 428
Table 9.3: Examining Differences in Output Activity 429
Table 9.4: Analysis of Efficiency Scores by Functional Classification 431
Table 9.5: Analysis of Scenario Three by Functional Classification 432
Table 9.6: Number of Acute Hospitals in Each Health Board Region 440
Table 9.7: Number of Acute Hospitals in Each Functional Classification 443
Table 9.8: Performance Indicators on Costs for Gynaecology Inpatient Data 459
Table 9.9: Cost per Case Indicators Compared with DEA Efficiency Scores 462
Table 9.10: Evaluation of Input Slacks for DMU *1 465
Table 9.11: The Factor Values for DMU *1 and its Reference Set 465

CHAPTER ONE

INTRODUCTION: MOTIVATION, KEY THEMES AND OVERVIEW

1.1 Introduction

The National Health Service in the United Kingdom is a vast organization, with a complex managerial and operational structure, encompassing an array of services provided across the country. Unlike the situation in the United States, the vast majority of health care in the UK is funded through general taxation and administered by central government. It is one of a number of public sector services to be financed in the UK in this manner, although others such as schools are administered through local government to a much greater extent. The NHS was established in 1948 by the Labour Government, guided by the Secretary of State for Health, Aneurin Bevan. Throughout its history, one of its central features has been that the treatment it provides is 'free' at the point of need.

The range and quantity of services provided by the NHS has steadily increased and, although demand for treatment may be unlimited, the resources allocated by the government to the NHS have not always satisfied this demand. This has resulted in the rationing of particular services or treatments, either directly or implicitly, through the lengthening of waiting lists.

It is in this context that the role of performance evaluation in the public sector has been developed and become an important part of the NHS in an attempt to provide value-for-money. The efficiency with which services have been provided has been central to this debate, such that the limited resources can be deployed to produce the maximum amount of services, which is the definition of technical efficiency given by Levitt and Joyce (1987).

In this chapter, the motivation for research of this nature and its objectives is presented, focusing on the requirement within the NHS for performance evaluation. This is followed by a brief discussion of the research methodologies to be applied. Finally, in order to emphasise the key themes to be addressed, an overview of each of the chapters in the thesis is also included.

1.2 Motivation

The assessment of performance has become endemic in almost all levels of government and business, ranging from the financial statements published by multinational corporations to the introduction of performance-related pay for many public sector employees in the United Kingdom. All segments of the public sector are now subject to numerous methods to assess their performance, covering efficiency, effectiveness, economy and quality, amongst others, under the general umbrella of value-for-money (VFM). Libraries, schools and police forces have all been evaluated using a selection of performance indicators, published annually in the form of league tables. Poor performers are highlighted and can be subject to a variety of methods to improve their situation.

This emphasis on performance evaluation has been particularly apparent and important in the context of Health Services in the UK: the National Health Service (NHS) is the largest employer in the country and has an annual budget of approximately £40billion (The Treasury, 1999). Upon its establishment in 1948, the key emphasis of the NHS was on providing a comprehensive service available to everyone, which would be free at the point of need (Pater (1981) reported by Rivett (1997).

Addressing efficiency or measuring performance were not highlighted as critical elements during this early period of its development. As the NHS has evolved since its inception in 1948, however, a great deal of attention has been given to developing strategies for improving all aspects of the services it provides. This has led to a series of reforms, introduced by successive administrations. These have focused on changing its organisational structure or identifying new approaches for the assessment and improvement of its performance.

Therefore, there are now over 2000 measures or indicators of performance collected annually for each hospital, recorded by various Government departments and agencies, including the Department of Health, the Scottish Office and the NHS Executive. Additionally, the structure of the NHS is very different today, in both managerial and organisational terms.

As the NHS has developed, so too have the methods available for performance assessment in the public sector. Over the last twenty years in particular, techniques of all types have been employed to investigate all aspects of performance, from a variety of perspectives and with a number of objectives. These include performance indicators, ratio analyses, econometric regression techniques, clinical audit, benchmarking and cost-benefit analyses.

During this same timeframe, the technique of Data Envelopment Analysis (DEA) has also developed as a tool for efficiency analysis, following its introduction by Charnes, Cooper and Rhodes (1978). It has always been traditionally associated with the assessment of public sector services (Ball and Roberts, 1998).

Ersoy *et al.* (1997) proposed that the DEA technique is 'more informative than the other efficiency measurement methods' and can be part of the process for making 'hospital and health systems rational and efficient'. Similarly, Salinas-Jiménez and Smith (1996) determined, following an application of DEA to primary health care data in the UK, that DEA 'offers a useful means of assessing performance in health care'.

The motivation for this research can be summarised by two themes. The first is the priority attached to performance assessment in the NHS today by Central Government. The second is the availability of a myriad of assessment alternatives, particularly Data Envelopment Analysis. With these in mind, the main objectives for the research are now addressed.

1.3 Research Objectives

In the previous section, the emphasis placed on performance assessment in the NHS has been firmly established. Additionally, the availability of a number of techniques to investigate various aspects of public sector performance has been verified. In previous studies to investigate performance in the NHS and also in the application of the DEA methodology, the techniques themselves have often been considered the most important aspect. The central tenet of this study, however, was based on an understanding of the services being evaluated. The views and perspectives of health care managers, evaluators and medical personnel have thus influenced all aspects of the research process. This approach was used to ensure that appropriate and useful information was provided by the analysis.

Within this framework, four key objectives for this research have been identified and are now discussed.

1.3.1 To Determine the Reasons for the Importance of Performance Assessment in Health Services in the UK

Performance evaluation is now a central aspect of public sector management. The reasons for this are investigated, with a specific emphasis on the NHS. This included an understanding of the complexities involved in performance assessment in this environment, identifying the key elements of public sector performance and highlighting the factors that can influence upon it.

1.3.2 To Identify the Methods Used for Performance Assessment in the NHS

As performance assessment has become more important, a number of methods have already been introduced in the NHS, in order to bring about improvements in its overall performance. The benefits and limitations of these approaches are sought both from the literature and from the perspective of NHS personnel, in order to determine the necessary characteristics of any alternative techniques.

1.3.3 To Investigate Alternative Methods for Performance Assessment

Focusing on the Data Envelopment Analysis (DEA) approach to performance assessment, the benefits of this alternative are investigated. This relates specifically to applications in the NHS, linking together the information obtained in the investigation of the previous two objectives.

The focus is on the potential for DEA to overcome limitations identified in the existing techniques and its applicability to the complexities of the Health Care environment.

1.3.4 To Develop New Approaches for Performance Assessment

The final objective relates to the practical applications of performance assessment techniques, as it is essential to determine if the methods employed provide useful information that can be used directly in the improvement of particular aspects of performance. This involves understanding the requirements of the Health Care Manager, the nature of the environment and the specifics of the techniques. This objective is possibly the most important, as it draws together all the other aspects of the research. It requires insights from health care managers, in order to look at performance assessment from their perspective and develop an acceptable application procedure for efficiency measurement techniques.

1.4 Research Methodology

The methodological approaches to be followed in order to meet the objectives of the research are threefold: investigation of the appropriate literature, application of the DEA technique and discussions with significant individuals.

1.4.1 Reviewing the Literature

The literature requirements covered two main areas: (i) Data Envelopment Analysis and its application and (ii) the NHS and issues of performance assessment.

The discussion on the theoretical development and the practical application of DEA is based on conventional literature sources. These are very extensive, as is demonstrated by the DEA Bibliography compiled by Seiford (1994), included in the comprehensive DEA text by Charnes *et al.* (1994).

In addition to the traditional sources of information, including management science and health service journals and books, the Internet has been a vital source of material on the nature of the NHS and performance assessment issues. The Department of Health, the Scottish Office, the NHS Executive (England and Scotland) and the Government's statistical services are all accessible via the Internet. A variety of materials are available, including press releases, statistical information, health service activity and expenditure data and government documentation, such as public expenditure reports and NHS annual reports.

1.4.2 Including the Views and Perspectives of Individuals

A number of contacts within the NHS were identified and their views on performance evaluation sought. Several employees of the Blackburn, Hyndburn and Ribble Valley NHS Hospitals Trust, including the Directors of Finance and Corporate Development, were interviewed informally. Further interviews were also conducted with several members of the Performance Directorate of the NHS Executive in Leeds. Discussions on a more informal basis were also undertaken with numerous other sources connected with the NHS, providing an insight into the inner-workings of a health care unit without generating specific perceptions or views. This is useful in providing context for the application of performance evaluation techniques.

Informal discussions on the subject of Data Envelopment Analysis at a number of conferences and seminars provided additional insights into the technique and its application. DEA symposia held in France (Marseille, 1997) and Germany (Wernigerode, 1998) provided an opportunity to discuss matters of theory and practice with experienced researchers, including Lawrence Seiford, Rajiv Banker and Henry Tulkens.

1.4.3 Applying the DEA Methodology

A significant part of this investigation is concerned with the analysis of the results obtained from a number of models developed using the DEA methodology. This includes various approaches for illustration and interpretation, using numerical, tabular and graphical forms.

In order to solve the models developed in the research, specially written DEA software was employed, specified by Dr. Richard Thomas of the Department of Management and Organization at the University of Stirling. The software was capable of solving the majority of DEA models defined, including those with a large variable set.

1.5 Thesis Overview

Using the methodological approaches identified above and the objectives identified in section 1.3 as the focus, the structure of the thesis is illustrated in figure 1.1 below. A brief outline of each of the remaining chapters follows, with the key research questions for each chapter clarified. The links between the chapters are emphasised.

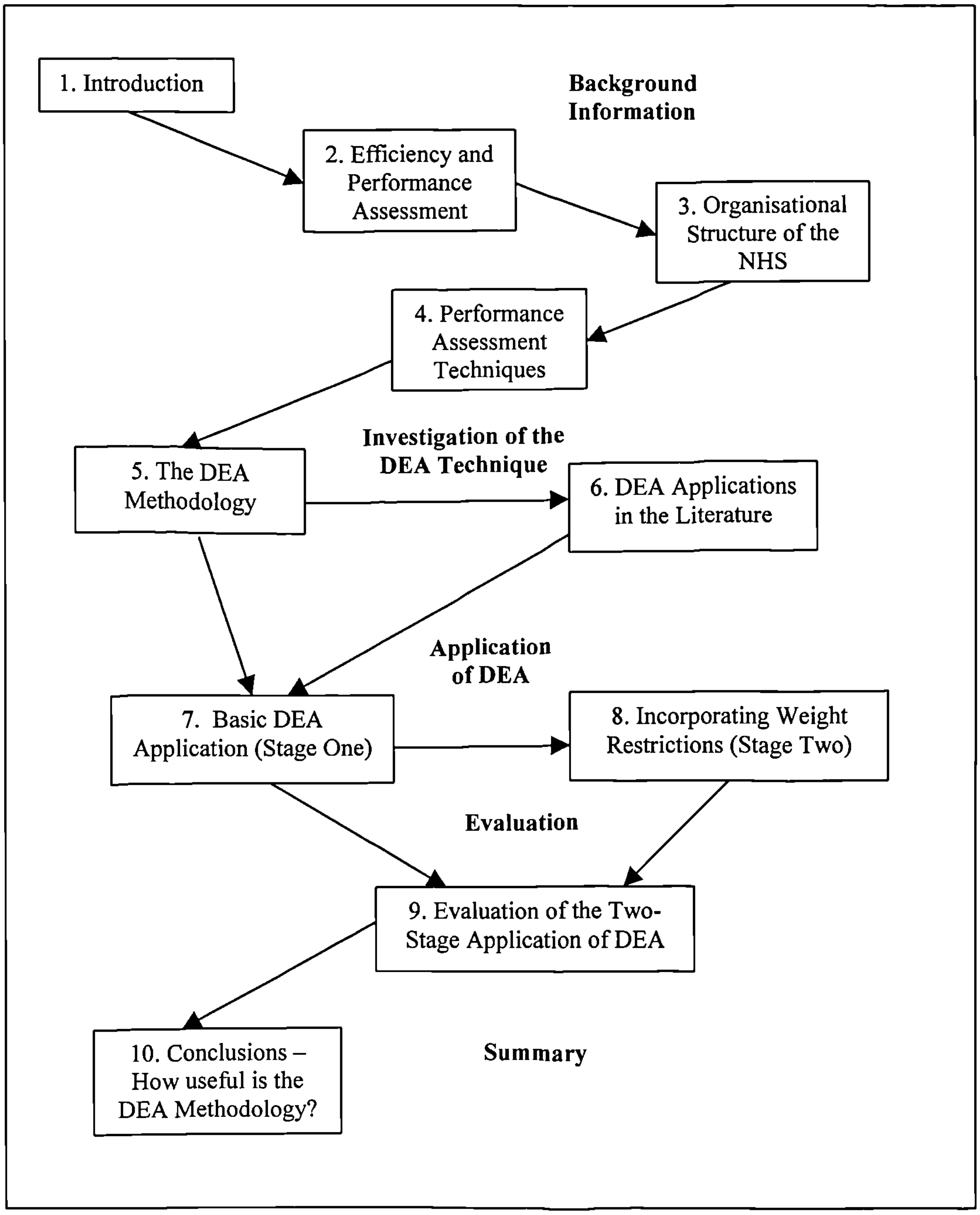


Figure 1.1: Overview of Thesis Chapters

1.5.1 Background Information

Chapters two to four are intended to provide the background information required for the development of alternative approaches to performance assessment discussed in later sections. In relation to the research objectives, these chapters are primarily concerned with addressing the first two (sections 1.3.1 and 1.3.2).

Chapter Two: Can performance in the public sector and the NHS be investigated in such a way as to make use of traditional definitions of its key elements?

The key elements of public sector performance are discussed, with particular reference to their application in health services. There is also an introduction to the Data Envelopment Analysis technique, as well as the other approaches applicable for performance assessment.

Chapter Three: How important is the nature of the NHS (history, structure and management philosophy) in the definition and application of performance assessment techniques?

This discussion serves to provide a context for the investigation of performance in a health service environment. It also provides a more detailed analysis on the nature of health services in the UK and the impact this can have on the evaluation of performance, briefly discussed in more general terms in chapter two. Observations made on the changing nature of the services provided are also included. These are linked to the DEA application process in chapter seven.

Chapter Four: Are the methods currently applied to performance evaluation appropriate and do they provide useful information?

The chapter contains a detailed evaluation of each of the main approaches (performance indicators and clinical audit). This includes reference to the rationale behind their development, the particular areas where they have been implemented and an evaluation of their strengths, weaknesses and limitations. Information gained from the interviews with Health Care Managers is introduced to illustrate the theoretical observations gathered from an array of literature sources.

1.5.2 Investigation of the DEA Technique

The second phase of the thesis is found in chapters five and six, which are concerned with an investigation into the DEA technique and related to the third of the research objectives (section 1.3.3). The main source of material is from an extensive array of literature on DEA, although observations form a number of health care professionals are included.

Chapter Five: Does DEA provide an alternative for the evaluation of efficiency in a health care environment?

The theoretical development of the DEA methodology, from its inception in the work of Charnes, Cooper and Rhodes (1978) through to the most recent developments, is considered. The use of frontier methodologies in relation to health care efficiency assessment is debated, linking back to the discussions in chapter two. The numerous developments to the original methodology are debated, with particular attention given to the theory of restricting the free allocation of the factor weights. Data relating to the provision of gynaecology inpatient services at a small number of hospitals in Scotland is used to illustrate the discussion.

Chapter Six: What information can be learned from the literature on health care applications of the DEA methodology?

A comprehensive evaluation of a small number of papers in the literature, each of which reports on the use of frontier-type methodologies in the evaluation of hospital efficiency, is included. This is contrasted with the view of a number of health care managers. A bibliography of health care applications of DEA is included, illustrating the numerous areas of health services that have been evaluated using the technique.

1.5.3 Application of the DEA Technique

This section of the thesis is concerned with application, developing the theoretical observation found in the preceding chapters into a two-stage application of the DEA technique using data on the provision of health services in Scotland. Material from the DEA literature and the views of a number of health care professionals are incorporated into the modelling process.

Chapter Seven: Can a model-building process be developed for a large-scale application of the DEA methodology?

Information obtained on the structure and operating practices of hospitals within the NHS, first discussed in chapter three, is also included to provide context for the evaluation process and assist in the development of the DEA models. The models developed and discussed in this chapter are intended to comprehensively illustrate the basic features of the technique and also the lengthy processes involved in its application. Data on the provision of acute hospital services in Scotland is used in the analysis. This chapter presents the first stage of the application process.

Chapter Eight: What impact does the inclusion of weight restrictions have on the applicability of the DEA methodology?

A fundamental change to the basic methodology is considered, that is, restricting the free allocation of the factor weights. The proposed reasons for the change are considered, drawing upon the discussions found in chapter five. The theoretical implications of this change are discussed in chapter five, whereas the focus of this chapter is on the practicalities involved. The approaches used to select and apply the weight restrictions are considered.

The relationship between weight restrictions and policy objectives is investigated, as are the benefits and limitations involved in making use of this development. The addition of weight restrictions introduces the concept of the efficiency strategy. Chapter eight presents the second stage of the application procedure.

1.5.4 Evaluation

The fourth phase of the thesis is the evaluation of two-stage application procedure presented in chapters seven and eight and also on the features of the DEA methodology itself. This relates to both the third and fourth of the research objectives (sections 1.3.3 and 1.3.4).

Chapter Nine: Does the DEA methodology have actual benefits in practice and how can it be enhanced?

The primary causes for the limited use of the DEA technique in the health service environment to date are examined and the changes that are required to the two-stage application process are investigated. The observations of the health care managers and evaluators (discussed in chapter six) provide the foundation for the evaluation. The analysis found in chapters five, seven and eight is used to illustrate the discussions. The relationship between the existing approaches used for performance evaluation and the results from the DEA methodology are also investigated, with reference to the discussion in chapter four. The possibility of extending the approach to the evaluation of other aspects of health service activity is considered. The evaluation of the approach is taken from the perspective of the health care manager, identifying how the DEA methodology can be further enhanced.

1.5.5 Summary

The final section provides a summary of the research and draws together the objectives of the research as a whole.

Chapter Ten: What is the future of DEA within the NHS, both short-term and long-term?

This chapter evaluates the value of the DEA approach in its enhanced form and determines future strategy for research in this field. The final phase of the research draws together the main points identified in each of the previous chapters and the original aims of the research are addressed. A brief summary of the main points of the research is included, using the research questioned identified for each of the chapters as a guide. Limitations to the approaches used in the research process will also be considered. Future research possibilities are also debated, focusing on the potential for the DEA approach to be accepted as an important tool for efficiency assessment in the health service in the UK.

1.6 Summary

This chapter has provided a very general introduction to all of the themes to be addressed in greater detail in the remainder of this thesis. The raised profile of performance assessment in the NHS has been highlighted and the numerous methods available have been briefly discussed. The technique of DEA has been introduced as an alternative method for performance evaluation, requiring further investigation.

Four key objectives have been identified, which are addressed by the research in the remainder of this thesis. In the final chapter, the success or otherwise of achieving these objectives is discussed. Additionally, a research question has been identified for each of the main chapters.

The chapters to follow are developed according to the overview defined in the previous section, with attention given first to obtaining appropriate definitions for the key elements of performance.

1.7 References

1. **Ball, R. and Roberts, E.** (1998) The Relevance of DEA in the Public Sector - A Discussion Paper for Young OR 10, Young OR 10 Keynote Papers, Operational Research Society.

2. **Charnes, A., Cooper, W. W. and Rhodes, E.** (1978) Measuring the Efficiency of Decision- Making Units, *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429 - 444.

3. **Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (eds.)** (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

4. **Ersoy, K., Kavuncubasi, S., Ozcan, Y. A. and Harris, J. M.** (1997) Technical Efficiencies of Turkish Hospitals: DEA Approach, *Journal of Medical Systems,* Vol. 21, No. 2, pp. 67 – 74.

5. **Levitt, M. S. and Joyce, M. A. S.** (1987) The Growth and Efficiency of Public Spending, Cambridge University Press.

6. **Pater, J. E.** (1981) The Making of the NHS, King's Fund Publishing.

7. **Rivett, G.** (1997) From Cradle to Grave, Fifty Years of the NHS, King's Fund Publishing.

8. **Salinas-Jiménez, J. and Smith, P.** (1996) Quality in Primary Health Care, *Annals of Operations Research*, Vol. 67, pp. 141 – 161.

9. **Seiford, L.** (1994) A DEA Bibliography: 1978 – 1992, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (eds.) (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

CHAPTER TWO

THE MEASUREMENT OF EFFICIENCY AND PERFORMANCE IN THE PUBLIC SECTOR

2.1 Introduction

Flynn (1986) proposed that the search for adequate and appropriate methods for the measurement of efficiency and performance had become a 'central preoccupation for public sector organisations'. The reasons for this change, in relation to the evolving structure of the NHS and the current political and economic climate, are considered in this chapter.

However, prior to this, the concepts of efficiency and other related concerns, are discussed in general terms, focusing primarily on public sector organisations as a whole. The rationale for such an approach is that it is vital to gain a clear understanding of the key elements of the public sector and public sector performance before applying performance assessment techniques.

Therefore, this chapter contains a general introduction to public sector services, followed by definitions of the key elements of performance, including efficiency, economy and effectiveness. These are also related to the provision of health services, with the attendant problems identified.

To conclude the chapter, the methods that may be used to examine these concepts are then presented, with an introduction also given to the theory and applicability of data envelopment analysis, which forms the basis of the subsequent analysis.

2.2 The Public Sector

The public sector services in the UK are those services that are funded primarily with finance gained from taxation and other sources of income, which is then allocated through central and local government. The organisations that come under the heading of 'the public sector' are local authorities (schools, the police, ambulance and fire services), the judiciary, the NHS and the health services, environmental agencies and the armed forces.

Public sector employees include civil servants, teachers, doctors and police officers. These public sector employees are distinct from those employed in the old 'nationalised industries', which have been greatly reduced over the past fifteen years. This was in accordance with the privatisation policy of the Conservative party, which formed successive governments between 1979 and 1997. Some of the organisations privatised during this period were British Rail, British Telecom and the utility companies supplying gas, electricity and water. (One of the reasons given for the privatisation of these organisations was that it would bring about a greater level of efficiency.)

However, each of the public sector services has different operating practices, serving different sections of the population and working towards different objectives. The following diagram has been proposed as a simplified model of how public sector services operate, showing that each public sector unit, be it a hospital, a university or a police force, is given a financial contribution which is to be transformed into a benefit for the population it serves.

Costs ⇒⇒⇒⇒ Public Sector Unit ⇒⇒⇒⇒ Benefits

Physical Inputs Physical Outputs

SMITH (1995)

Figure 2.1: Model of Public Sector Performance

Using the diagram as a guide, the cost inputs for a school would be used to employ teachers, purchase books and other necessary equipment and provide the finance to operate the school itself. The process of the school is the provision of the education services, with the physical outputs being the pupils leaving the school after the proscribed number of years. The ultimate benefit of the school system would be an educated population.

For a hospital, a variety of relevant personnel would be employed to provide all the required treatments and services, using medical equipment, pharmaceuticals and other facilities in the hospital unit, whose running costs were provided from the financial inputs. The physical outputs would be the patients leaving hospital following their treatment, with the eventual benefit being a healthy population.

Unfortunately, the transformation of cost to benefit is an abstract idea that is not easily quantified or examined. The costs of a public sector unit such as a school can be measured in financial terms and a 'total cost' calculated. However, it is much more difficult to explain in quantifiable measures, the overall benefits of investing public money in a school. Clearly, the children who attend each school should benefit through improved educational standards.

However, it is not a straightforward process to directly measure the educational standards of the population. Hence the concepts of inputs and outputs are introduced as an intermediate stage in the model of public sector performance. As was discussed above, the physical inputs are the different elements provided using the financial inputs to allow the process of the public sector unit to be carried out. The physical outputs are seen as substitute measures for the overall benefits. These concepts are discussed in greater detail in the following sections.

2.2.1 Physical Inputs

Physical inputs into a public sector unit are resources, personnel, capital and supplies, each of which can usually be reduced to a monetary value. For example, in a hospital the inputs would be the doctors, nurses, administrators and other categories of staff, the beds and facilities used for the treatment of patients and the resources used to carry out medical procedures such as pharmaceuticals, medical supplies and electronic equipment.

2.2.2 Physical Outputs

It is more difficult to define physical outputs in general terms, as they are dependent on the type of unit under examination, and cannot be readily quantified in economic terms. They are the results of the unit's work and represent the impact that the unit has on the population it serves. For example, the outputs for a hospital could be the number of people treated in a range of categories for inpatients and outpatients or the number of procedures carried out. For a school, the outputs are the number of pupils being taught.

Physical outputs do not directly measure the success with which a unit operates, as they do not record if the population is healthier or more educated as a result of the processes carried out by each public sector unit. This is often investigated by introducing the additional concept of outcomes, particularly in the health sector.

Clearly, public sector units are much more complicated than the simple model described in figure 2.1 above. There are often factors beyond a unit's control which affect the way it operates and the outputs it produces - these can be environmental factors or externally applied restrictions on cost or procedures.

These issues become important when methods are proposed for the measurement of efficiency and will be discussed again in subsequent chapters. However, it is the concepts of inputs and outputs that are the main factors in the definition and examination of efficiency that now follows.

2.3 Defining Efficiency

The problem of defining efficiency is almost as complex as the question of how to achieve it, particularly in public sector organisations. Efficiency is a measure of the relationship between the inputs into a unit or organisation and the outputs from it. Alternatively, it is an examination of the process by which inputs are transformed into outputs. Inefficiency occurs where resources are wasted, poorly utilised or not used to their maximum potential. According to Levitt and Joyce (1987), the aim of management within a public sector organisation should be to improve the efficiency with which resources are used.

In relation to the provision of health services specifically, Allsop (1984) provided the following as a description of efficiency:

> "Efficiency can mean simply a service or procedure which costs less, irrespective of quality. However, the term may be taken to mean at the least the maintenance of a given standard of service at a lower cost or an improved standard at the same cost."
>
> ALLSOP (1984)

The above definition suggests that efficiency is connected to the provision of services at the least cost. However, it also shows that efficiency in relation to public sector services is much more complicated, with several elements involved in its measurement. The various types of efficiency are now introduced.

2.3.1 Types of Efficiency

Efficiency is not just a measure of the process by which inputs are transformed into outputs at the least cost, as the above definition has suggested. Metcalfe and Richards (1990) suggested that there are actually many types of efficiency, as listed and explained below:

- **Technical Efficiency**: the physical use of resource inputs in relation to physical outputs.
- **Economic Efficiency**: the cost of using inputs in relation to the value of outputs (also called price efficiency).
- **Allocative Efficiency**: the optimal distribution of resources, guided by process where possible, to ensure that resources are distributed among producing units to consumer wants in ways that reflect costs of provision.

- **Productive Efficiency**: productivity in relation to resources or producing at minimum cost.
- **Operational Efficiency**: cost-consciousness in performing existing functions.
- **Adaptive Efficiency**: increased flexibility by the reformulation of objectives and a speedy adjustment to environmental change.

Levitt and Joyce (1987) proposed a more condensed view of efficiency, suggesting that there are three requirements for a process or organisation to be classified as efficient:

1. The maximum possible amount is produced with the resources used i.e. it is impossible to reduce the volume of any input without reducing the volume of output (technical efficiency).
2. The cost of any given level of output is minimised by combining inputs in such a manner that one input cannot be substituted for another without raising the total cost (economic efficiency).
3. The mix of outputs of different goods or services produced from the given resources maximises the benefit to the consumers (allocative efficiency).

In fact, any consideration or measure of efficiency tends to focus on just two strands, corresponding to the elements below:

> ".. the first is doing more with the same, or fewer, inputs (efficient day-to-day use of resources) and the second, a rigorous scrutiny of alternative services (efficient investment).." HAYWOOD (1974)

These two elements are referred to as technical and allocative efficiency.

Addressing allocative and technical efficiently independently is seen to be important as they 'clearly have quite different policy implications' (Barrow and Wagstaff, 1989).

2.3.2 Technical Efficiency

As defined above by Metcalfe and Richards (1990), technical efficiency is a measure of the physical use of resources. Mooney and Ludbrook (1995) have related the search for technical efficiency in terms of devising the least costly solution to meeting a particular accepted objective and as a question about *'how to?'* rather than *'whether to?'*

Therefore, technical inefficiency, according to Barrow and Wagstaff (1989), is the result of too little output being produced by a given set of inputs. Such a situation occurs if the amount of output produced from a given set of inputs is regarded as too low, or the amount of inputs required to produce a given output is too high.

The Audit Commission is the public body responsible for the investigation of performance in the NHS in England and Wales. The definition it uses for an efficient public sector organisation, given below, can be seen to correspond to the definitions of technical efficiency included in this section.

> "An efficient operation produces the maximum output for a given set of resource inputs, or it uses the minimum of inputs to produce a given quantity and quality of service provided."
>
> GODDARD (1989)

2.3.3 Allocative Efficiency

Allocative efficiency is a consideration of the way in which resources are allocated, combining the needs of the users, the costs of the resources and the nature of the process. According to Crush and Ball (1995), allocative efficiency comes with deploying resources to particular service elements in such a way that it is impossible to achieve greater resource efficiency by re-deploying these same resources to other areas, or in a different proportion.

Mooney and Ludbrook (1995) suggested that the pursuit of allocative efficiency centres on the objective of using the resources available in the most beneficial way for society, whilst being aware that the scarcity of resources means that not all objectives can be met.

Barrow and Wagstaff (1989) reported that allocative inefficiency therefore occurs where inputs are employed in the wrong proportions given their prices and productivity, such as employing highly trained staff to perform simple tasks. Allocative efficiency, thus, is not simply related to the relationship between physical inputs and outputs, as was determined to be the case for technical efficiency. There are other factors to be considered, including the differing financial costs of the inputs and the overall benefits to the public sector system.

Since the elements of efficiency have now been appropriately defined, the debate switches to the processes involved in achieving efficient practice in a public sector service.

2.4 Achieving Efficiency

Technical efficiency has been defined to be the measure of the relationship between inputs and outputs. A process or organisation, therefore, achieves efficiency when:

> "..output is maximised from a given input of resources."
>
> BAGGOTT (1994)

However, this definition relied on the assumption that it is actually possible to achieve efficiency, a view that has been questioned:

> "Efficiency has been likened to a state of grace: something to which all aspire, but which few achieve and then only fleetingly."
>
> BROOKS (1985)

A similar theory is that proposed by Kelly and Glover (1996), who suggested that is unlikely that the NHS can ever be managed efficiently:

> "Our argument is with the notion that in a service so large and so complex and diverse as the NHS, containing conflicting professional groups, a wide variety of functions and dealing with the general public as its client group, there is an ultimate answer to the efficiency problem. The best that can probably be attained is to minimise inefficiency." KELLY AND GLOVER (1996)

Therefore, if attaining efficiency is not a realistic proposition, one of the aims of public sector organisations has thus become to improve efficiency, as was in fact proposed earlier as one of the aims of management within a public sector organisation. The following five points have been suggested as a course of action to achieve improvements in efficiency:

- Define the intended outputs and their associated costs;
- Gauge the impact on output of a change in inputs after allowing for factors beyond the control of the management concerned;
- Establish whether input minimisation or output maximisation is the objective in particular services;
- Assess the scope for improving technical efficiency: given the existing mix of inputs, can output be increased or alternatively can total resources be reduced for a given output?
- Assess the scope for improving allocative efficiency: given existing output, can inputs be substituted for one another so as to reduce total costs? LEVITT AND JOYCE (1987)

The general points discussed are now related to a health service context directly.

2.5 Efficiency and Health Services

Previously, the discussion on efficiency has been in general terms, related to public sector services as a whole. In this section, the concepts of efficiency introduced above are linked to the health service environment, in terms of profile, response, practice and measurement difficulties.

2.5.1 Profile of Efficiency

Efficiency has come to the forefront of discussion relating to the provision of public sector services because there are now severe restrictions on the level of finance available and constraints are being placed on all areas of spending. In fact, some of the primary responsibilities of the Audit Commission are connected to efficiency:

> "..to promote 'best practice' in local government and NHS bodies, encouraging economy, efficiency and effectiveness in both the management and delivery of services."
>
> AUDIT COMMISSION (1997)

This move towards focusing on efficiency improvements would appear to be well warranted:

> "Decision-making in the NHS is seen to be political, irrational, and consequently not conducive to efficiency. There is also a lack of clarity in setting objectives and determining priorities, and so in pursuing efficiency. Incentives to be efficient are lacking."
>
> MOONEY AND LUDBROOK (1984)

As early as 1971, in a review of efficiency in the NHS, Cochrane (1971) identified that inefficiencies existed within the Health Service, and that these could be divided into four main groups:

- The use of ineffective therapies;
- The inappropriate use of effective therapies;
- The inappropriate use of health care setting;
- Incorrect lengths of stay in treatment facilities.

The profile of efficiency has also been raised in the USA, where experts have reported that rising hospital costs may actually be the result of internal inefficiency and waste in the use of hospital services (Chilingerian and Sherman, 1990). Indeed, the reforms currently being debated have focused in this area:

> ".. the productive efficiency of US hospitals operating under different constraints remains under increased scrutiny."
>
> BURGESS AND WILSON (1998)

2.5.2 Managerial Response to Efficiency

The response to the elevation of efficiency to a central concern of management has been mixed. In the UK, efficiency drives were introduced in the NHS in the 1980's, portrayed by the government as a necessary and worthwhile process for improving the service for everyone. This type of approach to efficiency has been summarised thus:

> "Efficiency, portrayed as a purely technical, instrumental means to politically approved ends is often presented as an unqualified good like apple pie or motherhood."
>
> METCALFE AND RICHARDS (1990)

However, the reaction within the NHS to efforts to improve efficiency has been very different and resentment existed, according to Mooney and Ludbrook (1995). This may be because efficiency was equated with reductions in funding levels and changes were enforced rather than introduced with co-operation on all sides. The way in which efficiency has been perceived, therefore, has not helped the attempts to improve it:

> "Because efficiency is too often equated with cheapness or penny pinching, the desirability of being efficient is lost."
>
> MOONEY AND LUDBROOK (1995)

The same authors, however, believed that the reaction of the NHS had been unfortunate because of the expressed viewpoint that efficiency 'is not something to be despised or feared' (Mooney and Ludbrook, 1995).

Additional concerns with the introduction of efficiency strategies have centred on the view that they have concentrated on short-term gains rather than long-term planning. In many regards, therefore, efficiency in relation to the provision of health care has become a political issue rather than a neutral concept for improving service. This debate on long-term versus short-term efficiency strategies is discussed again in later chapters, following the introduction of alternative methods for the assessment of efficiency, notably Data Envelopment Analysis. These strategies will focus on how efficiency can be investigated under a long-term perspective, with targets identified for each stage of improvement.

2.5.3 Efficient Practice

The simplest way in which efficiency is related to the provision of health care is that, efficiency:

> ".. is about organisation, i.e. the internal day-to-day running of health services. It can simply be a matter of comparing 'inputs' with 'outputs', assuming that an efficient service has more outputs for given inputs than a less efficient service; for example, a higher number of patients treated in a lower number of beds."
>
> FLYNN (1986)

Recent NHS documentation (Department of Health, 1997), however, suggested that efficient practices in the NHS go far beyond simply ensuring a greater number of outputs are to be produced from the given inputs. 'The New NHS: Modern, Dependable' promotes the view that patients will suffer if resources are not used efficiently and effectively.

This is in accordance with the view expressed by Culyer (1991) that efficiency in the provision of health services is exceedingly complicated and actually falls into four distinct categories:

1. Providing only services that are effective (i.e. where there is clear evidence that patients enjoy better health as a result of care);
2. Providing effective services at minimum cost;
3. Concentrating resources on effective services, provided at minimum cost, that offer the most benefits in terms of health;
4. Providing a mix of effective services at minimum cost and on such a scale that the benefits to society of providing more services are outweighed by the additional costs.

However, the pursuit of efficiency in the health service cannot be carried out in isolation, as the above discussion has suggested. The complex nature of efficiency, furthermore, creates difficulties not just in definition but also in measurement, which is considered in the next section.

2.5.4 Difficulties with Efficiency Measurement in Health Services

Standard approaches to efficiency measurement have often been found to create problems when they have been applied to the examination of health services. This can be illustrated by comparing the results from a standard method for efficiency assessment for a hotel and a hospital. A frequently used measure of the internal efficiency of a hotel is the room occupancy rate, with the equivalent of this in a hospital being the bed occupancy rate. A hotel aiming to be successful and hence profitable would require that the occupancy rate for a room be kept high.

However, the idea of a health care system is to ensure the population is healthy and thus it is better if fewer people are in hospital, equating to a low bed occupancy rate, which may be not be regarded as efficient performance. Therefore, keeping patients in hospital for unnecessarily lengthy periods of time would achieve a higher occupancy rate, but is clearly not a realistic approach to improving 'efficiency' in health services. Alternatively, reducing the number of beds available would maintain a high occupancy rate.

A similar example below considers the difficulty of determining appropriate measures of output:

> "If the output of the hospital service is defined as hospital beds and treatment facilities, efficiency may be defined as unit costs of throughput (meaning average number of patients per bed per year). However, if the throughput is defined as the ratio of discharges plus deaths to beds available, an increase in efficiency could apparently be achieved if there was an increase in the number of early deaths."
>
> FLYNN (1985)

The complex nature of efficiency, with its various components as defined in section 2.3, also clearly impacts upon the data required and the methods employed for its measurement. In order to measure allocative efficiency, data on the prices of both inputs and outputs is needed, which can be 'unobservable or have severe measurement errors' (Burgess and Wilson, 1998). However, it has been suggested that the data that is required for the measurement of technical efficiency in a health care environment (measures of physical inputs and outputs) is more readily accessible and consistent than the corresponding price data (Burgess and Wilson, 1998).

All the different components of efficiency must also be measured in order to determine the overall efficiency of an organisation, further complicating the process of efficiency measurement. For this reason, assessments of efficiency have frequently concentrated on one particular element, notably technical efficiency (Nyman *et al.* (1990), Borden (1988), Rosko (1990) and Chilingerian (1989)).

Therefore, the majority of the research in this thesis is focused on the measurement of technical efficiency. Whilst measuring only technical efficiency may underestimate the overall efficiency, in a health service context, the 'elimination of technical inefficiencies seems a desirable goal' (Burgess and Wilson, 1998).

The complex nature of efficiency and its measurement in a health service context, as described above, has introduced several other elements to the debate surrounding the performance of public sector services.

2.6 Efficiency in a Broader Context

Many of the above references to efficiency have also introduced other concepts with which efficiency appears to be inextricably linked. For example, the categories defined by Culyer (1991) in section 2.5.3 have suggested that, in a health service context, efficiency and effectiveness are very closely related.

However, efficiency can also be considered in a much broader context, in association with several other key factors, as follows.

2.6.1 Efficiency, Effectiveness and Economy

The Audit Commission, in the review of its main responsibilities, associated efficiency with economy, effectiveness and best practice (Audit Commission (1997)). This corresponds with the view of Butt and Palmer (1985), who asserted that a review of performance within the public sector must cover these three elements, denoted as the 'Three Es' by Goddard (1989), and defined as:

> **Economy:** acquiring resource in appropriate quantity and quality at the lowest cost.
>
> **Efficiency:** Making sure that the maximum useful output is gained from the resources devoted to each activity, or, alternatively that only a minimum level of resources are devoted to achieving a given level of output.
>
> **Effectiveness:** Ensuring that the output from any given activity (or the impact that services have on a community) is achieving the desired results. BALL (1996)

Economy, effectiveness and efficiency are said to be the three characteristics of a public sector service, related as shown in figure 2.2 below.

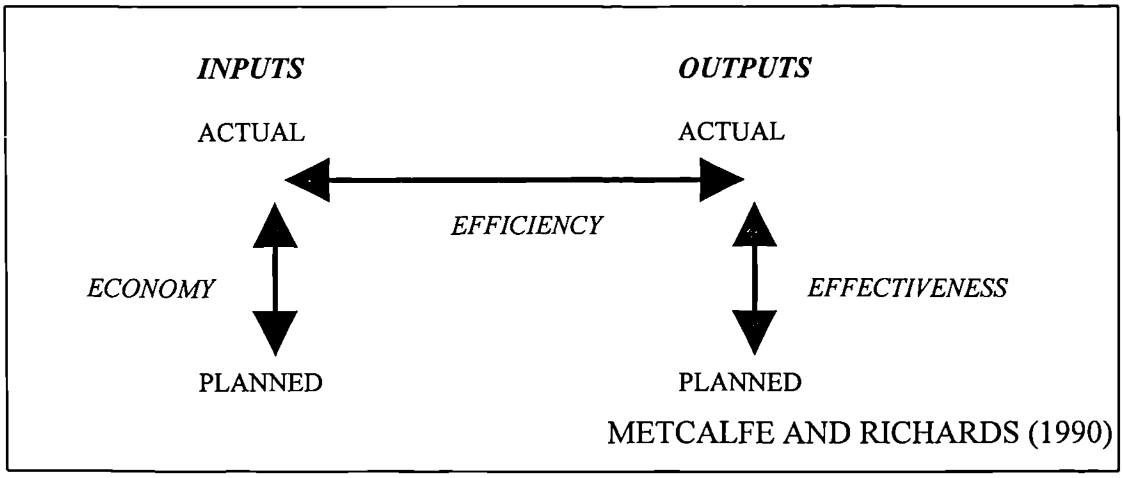


Figure 2.2: The Relationship between Economy, Efficiency and Effectiveness

In combination, economy, efficiency and effectiveness fall under the umbrella of 'Value for Money'. It can be seen that economy is the relationship between planned inputs and actual inputs, efficiency is the relationship between actual inputs and actual outputs and effectiveness is the relationship between planned outputs and actual outputs.

With reference to health services, it is the relationship between efficiency and effectiveness that is of particular interest.

> "The distinction between efficiency and effectiveness is as follows: a service is efficient if it does well; it is effective if it produces benefit." ST. LEGER ET AL. (1992)

Baggott (1994) ascertained that health services must be both effective and efficient, whilst Flynn (1986) suggested that the health service cannot be classified as efficient unless it is effective - resources are wasted, and thus inefficiencies exist, if treatments are ineffective. An efficient organisation that is ineffective would ultimately produce undesired outputs and outcomes. Effectiveness, however, is much more difficult to classify as it takes the discussion beyond the quantifiable outputs.

Outputs were defined as surrogate measures for the benefits achieved by resource input into a particular public sector unit. Outputs should produce outcomes in the long term, which become benefits if the outcomes are positive. Taking a simple example of an orthopaedic unit within a hospital, where financial resources are transformed into replacement hips, a useful output measure could be the number of patients discharged after a hip replacement operation.

Efficiency could be improved by, for example, reducing the cost of each operation to the lowest possible level. The effectiveness of the department is measured in terms of how many of these hip replacement operations actually result in full or partial recovery (measured by a mobility index), and how many require additional hospitalisation - the long term outcomes of the financial inputs.

2.6.2 Equity and Equality

In the 1980s, much of the emphasis of health care management was on providing efficient services. However, recently other ideas have become much more relevant, including equity and equality, which are central to recent proposals by the Labour Government (Department of Health, 1997).

Equity in health care means 'equal access to health services' - all patients should receive the same level and quality of service, with access based only on need. It has been described as 'fairness through co-operation' (Department of Health, 1997a) and requires the overall quality of the health service to move up towards 'best practice' rather than down to the lowest common level.

Equality is a slightly different concept, requiring that everyone should have the same opportunity to be healthy. The focus of improving equality in health is on tackling the social and environmental problems that are linked to poor health. These include such things as low income and unemployment, poor housing and pollution (Department of Health, 1997b). Achieving equality involves co-operation across all departments and an integrated social policy.

It was stated earlier that effectiveness and efficiency should not be considered independently within the health service, that is, efficiency without effectiveness is not actually efficiency. Other concepts such as equity and equality are now being treated as equally important, to be pursued alongside efficiency and not be neglected in order to achieve efficiency targets.

2.6.3 Quality of Care

It has also been suggested that efficiency is just one of a number of factors, listed below, which should be considered in the broader context of achieving quality of care, with all of them to be regarded as equally important.

St. Leger et al (1992) gave the seven aspects of quality care as:

- Access to services;
- Relevance to need for the whole community;
- Equity;
- Social acceptability;
- Effectiveness;
- Efficiency;
- Economy;

'The New NHS: Modern, Dependable' (Department of Health, 1997) also puts forward the link between quality and efficiency, saying that they go 'hand-in-hand.' The focus of efficiency drives will be to ensure 'that every pound in the NHS is spent to maximise the care for patients'.

In addition, encouraging a move away from the efficiency drives that characterised government policy for health care in the early 1980s (see chapter three), the emphasis of the NHS will change:

> ".. to shift the focus into quality of care so that excellence is guaranteed to all patients, and quality becomes the driving force for decision-making at every level of the service."
>
> DEPARTMENT OF HEALTH (1997)

Clearly, the relationship between quality and efficiency is complex, with a comprehensive amount of research devoted to the assessment of quality in public sector services, beyond the scope of this investigation, where the impact of quality is only to be considered in more general terms.

The links between efficiency and all these factors under the broad spectrum of public sector performance as discussed above, however, are important as they can be seen to have some impact on the methods employed for their measurement. In the example presented earlier relating to the provision of services at an orthopaedic department in a hospital, measuring the efficiency in terms of procedure cost is relatively straightforward. Measuring effectiveness in relation to the number of successful operations is slightly more complex, although not impossible.

This discussion has introduced the key elements of performance in a health service context, related particularly to the complexities involved in its measurement. The following section discusses the main approaches that have been applied to date.

2.7 Methods for Measuring Efficiency and Performance

There are several methods used to measure the efficiency of public sector organisations and the use of resources within them. A brief introduction to each of these methods is given below, together with an evaluation of their strengths and weaknesses, before the key technique of Data Envelopment Analysis is introduced.

2.7.1 Ratio Analysis

One of the simplest and most frequently used methods for assessing performance and efficiency is that of ratio analysis. A single input can be related to a single output in the form of a ratio, and efficiency is calculated according to the following equation:

$$Efficiency = \frac{Output}{Input}$$

Figure 2.3: The Efficiency Ratio

The ratio definition above relates one input and one output in order to measure efficiency. Other aspects of performance can be investigated by relating two inputs in the form of a ratio. Outputs can be examined in relation to the population as a whole. Examples of the types of ratios that are calculated in relation to the health service are:

- Cost per patient day;
- Deaths per 1000 patients;
- Bed occupancy rate;
- Number of nurses per consultant;
- Patients treated per thousand of the population.

The appeal of ratio analysis as a means for efficiency and performance assessment lies in its apparent simplicity - ratios are easy to calculate and would appear to be easy to understand. The amount of data available that can be interpreted as ratios is plentiful, especially with the widespread use of computers and the growth of information technology within the health services.

Ratios are interpreted, when comparing performance across several units, by looking at the value of the ratio for each unit alongside the mean value across all units. A value that differs significantly from the mean, being either abnormally high or low, is suggestive of unusual, or potentially inefficient, behaviour. This level for 'abnormality' is usually defined in relation to the standard deviation, that is, either one or two standard deviations from the mean. In this regard, Sherman (1986) suggested that ratios could be useful in identifying 'extremely good or extremely poor' operating relationships. Using ratios as a means for comparison with past performance, or the performance of other units, is also often referred to as benchmarking.

There are however many drawbacks with using ratios as a means for examining efficiency. The most serious limitation of ratio analysis is the essence of its simplicity - they only relate one input to one output. In complex organisations such as the NHS, this does not give a complete picture of how the service is operating:

> "They cannot constitute a single measure of hospital efficiency whereby it is immediately possible to focus upon hospitals which are classified as inefficient and identify, with ease, the areas in which these inefficiencies are likely to be occurring."
>
> CRUSH AND BALL (1995)

Problems can also be encountered when examining particular ratios in isolation, as with some of the results published in the form of league tables. A frequently quoted example for this is the ratio 'death rate per 1000 patients' (see, for example, Crush and Ball, 1995). Interpreting this value is fraught with danger - there are no guidelines to say which values are 'too high' and no account is taken of environmental factors or the complexity of cases undertaken.

The above problem has been counteracted by the calculation of large sets of ratios. However, this presents an additional and equally significant problem:

> "One hospital may appear relatively efficient on one group of ratios and inefficient on another group. Another hospital may have the opposite result for the same ratios... Consequently, it is difficult to conclude which hospitals are inefficient using ratio analysis."
>
> SHERMAN (1984)

Table 2.1 summarises the strengths and weaknesses of ratio analysis as a tool for efficiency assessment discussed above:

Strengths	Weaknesses and Limitations
• Useful in identifying which particular aspects of hospital operations deviate from the norm. • Involves simple and easily understandable mathematics.	• No objective way of pinpointing inefficient hospitals in a comparison or of defining what efficiency is in comparison to the mean. • Different results on performance can be obtained depending on the ratios used. • Difficult to agree on what is an abnormally high or low ratio. • Can only handle single inputs and outputs.

Table 2.1: Strengths and Weaknesses of Ratio Analysis

Despite the many reservations, ratios are in fact used extensively as tests of efficiency in the NHS and the Health Care systems in the United States. In the UK, performance indicators, many of which are ratios of inputs to outputs, have become the most accepted method for examining efficiency and performance following the introduction of the first performance indicator package in 1981. Performance indicators are discussed comprehensively in chapter four.

2.7.2 Econometric Regression Analysis

Regression techniques have been used to examine performance within health services, particularly in the USA, but to a lesser degree than ratio analysis. It has been discussed in relation to the NHS in theoretical terms, although not used in many practical investigations. Sherman (1984) suggested some of the different types of regression analyses which have been carried out in relation to health services:

- Marginal cost per patient;
- Breakdown of fixed cost versus variable cost;
- Existence of economies of scale.

Regression analysis is a technique designed to explain the relationships between specific variables and to predict the value of one dependent variable using the values of the numerous independent variables. The equation calculated fits a straight line through the data, which minimises the square of the distance between the observations and the regression line. Smith and Mayston (1987) suggested that regression analysis is generally used to quantify the relationship between any number of inputs and a single output.

However, this can be reversed, with a single input related to numerous outputs as in an examination of the impact of case mix on total cost. In the first case, with one output related to multiple inputs, the procedure for estimating efficiency would be as follows:

> 1. A measure of output is taken as the dependent variable.
> 2. Inputs are identified, which can be environmental factors and resource inputs, to be used as independent variables.
> 3. The significant inputs (the determinants of output) are found using step-wise regression techniques - all the factors are examined and the insignificant inputs are removed one by one until the final regression equation is obtained.
> 4. The measure of efficiency is the residual error - the difference between the actual output and the output predicted by the regression model. SMITH AND MAYSTON (1987)

The primary advantage of regression techniques over ratio analysis is that they can handle multiple inputs measured against a single output, or vice versa. However, regression analysis also has some important and well-observed limitations. Due to the nature of the technique, as with the application of ratio analysis, performance is related to the mean rather than a frontier:

> "The use of least-squares regression techniques results in estimates of average (or central tendency) relationships, which are not necessarily efficient relationships." SHERMAN (1984)

In cases where regression is used to relate multiple inputs to a single output, the extent to which the objective associated with the particular output is being met is indicated.

However, the inter-dependence between all the outputs and subsequent objectives is not considered (Smith and Mayston, 1987). In other words, regression techniques can not be easily applied to the evaluation of performance in a broader spectrum, including the investigation of effectiveness.

The question of interpretation is also a problem. If costs are examined against case mix, as pointed out by Sherman (1984), hospitals with costs some arbitrary distance from the mean are labelled as potentially inefficient. There are, again, no guidelines to determine how far this distance should be (before a label of inefficient is attached to a given hospital), although the notion of 'two standard deviations from the mean' is again frequently applied. Table 2.2 summarises the strengths and weaknesses of regression analysis:

Strengths	Weaknesses and Limitations
• Can handle multiple inputs/outputs to a certain degree. • Useful in understanding the characteristics of what impacts on costs, for example case mix.	• Uses estimates of relationships that may not be efficient relationships to assess efficiency. • Mean relationships found do not directly locate inefficient hospitals. • Would only provide insights into efficient hospital behaviour if all the hospitals in the study were known to be efficient.

Table 2.2: Strengths and Weaknesses of Regression Analysis

2.7.3 Clinical Audit

Both regression and ratio analysis are focused on the evaluation of performance at a general level. For example, ratios such as 'cost per patient discharge' are calculated for a whole specialty or categories of procedures.

However, it is possible to evaluate performance using a much more detailed approach, focused on the identification of very specific examples of poor performance, in relation to the medical treatment provided and its outcome. Thus, in recent years, the evaluation of performance at the individual case or patient level has become prevalent in the health service in the UK.

This approach to evaluation, referred to as medical or clinical audit, was made compulsory in the NHS by the last major reforms introduced by the Conservative Government, discussed in detail in chapter three (Department of Health, 1989).

Prior to this point, the systematic reviewing of past cases that makes up the audit process at each hospital had been done at the discretion of the individual consultant, rather than as a specified part of their workload. However, many hospitals carried out extensive audits and several were also conducted on a national basis, such as the Confidential Enquiry into Peri-Operative Deaths (CEPOD) (Baggott, 1994).

The profile of the audit process has been raised considerably in the United Kingdom over the past few years, not just as a result of its compulsory status. Several instances of poor performance have highlighted the need for this type of performance evaluation. The treatment of young children with severe heart problems dating back to 1988 at the Bristol Royal Infirmary received a great deal of media attention. The investigation by the General Medical Council concluded that 'too many babies were dying' (The Daily Telegraph, 30/05/98). The public enquiry into this area is ongoing.

Another area of health care treatment where problems have been identified using audit is that of cancer care and cancer screening. For example, more than 150 patients were found to have received too much radiation during treatment at the Royal Devon and Exeter Hospital in 1988 (The Sunday Telegraph, 01/06/97). At the North Staffordshire Royal Infirmary, it was discovered in 1991 that over 1,000 patients had been given too little radiation (The Sunday Telegraph, 01/06/97). In these cases, the audit approach was used to identify instances where the treatment provided has not reached the required clinical standard, as identified by the medical authorities or the hospitals.

The limitations of the audit approach are seen to be focused in two main areas: its potential impingement on clinical innovation and its over-reliance on hard data rather than measures of quality and patient satisfaction (Allsop, 1995). An extensive review of the audit process, including an evaluation of its benefits and limitations is given in chapter four.

2.8 The DEA Approach to Efficiency Measurement

Data Envelopment Analysis is a method for measuring efficiency based on the idea of the efficient frontier. It was named and developed by Charnes, Cooper and Rhodes (1978), extending the ideas of efficiency first proposed by Farrell (1957). It is based upon the linear programming methodology and is used to measure the relative efficiency of a number of operating units, ideally having similar goals and objectives. It has become more widely used over the last fifteen years with assessments made of electricity and telecom companies, schools and universities, hospitals and health service programmes and banks, amongst others.

DEA has generally focused on the measurement of technical efficiency, to which it is more readily suited than the evaluation of allocative efficiency, or the other types of efficiency, and effectiveness. The focus on technical efficiency corresponds with the emphasis placed on this in the evaluation of health service performance by such organisations as the Audit Commission, as discussed in section 2.3. Technical efficiency is a measure of how resources are used to produce particular outputs, rather than the nature of the output mix itself or the relationship to the achievement of organisational goals and objectives.

It is the process of transforming inputs into outputs to which the DEA efficiency scores is naturally related and along which framework it has been developed:

> "Our measure is intended to evaluate the accomplishments, or resource conservation possibilities, for every decision making unit with the resources assigned to it. In golfing terminology it is, so to speak, a measure of "distance" rather than "direction" with respect to what has been (and might be) accomplished."
>
> CHARNES ET AL. (1978)

The measurement of allocative efficiency has been investigated through various adaptations of the basic DEA methodology, such as those proposed by Byrnes and Valdmanis (1994). The introduction of weight restrictions, which is discussed in chapter five and then applied in chapter eight, has also led to further exploration of allocative efficiency, although through more indirect means. The DEA approach can also be used to investigate other elements of performance, such as those introduced in section 2.6, including effectiveness and quality of care. This is through the introduction of specialised factors and the interpretation of DEA results.

Efficiency in the DEA model is calculated for each unit as the ratio of a weighted sum of outputs to a weighted sum of inputs, with the proviso that similar ratios for every unit in the sample be less than or equal to one with the same weights applied. Units that are evaluated as efficient can have different amounts of each of the inputs and outputs.

As a technique, DEA has several advantages over the previously described approaches. It is able to handle multiple inputs and outputs, which need not simply be resource values and output levels. Environmental, categorical and quality factors can also be included.

The DEA approach produces a comprehensive summary statistic for the efficiency of each unit, without the problems encountered when using ratios of looking at one value in isolation or obtaining varied results for different sets of ratios. This also removes the arbitrary element of determining efficiency, which is encountered with both ratio and regression analysis.

DEA is also proposed as a more appropriate method for examining the efficiency of public sector units, in that it is based on the idea of an efficiency frontier, made up of the units that are operating efficiently. The results of the analysis are based on extremal methods (at the efficiency frontier) and not average values (Chalos and Cherian, 1995). Also, even those units that are judged to be efficient can still identify scope for improvement by looking at the 'best practice' of some of the other units.

There are, of course, limitations with the technique of DEA, particularly as it has developed considerably. It involves complicated mathematical approaches that are not easily understood by those to whom efficiency analyses are of greatest concern. Some of the approaches involved, for example, in the selection of variables, are highly subjective. The results can also be highly sensitive to errors in the data, notably extreme values, which are common in public sector services.

The key strengths and weaknesses of Data Envelopment Analysis, summarised in table 2.3 overleaf, are further debated in chapter five, including a comprehensive description of all aspects of the technique. This includes a detailed investigation into the many stages involved in the application of the DEA methodology.

To conclude this section, the position of DEA in relation to other techniques is presented:

> "DEA is not, however, a panacea... Rather, it is proposed as a complement to ratio analysis and regression techniques to gain insights beyond the reach of these other techniques. Moreover, DEA results can actually increase the value of subsequent use of ratio and regression analysis." SHERMAN (1986)

	Strengths	Weaknesses and Limitations
Measuring efficiency	• Measures efficiency, compared with the best practice frontier, not mean performance. • Considers units in best possible light. • Produces a single performance measure from multiple inputs and outputs – less complex. • Indicates the general magnitude of the inefficiencies.	• Locates only relative inefficient units - cannot locate all, as all units in set may be inefficient. • Does not identify the actual efficiency frontier. • Units with unusual or extreme input/output mixes usually defined as efficient.
Weights	• Do not need to fix a common set of weights prior to analysis. • Weight restrictions can be introduced to include prior judgement, if required.	• Can set weights at virtually zero – units can ignore some inputs and outputs if unfavourable. • Weight restrictions are highly subjective – objectivity of technique is lost.
Selection of variables	• Can handle multiple inputs and outputs explicitly and simultaneously. • Inputs and outputs do not need to be commensurate, that is, measured in the same units. • Can include uncontrollable variables, such as environmental factors, and quality measures.	• Highly sensitive to missing or incorrect data - could falsely identify a unit as efficient. • Excluding important variables (or including irrelevant ones) can seriously affect results. • Results dependent on variable choice – subjective.
Robustness	• Units identified as inefficient will have inefficiencies as least as large as those identified.	• Results very sensitive to variables chosen.
Results	• Provides information on relative efficiency, reference sets and targets to be used to improve efficiency. • Provides new managerial and theoretical insights for organizing and analysing data.	• Does not locate operating practices that produce inefficiency or the optimal path to improve efficiency - directs attention only. • Danger of directing attention to a relatively unimportant variable.

Table 2.3: Strengths and Weaknesses of Data Envelopment Analysis

2.9 Conclusions

There are clearly many issues concerning the definition, measurement and assessment of efficiency in public sector services. It is, however, apparent that accurate and relevant techniques for the measurement of efficiency are required. Data Envelopment Analysis has been proposed as a suitable and valuable method to be used in the assessment of efficiency within health services, with this further investigated in later chapters.

However, as was discussed in sections 2.5 and 2.6, there are clearly specific elements of the health service in general and hospitals in particular that are likely to be important in the development and application of efficiency and performance measures. Therefore, the nature of the NHS in the UK, upon which the subsequent analysis is to be based, is investigated in the following chapter, focusing on its organisational structure and its impact on performance assessment. The impact on the methods currently employed for performance assessment, notably performance indicators and clinical audit, are then discussed in chapter four, prior to the extensive application of the DEA technique in the remaining chapters.

2.10 References

1. **Allsop, J.** (1984) Health Policy and the National Health Service, Longman.

2. **Allsop, J.** (1995) Health Policy and the NHS: Towards 2000, Longman.

3. **Audit Commission** (1997) The role of the Audit Commission, Internet Document.

4. **Baggott, R.** (1994) Health and Health Care in Britain, MacMillan.

5. **Ball, R.** (1996) Performance Measurement in the Public Sector, Paper presented at IFORS, July 8th-12th, 1996.

6. **Barrow, M. and Wagstaff, A.** (1989) Efficiency Measurement in the Public Sector: An Appraisal, *Fiscal Studies*, Vol. 10, Issue 1, pp. 72 – 97.

7. **Borden, J. P.** (1988) An Assessment of the Impact of Diagnosis-Related Group (DRG) - Based Reimbursement on the Technical Efficiency of New Jersey Hospitals Using Data Envelopment Analysis, *Journal of Accounting and Public Policy*, Vol. 7, No. 2, pp. 77 – 96.

8. **Brooks, R.** (1985) Efficiency in Health Care in Health Services Performance: Effectiveness and Efficiency, Long, A. F. and Harrsion, S. (Eds.) Croom Helm.

9. **Burgess, J. F. and Wilson, P. W.** (1998) Variations in Efficiency among US Hospitals, *INFOR*, Vol. 36, No. 3, pp. 84 – 102.

10. **Butt, H. and Palmer, B.** (1985) Value for Money in the Public Sector: The Decision-Maker's Guide Oxford, Basil Blackwell.

11. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

12. **Chilingerian, J. A. and Sherman, H. D.** (1990) Managing Physician Efficiency and Effectiveness in Providing Hospital Services, *Health Services Management Research*, Vol. 3, No. 1, pp. 3 – 15.

13. **Chalos, P. and Cherian, J.** (1995) An Application of Data Envelopment Analysis to Public Sector Performance Measurement and Accountability, *Journal of Accounting and Public Policy*, Vol. 14, No. 2, pp. 143 – 160.

14. **Charnes, A., Cooper, W. W. and Rhodes, E.** (1978) Managerial Economics - Past, Present and Future, *Journal of Enterprise Management*, Vol. 1, No. 1, pp. 5 – 23.

15. **Chilingerian, J. A.** (1989) Investigating Non-medical Factors Associated with the Technical Efficiency of Physicians in the Provision of Hospital Services: A Pilot Study, *Annual Best Paper Proceedings of the Academy of Management*, pp. 85-89.

16. **Cochrane, A. L.** (1971) Effectiveness and Efficiency: Random Reflections on Health Services, Nuffield Provincial Hospital Trust.

17. **Crush, J. and Ball, R.** (1995) The Measurement of Efficiency in Healthcare - The Benefits of a DEA Approach, University of Stirling Unpublished Working Paper.

18. **Culyer, A.** (1991) The Promise of a Reformed NHS: An Economists Angle, *British Medical Journal*, Vol. 302, pp. 1253 – 6.

19. **Department of Health** (1997a) Fairness and Equity for Hospitals, Press Release 97/169, 16.7.97.

20. **Department of Health** (1997b) Government Takes Action to Reduce Health Inequalities, Press Release 97/192, 11.8.97.

21. **Department of Health** (1997) The New NHS, Modern, Dependable, HMSO.

22. **Department of Health** (1989) Working for Patients, HMSO, Cmnd 555.

23. **Farrell, M. J.** (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A (General)*, Part III, Vol. 120, pp. 253 – 279.

24. **Flynn, N.** (1986) Performance Measures in Public Sector Services, *Policy and Politics*, Vol.14, No. 3, pp. 389 – 404.

25. **Goddard, A.** (1989) Are Three "E"s Enough?, *OR Insight*, Vol. 2, No. 3

26. **Haywood, S. C.** (1974) Managing the Health Service, George Allen and Unwin Ltd.

27. **Kelly, M. and Glover, I.** (1996) In Search of Health and Efficiency: the NHS 1948 - 1994 in Beyond Reason: the National Health Service and the Limits of Management, Leopold, J., Glover, I. and Hughes, M. (Eds.), Avebury.

28. **Levitt, M. S. and Joyce, M. A. S.** (1987) The Growth and Efficiency of Public Spending, Cambridge University Press.

29. **Metcalfe, L. and Richards, S.** (1990) Improving Public Management, SAGE Publications.

30. **Mooney, G. H. and Ludbrook, A.** (1984) The NHS: Efficiency need not be a Dirty Word, *British Medical Journal*, Vol. 288, No. 6433, pp. 1817 – 1818.

31. **Nyman, J., Brick, D. and Link, D.** (1990) Technical Efficiency in Nursing Homes, *Medical Care*, Vol. 28, No. 6, pp. 541-551.

32. **Rosko, M. D.** (1990) Measuring Technical Efficiency in Health Care Organizations, *Journal of Medical Systems*, Vol. 14, No. 5 pp. 307-322.

33. **St. Leger, A. S., Schnieden, H. and Walsworth-Bell, J. P.** (1992) Evaluating Health Services' Effectiveness, Oxford University Press.

34. **Sherman, H. D.** (1986) Managing Productivity of Health Care Organisations in (1986) Measuring Efficiency: An Assessment of Data Envelopment Analysis (No. 32 in New Directions for Program Evaluation Series), Silkman, R. H. (Ed.), American Evaluation Association, Jossey Boss Inc.

35. **Sherman, H. D.**(1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927 – 938.

36. **Smith, P.** (1995) Performance Indicators and Outcome in the Public Sector, *Public Money and Management*, Vol. 15, Pt. 4, pp. 13 – 16.

37. **Smith, P. and Mayston, D.** (1987) Measuring Efficiency in the Public Sector, *Omega*, Vol. 15, No. 3, pp. 181 – 189.

38. **The Daily Telegraph** (30/05/98) Too Many Babies were Dying (Electronic Telegraph).

39. **The Sunday Telegraph** (01/06/97) Women Mutilated by Cancer Treatment (Electronic Telegraph).

CHAPTER THREE

'SQUARING TWO CIRCLES' – THE CHANGING ORGANISATIONAL STRUCTURE OF THE NHS AND ITS IMPACT ON, AND RELATION TO, EFFICIENCY AND PERFORMANCE ASSESSMENT.

3.1 Introduction

The National Health Service, throughout its brief history, has faced a constant struggle to reconcile the tensions that have been a product of its organisational structure, described by Klein (1989) as 'trying to square two circles'. The problems for the NHS exist in determining who should have control of services and who should decide how the money should be spent. This can be summarised thus:

> "First it tries to reconcile central funding and Government accountability for national standards of service with the need for local autonomy to meet local need.... The second tension results from the compact struck between Government and the medical profession in 1948, which balanced central accountability for raising and allocating finance with clinical freedom to spend it."
>
> RANADE (1994)

Despite these tensions, the National Health Service has survived for more than fifty years, although with an ever-changing structure. However, according to Kelly and Glover (1996), its development in Britain has been based on a single line of evolution, as follows:

> "First, that Health Services can be managed in a rational way. Second, that health care can be delivered efficiently. Third, that efficiency will produce health for the population at large."
>
> KELLY AND GLOVER (1996)

The development of the NHS is presented, beginning with its introduction in 1948 through a series of reforms, which have affected its organisational structure and the methods employed for performance assessment.

This evaluation concludes with the most recent changes to the NHS, proposed by the Labour Government since its victory in the General Election in 1997.

In the previous chapter, the theory surrounding the measurement of efficiency and other elements of performance was introduced. It was observed that the nature of the Health Service environment has a critical impact on the definitions of these concepts and the methods employed for their measurement. Particular attention is given to how efficiency and performance have been measured during the development of the NHS and how the changing structure has been related to, or impacted on, these methods.

Throughout the chapter, indicators of performance are referenced in order to assess the impact of changes in organisational structure on the services provided by the NHS. The intention of such a discussion is to provide a context for the efficiency analyses that follow.

The financial position of the NHS in the UK is also examined, in relation to both the changing levels of expenditure during its fifty-year history and also in comparison with health care expenditure in other countries. The rationale for such an investigation is to determine if any funding problems within the NHS are caused by financial or economic problems.

To conclude the chapter, the long-term future of the NHS is assessed, in terms of any further possible changes to its structure and management and their impact on levels of performance and measures of efficiency.

3.2 The Origins of the NHS and its Early Development (1948 - 1974)

The NHS was set up after the Second World War, although planning had begun in the late 1930s, by the Acts of Parliament of 1946 for England and Wales and 1947 for Scotland, to ensure that everyone could have access to the best care available. The driving force behind the formation and initial development of the NHS, and the name most associated with it, is generally acknowledged to be Aneurin Bevan, the Minister for Health in the post-war Labour Government. At the time, this approach to health care was a pioneering concept. The Government's intention was:

> ".. to ensure that in future every man, woman and child can rely on getting all the advice and treatment and care they may need in matters of personal health; that what they get shall be the best medical and other facilities available; that their getting these shall not depend on whether they can pay for them, or any other factor irrelevant to real need." MINISTRY OF HEALTH (1944)

Bevan described the new NHS as a 'great and novel undertaking' (Rivett, 1997). The new Labour Government, looking back over fifty years of development, described the establishment of the NHS as 'the greatest single act of modernisation ever achieved by a Labour Government' (Department of Health, 1997).

Prior to 1948, access to Health Services had been dependent on geographical location, financial status and a certain amount of luck, especially for the poorest sections of the population. It was up to the individual to make decisions on health care based on how much treatment they could afford (either by paying directly or through insurance schemes), or depend on the services provided by voluntary organisations.

The changes that came in 1948 were at the end of a long process of development. Over the previous 100 years, the responsibility for health and welfare had gradually been taken over by the state, with the introduction of a series of important pieces of legislation: the Public Health Act in 1848, the 1929 Local Government Act, concerned with provision of hospital services, and the National Insurance Act, relating to primary health care.

However, as the twentieth century progressed, the need for improvements to be made in the provision of health care was highlighted in numerous reports, detailed by Ham (1992): the British Medical Association (1930, 1938 and 1942), the Sankey Commission on Voluntary Hospitals (1937), the Royal Commission Report (1926) and the Beveridge Report (1942), which proposed state funding of health services.

There were no universal answers to the problems, which led to disagreement over what should be the nature of any reforms: the BMA favoured an extended system of health insurance whereas the Royal Commission proposed that health care could be financed from general taxation. However, most agreed that there was a need for greater co-operation in the running of hospitals and access to services should be increased.

The Second World War heightened the need for immediate action and planning began in 1942 to introduce a national health care system as soon as the war was over. The NHS that finally came into being on the July 5th, 1948, was the result of much bargaining and negotiations.

In the discussion, the medical profession represented by the BMA was very influential, as was the Minister for Health, Anuerin Bevan:

> "The shape taken by the NHS was the outcome of discussions and compromise between ministers and civil servants on the one hand and a range of pressure groups on the other...the medical profession, the organisations representing the hospital service, and the insurance committees with their responsibility for general practice."
>
> HAM (1992)

Many of the elements of the health system were the possible, taking into account the views of the stakeholders and the institutions already in place, rather than the desirable.

The structure, therefore, of the NHS in England in this early period was as laid out in figure 3.1 below.

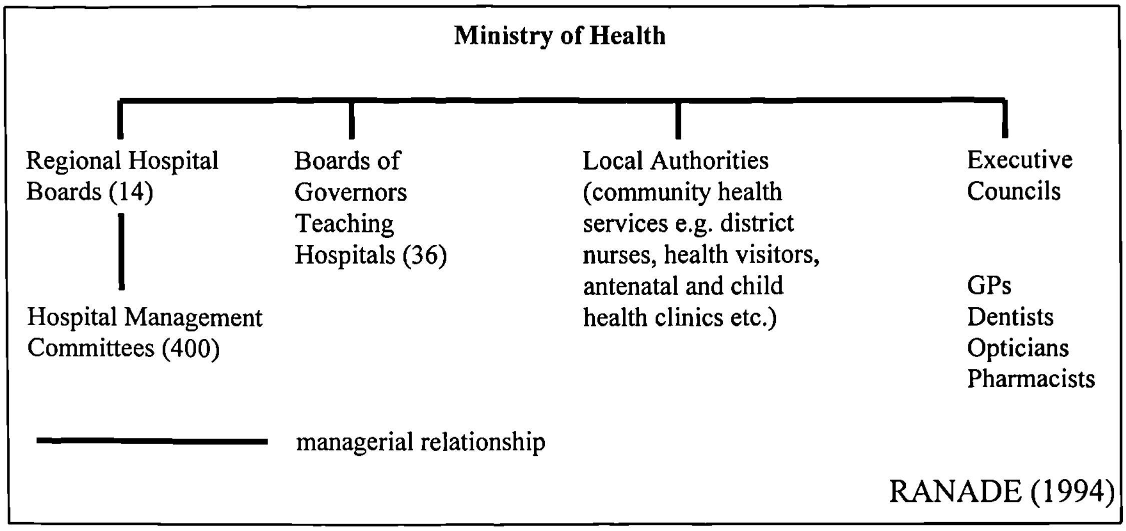


Figure 3.1: Structure of the NHS 1948-74 (England and Wales)

During this period, the Ministry of Health held overall control of the NHS in England, with slight differences in the services in the rest of Britain.

The Health Services in Wales, Scotland and Northern Ireland were under the control of their respective Government departments, as remains the situation today, although the other tiers of the organisation were consistent with the structure shown in figure 3.1.

Before nationalisation, hospital services came from three sources: voluntary hospitals run by charitable donations, municipal hospitals provided by local authorities and the emergency hospitals built during the second world war. There were about 500,000 hospital beds available by the transition to the NHS, when the responsibility for running all hospital services was co-ordinated by the Regional Hospital Boards. They took over the administration of all hospital services from the local authorities and charitable organisations, aiming to achieve the element of co-ordination as desired. Finance for the running of hospitals passed down from the Ministry of Health through the regional boards to Hospital Management Committees.

The teaching hospitals were treated separately, which gave them special recognition and more autonomy - they were financed directly from the Ministry of Health. It was intended that teaching hospitals would be distributed more equitably, so that more of the population could have access to the new methods of treatment, with at least one in each region. Previously, most of the teaching hospitals were located in London.

The executive councils, which administered the GP and associated services, were the replacements for the pre-nationalisation Insurance Committees, with funding coming directly from the Ministry of Health.

The Ministry of Health was also involved in the appointment of committee members, co-operating with local professionals and authorities. The function of the Executive Councils can be summarised thus:

> "In no sense were executive councils management bodies. They simply administered the contracts for family practitioners, maintained lists of local practitioners, and handled complaints by patients." HAM (1992)

Local authorities retained the responsibility for running the environmental and personal Health Services, with the funding coming from central Government grants and local rates, which had generally been the pattern prior to nationalisation. The most important individual was the Medical Officer of Health in each authority.

The perspective on health care generally held in the 19th Century was that medicine could only treat disease and illness to a certain degree and, as such, a Health Service had only a limited value. However, the founders of the NHS believed that, with the defined level of investment in service provision, it would be possible to improve the overall health of the population to such an extent that the NHS would not need increased funding every year - a healthy population was an attainable objective:

> ".. by the time the NHS was established in 1948, the possibility of being able to overcome the scourge of illness seemed to be within reach of the population at large. Some politicians even believed that the NHS would be so effective as to reduce absolutely the amount of illness and hence the costs of the NHS to the nation."
>
> KELLY and GLOVER (1996)

This led to a certain amount of complacency in the planning for the future health care needs of the country:

> "There would therefore be no difficult choices to be made in health care between some areas of need and others. The only important question was the length of time necessary for the 'pool of illness' to reduce." ALLSOP (1984)

The NHS in this era had changed very little since the period before nationalisation in terms of most of its organisational structure. Problems, such as variation in service according to location, staff shortages and inadequate hospital buildings, were still apparent and, in some areas, had worsened.

It was almost inevitable that there would be problems - need and demand has continued to increase throughout the last fifty years, with the NHS appearing to be in financial crisis almost from the outset.

New and more expensive treatments were also becoming the norm, with a greater amount of money to be spent on the technical aspects of health care, such as pathology. As early as 1949, there were major staff shortages and the first temporary ward closures. These early attempts at rationing services began almost immediately, through waiting lists, prioritising and the introduction of charges for dental and optical services in 1951.

The belief that demand for health care would gradually be reduced was also disputed, notably by Dr. Ffrangcon Roberts in 1952:

> "Medicine.. was expanding and new and more expensive methods of treatment were being developed. The result of this would not be the final conquest of disease but would leave doctors more difficult problems to solve - those relating to the treatment and care of degenerative and chronic illness." ALLSOP (1995)

The order of the day in this period was to find ways of containing the NHS budget, as spending on Health Services was spiralling and already exceeding the initial forecasts. However, there were inconsistencies in spending - the NHS total budget actually dropped as a proportion of GNP in 1953, although the proportion of the budget allocated to staff costs had increased.

Very few new hospitals were being built - twentieth century medicine was being practised in nineteenth century conditions. As today, the issue of NHS funding was highly contentious, with accusations of both under-funding and over-funding.

For example, the Guillebaud Committee reported in 1956 that there was no evidence of extravagance or inefficiency in the NHS, which required increased financial input in certain areas (Ministry of Health, 1956). This view was repeated in 1958:

> "It is high time to let the Health Service go on a spending spree."
> ECKSTEIN (1958)

The Conservative Government responded to the financial problems highlighted during the 1950s with the 1962 Hospital Plan (Ministry of Health, 1962), giving an additional £500 million over ten years to be used to improve hospital services, and incorporating a hospital building project.

The structure of the NHS was also under increasing criticism by the 1960s. Baggott (1994) summarised the problems as overlap, duplication and lack of co-ordination, which were the same problems identified in the pre-nationalised service. The local authorities, hospital boards and executive councils were expected to plan and work together by central Government but a lack of co-operation and communication was still a problem.

As the 1960s came to a close, issues of performance and efficiency were given prominence for the first time in the Health Service. In the early days of the NHS, the methods used to measure efficiency and performance were fairly arbitrary: total volume expenditure was used to reflect the efficiency with which costs were being contained. The performance of individual hospitals was not examined to any great extent and only a small amount of attention was given to any interpretation of outputs.

Data of various types, mainly relating to activity, has always existed in the NHS, such as the number of patients treated in each specialty, waiting times for admission, bed usage, length of stay and number of staff employed. However, this data was not really used to measure performance, the quality of service provision or efficiency.

3.3 Redeveloping the NHS (1974 - 1982)

Kelly and Glover (1996) described this phase of development as searching for efficiency through rational bureaucratic means. Ranade (1994) explained the re-organisation in terms of the twin principles of efficient management and rational planning, both to be achieved by unifying Health Services.

The Government's white paper in 1972 (Department of Health and Social Security, 1972) set out the intentions of re-organisation as 'maximum accountability upward, maximum delegation downward' (Ranade, 1994).

The first proposals for major re-organisation came in 1968 from the Labour Government of the day, at the same time as the system of Local Government was being reviewed. It was hoped that the revised Health Boards would be linked with the new local authorities. The new Health Boards were intended to operate as 'small executive management teams' with a reduced medical input into decision making.

The idea behind the reforms was to provide better co-ordination and management and eliminate the structural problems which existed, that is, the division of the Health Service into four separate strands under the Ministry of Health with little cross-communication.

When restructuring came in 1974, following the passing of the NHS Reorganisation Act in 1973 under a Conservative Government, the changes implemented had altered dramatically from the initial proposals of the Labour Government some six years earlier. The Health Boards grew in size from the small management teams envisaged in 1968, leading to increased professional involvement at every level. Figure 3.2 overleaf shows that the Health Service in England was split into three levels beneath the Department of Health: regional, area and district, with the link between local authorities and health provision broken.

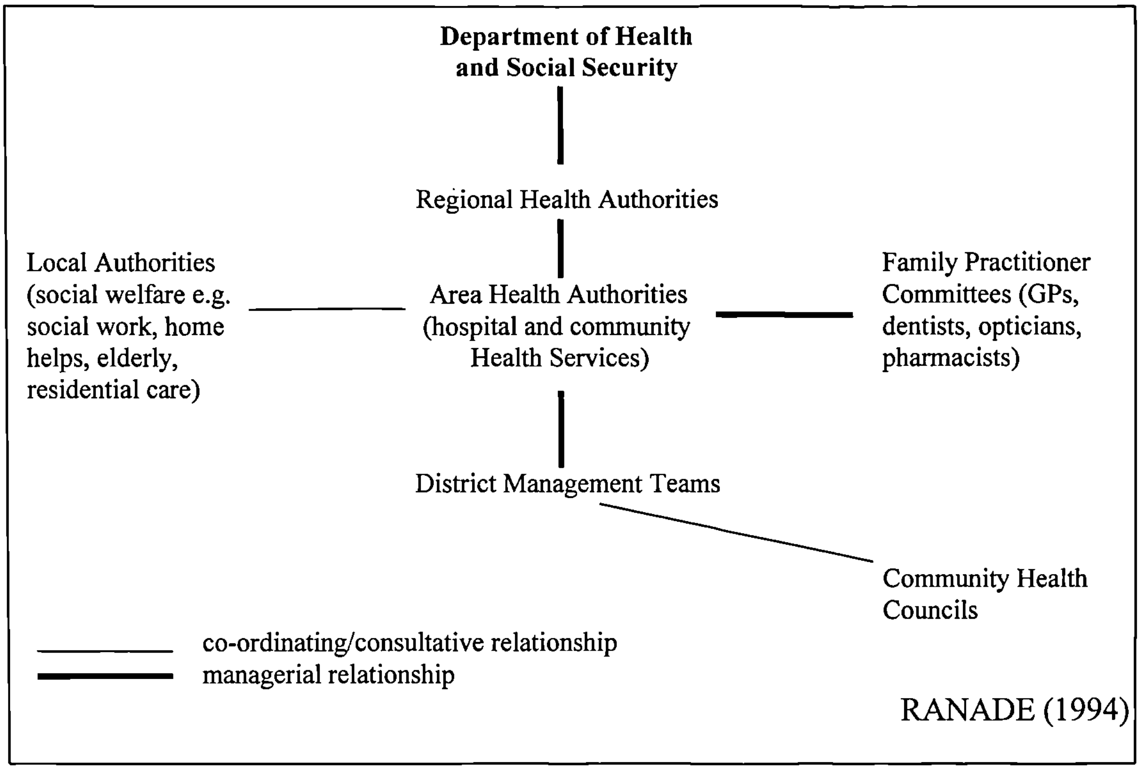


Figure 3.2: The structure of the NHS 1974-82 (England)

The Regional Health Authorities (RHAs) were intended to take over the roles of all the previous administrative elements of the Health Service beneath the Ministry of Health, in order to unify the service. The control of GPs passed from Executive Councils to the Family Practitioner Committees, with many of the same functions.

These changes affected the services in Wales, Northern Ireland and Scotland in a slightly different way. In Wales, the Welsh Office combined the role of Central Government and regional health authority, giving one less administrative tier.

In Northern Ireland, the system was also simplified in terms of the number of administrative levels.

In Scotland, all Health Services were controlled by the Secretary of State for Scotland, and the service was split into Health Boards, roughly equivalent to the Local Government regions (Scottish Home and Health Department, 1971).

These combined all the elements of the three levels in England (Regional Health Authorities, Area Health Authorities and District Management Teams) and also assumed managerial control for General Practitioner Services. The single-tier structure was created in Scotland with the intention of bringing more cohesion to the system and to reduce 'uncertainty about the division of responsibility' (Scottish Home and Health Department, 1971).

The links with the community would be maintained through the establishment of Local Health Councils, equivalent to the Community Health Councils in England. Additionally, the re-organisation in Scotland emphasised the need for a close working relationship with the local authorities, as 'the effectiveness of each authority's services will frequently depend on facilities provided by the other' (Scottish Home and Health Department, 1971).

A highly important factor of the re-organisation, according to Kelly and Glover (1996), was the exclusion for the first time of locally elected Government officials from a role in health care provision. The reforms reflected a more centralised approach to Health Service organisation in Britain, and, to a certain degree, the approach of the Conservative Party to government.

Power was transferred from the locally elected authorities to the Government-appointed Health Authorities or Health Boards. For example, the Secretary of State for Scotland appointed the members of each of the Scottish Health Boards.

The 1974 re-organisation was, according to many commentators, a complete disaster, failing to meet any of its objectives, operate successfully or improve the Health Service at any level.

Ranade (1994) described the re-organisation as 'seriously flawed' and the result of a series of political compromises and adjustments. Kelly and Glover (1996) noted that the three-tier system of organisation in England was 'dysfunctional' and 'one-tier too many'. The approach was far too bureaucratic, with many more people being employed who had no involvement in patient care. The number of staff involved in administrative tasks increased by about sixteen thousand.

In 1976, Merrison headed a Royal Commission to examine the problems that the re-organisation had brought, both for patients and staff. The findings of the committee were summarised in a review published in 1979:

> "There was a great deal of anger and frustration at what many regard as a seriously over-elaborate system of government, administration and decision-making. The multiplicity of levels, the over-elaboration of consultative machinery, the inability to get decision-making completed nearer the point of delivery of services and what some describe as unacceptably wasteful use of manpower resources were recurrent themes in most areas." MERRISON (1979)

The Health Service at this time was again facing financial problems. At the same time as the re-organisation came into effect in 1974, the Public Accounts Committee reported that the NHS was seriously under-funded:

> "It is the opinion of our committee that no Government has ever provided sufficient money to allow the Health Service to function and to react to growing needs effectively. As a result of the inadequacy of finance, the service is grinding to a halt."
>
> HOUSE OF COMMONS (1974)

This view was echoed by the BMA in the Royal Commission report in 1979:

> ".. for some years now the money allocated by the Government for the (health) service has been quite inadequate to meet the demands made upon it." MERRISON (1979)

An equivalent situation was observed in Scotland, where Health Boards in the period from 1977-80 were advised that the growth rate for expenditure would be reduced to just 1.5% per annum, with 'major implications for the maintenance and development of services' (Scottish Home and Health Department, 1976). Demand for health care and NHS facilities had continued to increase and was exceeding the levels of service available.

The period during the 1970s heralded the search for new measures of performance, as the efficiency with which resources were used and the effectiveness of health care policy became important areas of concern. The intention was to improve the financial control of Central Government over the NHS. In 1979, the first work on the Performance Indicator packages began at the University of Birmingham.

The major problem, though, was the type of data available, with most of it focusing on activity rather than outputs and quality. The influence of manpower costs on the performance and efficiency of the Health Services was noted in this period. Staffing costs were increasing at a faster rate than the number of patients treated and the manpower costs as a percentage of the total NHS budget reached 70%, reported by Allsop (1984). Thus, the relationship between staffing levels and quality of service became an important discussion point at this juncture.

3.4 A New Direction for the NHS (1982-90)

The NHS again came under review after the Conservatives regained power in the 1979 General Election. Ranade (1994) summarised the principles that underlined the new changes as 'small is beautiful' and 'making the decisions should be done as close to the patient as possible'. Levitt and Wall (1992) summed up the changes as, the abolition of the area tier, the simplification of the machinery for consultation and the introduction of unit management.

The reforms made were in accordance with the proposals published in 'Patients First' in 1979 (Department of Health, 1979). The major element of the changes in England was the removal of the Area Health Authorities. The new District Health Authorities (DHAs) combined the functions of the former areas and districts. Ham (1992) summarised the new emphasis of the DHAs: delegating power to units of management, ranging from single hospitals to the district-wide provision of particular services. There would be little change to the position of Community Health Councils and Family Practitioner Committees.

The new structure is illustrated in Figure 3.3 below:

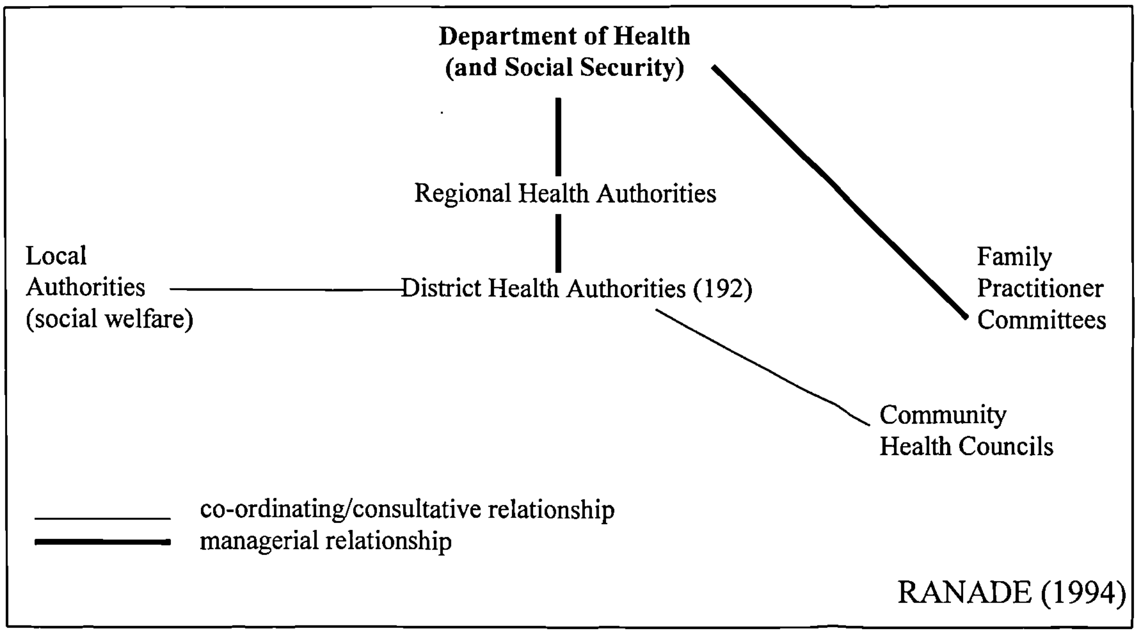


Figure 3.3: Structure of the NHS 1982 - 90 (England)

As with the other attempts at re-organisation, flaws in the changed system were identified almost immediately, and further developments were discussed. This led to vast changes in the NHS, as the 1980s was a period of almost continuous reform, labelled as 'an efficiency strategy for the NHS' by Ranade (1994).

The rationale behind many of the changes, as recognised by Pollitt (1990), was the assumption that better management could solve a range of social and economic problems, beginning with the provision of Health Services.

In 1983, Roy Griffiths, the then Chairman of Sainsburys, led a review of the NHS, focusing specifically on issues of management. The review identified serious flaws within the management of the Health Service, such that it was difficult to define who were the people in charge.

Griffiths determined that the NHS was suffering from 'institutionalised stagnation', in as much as no one was responsible for leading and instigating change and there was uncertainty as to what the objectives of the NHS were or whether they were being achieved (Allsop, 1995). The Griffiths report proposed a new management strategy, with clear and effective chains of command. Griffiths' idea was to incorporate aspects of business management, in order to change the organisational culture of the NHS. This was a hallmark of the approach to government by the Conservative Party during this period.

The main change would be to introduce general managers to replace the management teams at regional, district and unit levels, providing a driving force for developing management plans, increasing productivity and initiating new approaches. Changes would also be made to the Department of Health by dividing it into a Health Services Supervisory Board, to set policy, and a NHS Management Board, to oversee policy implementation.

The change to the management strategy, instigated by the Griffiths Review, has been interpreted as the most influential in the series of changes to affect the NHS during the last fifteen years. They were accepted by the Department of Health and introduced in 1984 despite much opposition, particularly from the medical profession:

> "The reaction to the report has not been very enthusiastic. It has been interpreted as an attack on NHS staff, as a threat to clinical freedom, a blow to nurse management."
>
> HOUSE OF COMMONS (1984)

Other important changes began in 1983, when competitive tendering for services such as cleaning and catering were introduced. In 1984, limits were made to the number of drugs covered by prescription charges. In 1985, annual reviews of the performance of Regional Health Authorities were instigated. By 1987, changes were introduced to improve the quality of information available for decision-makers and wide-sweeping reforms were proposed to GP services (OECD, 1992).

A key task for the new general managers, some of whom were appointed from outside the NHS and had management experience in the private sector, was to promote greater cost-efficiency and a better use of resources within the NHS (Baggott, 1994). This led to the imposition of Cost Improvement Programmes (CIPs), defined by the DHSS as:

> "Measures which are aimed at releasing cash or manpower used in providing a service by getting the same service output for a smaller input of resource; or improving productivity by getting a higher output for the same input (or for a less than proportionate increase in input)" NATIONAL AUDIT OFFICE (1986)

Over the three-year period from 1985 to 1988, the cost improvement programmes resulted in 'savings of around £150million per year... through the rationalisation of patient services, sub-contracting, supply cost savings and energy savings' (Robinson, 1988). Greater emphasis was also placed on the closer monitoring of the performance of the NHS. This led to the introduction of the Performance Indicator Package in 1983/4, reflecting the Conservative Government's move towards accountability, control, efficiency and economy.

Every health authority received a package of 147 indicators, to be used to compare their performance against others on a regional and national basis in five broad areas. They were intended to highlight areas for a more detailed investigation. Further additions to the package were made in 1985 and 1987/8. The development and impact of Performance Indicators is examined more closely in chapter four.

By the end of the 1980s, however, the public had little confidence in the NHS and was 'disenchanted that it was not getting what it wanted' (Rivett, 1997). The health care professions had, in general, not welcomed the string of reforms, concerned that 'the drive for greater efficiency has actually involved cuts in services and a deterioration in the quality of care' (Robinson, 1988). This perception of efficiency is consistent with the discussion in section 2.5 in the preceding chapter. Following the 1987 General Election, a further review of the NHS was undertaken, resulting in the most extensive changes, which were introduced in 1991.

3.5 Changing the Fundamental Structure of The NHS (1991 - 1997)

The changes introduced in 1991, following the publication of the White Paper 'Working for Patients' (Department of Health, 1989), were the most fundamental changes to the organisational structure of the NHS since it was founded in 1948. In the late 1980s, there were mounting financial pressures on the NHS, with wards closing and operations cancelled. There were also large variations in terms of resource usage, waiting times and productivity. The question of containing the cost of the NHS was also important to the Conservative Government.

The reforms were intended to deal with these issues, and negate the main weaknesses of the NHS, summarised in four points by Ranade (1994):

1. A poor matching of funding to workload - no account was taken in resource allocation of cross-district patient flows.
2. Inappropriate incentives for managers and clinicians - there was a wide variety in the range and quality of services offered affecting outcomes in an unknown manner.
3. Lack of responsiveness to consumers - the NHS needed to become more aware of the needs of the population.
4. Few incentives to innovate - the strong controls and accountability led to caution.

The objectives of the reforms, stated in the Working for Patients White Paper, were:

> ".. to give patients, wherever they live in the UK, better health care and greater choice of the services available; and greater satisfaction and rewards for those working in the NHS who successfully respond to local needs and preferences."
>
> DEPARTMENT OF HEALTH (1989)

The reforms brought about a variety of changes in every area of service provision, the most striking of which was the split of the Health Service into purchasers and providers, introducing the concept of the internal market. Health care would still be financed by general taxation, but there would changes in the way hospitals and GP practices received their funding, through the introduction of a contracting process. Under the new system, hospitals could apply to become self -governing trusts, making them independent of health authority control and responsible for raising their own budgets from service contracts.

GPs could take responsibility for their own budgets by taking fundholding status and purchase services from the hospitals by arranging contracts. The idea of the internal market approach to health care provision was suggested by Enthoven (1985), an American economist, as a method of improving the Health Service:

> ".. by separating the purchase of health care from its provision and management, and subjecting providers to an element of competition for contracts, providers would now have a financial incentive to cut costs, improve quality and be more responsive to what customers wanted. Purchasers in turn, since they would still be cash-limited, would have an incentive to bargain for improved value-for-money."
>
> RANADE (1994)

Enthoven also believed that, although the changes to the structure of the Health Service would be fundamental, the creation of an internal market would have little effect on the users of the service, the patients (Enthoven, 1985).

Prior to the introduction of the internal market, concerns had been raised as to whether this approach was the best way forward for the Health Service. For example, Burke and Goddard (1990) suggested that an internal market would actually lead to services being delivered less efficiently. Proponents of the internal market, including Enthoven himself, also expressed some reservations, suggesting that it should be introduced gradually and tested 'through demonstration or pilot project' (Robinson, 1988). Robinson (1988) suggested that there was 'insufficient evidence to warrant adoption of the idea throughout the service in a single move but sufficient evidence to support the case for experiment.'

Despite these reservations, however, the internal market was established in 1991 and its central feature, the purchaser/provider relationship, is illustrated in figure 3.4 below:

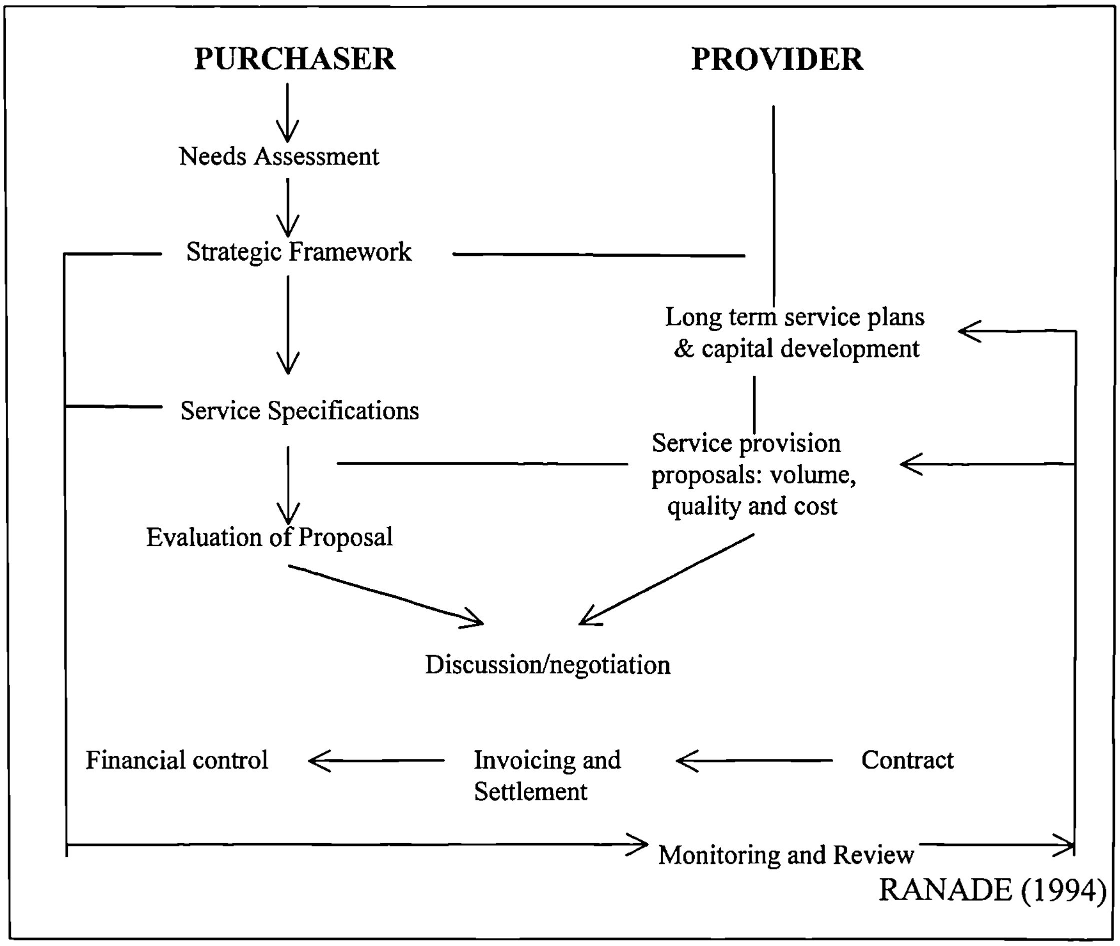


Figure 3.4: The Contracting Process

The changes were intended to extend the reforms of the 1980s by making the NHS more accountable and efficient, increasing its productivity through the introduction of competition. They were initially welcomed with scepticism and mistrust from the health professions and the general public. Detractors of the reforms accused the Conservative Government of preparing the NHS for privatisation, which the Government rigorously denied. The changed structure is illustrated in figure 3.5 overleaf:

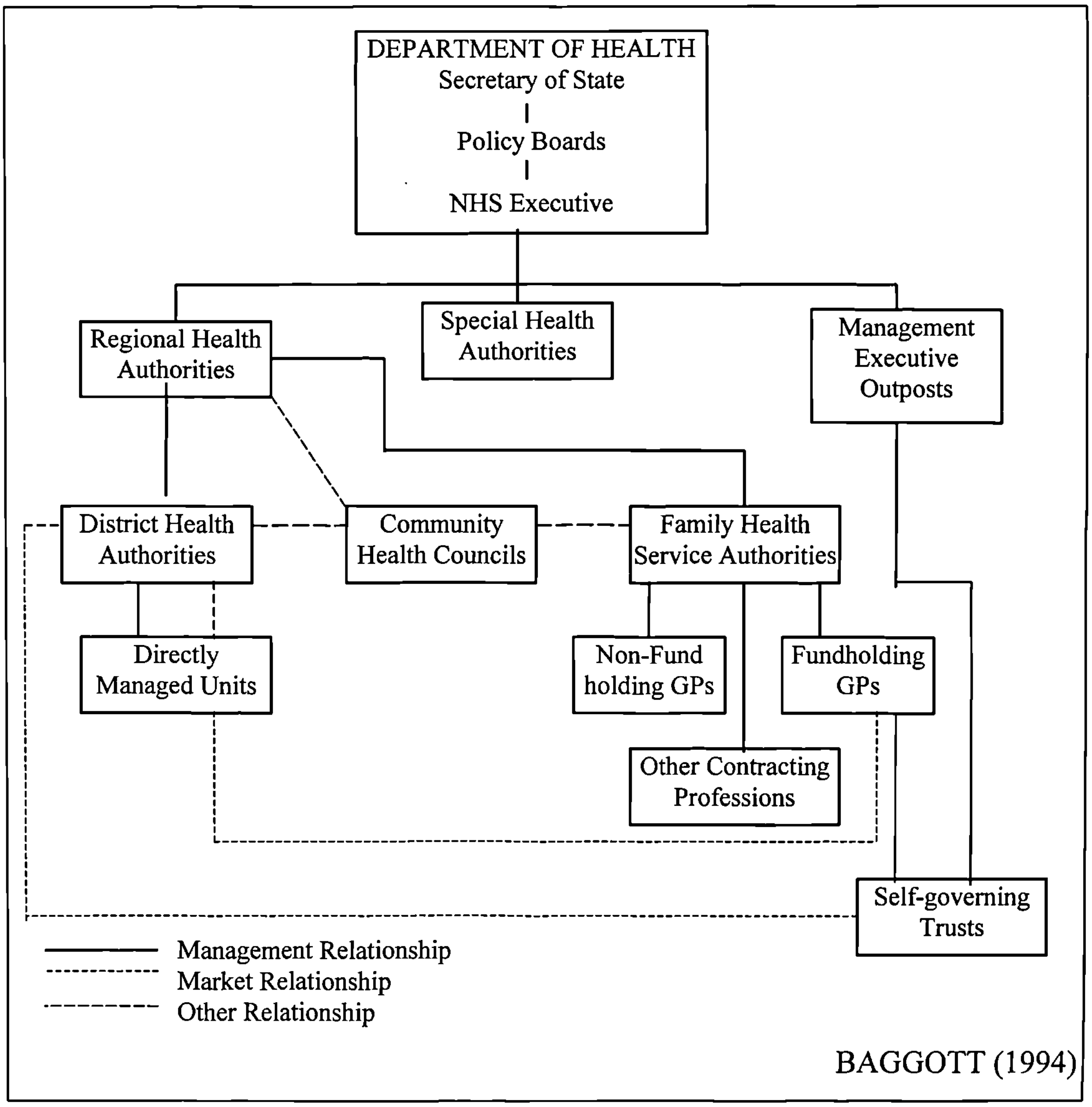


Figure 3.5: Structure of the NHS (England) 1991

The ideas of the internal market were also adopted in Scotland, Wales and Northern Ireland in much the same way as in England, with the central role still played by their respective Government departments, although the pace of implementation tended to be much slower (Baggott, 1994). In Wales, there was also no regional tier of health authorities, whilst the Welsh Health Technical Services Organisation provided the district health authorities with equivalent services to those provided by the RHAs in England.

In Northern Ireland, Health Services were to be administered through Health and Social Services Boards, also responsible for social services. In Scotland, the Common Services Agency continued to carry out an equivalent function to that of the Regional Health Authorities in England, whilst the Health Boards maintained their managerial responsibility over family practitioner services (Baggott, 1994). A more detailed discussion on the structure of the NHS in Scotland is given in chapter seven, prior to the detailed analysis of data relating to acute hospitals in Scotland.

The measurement of performance and efficiency also continued to develop during this period. The 1991 reforms introduced, as a part of the performance assessment process, the technique of medical (or clinical) audit. Medical audit was expected to become part of the regular activities of every consultant, in an attempt to improve the overall quality of care that every patient receives. The role played by medical audit in improving the performance of the NHS will be further debated in chapter four.

The Performance Indicator package was also extended, with seven topic areas and over 2000 indicators relating to the purchasers and the providers of Health Services included within it.

In addition, some of the indicators used to assess performance were also published in the form of league tables, referring to both hospitals and health authorities/boards. The tables listed indicators such as the total number of cancelled operations per year and the percentage of patients treated within certain time frames.

During the period of the last five years of the Conservative Government, there were no more major changes to the managerial and organisational structure of the NHS. However, there were further minor changes, which included changes to community health care provision and the improvement of primary health care facilities.

In relation to structural reform, the additional changes were related to the position of the NHS Management Executive (NHSME), with its headquarters in Leeds. This was absorbed into the Department of Health, taking responsibility for managing the purchasing of Health Services. Eight regional offices within the NHSME replaced the Regional Health Authorities. District Health Authorities and Family Health Authorities were integrated and took over the primary responsibility for the purchasing of Health Services, through GPs. Hospital services, in the majority of cases, were provided through the self-governing NHS Trusts. The organisational structure at the end of this period is illustrated in Figure 3.6 below:

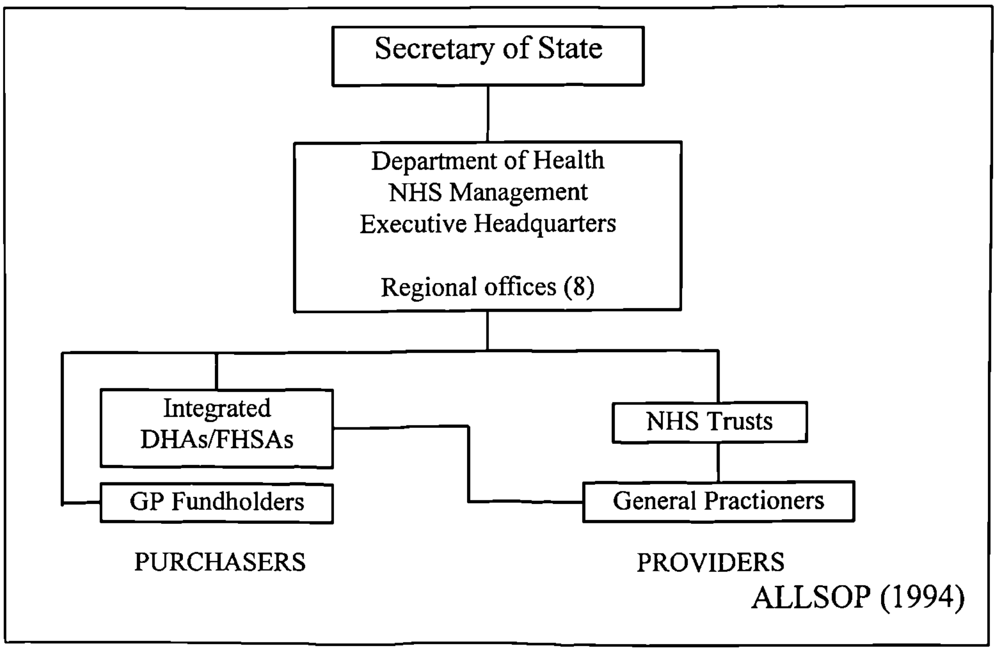


Figure 3.6: Structure of the NHS (1994)

In relation to the level of services provided during this period, statistics published just prior to the General Election in 1997, showed increases in the number of patients being treated, claimed as evidence of continuing improvements in the NHS:

> "The figures.. are further proof that the NHS is in excellent working order. They highlight significant increases in the total number of in-patient cases treated in 1995-96 - up by 6.5 per cent to 11.2 million - and particularly in day cases, which rose by 15 per cent to 2.8 million." DEPARTMENT OF HEALTH (1996a)

There were more NHS Trusts achieving five-star rating in the NHS performance League Tables for 1995-96 than ever before (Department of Health, 1996b). There were many more consultants being employed and a greater percentage of staff employed in direct patient care (Department of Health, 1996c, 1996d).

The Conservative Government also claimed significant reductions in waiting lists, as a demonstration of the effectiveness of their reforms:

> "Over the last five years there have been dramatic improvements in NHS waiting times. The number of patients waiting more than two years for hospital admission was 81,000 in March 1990. Now there are none, and 18-month waits have been virtually eliminated."
> DEPARTMENT OF HEALTH (1996e)

In 1995, the Conservative Government also attempted to tackle the issue of increasing management costs within the NHS. Trusts, Health Authorities, GPs and the Department of Health itself were required to make cuts of 5% in their management budgets, in order to release extra cash to be used for patient care.

Department of Health figures claimed that the savings would have released £300million by 1997 (Department of Health, 1996f). These approaches were similar to the Cost Improvement Programmes established during the 1980s.

However, the picture of the NHS as an improving service was hotly disputed at this time, especially in the political debate leading up to the General Election in 1997. There were fears raised within the medical professions that the quality of patient care was seriously affected by the drives for efficiency (The Guardian, 28/05/96). The National Association of Health Authorities and Trusts echoed this view:

> "There are very wide gaps this year on contracts. We have been producing efficiency savings year on year and there must come a point at which we have cut out all the excess padding and we end up cutting services." THE GUARDIAN (07/05/96)

There were also frequent reports of cancelled operations, ward closures and staff shortages. The following are examples of some of the problems highlighted in Scotland during 1995, taken from newspaper reports:

> ***"No bed for intensive care patient":*** a shortage of intensive care beds and staff in a Glasgow hospital were found to be highly significant factors in the death of a patient.
>
> THE HERALD (18/01/95)
>
> ***"Lack of doctors closes ward":*** Temporary ward closures due to a lack of junior doctors has led to children in a Glasgow hospital waiting longer for operations. THE HERALD (18/05/95)
>
> ***"Cancer ward crisis over bed shortage":*** Not enough beds to cope with seriously ill cancer and kidney patients at a hospital in Dundee.
>
> THE HERALD (01/09/95)

Complaints about the NHS more than doubled during the five-year period either side of the introduction of the reforms in 1991, according to annual report of the Health Service Ombudsman, with approximately 60% of complaints found to be justified. This report also found that health workers were more likely to become 'case-hardened' and forget the human element of their job, due to the 'remorseless drive to hit efficiency and performance targets' (The Guardian, 14/7/95).

The data analysis discussed in chapter seven and beyond relates to hospital activity in this time period. However, in order to complete the story, the current status of the NHS is now examined, with a discussion on the reforms introduced by the Labour Government following their victory in the General Election in 1997.

3.6 The 'NEW NHS' under NEW LABOUR (1997 and beyond)

The NHS has entered yet another period of transition, following the Labour victory at the General Election in May 1997. As it celebrated its 50th anniversary in 1998, the direction of those changes has been settled for the foreseeable future. The new focus of the NHS, reflected by the changing emphasis on the concept of efficiency discussed in chapter two, can be clearly observed in the titles of the most recent documents produced relating to its future. These included 'Designed to Care' (Scottish Office, 1997) and 'A First Class Service - Quality in the NHS', (Department of Health, 1998a). The most significant document, the Government white paper on NHS reforms published in 1997, describes the New NHS as 'modern and dependable' (Department of Health, 1997a).

The NHS, under the development and restructuring presented in the previously mentioned white paper, is to be based upon a 'third way', a popular phrase with New Labour. In relation to the NHS, the 'third way' means 'building on what has worked, but discarding what has failed' and the development of a system based on 'partnership and driven by performance, moving away from outright competition towards a more collaborative approach' (Department of Health, 1997a). In terms of organisational structure, the two approaches comprehensively rejected for the future of the NHS were:

> - The centralised approach to health care in the 1970s, where innovation was stifled and the needs of the institutions were considered ahead of patient care and
> - The divisive internal market system of the 1990s, intended to bring about efficiency but actually producing fragmentation in the decision-making process and distortion in the incentives it offered to such an extent that it was defined by unfairness and bureaucracy. DEPARTMENT OF HEALTH (1997a)

One of the first acts of the new Government in relation to health care, therefore, was the dismantling of the internal market, introduced by the Conservative Government in 1991 and described in detail in section 3.5.

The rationale behind the removal of the internal market was to move funding away from red tape and into direct patient care, as well as ending the inequalities the internal market was thought to produce, as stated by Frank Dobson, the Secretary of State for Health at this time:

> "There are examples galore where the so-called 'internal market', in which GP fundholders and health authorities buy services from NHS trusts, has led to a two-tier Health Service.... We remain committed to ending the internal market in health care, which has placed so many patients at a disadvantage. We must also undo the damage the market has created - the never-ending paperchase of invoices..." DEPARTMENT OF HEALTH (1997b)

As the notion of the third way implied, some elements of the NHS would remain unchanged by the reforms. In fact, the removal of the internal market has been the one of the only major organisational changes to date. Additionally, the number of trusts has been reduced.

However, in the long-term, further changes have been proposed, which include the development of primary care groups, intended to extend the opportunities offered by fundholding, and a more integrated approach to health care, involving integration across a range of Government agencies (Baker, 1998). The key aspects and strategies to be adopted in the 'New NHS' to encourage long term stability and forward planning are:

1. Maintaining the separation between the planning of hospital care and its provision - by empowering local doctors, nurses and health authorities to plan services, the NHS will be built around the needs of patients.
2. Building on the increasingly important role played by primary health care strategies - the role of the family doctor will be strengthened as they will be able to continue to influence the use of resources to improve patient care and community services will be emphasised.

3. Recognising the intrinsic strength of decentralising responsibility for operational management - NHS trusts will continue to have control over the key decisions relating to local services and patient care, in conjunction with the other stakeholders, leading to interdependence and not independence.

DEPARTMENT OF HEALTH (1997a)

Therefore, whilst the structure of the NHS is unlikely to be radically altered in the immediate future, the principles upon which the Government has proposed it should be run are very different and can be summarised in six key points, as follow:

- The renewal of the NHS as a truly national service - patients will get fair access to consistently high quality, prompt and accessible services across the whole country;
- The delivery of health will be a matter of local responsibility, guided by national standards but driven by the needs of the local community;
- Organisational barriers will be broken down, with strong links established with Local Authorities, so that the NHS will work in partnership, with the needs of the individual patients at the centre of the care process;
- Every pound in the NHS should be spent to maximise patient care, by cutting bureaucracy and encouraging efficiency by developing a more rigorous approach to performance;
- Excellence should be guaranteed by focusing on quality of care, with quality as the driving force in all decision-making;
- Public Confidence in the NHS will be rebuilt, as it seen as a public service, accountable to patients, open to the public and shaped by their views. DEPARTMENT OF HEALTH (1997a)

In chapter two, the revised approach to efficiency assessment was introduced, stressing the importance of quality and many other factors under the umbrella of performance assessment. In the light of this new ethos, five approaches have been proposed for ensuring efficiency:

- Aligning clinical and financial responsibility and developing responsibility for a single unified budget to primary care groups;
- Management costs will be capped in health authorities and primary care groups and reduced in trusts;
- The Government will publish reference costs for individual treatments and will require trusts to make known, and to benchmark, their own costs;
- There will be cash incentives to improve performance and efficiency for health authorities, trusts and primary care groups;
- Sanctions can be imposed on poor performers including withdrawal of freedoms and the right to move services between providers is retained. Direct intervention by the NHS Executive is also possible. BAKER (1998)

The knock-on effect of the changed approach to performance assessment is in the methods employed for measurement, although this is still in the very early stages of development. It has been suggested that, in the past, attention was directed towards the 'measured elements rather than the important things, and the wrong measures produced the wrong results' (Baker, 1998). At the present time, performance is primarily still assessed using the same approaches as previously, such as audit and performance indicators, with attention given to waiting lists and other similar measures, to be discussed now.

3.6.1 Performance Assessment for the 'New NHS'

In section 3.5, the performance of the NHS at the end of the period of Conservative administration, in relation to the size of the waiting lists, levels of patient activity and the meeting of performance targets, was detailed. Whilst figures showing increased activity and reduced waiting times were presented, these were tempered by claims of staff shortages, a lack of facilities and an increased number of complaints. For example, claims of reduced waiting lists were countered by accusations of 'fiddling the figures' (The Daily Telegraph, 24/11/95). In other words, the performance of the NHS could be described during this period as a 'mixed bag'.

In the light of the changes in structure and emphasis discussed above following the election of the Labour Government, the performance of the NHS can be reassessed. Frank Dobson (the Secretary of State for Health until October 1999) described provisional waiting list figures for the three-month period up till December 1997 as 'bad'. He felt that they reflected 'the scale of the challenge' faced by the Labour Government and the rise in waiting lists was caused by the priority to 'avoid a winter crisis' (Department of Health, 1998b).

Looking at the latest available figures (March 1999), the situation in England was that no one had waited longer than eighteen months for inpatient treatment, a repeat of the situation from the previous year when this was hailed as a 'massive achievement' (Department of Health, 1998c, 1999). The number of patients waiting over twelve months fell to its lowest figure since June 1997 and was recorded at 48,800, a reduction of 25,000 over a nine-month period (Department of Health, 1999).

The number of patients waiting for admittance to NHS hospitals decreased 'by a record 47,000 in March, taking the total down to 1,073,000 – a fall of over a quarter of a million since April 1998 and the lowest figure since September 1996' (Department of Health, 1999). The Secretary of State for Health, Frank Dobson, claimed that the decrease was 'an absolutely phenomenal achievement by NHS staff', brought about through 'a huge amount of work' (Department of Health, 1999).

The most recent figures available for Scotland (December, 1998) showed a very similar situation, with 78, 526 patients on the waiting lists, a fall of 6,405 in the last quarter, which was the 'largest quarterly fall for a decade' (Scottish Office, 1999).

In relation to some of the recognised performance targets, figures from 1998 showed that there had been 'a 12% drop in cancelled operations' in England (Department of Health, 1998d). Additionally, there was 'an 18% fall in the number of breaches of the patients' charter, where hospitals fail to readmit these patients within one month' (Department of Health, 1998d).

According to all the above figures, therefore, the situation in relation to waiting lists has improved over the last year, although these figures do not take into account the patients waiting for their first outpatient appointment with a consultant, referred to as the 'hidden waiting list' (Scottish Office, 1999). New performance initiatives in the whole of the UK will be centred on this area in the future, according to Sam Galbraith, the Scottish Health Minister:

> "Health Boards and NHS Trusts are now working on the implementation of over 190 strategic projects, funded from waiting list resources, which will tackle outpatient waiting times by; improving the interface between primary and secondary care; establishing more 'one-stop' and nurse-led clinics; and redesigning services around the needs of the patient. In addition, outpatient waiting times will benefit from the 'silent revolution' in information technology which is taking a fresh look at how we manage patients into, during and out of hospital." THE SCOTTISH OFFICE (1999)

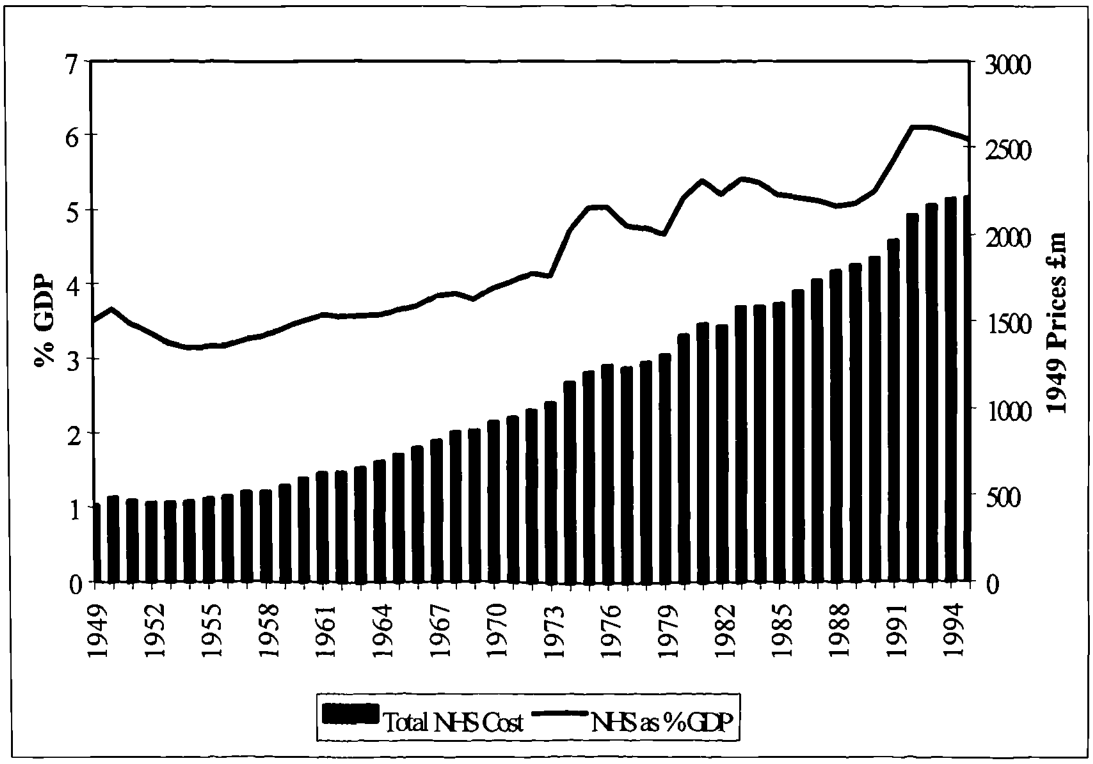
3.7 The NHS - Issues of Finance and Economics

Throughout its history, the question of the level of finance that the NHS should receive from general taxation has been a difficult one to answer. It has been recognised that the NHS is capable of consuming all the resources that are made available:

> "... the demand for health care is always likely to outstrip supply and ... the capacity of Health Services to absorb resources is almost unlimited." MERRISON (1979)

However, as about 80% of the financial input into the NHS comes from general taxation (Baggott, 1994), it has been a necessary part of the funding process for certain limits to be made to the level of financial input received by the NHS. The difficulty has come in deciding how much of the total Government spending should be spent on the NHS and how it should be allocated.

Figure 3.7 illustrates the changing levels of NHS spending since 1949, showing NHS expenditure both in monetary terms at 1949 prices and as a percentage of Gross Domestic Product (GDP). The trend for both is a steady increase. For certain periods, notably 1950 - 1955 and 1975 - 1980, NHS expenditure as a proportion of GDP dipped, mirrored by falls in actual expenditure as illustrated.

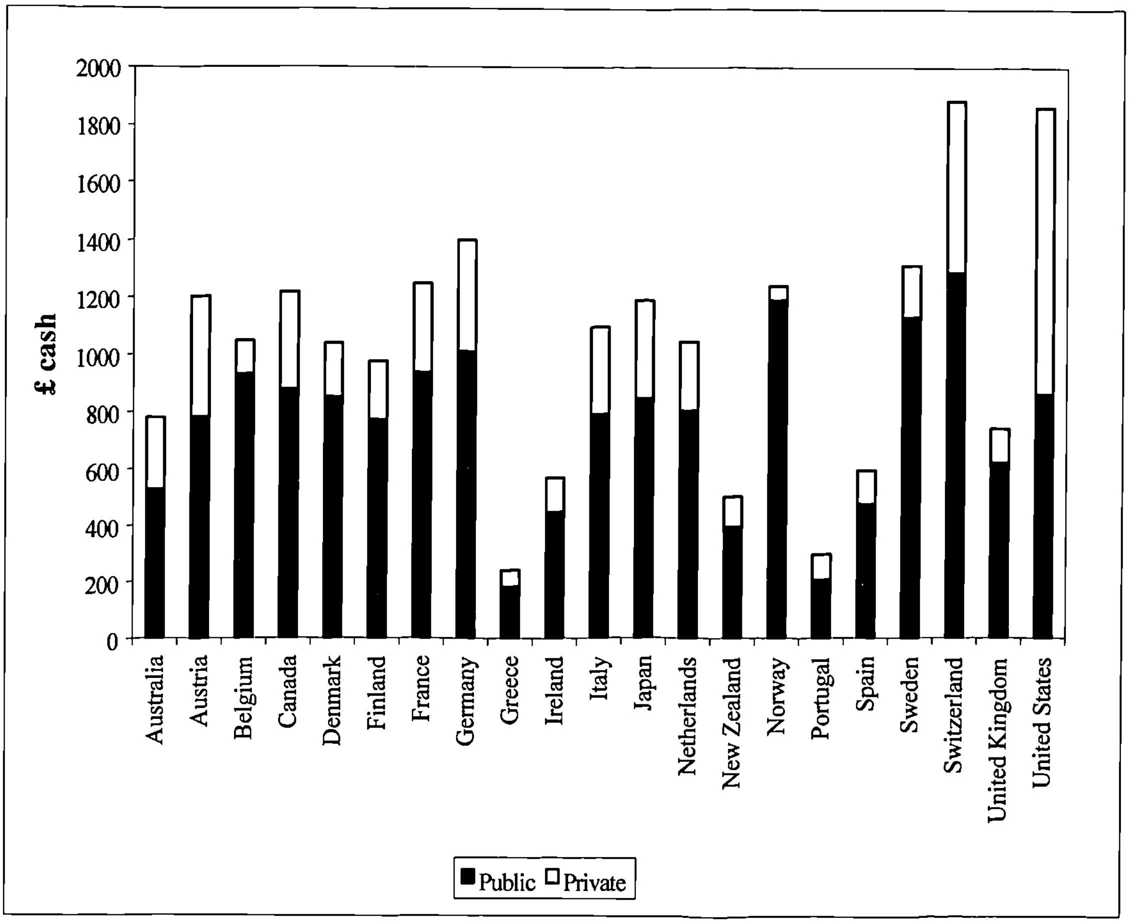


OHE (1995)

Figure 3.7: UK NHS Expenditure 1949 – 1995

Figure 3.8 overleaf shows the position of health expenditure in the United Kingdom, compared with spending in other OECD (Organisation for Economic Co-operation and Development) Countries, referring to the per capita spending. The UK level is lower than both the OECD average of £1,229 per person and the EU (European Union) average of £993 per person.

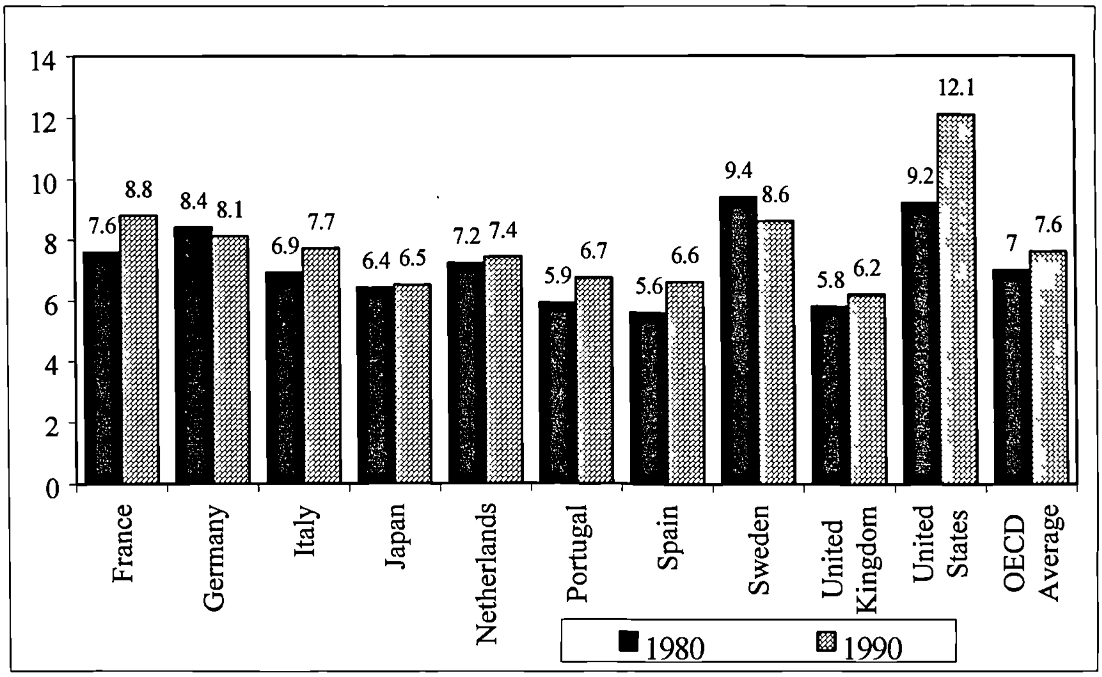
The chart also shows the expenditure divided into public and private contributions, with some variation clearly apparent. Focusing on public expenditure, the UK figure remains below the OECD and EEC averages (£614 compared with £750 and £754 respectively), although the OECD average private expenditure per person is much greater than the EEC figure (OHE, 1995).



OHE (1995)

Figure 3.8: Total Health Expenditure Per Person 1992

However, the variations in per capita expenditure could be related to the economic position of each of the OECD countries. Therefore, health care expenditure in a sample of the countries included in the figure above has been expressed as a percentage of GDP, as shown in figure 3.9 below, for 1980 and 1990.



DRUMMOND AND MAYNARD (1993)

Figure 3.9: Health Care Expenditure as a Percentage of GDP

As the figure above demonstrates, the figure for the UK was well below the OECD average, for both 1980 and 1990. In 1980, Spain was the only country of those included in the sample with a lower figure for health care expenditure as a percentage of GDP than the UK. By 1990, the UK level was the lowest of the ten countries, at just 6.2%, compared with the OECD average of 7.6%

Therefore, expenditure levels per capita and overall, as a percentage of GDP, on health care appear to have been relatively low for the UK, at least at the time frames included in the figures above. This suggests that the claims of under-funding discussed in section 3.2 may have been warranted. The financial picture for the NHS is brought up-to-date in the following section.

3.7.1 The Financial Position of the NHS in 1998

In July 1998, Gordon Brown, Chancellor of the Exchequer in the Labour Government, announced the results of the comprehensive spending review, a systematic review and reorganisation of the Government's financial strategy instigated following the Labour Party's victory at the last General Election. This resulted in a projected increase in health care expenditure of more than £20 billion over a three-year period, over and above all previously promised spending levels.

This was determined to be 'the biggest cash injection in the history of the NHS' (Department of Health, 1998e) and is to be used to deliver 'modernisation and reform' leading to 'demonstrable year on year improvements in all parts of the health and social care system' (Department of Health, 1998f).

The main areas for the spending of this cash boost can be summarised as follows:

1. The provision of more than 15,000 new nurses and 7,000 new doctors so that an extra three million patients can be treated, as well as extra funding for the training of new nurses and doctors in the coming years;
2. The establishment of the NHS Modernisation Fund, to distribute cash for projects leading to targeted improvements in services;
3. Funding the assurance that no new patient charges will be introduced in the lifetime of the current parliament;
4. Capital investment for the building and rebuilding of hospitals, clinics and GP premises;
5. The strengthening of links with social services to improve care for the elderly and the services for people with mental health problems;

6. The continuation of an integrated health strategy, tackling inequality in health and the causes of ill health, such as poverty, low pay, unemployment, poor housing, environmental pollution, crime and disorder.

DEPARTMENT OF HEALTH (1998e, 1998f, 1998g)

3.8 The Future of the NHS

It is difficult to predict the future of the NHS. The Chief Executive of Severn NHS Trust, Richard James, claimed that the NHS was on the verge of 'meltdown' in several areas and described its position as 'extremely grave' (The Guardian, 07/05/1996). Harriet Harmon, the shadow Health Secretary prior to the Labour Party's victory in the General Election in 1997, described the crisis in NHS funding as 'very severe and immediate' (The Guardian, 07/06/1996). However, the Department of Health reported that NHS spending was 'at an all-time high at £42.6billion in 1996/97' and the service was continually expanding and treating more patients (The Guardian, 25/06/1996). In the summer of 1998, Frank Dobson announced that the indications were that waiting lists would be reduced to levels below those in May 1997 before the end of the current parliament, whilst spending levels have dramatically increased (Department of Health 1998a, 1998e).

The above represent a confused picture of the NHS: is it either on the brink of collapse or facing a bright future? There is certainly the potential for change in all levels of the Health Service. Allsop (1995) suggests that primary health care will expand and the role of hospitals will diminish.

Proposals for re-introducing the idea of cottage hospitals and extending the range of services offered in General Practice correspond to this viewpoint. Schemes intended to improve the overall health of the population are also likely to be more prevalent in the future, such as the Health of the Nation project launched in 1992. This emphasises health promotion and prevention as well as treatment in the Health Service. 'Health of the Nation' targets exist for illness and disease in particular areas, such as heart disease and breast cancer, which, if achieved, may reduce some of the strain on the Health Service.

An integrated approach to health care is likely to be adopted, establishing stronger links between the Department of Health and other Government departments. This has been a policy favoured by the Labour Government, with the Health Services related to education, social services, employment and environmental policies and schemes. 'Quality of Care' and 'effective health care treatment' have replaced 'efficiency drives' and 'cost improvement programmes' as the 'buzz' words in the NHS.

Additionally, traditional approaches to treatment, requiring lengthy inpatient care, are gradually being replaced by day case procedures, where patients requiring some basic surgical procedures are admitted, treated and released within a single day. Several key procedures have already been identified within the NHS and targets have been set for the number of procedures health care providers must carry out as day cases (ISD, 1998). This approach, in conjunction with an increased role for primary care treatment programmes, may have a dramatic effect on hospital-based services in the long-term.

It is difficult to predict exactly what further changes will be made to the NHS in the future - the most fundamental factor shaping its development is the outcome of the General Election in May 1997, when the Labour Party formed the new Government. The changed emphasis and approach to health care has already been debated and this is likely to shape a new direction in the development of the NHS, the success of which will be determined by the passing of time. However, Ranade (1994) lists a series of questions that will need to be answered by successive Governments in the future, to guide the development of the Health Service:

> "How can Health Services be restructured to meet the needs of ageing populations more appropriately? What should the division be between public and private spending on health care? Can publicly funded systems continue to offer a comprehensive range of services? What are the ethics of the different ways of rationing health care? How relevant are medical definitions of health to current health care problems?" RANADE (1994)

3.9 References

1. **Allsop, J.** (1984) Health Policy and the National Health Service, Longman.

2. **Allsop, J.** (1995) Health Policy and the NHS: Towards 2000, Longman.

3. **Baggott, R.** (1994) Health and Health Care in Britain, MacMillan.

4. **Baker, M.** (1998) Making Sense of the New NHS White Paper, Radcliffe Medical Press Ltd.

5. **Burke, C and Goddard, A.** (1990) Internal Markets - the Road to Inefficiency, Public Administration, Vol. 68, No. 3, pp. 389-396.

6. **Department of Health** (1989) Working for Patients, HMSO Cm 555.

7. **Department of Health** (1996a) Increase in NHS Hospital Treatment Shows NHS is Working Better, says John Horam, Department of Health Press Release 96/326.

8. **Department of Health** (1996b) More Five Star Performers than Ever Before in 1996 NHS 'League Tables', Department of Health Press Release 96/218.

9. **Department of Health** (1996c) NHS Census Reveals Increase in Front Line Staff, Department of Health Press Release 96/279.

10. **Department of Health** (1996d) More Hospital Consultants, Better Patient Care - Malone, Department of Health Press Release 96/263.

11. **Department of Health** (1996e) New Citizens Charter for Commitments for the 21st Century, Department of Health Press Release 96/297.

12. **Department of Health** (1996f) NHS Trust Management Cost Cuts on Target = £103M Extra for Patient Care, Department of Health Press Release 96/269.

13. **Department of Health** (1997a) The New NHS, Modern, Dependable, HMSO.

14. **Department of Health** (1997b), Department of Health Press Release 97/091.

15. Department of Health (1998a) High Standards of Fast and Convenient Care, Department of Health Press Release 98/271.

16. Department of Health (1998b) Waiting List Figures Will Be Delivered Says Frank Dobson, Department of Health Press Release, 98/064.

17. **Department of Health** (1998c) Provisional Waiting List Figures: 31 March 1998, Department of Health Press Release, 98/199.

18. **Department of Health** (1998d) Minister Hails Drop in Cancelled Operations, Department of Health Press Release, 98/221.

19. **Department of Health** (1998e) £20 Billion Boost for the NHS - Biggest Cash Injection in the History of the National Health Service, Department of Health Press Release.

20. **Department of Health** (1998f) Biggest Programme of Renewal and Modernisation Since NHS was Founded - Dobson, Department of Health Press Release, 98/294.

21. **Department of Health** (1998g) Government Committed to Public Health Crusade Against Inequalities -Dobson, Department of Health Press Release, 98/298.

22. **Department of Health** (1999) Dobson Hails Biggest Waiting List Fall Yet as NHS Delivers Even More than Promised, Department of Health Press Release, 1999/0278.

23. **Department of Health and Social Security** (1972) National Health Service Reorganisation: England, HMSO Cmnd 5055.

24. **Department of Health and Social Security** and **Welsh Office** (1979) Patients First: Consultative Paper on the Structure and Management of the National Health Service in England and Wales, HMSO.

25. **Drummond, M. F. and Maynard, A. Eds.** (1993) Purchasing and Providing Cost-effective Health Care, Churchill Livingstone.

26. **Eckstein, H.** (1958) The English Health Service, Harvard University Press.

27. **Enthoven, A.** (1985) Reflections on the Management of the NHS, Nuffield Provincial Hospitals Trust.

28. **Ham, C.** (1992) Health Policy in Britain, MacMillan.

29. **House of Commons** (1984) (HC209) 1st Report 1983/4. Griffiths NHS Management Inquiry Report, Social Services Committee, HMSO.

30. **House of Commons Expenditure Committee (Social Services Sub-Committee), Fourth Report Session** (1974) Expenditure Cuts in Health and Personal Social Services, HMSO, HC 307.

31. **Information and Statistics Division** (1998) Scottish Health Statistics 1998, ISD, NHS in Scotland (Scottish Health on the Web Internet Document).

32. **Kelly, M. and Glover, I.** (1996) In Search of Health and Efficiency: the NHS 1948 - 1994 in Beyond Reason: the National Health Service and the Limits of Management, Leopold, J., Glover, I. and Hughes, M. (Eds.), Avebury.

33. **Klein, R.** (1989) The Politics of the NHS, Longman.

34. **Levitt, R. and Wall, A.** (1992) The Reorganised National Health Service, Chapman and Hall.

35. **Merrison, A. (Chairman)** (1979) Royal Commission on the National Health Service: Report, HMSO Cmnd 7615.

36. **Ministry of Health** (1944) A National Health Service, HMSO Cmnd 6502.

37. **Ministry of Health** (1956) Report of the Committee of Inquiry into the Cost of the National Health Service (The Guillebaud Report), HMSO Cmd 9663.

38. **Ministry of Health** (1962) A Hospital Plan for England and Wales, HMSO Cmnd 1604.

39. **National Audit Office** (1986) The Financial Management Initiative, HMSO HC588.

40. **OECD** (1992) The Reform of Health Care: A Comparative Analysis of Seven OECD Countries, OECD.

41. **OHE** (1995) Compendium of Health Statistics, Office of Health Economics.

42. **Pollitt, C.** (1990) Managerialism in the Public Services: the Anglo-American Experience, Blackwell, Oxford.

43. **Ranade, W.** (1994) A Future for the NHS? Health Care in the 1990s, Longman.

44. **Rivett, G.** (1997) From Cradle to Grave, Fifty Years of the NHS, King's Fund Publishing.

45. **Robinson, R.** (1988) Efficiency and the NHS: A Case for Internal Markets, The IEA Health Unit.

46. **The Daily Telegraph** (24/11/95) NHS Waiting List Figures 'Might Have Been Fiddled' (Electronic Telegraph).

47. **The Guardian** (07/05/96, page 1) Warning of 'Meltdown' in NHS Market as Trusts Resist Squeeze.

48. **The Guardian** (07/06/96, page 12) NHS Facing 'Winter of Tough Tests'.

49. **The Guardian** (14/7/95, page 10) Efficiency 'can be Bad for Patients'.

50. **The Guardian** (25/06/96, page 6) NHS 'Needs an Extra £6bn'.

51. **The Guardian** (28/05/96, page 10) Healthcare: NHS and Private.

52. **The Herald** (01/09/95, page 8) Cancer Ward Crisis over Bed Shortage.

53. **The Herald** (18/01/95, page 1) No Bed for Intensive Care Patient.

54. **The Herald** (18/05/95, page 9) Lack of Doctors Closes Ward.

55. **The Scottish Home and Health Department** (1971) Reorganisation of the Scottish Health Service, HMSO.

56. **The Scottish Home and Health Department** (1976) The Health Service in Scotland: The Way Ahead, HMSO.

57. **The Scottish Office** (1997) Designed to Care - Renewing the National Health Services in Scotland, HMSO.

58. **The Scottish Office** (1999) Massive Fall in Scots Hospital Waiting Lists, Scottish Office Press Release, 0326/99.

CHAPTER FOUR

THE MEASUREMENT OF EFFICIENCY AND PERFORMANCE IN THE NATIONAL HEALTH SERVICE

4.1 Introduction

The search for adequate and appropriate methods of measuring performance within the NHS has been ongoing since its foundation in 1948:

> "The attempts to find satisfactory measurements of yardsticks of performance have been persistently baffled. Enormous effort has been lavished during the twenty years of the National Health Service on the collection of statistics of hospital activity, and on the search among them for means of making valid comparisons, within the service itself and between the service and the other systems. ... The most carefully constructed parallels between one hospital and another dissolved on closer examination into a baffling complex of dissimilarities. Every attempt to apply a common standard had the effect of disclosing a deeper level of individual differences and incommensurables." POWELL (1966)

Indeed, the issue of how to measure the performance of health services has been examined as far back as the 18th Century, many years before health services began to be organised formally, when Dr Clifton in 1732 first proposed that health care should be evaluated (Jowett and Rothwell, 1988). It has not been until recently, however, that the evaluation and assessment of health services has become important to both the Government and the NHS, coinciding with a move to accountability and value for money in all aspects of the public sector. According to Carter *et al.* (1992), the most critical terms for performance assessment is the impact of the NHS on the health of the population. However, this was seen as inherently difficult, as was observed in chapter two, due to the difficulty of defining appropriate measures.

The nature of the NHS, with the volume of its staff and diversity of its procedures, also make it difficult to define comprehensive methods for performance measurement. Furthermore, the NHS is unlike most other large organisations in that profit is not a sensible yardstick against which to assess performance. However, especially over the last twenty years, attempts have been made to measure the performance and efficiency of the NHS using a variety of techniques, particularly as the value of performance assessment has been recognised and given greater prominence:

> "The collection of reliable data on costs, quality and outcomes can allow comparisons to be made between treatments, clinicians and hospitals. Such information is valuable to patients and their agents so that they can make informed choice." ALLSOP (1995)

The theoretical issues relating to techniques for the measurement of all aspects of performance is briefly discussed, as an extension to the introduction given in chapter two. Some key points in this section have been highlighted through discussion with Health Care professionals, as identified in chapter one, in order to provide an extra perspective to the theoretical debate. This approach is used throughout the chapter. The general points identified are extended in relation to the specifics of each of the performance measurement techniques introduced later in the chapter.

Following this brief opening discussion, the methods themselves are debated in great detail, with the focus on assessing their strength and weaknesses, introduced briefly in chapter two. The two main approaches under discussion are performance indicators and clinical audit, the two most widely used techniques in the NHS. Other methods will be discussed briefly prior to the summary.

4.2 Performance and Efficiency Measurement

The majority of performance assessment exercises in the NHS take data from a number of health service units and compare the performance of each unit against the others in the data set. The test is either against some average standard of performance or efficiency obtained through the sample, as with ratio analysis and regression techniques, or against some notion of an 'efficiency' frontier, in the case of Data Envelopment Analysis.

Despite the differences in approach, however, a common set of problems and issues have been proposed, which exist with any comparative techniques and must be addressed if they are to be successfully applied. In later sections, these themes are related to each method employed in the Health Service but are briefly presented first in relation to all methods used for the evaluation of performance. There is clearly some overlap between these themes, which are emphasised as the chapter progresses.

4.2.1 Data Availability and Accuracy

In most large organisations, the problem of collecting the relevant data to carry out relevant performance assessment exercises has been observed. However, this problem is heightened in the NHS due to its diversity, size and different organisational levels. It has been observed that there could be inconsistencies in the collection and recording of all types of data. Managers within the NHS have very little confidence in the data collection procedures of any other hospitals. Additionally, data could be missing, wildly inaccurate or simply wrong. The successful application of any technique is reliant upon the quality of the data that is available.

4.2.2 The Politics of Performance Evaluation

Performance evaluation, whatever form it takes, is often viewed with mistrust, because it has been externally imposed or the results are being made public, with some political objectives in mind. As was noted in chapter two, the definition and interpretation of the elements of performance that require assessment have often been determined by the political culture at any given time. A further problem in this area is that the nature of the performance assessment techniques can often be transformed if the political environment changes.

4.2.3 The Nature of the Health Service

As has been demonstrated in chapter three, the organisational structure of the NHS has been extremely complex throughout its fifty-year history. It also has the appearance of constantly being in a state of flux, in that a new series of reforms is proposed before the previous set have been fully implemented. There are a number of implications of these two aspects of the NHS on performance assessment.

Primarily, there are so many areas for which performance could be assessed, it is often difficult to focus on the most important aspects of the service. Also, the changes to the structure, and the resulting impact this has on management, can affect the way in which performance assessment techniques are implemented and controlled. Finally, the impact of technology in all aspects of the health service has affected performance assessment at all levels, particularly as data is now much more accessible.

4.2.4 Implications of Environmental Differences in Comparative Assessments

It is widely recognised that environment, both geography and demography, plays an important part in the prevalence of certain diseases and illnesses, as well as influencing access to health services and treatment options and opportunities. There is a concern expressed by many health care professionals that comparative performance assessment, based on a number of hospitals for example, does not take account of these differences. It is also suggested that 'deviation' from the norm or average does not necessarily imply that a particular hospital is demonstrating a poor level of performance.

4.3 The NHS Performance Indicators

Performance Indicators are data relating various elements of the performance, usually in the form of ratios such as 'cost per patient day' and 'admissions per thousand of the population.' Figure 2.1 in chapter two illustrated a model of public sector performance, where costs were transformed into benefits by some process, with physical inputs and outputs as the intermediary stages.

In general, performance indicators are related to this model, with organisations such as the health service developing indicators 'that are based on ratios of inputs, outputs and outcomes' (Carter *et al.*, 1992). Figure 2.2 introduced a more complex model of performance, showing the relationship between economy, efficiency and effectiveness. Performance indicators can be utilised when these, and other characteristics of performance, cannot be measured precisely (Jackson and Palmer, 1992).

Ball (1998) emphasised, however, the importance of correctly defining performance indicators, stressing that they must not be confused with management statistics, with the result that 'anything that can easily be measured becomes a performance indicator' (Elton, 1987). In the NHS, as will be evidenced by the debate to follow, this has often been the case. For example, the case-mix index for any given specialty is a management statistic, whereas 'average length of stay' is a performance indicator based on the following guideline:

> "When an indicator shows a difference in one direction this means that the situation is better whereas if it shows a difference in the opposite direction this means that the situation is less favourable."
>
> (CUENIN, 1987)

During the early years of the NHS, and up until the late 1970s, there was an arbitrary approach to the measurement of performance. Much of the information that was to form the core of the Performance Indicator packages was available, but was not examined or utilised with any real purpose. As the reforms to the structure of the NHS continued, coinciding with the Conservative Party regaining power at the 1979 General Election, the need for effective measures of performance was highlighted and became a priority.

4.3.1 Introduction and Development of Performance Indicators

The idea of a nation-wide system for performance measurement was first proposed in the late 1970s, with research carried out at the University of Birmingham amongst the earliest in this field.

At the beginning of the 1980s, there were calls from the Public Accounts Committee (1981, 1982) and the Körner Working Party, for greater financial control and accountability in the NHS and effective methods to monitor its efficiency (Carter *et al*, 1992).

The Department of Health responded by starting the development of the Performance Indicator package in collaboration with the NHS in 1981. In 1983, the first set of indicators was published, with about 145 indicators included in the set. The package was developed in response to the above mentioned criticisms and the political concerns of the day: the efficient use of resources, value for money and increased access to services (Carter *et al.* 1992), in a period of 'stringent Government economy' (Pollitt, 1984).

The indicators, therefore, reflected this - there were Performance Indicators relating to length of stay, throughput, cost of treatments, admission rates and waiting lists. They were not wide reaching - Pollitt (1985) suggested that these early indicators were based on a financial view of performance. The purpose behind the first indicator package was summarised thus:

> "The original aim was to articulate a set of indicators that would make visible the links between clinical activity and the use of money and manpower." POLLITT (1984)

The indicators, which the Regional Health Authorities were required to produce as part of their annual review for the fist time in 1983, were split into four major categories:

- Clinical indicators - average length of stay, discharges by speciality per population.
- Financial indicators - cost per case by category, cost per vehicle mile of ambulance services.
- Manpower indicators - cost per FTE by staff groups, ratio of nursing staff to doctors.
- Estate management indicators - ratio of population to land owned or occupied, maintenance expenditure per land occupied per annum. POLLITT (1984)

This first set of performance indicators was severely limited. It was based on data that was already routinely collected, to ensure that all the information required was readily available, and so provided little in the way of new insights. The immediate impact of the performance indicators was mixed: 'neither health authorities nor professionals took much notice of them' whereas 'finance and planning staff took them more seriously' (Baggott, 1994).

In 1985, the Joint Group on Performance Indicators, set up by the Department of Health and with the members drawn from a variety of sources including academics and clinicians, produced a new package of indicators. The number of indicators increased to in excess of 400 and they were sub-divided into eight groups covering all aspects of health provision. These categories were listed by Roberts (1990) as acute services, children's services, the elderly, mental illness, mental handicap, support services, estate management and manpower. The focus of these early sets of indicators was on activity, finance and manpower, with only a few indicators relating to outcomes.

An article in the British Medical Journal in 1988 recognised the growing importance of performance indicators and predicted their future development:

> "There is now so much emphasis on saving money and providing an efficient service that inevitably performance indicators are here to stay and probably their scope will increase." LOWRY (1988)

The latest development of the NHS performance indicator package prior to the 1997 general election was the Health Service Indicators (HSI) package, first introduced in 1989 and fulfilling the prediction made by Lowry (1988) above. The package had more than 2000 indicators, covering seven major topic areas: Health of the Nation, Outcome, Purchaser, Provider, Secondary, Background and Direct Access (NHS, 1996).

Indicators related to all elements of the NHS: in-patient and outpatient services, mental health care, accident and emergency treatment and primary care. Over the last few years, several key indicators have been published annually in the form of league tables. Every hospital in the country received a star rating for each of about fifteen indicators, which focused on waiting times, cancelled operations and other highly visible indicators seen to reflect the quality of service provided.

Following the Labour Party's victory at the General Election in 1997, the position of performance indicators looked set to change again, in accordance with a new approach to health care introduced in 1997 (Department of Health, 1997). This is further discussed in section 4.3.8, on the future development of performance indicators.

4.3.2 Use of Performance Indicators

Performance indicators were intended to be used in a variety of ways for assessing the performance of the NHS. The Health Service Indicators (HSI), according to the NHS documentation, were intended to be used in the following ways:

1. To raise questions and highlight issues for further discussion and investigation and not simply provide answers;
2. To place local activities in a national context;
3. To make comparisons over time;
4. To monitor effectiveness of actions and interventions;
5. To be used in conjunction with each other to tell a complete story. NHS (1996)

Using the example of the orthopaedic unit introduced in chapter two, the number of patients requiring additional surgery following an unsuccessful hip replacement operation could be recorded, in accordance with point four above. The result would then be compared with other units, with any major differences identified, linking in with the approach suggested by point one. Taking the investigation further, other indicators and statistics could be identified, relating to the age structure of the patients treated and the case mix for the department for example, using point five as a guide.

Performance indicators, according to a survey of their use by Jenkins *et al.* (1987), have been used to support or confirm preconceived views on performance, provide evidence to add weight to arguments and generate interest in particular issues.

Additionally, they have been applied to identify priorities for investigation, review and plan for particular services, identify norms and targets and provide information to be used to bid for extra resources.

Jenkins *et al.* (1987) identified the most frequently used performance indicators as those that related to bed usage and staffing levels. Three key points emerged from their study:

> 1. Performance indicators provided only part of the information used in any given debate;
>
> 2. The use of performance indicators was reactive - information was generally sought from them once an issue has been raised for discussion by other means;
>
> 3. Performance indicators tended not to introduce new direction into any debate. JENKINS *et al.* (1987)

Taking points one and three above in conjunction suggests that performance indicators at that time were not revealing anything startling. For example, the indicator measuring 'average length of stay' for the orthopaedic unit at one of the hospitals was measured at 12.3 days compared with the Scottish average of 7.3 days (ISD, 1996). The exact scale of the difference between this hospital and the national average might be a new piece of information for the hospital's management. However, they would probably have been aware that 'length of stay' for the orthopaedic department was 'too high' in relation to other departments around the country.

The points given below relate to actual rather than intended usage, highlighting how indicators have been used to instigate further investigations:

(i) They highlight areas of enormous variation in resource use that indicates areas where further investigation is warranted;

(ii) Where several of a hospital's indicators show similarly atypical values, they suggest the management of the hospital as a whole warrants investigation;

(iii) They provide objective support for existing suspicions or fears concerning performance, thus providing a trigger for action.

BIRCH AND MAYNARD (1986)

Other uses of performance indicators, particularly in relation to efficiency assessment, are discussed in section 4.3.6.

4.3.3 Users of Performance Indicators

Research into the use and value of Performance Indicators in 1987 by CASPE (an independent health care research organisation) determined who was actually using the existing indicators. They found that a substantial proportion of those with managerial responsibilities claimed to use Performance Indicators at some point in their work.

However, it was found that the different categories of management and clinical staff made varying use of Performance Indicators, with some hardly using them at all (Jenkins *et al.*, 1987). For example, almost all information specialists, District General Managers and planners claimed to have used Performance Indicators at some point, but only about half of the financial managers repeated this claim.

At Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust, the Director of Corporate Development felt that the nationally published Performance Indicators provided a useful summary of the quality of the services provided by the Trust. He made much more use of these indicators than the Director of Finance, who regarded much of the data provided in this form as misleading. Usage of the indicators at the trust amongst the various departments and managers was very much a matter of 'personal preference' (personal communication).

4.3.4 Interpretation of Performance Indicators

The interpretation of performance indicators has been identified as one of the most critical elements of their usage. Numerous reasons can be suggested as to why an indicator may differ from the average or normal level expected, and these need to be taken into account in the interpretation process. Some of the reasons why a specific indicator, such as 'average cost per discharge', may differ from the norm or national average are given below:

> (i) It indicates past as opposed to present performance;
>
> (ii) It arises from atypical demand factors as opposed to atypical hospital performance;
>
> (iii) It arises from supply factors that are peculiar to the locality in question;
>
> (iv) It arises from the resource use indicated by national norms being inefficient;
>
> (v) It arises from the inefficient use of resources in the hospital that has remained unidentified or uncorrected.
>
> BIRCH AND MAYNARD (1986)

It has been firmly argued that performance indicators cannot be taken in isolation - extra information was needed to fully investigate the existing inefficiencies. These include the level of service being achieved by the use of resources and the potential and scope for alternative resource usage (Birch and Maynard, 1986). Also, interpretation of performance indicators is not a simple issue:

> "Because there is such a variety of potential users of PIs, with different objectives and interests, it is important to recognise that there is no single way of interpreting performance data."
>
> SMITH (1995)

Lowry (1988) points to the difficulty of interpreting performance indicators and their usefulness if this is done properly:

> "Each indicator is a crude statistic and there is no 'correct' result for any of them; rather their value is in the concept that a ranking at an extreme of the rational distribution may require further investigation. If several related indicators appear at an extreme, this may be even more suggestive." LOWRY (1988)

The difficulty of interpreting any given indicator was demonstrated by an analysis of the 'average length of stay' indicator for the specialty of General Surgery, given by Lowry (1988). A higher than average value for this indicator may be caused by a high proportion of elderly patients or poor housing conditions in the locality, a complicated case-mix in the department, or inadequate primary care facilities. An examination of these factors, alongside investigating alternative discharge policies, would be required to determine if the high figure for length of stay reflected poor performance or the effects of the other factors.

This highlighted the importance of understanding environmental factors in the interpretation of performance indicators.

At BHRV NHS Trust, identifying hospitals operating in a similar environment was part of their approach to comparative performance assessment, in order that the indicators could be interpreted in a meaningful way. External consultants, using a number of characteristics specified by the trust, identified a cluster of 'similar' hospitals. This was found to be especially important in the examination of the treatment of certain illnesses, such as heart disease and cancer, which may be linked to environmental influences (personal communication).

To counteract difficulties caused by variations in data collection and the impact of environmental factors, the Director of Finance preferred to use 'benchmarking' against past performance levels at BHRV NHS Trust rather than comparing with current performance at other hospitals, operating outside his control. His perception was that performance indicators should generally only be used for internal comparisons.

4.3.5 Limitations and Criticisms of Performance Indicators

The view of Roberts (1990) was that the performance indicators used in the NHS do not actually measure performance - they could be used as tools for asking questions and identifying the means for improving performance but do not measure the actual performance of a particular element of the Health Service. However, as was observed at the beginning of the section on performance indicators, this was to be expected.

Several other major areas of concern regarding the use of Performance Indicators as measures of the performance of the NHS have been highlighted. The following six were given by McCarthy (1983):

- PIs pose more questions than they answer - any variations between indicators represent a variety of influences, such as resource availability and clinical policy, which need to be examined to fully understand any given situation;
- The sheer volume of indicators mean that a great deal of time is required to examine them and determine any useful information from them;
- PIs have been developed within institutional boundaries - they have not always looked at the whole picture of health care provision;
- No clear method exists for interpreting PIs to produce action;
- Data used to produce the indicators may not be accurate or consistent;
- The focus of PIs is on the process of health care and not on the outcome;

Performance indicators were found to be limited in terms of their use for operational management due to the time lag of up to two years that existed in their publication, as the situation under investigation could have clearly changed quite considerably.

A further criticism of performance indicators, emphasised by Barrow and Wagstaff (1989), focused on the relationship with efficiency. They have been described as singularly inadequate indicators of inefficiency as they fail to differentiate between inefficiency and the effects of other factors, which can be exogenous or endogenous.

Clearly, this would relate to the use of performance indicators in other areas within the public sector.

Jenkins *et al.* (1987) also found that, occasionally, particular indicators had been used selectively, creating only a partial picture of a given situation, to suit the ends of the managers examining the performance indicators. In other words, there was the potential for the misuse of performance indicators. Another limitation with the use of performance indicators in the NHS was the lack of clear objectives, for both the indicators and the NHS:

> ".. the objective of PIs must be to improve performance: to work effectively, indicators must show how to further the objectives of the organisation. But for the NHS objectives are not clear and the relationship between indicators and objectives is ambiguous. For example, a higher length of stay might better achieve the objective of prolonging life but would usually be taken as a sign of low efficiency." ROBERTS (1990)

Further limitations of performance indicators as a means of assessing performance for comparisons across hospital departments have been highlighted by an examination of the indicators relating to Accident and Emergency departments by Edhouse and Wardrope (1996). A widely used indicator of the standard of care provided by A&E departments has been: 'the time to immediate assessment.' This was used as a means for comparative performance assessment and published in league table format by the Department of Health. Hospitals received a five-star rating if 95% of patients arriving in the emergency department were 'assessed' within five minutes.

The major flaw with this indicator, according to Edhouse and Wardrope (1996), was that there was no standard definition for what this initial assessment should have involved. In some hospitals, A&E departments operated an advanced triage system - on arrival, patients were examined by a specially trained nurse, basic treatment given and their medical history recorded before being referred to a doctor for immediate attention or returning to the waiting area. At the other end of the scale, assessment may be by a visual examination only.

They also noted further discrepancies in the measurement of the waiting interval - some hospitals measured the waiting time from when the patients entered the department whilst others used the time of registration as the starting point. The performance indicator 'time to immediate assessment' takes no account of these differences and their potential impact on the quality of care provided. Nor are they recorded in the League tables. Also, the need for hospitals to achieve as many five-star ratings as possible may impinge on the quality of service provided. It is clearly easier and less costly to assess 95% of patients within five minutes if 'assessment' takes only a few seconds rather than a few minutes, as would be the case if a full triage system was in operation.

This issue has also made the jump to hospital drama, as it was raised in the BBC series Casualty on 30/11/96, where a discussion between the Nursing Manager of the A&E department and a senior manager of the hospital led to the quote:

> "You surely don't want us taking on extra nurses whose only job is to say hello so we can claim that everybody's been seen in the first five minutes." (BBC, 1996)

The problems highlighted in the case of A&E departments led to the following criticism of the league table system as a whole:

> "The star rating system provides no useful information; it does not truly reflect the quality of care and is misleading if used to compare the performance between departments...unless data are collected in a standardised way, the national performance figures will continue to be uninterpretable... It is possible that the pressure to assess patients quickly will lead to hasty judgements, purely to satisfy arbitrary numerical standards." EDHOUSE AND WARDROPE (1996)

The Director of Corporate Development at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust recognised the potential for problems in indicators being used in league tables, leading to the following key points:

- Performance indicators are of limited use for comparing performance across hospitals because there are no standard definitions for terms such as 'assessment' and procedures for measurement are not uniform - it is difficult to be completely confident that comparisons are made on equivalent information.
- The indicators given prominence in the league tables may be given more attention than other areas of service provision, allowing standards to fall in these areas, simply because the indicators in league tables are more visible - this has happened particularly with the indicators relating to waiting times and waiting lists.
- Quality of care may suffer in order to meet specified targets and achieve as high a star rating as is possible.
- The use of just one indicator as a measure of the overall performance of a department may be misleading.

Skinner *et al.* (1988) also highlighted how poor data quality can seriously affect the results given in performance indicators. They investigated the orthopaedics services provided in the London health district of Camberwell and found that an 'inexact classification and grading of operations' had led to startling errors in two of the key performance indicators. There was a 19.8% error in the 'weighted number of operations' and a 34.5% error in the 'number of major operations per consultant' (Skinner *et al.*, 1988). The study determined that:

> "If performance indicators are to be of use in planning then accurate figures are essential.... inaccuracies were found at all levels and were principally errors in the collection, classification, grading, and interpretation of data." SKINNER *ET AL.* (1988)

A further criticism of performance indicators in the health service can be related to the amount of data collected for inclusion in the performance indicator packages. A member of staff involved in the processing of performance indicator data at BHRV NHS Trust believed that much of this information was not used internally in any managerial capacity. Certainly, many of the indicators collected were found to have little or no practical value. The relationship between performance indicators and outcomes is now investigated, followed by a discussion on their link with efficiency assessment.

4.3.6 Performance Indicators and Outcomes

One of the major criticisms of performance indicators, particularly in the early period of their development and noted by McCarthy (1983) above, related to their failure to address the results or outcomes of treatments.

They did not relate to the success, or otherwise, of the NHS in improving the health of the population in general. This has gradually been addressed, most dramatically with the introduction of the Health of the Nation Strategy in 1992 and the most recent proposals for NHS reform, introduced by the Labour Government (Department of Health, 1997).

The focus of the Health of the Nation strategy has been to 'achieve a continuing improvement in the general health of the population..., with an emphasis on prevention and health promotion in addition to treatment' (Department of Health, 1996).

The success of the Health of the Nation Strategy, according to the Department of Health, was dependent upon Government, voluntary organisations, communities, families and individuals as well as the health agencies taking a 'shared responsibility for health' (Department of Health, 1996).

Some of the main health areas covered by the strategy are coronary heart disease, strokes, sexual health, mental illness, cancer, AIDS and HIV, accidental injuries, smoking, drinking and obesity. The link between the areas included in the strategy was that they are generally:

> ".. all major causes of premature death or avoidable ill health and offered significant scope for improvement in health."
>
> NATIONAL AUDIT OFFICE (1996)

The indicators relating to these areas were found in the HSI packages published annually until the recent reforms took effect. They related the progress that had been made towards a series of targets for each element of the strategy, most of which are long-term. Figures were published on a national basis and could also be broken down to examine progress regionally and to identify if regional trends existed for certain diseases. Examples of some of the targets were: reducing the incidence of breast cancer, reducing the proportions of obese men and women in the population, cutting the number of deaths from accidents in certain age categories and reducing the number of women smoking whilst pregnant. The most recent approaches to link together performance indicators and outcomes are discussed in section 4.3.8.

4.3.7 Performance Indicators and Efficiency

In the context of performance measurement using performance indicators as applied in the NHS, efficiency has been regarded as just one aspect of performance to be investigated along with many of the others identified in section 2.6. However, during the early stages of development for the performance indicator packages, improving efficiency was identified as one of the central policy objectives, as referenced in section 4.3.1.

Within the NHS itself, the performance indicators on activity and expenditure levels have been used to look for any improvements in efficiency. A weighted sum representing annual activity was compared with changes in expenditure to give a broad measure of potential improvements in efficiency.

This was in the form of a ratio, as defined in figure 2.3 (NHS, 1996). This performance indicator was often referred to as the 'efficiency index'. If activity had increased at a higher rate than expenditure, this was seen as evidence of improved efficiency.

This approach for identifying inefficiencies has been seen to be quite crude, according to the Director of Corporate Development of the BHRV NHS Trust in reference to the 'efficiency index'. It was felt that using this approach to improving efficiency could be counter-productive in terms of the hospital's long-term strategic planning.

Smith and Mayston (1986) discussed a potential benefit of performance indicators in the assessment of efficiency, suggesting that they could be used to 'point towards area of apparent inefficiency', with the intention of finally eliminating that inefficiency. Birch and Maynard (1986) proposed that they could be used to consider the efficiency of performance. Further discussion on the assessment of efficiency is found in the next section.

4.3.8 The Future Development of Performance Indicators

Performance Indicators have developed substantially over the last fifteen years, particularly in the move from being process-oriented to the inclusion of more indicators relating to health outcomes in the Health of the Nation Indicators. This is likely to continue, particularly in conjunction with medical audit and in the light of the recent reforms to the NHS (Department of Health, 1997).

Data to be included in the performance indicators has always been collected internally. A report in 1993 suggested that, if performance indicators are to be used more widely and be more readily accepted, there would need to be a move towards independent data collection and verification (Times Health Supplement, 1993).

Birch and Maynard (1986) suggested that there were four main areas for improvement in the development of performance indicators:

1. Hospital objectives - consider the mix of processes being performed, use data on quality, use data about patients treated, use additional data sources, classify data by clinician, consider all resources used

2. The ambiguity of indicators - consider not just inputs but outputs and the input-output relationships

3. Indicator specificity - recognise and respond to the fact that hospitals differ in mix of activities, environments, mix of patients, nature of costs incurred per patient for particular activities.

4. Indicator sensitivity - ensure that inefficiencies identified are accurate and genuine.

Some of the above points have been addressed during the last ten years, particularly in relation to health outcomes, and the most recent reforms to the structure of the NHS. These changes have introduced a 'new' approach to performance, with the identification of a 'National Performance Framework' intended to be useful to, and used by:

1. The general public - to inform them about their local NHS and as they make decisions about their own care;

> 2. NHS agencies – to inform and improve performance and to use in planning;
>
> 3. Ministers and the NHS Executive – to drive improvements in performance of health authorities and to demonstrate public accountability for the use of NHS resources. BAKER (1998)

The framework was designed to incorporate six dimensions, with a set of 'high-level' performance indicators to accompany it. The framework has been designed to reflect the new approaches to performance and efficiency assessment proposed by the NHS reforms, discussed in chapters two and three previously. Table 4.1 overleaf includes a description of each of the six dimensions, with examples given of some of proposed indicators.

Future changes to this list are inevitable, with clinical and primary care indicators the most likely area for further investigation (Baker, 1998). New indicators under this approach will be developed through consultation, requiring a long-term strategy for finding the most valuable performance indicators in the future:

> "The list is not, and never will be, perfect, but significant progress is being made in defining the appropriate products of health care in a comprehensive system... Few of the indicators offer anything new... Some of the indicators... are worthless as they have no scientific value or consistency... Nonetheless, unless measures such as these are used to change the way the NHS is managed, no effort will be put into improving the ones now on offer. Necessity may prove to be the mother of invention." BAKER (1998)

Dimension	Description	Indicators
Health Improvement	Reflecting the overall aim of improving the general health of the population, influenced by external factors.	- Deaths from all causes (age groups: 15-64 and 65-74) - Cancer registrations (e.g. stomach, skin, breast)
Fair Access	The NHS contribution must offer fair access to services in relation to need, irrespective of ethnicity, class, age, sex or geography.	- Surgery rates - (Hip and Knee replacements for the over 65s and Cataracts) - Conceptions below age 16 - % population registered with a dentist - Early detection of cancer
Effective Delivery of Appropriate Health Care	Care must be effective, appropriate and timely, and comply with agreed standards.	- Disesase prevention and health promotion (vaccination targets) - Inappropriate Surgery - Acute Care Management (Age and Sex standardised admission rated) - Cost-Effective Prescribing
Efficiency	How the NHS uses its resources to achieve value for money.	- Day Case Rate - Length of Stay in Hospital - Unit Costs - Generic Prescribing
Patient/Carer Experience	Measuring how patients view the quality of the treatment and care they receive and establishing a new national patient survey and NHS charter.	- Patients who wait more than two hours for Emergency Admission - Number of Cancelled Operations - Non-attendance at Outpatient Appointment - Outpatients seen within 13 weeks of Referral by GP - Inpatients admitted within 3 months
Health Outcomes Of NHS Care	Assess the direct contribution of NHS care to improvements in health.	- Avoidable diseases - Complications in Treatment - Emergency psychiatric readmission rate - Infant deaths - Survival Rates for Breast and Cervical Cancer

Adapted from DEPARTMENT OF HEALTH (1997) and BAKER (1998)

Table 4.1: National Performance Framework in the New NHS

4.3.9 Evaluation of Performance Indicators

The view of the Department of Health is that Performance Indicators are only a part of the process of attempting to improve the management of the NHS - their value comes in providing additional information to be used in the overall evaluation of performance. In 1988, the then Secretary of State for Social Services, Mr John Moore, reported favourably on the role of performance indicators:

> "People from outside the NHS can look at and evaluate how individual health authorities are doing.... The indicators place the emphasis on the critical examination of services. They raise questions, provide a means of helping to diagnose problems and then suggest possible solutions to those problems." LOWRY (1988)

However, this is not universally accepted:

> "In order that performance indicators can monitor effectively the performance of a hospital, the indicators should exhibit four main characteristics; they should relate to the objectives of the hospital; they should provide unambiguous information concerning performance; they should be specific in their implications; and they should be sensitive to what they are supposed to identify. The performance indicators which underlie the present evaluation of health services are lacking in all four characteristics."
>
> BIRCH AND MAYNARD (1986)

The problem with analysing the impact of Performance Indicators on the Health Service, according to Carter *et al.* (1990), has been that they are not actually providing any new information - the label on the package has changed but the contents were nothing new.

The growth of performance indicators, and other methods of performance assessment, reflected the importance placed upon comparing the performance of different health authorities and units by central Government. The position and usefulness of performance indicators was summarised thus:

> "If their limitations are properly understood, the indicators become valuable management tools to highlight topics that need further inquiry." LOWRY (1988)

4.4 Clinical Audit

Audit is a means for reviewing the activities of the past. In business, it usually involves a review of a company's accounts. In the National Health Service, clinical audit refers to a review of the patients treated, often case by case. It has become an essential part of the work of almost all members of the medical profession and an important element of the performance assessment process. It is often also referred to as medical audit.

As was discussed in chapter two, the audit approach looks to highlight specific example of poor performance, as opposed to the general observations that result from the application of other techniques such as performance indicators and regression analysis.

4.4.1 Definition of Clinical Audit

The most frequently cited definition of audit, as it relates to the NHS, is that given in the Working for Patients White Paper:

> ".. a systematic critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcome for the patient."
>
> DEPARTMENT OF HEALTH (1989)

Reviewing this definition shows that an audit, therefore, should have at least four main elements or concerns: the notion of medical quality, the differentiation between medical procedures, the examination of resource usage and the outcome of the treatment on patients (Packwood *et al.*, 1994). Robinson (1994) suggested that it was also essential to have some precise definitions of the objectives of the activity being audited.

Audit is an all-encompassing term for what can actually be a very varied activity. Packwood *et al.* (1994) identified that this process could be broken down into eight aspects, all of which have a variety of possible alternatives, in order to give a full picture of what audit is in practice: participants, organisers, subjects, methods, style, costs, impact and purposes.

4.4.2 The History and Development of Audit

The idea of audit was not a new approach prior to its elevation to compulsory status (see section 2.7.3) as a tool for performance assessment - doctors, and other members of the clinical professions, have always reviewed and assessed the patients they have treated, and the success of the methods used.

However, in the past the responsibility for audit lay solely with the medical profession, with no methods for control or enforcement applied. Hence, audit was arbitrary - some doctors were interested in it, others were not. Packwood *et al.* (1994) described the situation as 'fragmented'.

However, there were audit initiatives organised on local and national scales, one of the earliest being the Confidential Enquiry into Maternal Deaths, carried out on a national scale since 1952. A further example was the National Confidential Enquiry into Perioperative Deaths (NCEPOD), which was first completed in 1986 (Baggott, 1994). These national audits were organised by professional bodies such as the Royal College of Physicians and participation was entirely voluntary, although the two highlighted above achieved almost full coverage.

Local audits, often organised at hospital level by individuals or small teams, were in evidence, but not in any comprehensive manner. Even prior to medical audit being made compulsory, there was no universal agreement as to the value of it as means for performance assessment.

Also, some clinicians and medical practitioners were opposed in principle to any further development of the existing arbitrary approach to audit, claiming that any system of review infringed on their clinical freedom. Others were concerned with the amount of time that audit would require, taking clinicians away from the process of treating patients.

However, during the 1970s the position of audit gradually became more prominent, particularly with the Royal Commission Report in 1979, which highlighted the need for a more comprehensive and systematic development process for ideas such as audit (MERRISON, 1979). By 1989, the medical profession itself was beginning to recognise the role of audit:

> ".. medical audit should quickly become established practice for all physicians." ROYAL COLLEGE OF PHYSICIANS (1989)

The elevation of audit to the political stage also coincided with the rise of the performance indicator as a means for performance measurement - both relate to the Government's requirements for value for money and accountability in public sector services. Roberts (1990), amongst others including Griffiths (1983), highlighted several key areas where audit was seen to be able to contribute to these aims, and thus a popular option within Government:

- ⇒ Reducing the high cost of inappropriate treatment and mistakes.
- ⇒ Improving the effectiveness of resource usage.
- ⇒ Increasing clinical accountability.
- ⇒ Increasing patient satisfaction as consumers of health care.
- ⇒ Improving outcomes and thus the general health of the population as a whole.

The development process culminated in 1989 when the Government, in the White Paper Working for Patients, determined that medical audit was to become a compulsory element of the work of the medical profession, in particular of consultants and general practitioners (Department of Health, 1989).

A five-point plan, set out in the Working Paper, was the basis of the proposals for the introduction and development of medical audit, as below:

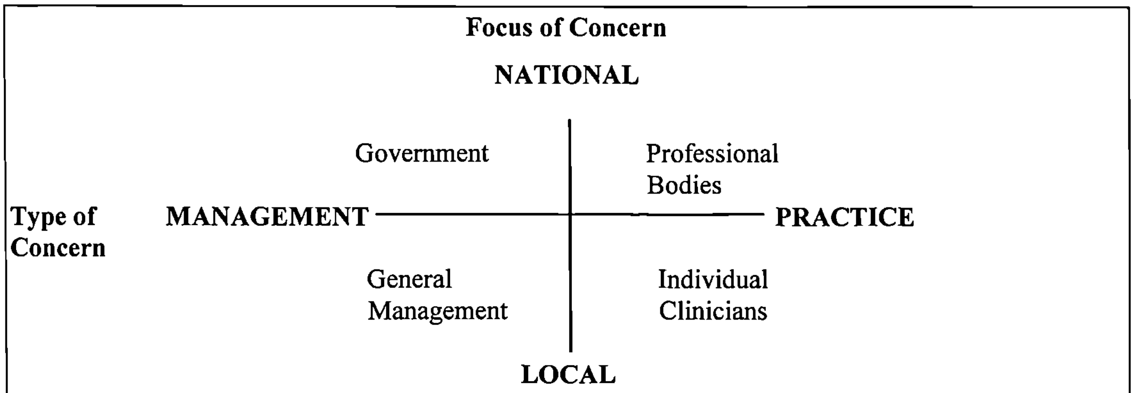
- Every consultant would be required to participate in medical audit in a form to be agreed locally between management and professions.
- The system would be medically led, with an advisory committee to be chaired by a senior clinician.
- Responsibility for ensuring effective systems of audit are in operation would lie with District Management teams.
- Findings from individual cases should remain confidential but any general issues should be reported to local management and beyond.
- Management does have the authority to instigate independent audit, if the quality or cost-effectiveness of a service is questioned. DEPARTMENT OF HEALTH (1989)

After the initial proposals, slight changes were made in that junior and senior hospital-based doctors were also expected to participate in the audit process. Since 1989, the Government has contributed in excess of £200 million to aid the development of audit in the National Health Service. This was particularly through allocations to regional health authorities and in 1993-4, over 20,000 audit programmes were completed in the hospital and community health services in England (National Audit Office, 1995). The National Audit Office has subsequently reported that the NHS Executive believed that 'the initial phase of stimulating the introduction of clinical audit' has been completed (National Audit Office, 1995) and the place of audit as a means for performance assessment is firmly established.

4.4.3 Audit in Practice

The definitions for audit, as given above, only give a very general idea as to what audit actually involves in practice, especially as there are so many possible approaches to the process. Audit has been local, national or regional; organised, implemented or evaluated by Government, professional bodies such as the Royal Colleges, hospital management or individuals; and concerned with issues of resource management, medical practice or quality.

The diagram below illustrates how some of these different aspects of audit have fitted together:



PACKWOOD *ET AL.* (1994)

Figure 4.1: Interests in Medical Audit

Government documents stressed that medical audit must be 'developed and implemented with care' (Department of Health, 1989). Several key issues relating to the implementation of audit in practice have been identified from the literature and through interviews with managerial staff involved in audit at Blackburn, Hyndburn and Ribble Valley NHS Trust.

4.4.3.1 Participation in Audit

Participation in audit has been found to vary enormously across every project initiated. It can range from clinicians carrying out an individual review to teams made up of consultants, junior doctors, nurses and administrative staff reviewing procedures according to specialty. In general, the majority of participants in audit are medical professionals.

At BHRV NHS Trust, the non-clinical staff members were being encouraged to participate to a greater extent in the audit process, in particular the nursing staff. The trust believed that this gave a more balanced approach to the evaluation procedure, as each group potentially had a diverse outlook on the cases being audited.

Packwood *et al.* (1994) suggested that staff members have benefited from participation in audit in a number of different ways. Junior doctors have used audit as an extension to their education, helping them to learn new approaches from senior consultants. Alternatively, nurses have perceived it to be an opportunity to increase their involvement in the medical decision-making process.

Although the initial target of all clinicians participating in audit had not yet been achieved, 'substantial progress had been towards that goal. About 60% of specialties in Scotland were participating in audit, and about 50% of general practitioners were involved in audit projects (Payne, 1995).

4.4.3.2 Scope of Audit

The scope of an audit has also been identified as an important part of implementation process in a hospital situation. Most audits were carried out at the local level, within hospital departments. National schemes of audit also existed, but were organised very differently. Within BHRV NHS Trust, there was a mix of local and national audit schemes, as well as participation in regional programmes if they were of recognised importance. Each type of audit can come with their own set of problems and benefits, identified in table 4.2 below.

Scope	Benefits	Problems
Local	Wider involvement, with nurses and administrators more able to contribute.	Hindered by hidden agendas of the participants – each clinician has own pre-conceived view on services provided.
	Focus on specific issues affecting the particular trust (may not be problems in other areas) and environmental variations are more readily recognised.	Threatening, as there are frequently only two or three clinicians working in some of the smaller specialties.
	Use a variety of approaches, ranging from peer review and case-by-case evaluation, monthly audit meetings to statistical analysis of several hundred cases.	Much greater level of commitment, particularly from the organising clinician, but also from all participants.
	Easier to operate long-term audit schemes and monitor the progress of any changes.	Shortage of data, as some procedures may be performed infrequently.
National	Large groups of study, giving a wider experience of different approaches and greater access to data.	Do not usually allow for multi-disciplinary involvement.
	No problems with hidden agendas and direct confrontation, as audit is carried out by external organisations.	Difficulties in taking into account any contributory contextual and environmental factors.

Table 4.2: Perspectives on the Scope of Audit from BHRV

4.4.3.3 Initiation of Audit

Audits have been instigated at BHRV NHS Trust, usually, in one of two ways. Clinical staff have chosen to examine a particular element of their caseload, either internally or by involvement with a national audit scheme. Alternatively, the hospital management proposed certain areas where attention was needed, highlighted by them or some external agency, such as a General Practitioner, Regional Health Authority or Government directive.

A certain percentage of the audits undertaken at the Trust each year were instigated by the management, in order to prevent clinicians having absolute control over the choice of subject for every audit. For example, variations in the treatment process for patients referred for the treatment of breast cancer, according to the nature of their referral, were found by an audit suggested by the local GPs.

4.4.3.4 Reasons for Audit

The motivations for carrying out an audit have been numerous. Looking first at clinical reasons, audit has been used to ensure patients are diagnosed correctly and subsequently receive the best treatment. A case by case review of patients treated over a certain time period was used to find mistakes, with steps made to ensure that these were not repeated in the future. Audit was also used as a method for assuring that the quality of medical treatment met an acceptable standard, set either externally or internally. At BHRV NHS Trust, the results of audits were being used to produce guidelines for junior and senior doctors relating to the treatment of certain procedures. In other words, audit was used in the process of medical education.

Alternatively, there could be managerial reasons for carrying out audits. Roberts (1990) gave several suggestions as to how audit can be used by Trust managers to improve the information they have at hand for assessing performance, allocating resources and addressing inefficiencies:

> "It can identify process variations which would not otherwise be visible. These include the use of unnecessary materials, the inappropriate use of services in automatic reattendances, number of inpatient days spent waiting for treatment, inappropriate admissions or care which is of no direct value to the patient concerned."
>
> ROBERTS (1990)

4.4.3.5 Information from Audit

Audits on a national scale have been used to provide information on a variety of important issues. For example, the NCEPOD survey in 1992, an audit of cases involving peri-operative deaths, found that a quarter of all deaths in the areas of trauma and orthopaedics occurred at weekends and bank holidays. It was suggested that 5% of these deaths could be linked with personnel shortages and a further 4% connected to limited resources (Baggott, 1994). In this case, the audit process drew attention to a lack of staff and facilities at critical times and the information could be used to improve performance by a re-organisation of resource allocation policies.

4.4.3.6 Impact of Audit

The influence of an audit depended upon the support, or otherwise, of clinicians and managers alike. If support for change existed, then audits were likely to have a much greater impact on services. Audit was also found to be an on-going process, requiring a long-term perspective to be effective.

In all cases, some monitoring method was required to determine if any changes recommended by an audit were implemented. A report by the National Audit Office (the Scottish equivalent of the Audit Commission) found that Scottish Office investment in audit was 'contributing to change in clinical practice, organisation and management which were leading to improvements in patient care' (Payne, 1995).

4.4.4 Evaluation of Audit

As with all methods for performance assessment, there is disagreement as to the value and effectiveness of medical audit. Maynard (1991) described audit as a 'black hole' and Robinson (1994) reported on the doubts that have been raised as to the effectiveness and cost-effectiveness of spending money on developing audit as a means for improving performance in the health service and believes that:

> ".. realising the benefits of audit requires an act of faith."
>
> ROYAL COLLEGE OF PHYSICIANS (1989)

In addition, Allsop (1995) suggested that there were three major problems with medical audit, summarised below:

- It runs counter to the principle of clinical autonomy and so its implementation may be resisted by clinicians.
- In many areas of medical practice, there are differences that offer valid alternatives for particular patients. If approached too rigidly, protocols can inhibit innovation.
- The evaluations of treatments have tended to rely heavily on clinical criteria which have emphasised hard data, excluding 'soft' measures such as quality of care and patient satisfaction.

However, countering the first point of Allsop (1995) is the view that medical audit has not really challenged the autonomy of the medical profession at all. They have generally accepted the idea where they have been able to retain control of the audit process, as was observed at BHRV NHS Trust.

In fact, Baggott (1994) suggested that audit has been set up in such a way as to ensure that medical interests have not been offended at all, at least in the short term, and audit could be used to actually further medical interest. Packwood *et al.* (1994) concurred, believing that audit was originally driven and shaped by the interests of the medical professionals at a local level, with managerial and national concerns of little importance.

This view was confirmed by Government documentation on audit, which described it as 'essentially a professional matter' (Department of Health, 1989). Further, it has been suggested that audit was actually established in such a way as to exclude management and patients (Pollitt, 1992 reported in Baggott, 1994).

This approach to audit allowed for a certain degree of innovation and individualism. However, it could place the future of audit in a very fragile position:

> ".. it can be readily ignored or omitted, its results argued away as idiosyncratic, its insights seen to be duplicated by other sources, its purposes conflicting, with no perception of any serious detriment to medical practice resulting from its absence."
>
> PACKWOOD *ET AL.* (1994)

In order for medical audit to be used to its full potential a balance was therefore required: Roberts (1990) recognised that management support was crucial and some degree of clinical accountability to management unavoidable but the audit process should not be 'owned, planned or controlled' by Trust managers. It required input and interest from all members of a hospital's personnel.

There have been conflicting views, also, as to the impact of audits and disagreement over whether anything was actually achieved by them. According to figures from the United States, for example, only 20% of a group surveyed attributed any changes in clinical practice to audit (Cassanova, 1990).

Packwood *et al.* (1994) found that if an audit suggested that a change in practice was required, the change was less likely to be made. Thus the audit would only be successful if more than one department was involved in the situation. They also been found that some audits 'sink without trace.'

However, Roberts (1990) reported on the Lothian Work, when data on operations and deaths was collated and reviewed annually. This reported significant improvements in performance over a five-year period, particularly in the number of patients requiring further surgery following postoperative complications. Packwood *et al.* (1994) related that a survey of those involved in audit showed that support staff believed that the positive outcomes of audit were: changes in practice and policy and better-informed doctors. 68% of junior staff believed that audit meetings had led to improvements in their work.

Audit has clearly been viewed in a variety of ways and has been used to provide answers to a host of questions relating to the provision of health care. These range from achieving quality in health care to improving the use of resources and staff, leading to greater efficiency, as the Government set out in its initial proposals. Alternatively, it was a waste of both time and money and beset with too many problems and ambiguities to be effective in assessing performance, as suggested by opinions expressed earlier, particularly in Maynard (1991) and Robinson (1994).

4.4.5 The Future of Clinical Audit

Clinical Audit is clearly a well-established part of performance assessment in the health services, as the case of BHRV NHS Trust suggested. However, the position and role of audit has been constantly developing, as audit was used in different ways by Trust managers and clinical staff to examine new areas of performance.

The position of clinicians as the instigators of audit may have changed. Baggott (1994) believed that the strengthened position of the purchaser of health care will lead to audit being used by the managers of health care providers to 'improve to some extent the quality of medical practice.' This perception could be amended following the removal of the Internal Market in the NHS, although the tensions between purchasers and providers could remain. Packwood *et al.* (1994) also believed that there would be a shift away from audit being driven by clinical views and it would become a much higher priority for Trust managers as well:

> "It is likely that both local management and the national professional bodies will play a stronger role in medical audit in the future.... there will be more of a balance between the different interests, which means that the room for discretion in audit by individual clinicians is likely to diminish."
>
> PACKWOOD *ET AL.* (1994)

In the future, audit may be used in conjunction with other methods for performance assessment, such as performance indicators, as hospital trusts and GPs are required to meet targets and reduce expenditure. Baker (1998), in his review of the reforms to the NHS introduced by the Labour Government in 1997, reported that 'every NHS trust will have to embrace the concept of clinical governance', of which clinical audit will be a central feature, used for quality improvement.

In conclusion, seven key points are presented as criteria for undertaking audit in the future, identified by a committee established by the Scottish Office:

- The issue addressed is a common problem;
- It is a significant or serious problem;
- Change following audit is likely to benefit patients;
- Change is likely to lead greater effectiveness;
- The issue is relevant to professional practice or development;
- There is a realistic potential for improvement;
- The end result is likely to justify the investment of time and effort involved.

CLINICAL RESOURCE AND AUDIT GROUP (1994)

4.5 Additional Approaches to Performance Assessment

The methods identified above are the two most commonly applied techniques for performance assessment in the NHS. Performance indicators have been generally applied to investigate performance at a general level, focused mainly on the processes involved in transforming inputs into outputs with some reference to outcomes. Audit, alternatively, has been used to investigate performance on a much smaller scale, often case-by-case, to identify specific instances where the treatment received by a patient did not reach a given standard, was ineffective or unnecessary. However, there are a number of other approaches used to investigate the various elements of performance, although to a much lesser extent. Two of these, QALYs and Regression, are discussed in this section to illustrate the scale of performance measurement in health services, although there are many other alternatives that could have been included, such as clinical governance and cost-benefit analysis.

4.5.1 QALYs

QALYs (Quality Adjusted Life Years) were developed to establish criteria for 'priority-setting and rationing' to investigate the 'relative cost-effectiveness of different treatments for the same illness' (Baggott, 1994). For a range of treatments or a health education programme, a cost per QALY could be calculated, where 'a year of healthy life was taken to be worth one; the value was lower if health was poorer or life expectation shorter' (Rivett, 1997). Baggott (1994) suggested that the QALY could be used in the future to determine if some low priority treatments, such as fertility treatment and the removal of wisdom teeth, should have their funding withdrawn.

According to Chisholm *et al.* (1999), QALYs 'have been heralded as important aids to planning and priority setting, but also criticised on technical and ethical grounds'. The main objection to the QALY has come from the medical profession, as they challenge 'the clinical freedom to carry out any treatment' and using this approach was like 'comparing apples and oranges' (Rivett, 1997).

4.5.2 Regression Techniques

Regression analysis has also been proposed as a measure of efficiency, with Sherman (1984) and Smith and Mayston (1987) reporting on its application to health service evaluation (see section 2.7.2). Sherman (1984) suggested that regression techniques have been 'particularly useful in understanding characteristics that impact on cost'. For example, Johnson *et al.* (1999) used ordinary least squares regression to 'determine the effect of clinical and demographic variables on individual cost of care' in an evaluation of the treatment of patients with cystic fibrosis.

Alternatively, Siddiqui *et al.* (1999) used a Poisson random-effects regression model to 'assess the effects of parental and peer approval of smoking on adolescents' current level of smoking'. They also used a number of qualifying factors in their models, such as ethnicity and gender, to find other significant influences. Regression techniques have also been used to identify the factors influencing the success of laser treatment for myopia and astigmatism (Huang *et al.*, 1999) and for 'risk stratification in cases of acute lymphoblastic leukaemia in children' (Groves *et al.*, 1999).

4.6 Summary

In the previous sections, the most frequently used methods for the measurement of efficiency and performance in the health service have been introduced and evaluated. With each of these methods, serious technical deficiencies and reservations have been identified and discussed. These include the partial nature of performance indicators, providing only a snapshot of performance not a complete picture and the difficulty of accounting for environmental influences and differences. Regression analysis and performance indicators have also tended to identify 'average' performance, rather than 'best practice'. This may encourage health service institutions to cluster around the average rather than aspire to reach new levels of efficiency (although benchmarking promotes best practice to a much greater degree). These methods also do not give a comprehensive measure of the efficiency with which services are provided, nor a single summary statistic that can easily be used for comparison.

Therefore, it seems clear that there is a potential for alternative techniques for performance and efficiency assessment to be used within the Health Service. With this in mind, the technique of data envelopment analysis is presented in great detail in chapter five before a comprehensive example of DEA in practice is discussed in the subsequent chapters.

4.7 References

1. **Allsop, J.** (1995) Health Policy and the NHS: Towards 2000, Longman.

2. **Baggott, R.** (1994) Health and Health Care in Britain, MacMillan.

3. **Baker, M.** (1998) Making Sense of the New NHS White Paper, Radcliffe Medical Press Ltd.

4. **Ball, R.** (1998) Performance Review in Local Government, Ashgate.

5. **Barrow, M. and Wagstaff, A.** (1989) Efficiency Measurement in the Public Sector: An Appraisal, *Fiscal Studies*, Vol. 10, Issue 1, pp. 72-97.

6. **Birch, S. and Maynard, A.** (1986) Performance Indicators and Performance Assessment in the UK National Health Service: Implications for Management and Planning, *International Journal of Health Planning and Management*, Vol. 1, pp. 143-156.

7. **Carter, N., Klein, R. and Day, P.** (1992) How Organisations Measure Success: the Use of Performance Indicators in Government, Routledge.

8. **Cassanova, J. E.** (1990) Status of Quality Assurance Programs in American Hospitals, *Medical Care*, Vol. 28, pp. 1104-1109.

9. **Chisholm, D., Healey, A. and Knapp, M.** (1997) QALYs and Mental Health Care, *Social Psychiatry and Psychiatric Epidemiology*, Vol. 32, No. 2, pp. 68 – 75.

10. **Clinical Research and Audit Group** (1994) The Interface between Clinical Audit and Management, The Scottish Office.

11. **Cuenin,** S. (1987) The Case of Performance Indicators in Universities, An International Survey, *International Journal of Institutional Management in HE*, Vol. 11, No. 2.

12. **Department of Health** (1989) Working for Patients, HMSO, Cmnd 555.

13. **Department of Health** (1996) Health of the Nation Update (Internet Document), Department of Health.

14. **Department of Health** (1997) The New NHS, Modern, Dependable, HMSO.

15. **Edhouse, J. A. and Wardrope, J.** (1996) Do the National Performance Tables Really Indicate the Performance of Accident and Emergency Departments, *Journal of Accident and Emergency Medicine*, Vol. 13, No. 2, pp. 123-126.

16. **Elton, L. B.** (1987) 'Warning Signs', The Times Educational Supplement, 11/09/87.

17. **Griffiths, R.** (1983) NHS Management Enquiry, Department of Health and Social Services.

18. **Groves, D. J., Smye S. W., Kinsey S. E., Richards S. M., Chessells J. M., Eden, O. B. and Bailey C. C.** (1999) A Comparison of Cox Regression and Neural Networks for Risk Stratification in Cases of Acute Lymphoblastic Leukaemia in Children, Neural Computing and Applications, Vol. 8, No. 3, pp. 257 – 264.

19. **Huang, D., Stulting, R. D., Carr J. D., Thompson K. P. and Waring G. O.** (1999) Multiple Regression and Vector Analyses of Laser In Situ Keratomileusis for Myopia and Astigmatism, *Journal of Refractive Surgery*, Vol. 15, No. 5, pp. 538 – 549.

20. **Information and Statistics Division (ISD)** (1996) Scottish Health Service Costs 1995-96, ISD, NHS in Scotland.

21. **Jackson, P. and Palmer, B.** (1992) Developing Performance Monitoring in Public Sector Organizations, Leicester, University of Leicester.

22. **Jenkins, L., Bardsley, M., Coles, J., Wickings, I. and Leow, H.** (1987) Use and Validity of NHS Performance Indicators - A National Survey, CASPE.

23. **Johnson, J.A., Connolly, M. A., Jacobs, P., Montgomery, M., Brown, N.E. and Zuberbuhler, P.** (1999) Cost of Care for Individuals with Cystic Fibrosis: A Regression Approach to Determining the Impact of Recombinant Human DNase, *Pharmacotherapy*, Vol.19, No.10, pp.1159 – 1166.

24. **Jowett, P. and Rothwell, M.** (1988) Performance Indicators in the Public Sector, MacMillan.

25. **Lowry, S.** (1988) Focus on Performance Indicators, *British Medical Journal*, Vol. 296, Issue 6627, pp. 992-994.

26. **Maynard, A.** (1991) A case for Auditing Audit, *Health Services Journal*, Vol. 18, No. 26.

27. **McCarthy, M.** (1983) Are Efficiency Measures Effective, *Health and Social Services Journal*, 15/12/83, pp. 1500-1501.

28. **Merrison, A. (Chairman)** (1979) Royal Commission on the National Health Service: Report, HMSO Cmnd 7615.

29. **National Audit Office** (1995) National Health Service (England): Clinical Audit in England, National Audit Office, Press Notice 64/95.

30. **National Audit Office** (1996) Health of the Nation: A Progress Report, National Audit Office, Press Notice 62/96.

31. **National Health Service** (1996) Health Service Indicators Information Sheet, NHSME.

32. **Packwood, T., Kerrison, S. and Buxton, M.** (1994) The Implementation of Medical Audit, *Social Policy and Administration*, Vol. 28, No. 4, pp. 299-316.

33. **Pollitt, C.** (1984) Blunt Tools: Performance Measurement in Policies for Health Care, *Omega*, Vol. 12, No. 2, pp. 131-140.

34. **Pollitt, C.** (1985) Measuring Performance - a New system for the National Health Service, *Policy and Politics*, Vol. 13, No. 1, pp. 1-15.

35. **Pollitt, C.** (1992) 'The Struggle for Quality: The Case of the NHS', Paper Presented at the Political Studies Association Annual Conference, Queens University, Belfast.

36. **Payne, B.** (1995) Clinical Audit: is it Value for Money? in Evaluating Clinical Audit: Past Lessons, Future Directions, ed. Walshe, K., The Royal Society of Medicine Press Limited.

37. **Powell, J. E.** (1966) A New Look at Medicine and Politics, Pitman Medical.

38. **Rivett, G.** (1997) From Cradle to Grave, Fifty Years of the NHS, King's Fund Publishing.

39. **Roberts, H.** (1990) Performance and Outcome Measure in the Health Service in Output and Performance Measurement in Government: The State of the Art, Cave, M., Kogan, M. and Smith, R. (Eds.), Jessica Kingsley Publishers.

40. **Robinson, M. B.** (1994) Evaluation of Medical Audit, *Journal of Epidemiology and Community Health*, Vol. 48, No. 5, pp. 435-440.

41. **Royal College of Physicians** (1989) Working Party on Medical Audit: A First Report, Royal College of Physicians.

42. **Sherman, H. D.** (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927 – 938.

43. **Siddiqui, O., Mott, J., Anderson, T. and Flay, B.** (1999) The Application of Poisson Random-effects Regression Models to the Analyses of Adolescents' Current Level of Smoking, *Preventative Medicine*, 1999, Vol.29, No.2, pp.92 – 101.

44. **Skinner, P. W., Riley, D. and Thomas, E. H.** (1988) Use and Abuse of Performance Indicators, *British Medical Journal*, Vol. 297, Issue 6658, pp. 1256-1259.

45. **Smith, P.** (1995) Performance Indicators and Outcome in the Public Sector, *Public Money and Management*, Vol. 15, Pt. 4, pp. 13-16.

46. **Smith, P. and Mayston, D.** (1987) Measuring Efficiency in the Public Sector, *Omega*, Vol. 15, No. 3, pp. 181 – 189.

47. **Times Health Supplement** (April 1993, pp. 7 – 10) Making the Social Market Work.

4.8 Additional Sources

1. Interviews with Senior Management at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust (conducted in July, 1996).

2. BBC Television Drama Series, Casualty, 30/11/96.

CHAPTER FIVE

THE DATA ENVELOPMENT ANALYSIS APPROACH TO EFFICIENCY ASSESSMENT

5.1 Introduction

As has been demonstrated in earlier chapters, the measurement of the efficiency with which health services are provided has become a central concern of health care managers, particularly over the last decade:

> "High health care costs in the United States (11.7% of GNP in 1988) and the effect of these costs on governmental budgets and private industry has stimulated cost consciousness among the purchasers of hospital care." BYRNES AND VALDMANIS (1994)

The same situation is also repeated in the United Kingdom, despite the fact that a much smaller proportion of the GDP is allocated to health care expenditure in comparison with the United States and other countries (see figure 3.9). The following quote, taken from the white paper proposing the Labour Government's plans for the new NHS, expresses the government's commitment to the assessment of performance:

> "...Performance... will be assessed against new broad-based measures reflecting the wider goals of improving health and health care outcomes, the quality and effectiveness of service, efficiency and access. Performance will be judged by greater use of comparative information." DEPARTMENT OF HEALTH (1997)

The methods used to carry out these efficiency evaluations, such as performance indicators and regression techniques, have been discussed in previous chapters. The major limitations and deficiencies of these techniques have also been identified and interpreted. The following criteria have been suggested as the requirements for any technique to be an effective method for the measurement of efficiency:

1. Derive a single summary measure of the relative efficiencies for a set of DMUs;
2. Handle noncommensurate multiple outputs and input factors;
3. Not be dependent on a set of *a priori* weights or prices;
4. Handle qualitative factors such as 'sensitivity to public needs' in addition to quantitative factors;
5. Be theory-based, transparent and reproducible;
6. Be equitable and defensible.

Adapted from LEWIN AND MOREY (1981)

The technique of data envelopment analysis (DEA) is therefore presented as an alternative, or as an addition, to be used in the measurement of efficiency, counteracting the deficiencies of the other techniques and containing 'all the desirable features' required by Lewin and Morey (1981) as listed above.

This chapter provides an introduction to the methodology of efficient frontiers, with a brief summary of the application of the DEA technique to the evaluation of health service efficiency, which is covered in much greater detail in chapter six. This is followed by an extensive definition of the DEA methodology and its application, illustrated using data generated from gynaecology inpatient services in Scotland.

The remainder of the chapter is devoted to extending the basic DEA methodology, introducing some alternative formulations of the models and other ways for analysing the data, including statistical analysis of the results from DEA models. The chapter concludes with a brief summary of the benefits of the DEA approach, prior to a large-scale application of the methodology in chapters seven and eight.

5.2 The Development of Efficient Frontiers Methodology

Throughout the latter part of this century, a great deal of emphasis has been given to the development and use of models for the measurement of efficiency and performance. The motivation for such research, particularly into the use of frontier approaches to efficiency measurement, was summarised thus:

> "The problem of measuring the productive efficiency of an industry is important to both the economic theorist and the economic policy maker. If the theoretical arguments as to the relative efficiency of different economic systems are to be subjected to empirical testing, it is essential to be able to make some actual measurements of efficiency. Equally, if economic planning is to concern itself with particular industries, it is important to know how far a given industry can be expected to increase its output by simply increasing its efficiency, without further resources." FARRELL (1957)

This key text in the development of frontier methodologies for efficiency assessment also identified the weakness of many existing techniques, in that 'they fail to combine these measurements into any satisfactory measure of efficiency' (Farrell, 1957). The approach followed by Farrell (1957), denoted as the Farrellian Efficiency Model by Wilkinson (1991), was to define efficiency as:

> ".. the ratio of the weighted sum of the outputs to the weighted sum of the inputs." WILKINSON (1991)

This is an extension to the simple ratio approach defined in figure 2.3. Therefore the efficiency of the j^{th} DMU would be calculated using the equation shown in figure 5.1 overleaf:

$$\text{Efficiency} = \frac{\sum_{r=1}^{r=s} U_r O_{rj}}{\sum_{i=1}^{i=m} V_i I_{ij}},$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),

U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,

V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,

s = number of outputs and

m = number of inputs.

Figure 5.1: Farrellian Definition of Efficiency

Feldstein (1967) acknowledged that the Farrellian efficiency model is 'theoretically more sound for efficient assessment' than the economic approaches found in most hospital studies, such as regression. This definition of efficiency has three underlying assumptions:

> 1. The production frontier is defined by the most efficient organisations - there is always at least one efficient organisation, which defines the frontier, and all inefficient organisations lie below the frontier.
> 2. There are constant returns to scale - equal proportionate increases in inputs leads to the same proportionate increases in all outputs along the efficient frontier. This implies that the production function is a ray, with a constant gradient and passing through the origin.
> 3. The production frontier is convex to the origin and has nowhere a positive slope - reducing the use of one input necessitates an increase (or no decrease) in the use of other inputs in order to maintain output levels. SMITH AND MAYSTON (1987)

A further paper by Farrell and Fieldhouse (1962) extended Farrell's original work.

However, the Farrellian approach to efficiency measurement was not widely utilised until it became the basis for Data Envelopment Analysis, as defined by Charnes, Cooper and Rhodes (1978).

5.3 Applicability of the DEA Methodology for the Evaluation of Health Services

The two major characteristics of DEA which make it particularly appropriate for public sector service evaluation have been defined as: its ability to handle multiple inputs and outputs and the fact that the production function, or input-output relationship, does not need to be known in advance (Sherman, 1984). DEA is particularly valuable in the examination of the efficiency of health services, for numerous reasons. Health service institutions, such as hospitals, general practices, health authorities and health programmes, are complex organisations. There are many contributing factors affecting efficiency and overall performance.

Taking a hospital providing acute services as an example, there are many elements to the services it provides. There are numerous categories of employees and a variety of contributing environmental and organisational influences. All of these have been shown to have an impact on the overall efficiency of service provision. The DEA model can incorporate multiple inputs and outputs, which may be measures of quality and environmental influences and can be incommensurate. The applicability of DEA to the investigation of health service efficiency is further supported by the numerous and extensive investigations carried out over the last twenty years.

The connection between DEA and public sector services was immediately apparent in the first paper, presented by Charnes, Cooper and Rhodes (1978). Early applications of DEA, such as Lewin and Morey (1981), actually concentrated on public sector services, particularly due to the fact that there was a lack of adequate alternative approaches for estimating public sector efficiency. Nunamaker (1983) and Sherman (1984) were some of the first papers to focus specifically on the application of DEA to health care efficiency measurement. Subsequently, many other aspects of the health service have been investigated using DEA.

Norman and Stoker (1991) investigated the applicability of DEA for measuring the efficiency of regional health authorities, whilst Chilingerian and Sherman (1990) focused on the performance of individual health professionals, specifically physicians.

Sherman (1984) selected the medical-surgical departments of hospitals for evaluation using DEA, whereas Hollingsworth and Parkin (1995) looked at hospitals as a whole. Attention has also been given to the related issues of the performance of pharmacies by Capettini *et al.* (1985), rural health care programmes by Huang and McLaughlin (1989) and nursing homes by Kleinsorge and Karney (1992).

The earliest studies related to the health sector in the USA. However, in recent years, DEA has been more readily applied to data from the National Health Service in the UK, by authors such as Thanassoulis *et al.* (1995) and Sczcepura *et al.* (1993), and also health services in other European countries.

Not all of the work in this area using data from US sources has readily translated to the UK environment, particularly those in the area of ownership and the differences in efficiency between for-profit and not-for-profit hospitals and nursing homes. However, the general principles adopted in the US studies, such as the importance of data quality, the selection of variables, the issues of quality and environmental influences and the interpretation of the results, are all relevant to any application of DEA to UK data.

Many researchers from across the world have presented DEA as a technique that should be invaluable in all aspects of health service efficiency measurement, clearly expressed below:

> "DEA offers researchers, policy analysts and managers an innovative methodology for examining efficiencies in complex health care organisations." OZCAN *ET AL.* (1992)

In one of the most recent papers on this subject, similar conclusions were reached:

> "This relatively new weapon in the researcher's arsenal opens up unlimited, heretofore unconsidered opportunities, for examination of production efficiencies at a great many levels and settings for health care delivery." HOLLINGSWORTH AND PARKIN (1995)

The application of DEA is no longer restricted to the public sector services, with recent investigations into the efficiency of banks, telecom companies, maintenance units, *inter alia*. Most of the illustrations given in the following sections will relate to the health sector and these alternative areas of application will only be referenced in theoretical discussions where necessary.

An extensive bibliography is given at the end of chapter six, highlighting the many examples of papers relating DEA to the investigation of health service efficiency, in addition to those mentioned in this chapter. Furthermore, as a means for introducing the DEA application in chapter seven, an extensive discussion of a small number of key papers is also included in chapter six.

5.4 The DEA Methodology

In this section, the original DEA methodology of Charnes, Cooper and Rhodes (1978) is presented, illustrated using a sample of data from the NHS in Scotland, taken from the Costs Book (ISD, 1996). The data sample relates to the provision of inpatient gynaecology services at acute hospitals in Scotland in the period 1995/96. Appendix 2A contains a list of the hospitals included in the sample and the data is presented in appendix 2B. Appendix 2C contains the output from a selection of the models used.

For each department, six factors have been identified: the four possible inputs are Total Costs (I_1), Total Allocated Costs (I_2), Total Direct Costs (I_3) and Average Number of Staffed Beds (I_4); the output factors are Total Patient Days (O_1) and Total Discharges (O_2). The factors have been selected for illustrative purposes only in this chapter and no further attention will be given at this stage to their relative merits, as this is covered in chapter seven. The hospitals included in the sample have been selected from all the acute hospitals in Scotland, according to three characteristics:

- No missing data;
- The total number of patient days is greater than 2000 and
- The total number of discharges is greater than 500.

These characteristics were applied so that missing values did not compromise the DEA calculations and the departments in the sample were of a similar size and nature. 23 hospitals fulfilled these criteria.

As referred to above, the DEA methodology was developed from the Farrellian definition of efficiency. The step proposed by Charnes, Cooper and Rhodes (1978) to develop the Farrellian model, which resulted in the development of DEA, was to allow each Decision Making Unit (DMU) to value inputs and outputs differently and thus adopt different weights. The weights adopted by each DMU would be those which show 'it in the most favourable light in comparison to the other units' (Dyson *et al.*, 1990). It is this 'free allocation of weights' that was presented as a strong argument for using this approach to efficiency assessment.

The DEA technique then 'builds an efficient production frontier allowing DMUs to be relatively efficient even though they might have different mixes of inputs and outputs' (Ehreth, 1994). The original DEA model is referred to as the CCR model and is named after the authors who developed the technique (Charnes, Cooper and Rhodes).

The model is able to produce an objective evaluation of overall efficiency, identify the sources of any inefficiencies and also estimate the amount of these inefficiencies identified (Charnes *et al.*, 1994). The DEA methodology 'compares a set of organisation's actual inputs used to produce their actual output levels during a common time period' (Sherman, 1984).

It identifies the best practice (or efficient) units in the observation set, those which produce the efficiency frontier and locates the relatively inefficient units by comparison with the best practice units.

Figure 5.2 illustrates the idea of a 'efficiency frontier' for a single input-single output data set (denoted by G1), using Total Costs (I_1) as the input and Total Patient Days (O_1) as the output. DMU #15 forms the efficiency frontier for this sample and is the only efficient DMU.

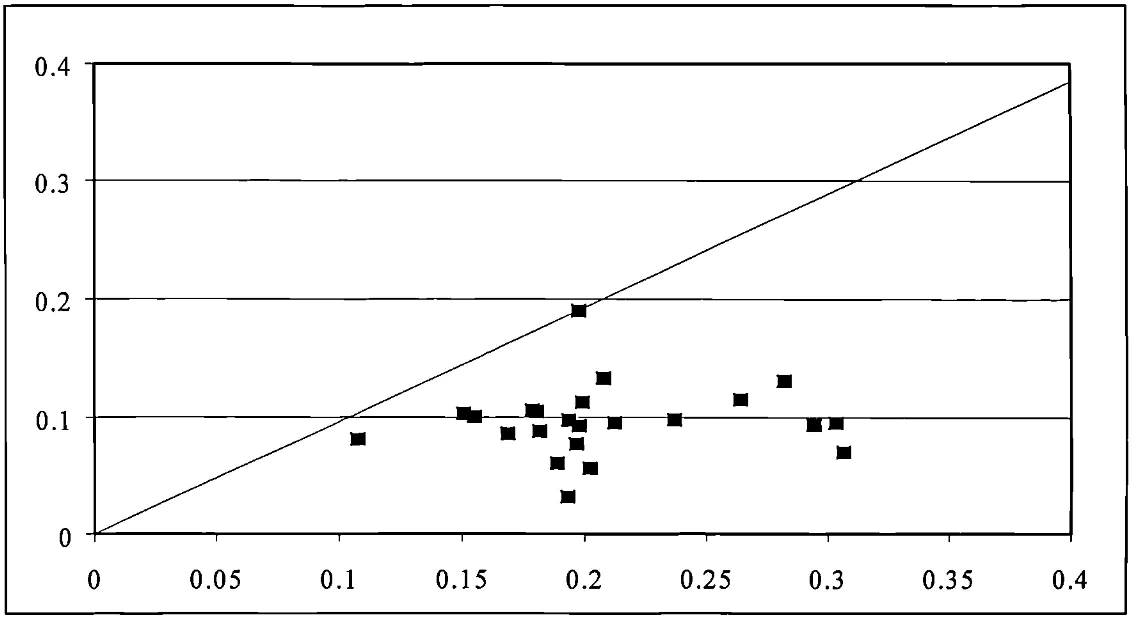


Figure 5.2: Locating the Efficiency Frontier for Model G1

In DEA, the CCR model can be orientated in one of two distinct ways, either output maximisation or input minimisation, depending on the motivation behind the research. With an input orientation, the focus is on 'maximal movement toward the frontier through proportional reduction of inputs'. For an output orientation, the movement towards the efficiency frontier is through the 'proportional augmentation of outputs' (Charnes *et al.*, 1994).

According to the orientation selected, the units labelled as 100% relatively efficient under DEA are done so if:

(a) None of its outputs can be increased without either
- Increasing one or more of its inputs or
- Decreasing some of its other outputs.

(b) None of its inputs can be decreased without either
- Decreasing some of its output or
- Increasing some of its other inputs.

CHARNES AND COOPER (1985)

The identification of units as 100% efficient, however, does not imply that they are absolutely efficient, such that they cannot improve their technical efficiency at all:

> "Units with an efficiency ratio of 1 (E = 1) are not necessarily absolutely efficient but rather represent the 'best practice' group of units, which means that they are not clearly inefficient compared with the other units in the set. This situation arises because the identity of the absolutely efficient units is not known because of lack of knowledge of the efficient input-output relationships. Hence a unit that is found to be relatively efficient may also be able to improve its operating efficiency." SHERMAN (1984)

The nature of the inefficiencies within each DMU classified as inefficient can also be determined. Each will have 'true inefficiencies at least as large as the amount located with DEA' and will theoretically have 'the ability to produce its same level of outputs with fewer inputs based on the actual output-input levels of *units* that were compared with the inefficient unit' (Sherman, 1984). This is due to the 'conservative' nature of the efficiency calculations within DEA.

There are two basic assumptions underlying the efficiency calculations in the original DEA model:

1. All points along the efficiency frontier are practically attainable production possibilities;
2. There are constant returns to scale, or linearity. EHRETH (1994)

The first assumption holds true for all DEA models because the efficiency frontier is identified from the 'best-practice' of all the DMUs in the sample and is only an approximation to the true efficiency frontier. The second assumption is more contentious. The original DEA model assumed that the production relationship is defined by constant returns to scale (CRS), that is, an increase in one unit of input produces an equivalent increase in each output, assuming all other things to be equal.

Smith and Mayston (1987) suggested that, if the sample selected for analysis contained homogeneous units, whilst the size and scale of activities between the DMUs may vary, the variations are not large enough to make the 'constant returns to scale' assumption unduly restrictive. For the time being, the validity of this assumption relating to constant returns to scale will be accepted, but will be addressed in greater detail later in the chapter and in the subsequent discussions.

The mathematical formulation and solution of the original DEA model is now described, using the same notation as was used in the definition of Farrellian efficiency in figure 5.1 above.

5.4.1 The Ratio Form of the DEA Model

The DEA model was first formulated in the form of a ratio, which was then transformed into a linear programming model, to be presented in the next section. The ratio forms of the DEA methodology had 'a strong intuitive appeal, since they extend the engineering ratio approach for efficiency measures' (Charnes *et al.*, 1994). The following figure gives the ratio form of the basic DEA model, with an output orientation, and can be compared with the Farrellian definition of efficiency in figure 5.1 and the simple ratio definition in figure 2.3.

$$\text{MAXIMISE } e_\alpha = \frac{\sum_{r=1}^{r=s} U_r O_{r\alpha}}{\sum_{i=1}^{i=m} V_i I_{i\alpha}},$$

$$\text{SUBJECT TO: } \frac{\sum_{r=1}^{r=s} U_r O_{rj}}{\sum_{i=1}^{i=m} V_i I_{ij}} \leq 1, j = 1 \ldots \alpha \ldots n$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),
U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,
V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,
s = number of outputs and
m = number of inputs.

Figure 5.3: The Ratio Form of the DEA Efficiency Equation

Essentially each DMU is allowed to vary the weighting attached to each input and output in order to maximise its efficiency, subject to the constraint that no other DMU should consequently obtain an efficiency score greater than 1. For a full complement of efficiency ratings, the above model has to be solved for each of the n DMUs.

Taking the single input, single output situation illustrated in figure 5.2, the efficiency score for DMU #1, for example, would be calculated by solving the set of equations in figure 5.4 below and, by this method, its relative efficiency within the small sample is determined to be 66.45%. The data can be found in appendix 2B.

$$\text{MAXIMISE: } \frac{17236U_1}{4905V_1}$$

$$\text{SUBJECT TO: } \quad \frac{17236U_1}{4905V_1} \leq 1 \qquad (1)$$

$$\frac{5547U_1}{1857V_1} \leq 1 \qquad (2)$$

$$\frac{8495U_1}{2330V_1} \leq 1 \qquad (3)$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\frac{5599U_1}{1725V_1} \leq 1 \qquad (22)$$

$$\frac{4559U_1}{1163V_2} \leq 1 \qquad (23)$$

Where: U_1 and $V_1 \geq \varepsilon$, with ε being a very small positive number;
U_1 is the weight attached to the output for DMU #1and
V_1 is the weight attached to the input for DMU #1.

Figure 5.4: The Efficiency of a DMU #1

The solution of the above model is fairly straightforward in this simple case, which has just small number of DMUs, one input and one output. As more variables are introduced into the DEA model, the solution process becomes much more complex. Therefore, the ratio form of the model has subsequently been translated into a straightforward linear programming form. This has simplified the solution process for the more complicated multi-input, multi-output problems.

5.4.2 Formulation of the DEA Model as a Linear Programme

In order to solve the DEA model, the fractional linear programme (or ratio form) has been transformed into a more straightforward linear programme, with various approaches applied to complete this procedure. (See, for example, Charnes and Cooper, 1962.) This form of the DEA model is generally denoted by the acronym CCR, in reference to Charnes, Cooper and Rhodes (1978) where the DEA methodology was first presented in this form, and this notation will be used subsequently.

Linear programmes take on one of two forms, either primal or dual and within the DEA methodology, the models can have either an output or input orientation. The input-orientation models focus on 'maximal movement toward the frontier through proportional reduction of inputs' and the output-orientation models are distinguished by the 'maximal movement via proportional augmentation of outputs' (Charnes *et al.*, 1994). The orientation of the model does not impact on the efficiency score allocated to each DMU, but on the way improvements in efficiency can be achieved.

There are four different linear programming forms of the basic DEA model, these being the input-oriented CCR primal, the input-oriented CCR dual, output-oriented CCR primal and output-oriented CCR dual. As noted by MacMillan (1987), in order to solve the DEA model, it has generally more appropriate to translate the primal form of the linear programme into its equivalent dual form, as the 'problem will have more constraints than variables'. This has generally been done using some form of specialist software.

Therefore, the efficiency scores for each DMU can be generated by the solution of either of the formulations below adapted from Charnes *et al.* (1994), where figure 5.5 represents an input-oriented model and figure 5.6 an output-oriented model.

$$\text{MINIMISE: } e_\alpha = \sum_{i=1}^{m} V_i I_{i\alpha} \,,$$

$$\text{SUBJECT TO: } -\sum_{r=1}^{s} U_r O_{rj} + \sum_{i=1}^{m} V_i I_{ij} \geq 0, \; j = 1....\alpha....n$$

$$\text{AND: } \sum_{r=1}^{s} U_r O_{r\alpha} = 1$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),
U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,
V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,
s = number of outputs and
m = number of inputs.

Figure 5.5: The Input-Oriented CCR Model (Dual) (Output Maximisation)

$$\text{MAXIMISE: } e_\alpha = \sum_{r=1}^{s} U_r O_{r\alpha} \,,$$

$$\text{SUBJECT TO: } \sum_{r=1}^{s} U_r O_{rj} - \sum_{i=1}^{m} V_i I_{i\alpha} \leq 0, \; j = 1....\alpha....n$$

$$\text{AND: } \sum_{i=1}^{m} V_i I_{i\alpha} = 1$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),
U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,
V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,
s = number of outputs and
m = number of inputs.

Figure 5.6: The Output-Oriented CCR Model (Dual) (Input Minimisation)

The linear programming forms, particularly the dual form, are more easily solved using computer-based linear programming packages or specially written DEA software, many of which have now been developed. (See Charnes *et al.*, (1994) for an extensive guide to many of the software packages available.) For the purposes of this investigation, however, software written by Dr. Richard Thomas at the University of Stirling has been utilised to carry out the analyses of the data presented. Using the simple model discussed previously (G1) the DEA model, solved as a linear programme using the computer software package, produces the following results:

Ranking	DMU	Efficiency Score	Ranking	DMU	Efficiency Score
1	#15	100	13	#6	65.0
2	#11	84.0	14	#22	61.4
3	#14	75.7	15	#8	60.6
4	#19	74.4	16	#2	56.5
5	#23	74.1	17	#21	55.5
6	#13	73.9	18	#20	50.2
7	#12	73.1	19	#16	49.8
8	#17	70.0	20	#4	48.7
9	#3	68.9	21	#10	48.7
10	#1	66.4	22	#18	47.4
11	#9	66.3	23	#7	45.8
12	#5	65.0			

Table 5.1: The Results from Model G1

The results from the model confirm the graphical observation made in figure 5.2, such that there is just one efficient unit in this simple model, that is, DMU #15. The spread of the efficiency scores is quite large, corresponding to the fact that the data points in figure 5.2 are widely distributed away from the frontier. With such a simple model, it is difficult to perform a much more detailed analysis. However, as the complexity of the technique is extended in subsequent sections, further information is required relating to the DEA model, to be discussed in the following sections.

5.4.3 Interpretation of the Factor Weights

In the DEA methodology, as was stated above, each DMU is allocated a set of weights that will maximise its efficiency score. Interpretation of the weights in the case of the simple model defined above (model G1 with one input and one output) is relatively straightforward: 'they can be interpreted as the amount of resources used to generate a unit of output' (Belton and Vickers, 1993). As the DEA model becomes multifaceted (discussed in the following sections), this situation becomes more complex: 'when there are multiple inputs and outputs then there is no intuitively appealing interpretation' (Belton and Vickers, 1993).

It is possible that the efficiency score for every DMU in the sample will be based on an entirely different set of weights (Boussofione *et al.,* 1991). In all cases, a multiple of the set of factor weights for any given DMU also produces an optimal solution, that is, the maximum efficiency score for that DMU (Boussofione *et al.,* 1991). Additionally, for each DMU, it may be possible to identify an alternative set of weights, which will also produce an optimal solution for that DMU, such that it will be classified with the same efficiency score as previously. In the case of the efficient DMUs, there are 'at least two optimal solutions and possibly many more, each solution indicating a potentially different set of weights' (Belton and Vickers, 1993). (This is emphasised in section 5.4.6, which introduces the reference set, and in section 5.8.5, which discusses cross-efficiency.)

In chapter eight, the implications for alternative patterns of the factor weights are discussed, through the introduction of weight restrictions, as defined in section 5.7.

5.4.4 Virtual Inputs and Outputs

For a clearer understanding of how the efficiency score for each DMU is derived and the contribution of each variable in the calculation of the efficiency scores, the concepts of virtual inputs and virtual outputs are introduced:

> "For a unit, the virtual output (input) attributable to a given output (input) is the product of that output (input) and its corresponding weight." THANASSOULIS *ET AL.* (1987)

In this investigation, virtual inputs and outputs are given in percentages and both total one hundred. This is slightly different to the approach defined in Thanassoulis *et al.* (1987), where the virtual input percentages summed to 100 and the virtual output percentages summed to the efficiency score for each DMU. The virtual inputs and outputs can be examined to determine the significant factors for each DMU, in terms of explaining its efficiency classification, and to identify good practice and different operating procedures. In some instances, the values of the virtual inputs and outputs have also been referred to as the virtual weights.

In terms of the simple model already presented, the concept of virtual inputs and outputs is of little relevance. However, if the simple case is extended to incorporate two inputs by dividing the 'total costs' into 'total allocated costs' and 'total direct costs', the solution of a new DEA model can be obtained (denoted by G2). (A full definition of what these different cost factors incorporate is given in chapter seven.) In the new model (input-oriented CCR), there are two efficient units (#11 and #15) and the average efficiency score is 69.14%.

Table 5.2 below shows the virtual inputs for a selection of the DMUs alongside the efficiency scores that were calculated from the application of the basic DEA model. A_1 is the virtual input corresponding to the first input 'total direct costs' and A_2 corresponds to 'total allocated costs'.

The virtual weights shown have been obtained from one application of the DEA methodology. As was discussed in the previous section, the weights shown for each of the DMUs are not necessarily unique – it is possible that the DMUs could achieve an equivalent efficiency score using a different allocation of the weights.

DMU Number	Efficiency Score	Virtual Inputs A_1	A_2
#1	68.74	49.4	50.6
#3	75.14	59.4	40.6
#8	63.08	50.5	49.5
#10	54.24	100	0
#11	100.0	77.7	22.3
#14	84.50	64.2	35.8
#15	100.0	100	0
#17	74.31	54.4	45.6
#19	75.44	45.8	54.2
#21	56.77	47.5	52.5

Table 5.2: Virtual Inputs for Model G2

As can be seen clearly from the small selection of DMUs given above, the efficiency scores for each DMU are not all calculated by attaching the same weight, and therefore importance, to each of the inputs. For example, one of the 100% efficient DMUs (#15) bases its efficiency classification on the contribution of just one of the inputs, 'total direct costs', as does DMU #10, which is not classified as 100% efficient. In the majority of cases, the virtual input for 'total direct costs' contributes most to the efficiency scores, that is, for 18 of the 23 DMUs.

In this example, as there is just one output, there is little value in analysing the virtual outputs – they would be designated to be 100% for each DMU. As is demonstrated in later applications of the DEA model that use multi-input, multi-output formulations, analysis of the virtual outputs can be undertaken in an identical manner.

This analysis of the virtual weights demonstrates one of the central facets of the DEA methodology, such that, the DMUs are assessed in their best possible light and the weights selected are done so in order to maximise the efficiency score. In this regard, this will generally always result in a range of virtual weight percentages across all the DMUs.

5.4.5 The Output from a DEA Model

As has already been shown, the DEA model gives each DMU an efficiency score and the DMUs can subsequently be ranked in order of their efficiency. However, the ranking of DMUs according to their efficiency score is just one of the results of the modelling process. Thanassoulis *et al.* (1987) identified some of the ways in which the output from a DEA model may be utilised by each of the DMUs investigated:

- To obtain an initial sorting of units into relatively efficient and inefficient units;
- To identify aspects of performance in relatively efficient units which may be useful in developing good operating practices;
- To detect comparable efficient units for each of the inefficient units and to set targets for improved inefficiency;
- To determine the areas of performance for the relatively efficient units that warrant additional investigation and show the potential for further improving their efficiency.

Lewin and Morey (1981) determined that, of all the possible uses of the DEA methodology, two of the most important are: 'its synthesis of a single summary measure of relative efficiency for each unit and an indication of the levels of improvement needed before an inefficient unit would be rated as efficient.'

The DEA methodology (using an output-oriented approach in this case) can be used to determine the improvements that are needed to bring about efficiency, for each of the inefficient units. To investigate this, a third model based on the gynaecology inpatient data is utilised (denoted by G3), with the same two inputs as used in model G2 and 'total inpatient discharges' as the single output.

There are three DMUs classified as efficient units (#11, #12 and #19) and the average efficiency score is 72.33%. Looking at the inefficient DMUs, the lowest efficiency score is 44.7% (DMU #7), which would need to more than double its output (2290 -> 5128 for 'total inpatient discharges') in order to achieve efficiency. DMU #10 has a relative efficiency score of 74.25% and could become efficient by increasing its 'total inpatient discharges' by approximately 600. Appendix 2C contains a detailed analysis of the results for model G3, illustrating the type of output produced by the models applied in chapter five.

The most obvious use of the DEA approach, as demonstrated above, is the allocation of efficiency scores and the separation of units into efficient and inefficient groupings. However, the analysis can also be used for a variety of other purposes, some of which are suggested overleaf:

1. To identify elements or characteristics of efficiency;
2. To seek out the inefficient operating units;
3. To group together units with similar operating procedures and efficiency structures;
4. To determine the effect of a major change in operating practices and procedures, or organisational structure;
5. To assess the impact of management on efficiency;
6. To determine the importance of a particular individual or group of individuals in organisational efficiency;
7. To evaluate the impact of external factors *e.g.* environment;
8. To examine the link between efficiency and other factors, such as effectiveness;
9. To make observations on the overall efficiency of a group of hospitals.

Some of the above approaches are utilised at some stage during the analysis in the following chapters, to fully illustrate as many of the various facets of the DEA methodology as possible, particularly points one, two and three. However, due to the nature of the data available and the focus of this analysis, it is not possible to comprehensively address each of them.

5.4.6 The Reference Set

For each inefficient unit, a reference set or peer group is identified, corresponding to the third point made above, such that 'each peer unit is efficient with the inefficient unit's weights' (Dyson *et al.*, 1990). The reference set describes a group of DMUs that have a similar focus in terms of their inputs and outputs. This can help to highlight the specific inadequacies in the performance of the inefficient units, contrasting with the operating practices of the efficient DMUs.

The peer group identified varies according to the DEA model formulation selected, as is further investigated in chapter eight. The DMUs in the reference group are weighted to form a composite unit, which lies on the efficiency frontier closest to the inefficient DMU, with usually one of the units in the reference group dominating.

The notion of the reference set can be used to provide insight into the efficient units, by examining the number of times an efficient DMU appears in the reference sets for the inefficient DMUs, investigated extensively by Doyle and Green (1995) and called the reference set count. DMUs that appear in the reference set for a large number of the other DMUs can be regarded as 'genuinely efficient' in comparison with the DMUs for which it forms the efficiency frontier (Smith and Mayston, 1987). However, an efficient DMU that only appears in a small number of the reference sets is 'unlikely to offer truly efficient performance' (Boussofiane *et al.*, 1991).

Referring again to model G3, the basic CCR model identified three efficient DMUs (#11, #12 and #19) and provides interesting insights into the data. #19 is used as a reference by two other DMUs, #11 and #12 are used by just one unit, whilst #12 and #19 form the reference set for the remaining 17 DMUs. Using the reference set count, #19 can be regarded as the 'truly' efficient DMU in this example because it appears in the reference set for all but one of the inefficient DMUs.

The reference groups can be used to provide targets for the inefficient DMUs. The results from model G3 reveal how much more output the composite units, made up from the reference group DMUs, have produced with equal amounts of input.

DMU #2, for example, has an efficiency score of 70.78% and two DMUs form its reference set, that is, #12 and #19. The composite output level is 2759 inpatients discharges compared to the 1953 discharges completed by DMU #2.

5.5 Extending the DEA Methodology

The simple cases presented above, using the original CCR model defined in figure 5.5, demonstrate the basic ideas of the DEA approach to efficiency measurement. However, there are many other facets of the technique that can be incorporated in order to represent adequately a health service environment. Clearly, a single output, single input model, or the extended double input model, do not accurately define all the activities of a gynaecology inpatient department, or any other areas of health service activity. The linear programming form of the CCR model presented in figures 5.5 and 5.6 are designed to handle much more complicated scenarios with multiple inputs and outputs, the accommodation of which is now discussed.

5.5.1 Variable Selection in DEA Models

In the original example (model G1), only two variables were used in the DEA model, these being 'total costs' and 'total number of patient days'. In this instance, only one of twenty-three gynaecology departments was rated as efficient (see Table 5.1). Subsequent models (G2 and G3) have introduced alternative inputs and outputs. The inclusion of 'average number of staffed beds' as another input is also a possibility. The impact on the efficiency scores of using several different variable mixes is shown in Table 5.3 overleaf. Results have been calculated in each case using the CCR model show in figure 5.5, which has constant returns to scale and an input orientation.

Model	Inputs	Outputs	No. of Efficient Units	Efficient Units	Average Efficiency Score	Minimum Efficiency Score	Least Efficient Unit
G1	I_1	O_1	1	#15	64.41	45.8	7
G2	I_2, I_3	O_1	2	#11, #15	69.14	48.9	7
G3	I_2, I_3	O_2	3.	#11, #12, #19	72.33	44.7	7
G4	I_1	O_2	1	#19	66.5	41.7	7
G5	$I_2, I_3, I_4,$	O_1	4	#11, #12, #15, #17	79.74	62.4	16
G6	$I_2, I_3, I_4,$	O_2	4	#10, #11, #12, #19	81.11	63.2	1

KEY: I_1: 'Total Costs' I_2: 'Total Allocated Costs'

I_3: 'Total Direct Costs' I_4: 'Average Staffed Beds'

O_1: 'Total Patient Days' O_2: 'Total Inpatient Discharges'

Table 5.3: Impact of Variable Changes on Efficiency Scores

In each of the six cases, with a different mix or number of variables, the efficiency scores calculated using the simplest form of the DEA model do not remain constant. In particular, unit #11 is efficient in all cases where the cost function is split into 'total allocated costs' and 'total direct costs', but not when a single 'total costs' input is utilised. Also, the average efficiency score is seen to increase with the number of variables used, as is the number of units classified as efficient. Further examination of units #7, #15 and #19 illustrates the variations in greater detail:

	Unit #7		Unit #15		Unit #19	
Model No.	**Efficiency Score**	**Ranking**	**Efficiency Score**	**Ranking**	**Efficiency Score**	Ranking
G1	45.8	23	100	1	74.4	4
G2	41.7	23	73.5	6	100	1
G3	48.9	23	100	1	75.4	7
G4	44.7	23	76.8	7	100	1
G5	89.3	6	100	1	77.5	16
G6	78.7	11	76.8	14	100	1

Table 5.4: Variation in Efficiency Scores

The results obtained clearly have some connection with the number and nature of the variables selected, particularly in relation to the efficiency score for each DMU and the number of efficient units. The above examples demonstrate the effect of up to four variables - there could actually be many more variables used to describe the gynaecology departments. For example, the 'total direct costs' input could be broken down into several different factors, covering the costs for medical, nursing, pharmacy, professions allied to medicine (PAMs), theatre, laboratory and other direct costs individually or in other groupings. Appendix 1 contains definitions for all key terms.

If the data were available, the outputs could be expressed in greater detail, by splitting up the 'total inpatient discharges' into several distinct categories. This can be illustrated using Model G7, which has five inputs ('total medical costs', 'total nursing costs', 'total theatre costs', 'total other direct costs' and 'total allocated costs') and 'total inpatient discharges' as the single output. The number of efficient units increases to nine and the average efficiency score is 89.0%. The additional data is included in appendix 2B.

The above results are not unexpected - the DEA methodology means that the inclusion of a greater number of variables gives each DMU an increased opportunity to be placed on the efficiency frontier. Whilst changes to the efficiency score for each DMU may be the expected result of including new variables, or changing existing ones, it would be hoped that the ranking for each DMU would remain constant, or vary only slightly.

Therefore, the problems to be addressed now are how to determine exactly how many variables should be included, what these variables should be and what methods can be used to identify them.

The DEA literature is fairly comprehensive on the importance of selecting the 'right' inputs and outputs, both in terms of number and type, to accurately describe a given health care situation. Examples of this can be found in Sherman (1984), Thanassoulis *et al.* (1987) and Beasley (1990). The importance of consultation is stressed, and many techniques, some of which are statistically based, have been suggested. Chapters six and seven present a comprehensive debate on this issue in relation to health care applications, with the general principle debated here.

The difficulty in variable selection comes in finding the balance between using all the variables necessary to describe the situation and having so many variables that the DEA model becomes less discriminating. This can lead to a majority of the DMUs being labelled as '100% efficient', which is difficult to justify. It implies either that the analysis has been a waste of time as most units are efficient anyway, or that the technique is flawed, as it is improbable that the majority of units are in fact efficient, especially if the data relates to the health service.

The definition of an appropriate set of inputs and outputs, which are the measures of performance, is a prerequisite for a successful application of DEA, as it is for any other technique (as was discussed in chapter four in relation to the data used in performance indicators).

If the correct inputs and outputs are not defined and used in the DEA study, the assessment will have little benefit as a means of performance and efficiency measurement - if only some of the relevant inputs/outputs are used then the DEA results will be less than comprehensive.

> "The inputs and outputs used in DEA define the basis to be used for assessing the units concerned, and so they must be determined with great care. This phase would normally involve wide consultation of those being assessed, to determine what they see as constituting the outputs of their function and what environmental factors and resources (inputs) affect those outputs"
>
> THANASSOULIS *ET AL.* (1987)

Sherman (1984), in a study of the medical-surgical areas of a set of teaching hospitals, also recognised the importance of understanding the processes of the units and the need for consultation in the selection of variables. The inputs and outputs utilised were selected in co-operation with a panel of hospital experts and their identification was based upon an understanding of what resources were used to provide the different services offered.

According to Wilkinson (1991), inputs and outputs should be included if and only if they are relevant to the performance assessment and excluded if and only if they are not. The effects of excluding relevant factors or including those that are not relevant are equally problematic. If a relevant input or output variable is excluded, the relative efficiency rating will be calculated without reference to that relevant factor and will be distorted.

Equally, if an input or output variable is included that is not relevant, the relative efficiency rating will also be distorted. A unit's entire efficiency score could depend on a factor that should not have been included and is totally irrelevant.

For example, using 'average number of staffed beds' in an investigation of inpatient services is a valid option. However, to include this variable in an examination of outpatient activity would be less appropriate - the number of beds available in an inpatient ward is not directly connected to the delivery of outpatient care.

The scale of the dilemma of variable selection can be illustrated by the example of DMU #7 from the gynaecology department example, as highlighted in table 5.4. If 'average number of staffed beds' is included along with 'total allocated costs' and 'total direct costs' as the inputs in the DEA model (denoted as model G5), DMU #7 scores an efficiency rating of 89.3% and is ranked 6th. Without this new input, DMU #7 is ranked 23rd in the list of DMUs, with an efficiency score of 48.9%, using model G3. The efficiency score changes dramatically upon the inclusion of an additional input factor.

The approach used by Beasley (1990) to select the variables in a study of university departments was to begin the process of variable selection by considering, conceptually, what are the inputs and outputs of the units. For example, in a university, the conceptual inputs would be the financial costs of the provision of equipment and academic staff and the conceptual output could be increased knowledge. In a health care situation, the conceptual inputs could be the people and resources employed. The conceptual output would be increased health and well being.

Actual inputs and outputs are then determined from the conceptual inputs and outputs by considering the data that is available to measure them.

There are several other approaches to the selection of variables than that used by Beasley (1990). A frequently used technique is to start with a large number of variables and reduce the list by eliminating those that are not relevant, appropriate or necessary. This can be done through a variety of methods. Statistical analysis, such as identifying pairwise correlations amongst the inputs or outputs, has been applied. Other suggestions include consultation with experts and heuristic methods, which involves examining the effects of removing particular inputs and outputs in turn on the efficiency score for each unit.

However, this issue is highly subjective – if a pair of outputs are found to be highly correlated, does this automatically mean that one of them should be excluded? If one is to be removed from the variable set, what criteria should be employed to determine which to exclude?

Alternatively, the DEA model can be run with a variety of inputs and outputs, and the different efficiency scores obtained with each variable set investigated and the most influential factors identified. This approach was used by Ehreth (1994), in an investigation of hospital performance. Nunamaker (1985) considered the implications of changes to the variable set on the efficiency scores of units that had previously been both efficient and inefficient. He stressed the importance of establishing the variables to be used prior to beginning study, using statistical methods and expert judgement.

Nunamaker (1985) found adding in new variables, or disaggregating existing ones, to impact on the efficiency scores of the inefficient units in particular. It was also suggested that an efficient DMU could not become inefficient if one of the factors was broken down into its component parts, such as dividing the 'total costs' function into 'total allocated costs' and 'total direct costs'.

However, a dilemma exists in that, while the inclusion of all relevant factors is important, there is the suggestion that the number of factors is also important.

> "In principle, all inputs and outputs relevant to the function of the units should be included. However, the larger the number of inputs and outputs in relation to the number of units being assessed, the less discriminatory the method appears to be.... Thus the number of inputs and outputs included in a DEA assessment should be as small as possible, subject to their reflecting adequately the function performed by the units being assessed."
>
> THANASSOULIS *ET AL.* (1987)

This point has also been emphasised by other studies, as below:

> "The larger the number of inputs investigated or outputs included, the more difficult it is for DEA to determine which facilities are technically efficient and which are inefficient. In the extreme.... all centres would end up on the best practice frontier."
>
> FINKLER AND WIRTSCHAFTER (1993)

The problem identified within the methodology is that the larger the set of variables included, the greater the chance that a unit will find a set of weights for the inputs and outputs that result in being identified as efficient:

> "With typical multi-input, multi-output problems applied to a dozen or so DMUs, the foundation CCR model is likely to find more of them efficient than not." LEWIN AND MOREY (1981)

Hollingsworth and Parkin (1996) concurred, pointing out that a large variable set could provide additional information about the performance of the inefficient DMUs.

Banker *et al.* (1989) proposed that the number of variables to be included should be approximated by the following 'rule of thumb': the number of DMUs in the sample should be greater than or equal to the number of variables, multiplied by three. The number of DMUs should be much greater than the product of the number of inputs and outputs, so that DEA can be of 'discriminatory value' (Boussofione *et al.*, 1991).

However, these approaches to variable selection that rely on statistical interpretation and pre-determined (and arbitrary) rules-of-thumb, fail to take into account the particular situations under investigation, treating one sample of DMUs exactly like any other. Ball and Roberts (1998) considered this to be contrary to a successful application of DEA, such that the situation under investigation should determine the methodological steps in the DEA model development and not pre-conceived statistical ideas. Their analysis focused on a data sample with 12 DMUs and 9 variables, all included following discussion with policy-makers from the situation being examined.

It may also be important to consider data availability and accuracy. There is little point in including a variable if half the data is missing or inaccurate. Huge errors may affect efficiency scores, as DEA is sensitive to data extremes.

In summary, from an application of DEA to schools by Chalos and Cherian (1995), the following criteria for the selection of variables is proposed:

> 1. Empirical precedent exists - previous studies have used the variables identified.
> 2. Expert opinion concurs with the variables chosen as appropriate measures of efficiency.
> 3. Statistical testing shows significant relationships between inputs and outputs.
> 4. The data for these variables has no missing, or suspect, values.
>
> CHALOS AND CHERIAN (1995)

5.5.2 The Nature of Variables

The actual variables chosen for inclusion in a DEA model will clearly depend upon the nature of the units under investigation. However, for the examination of efficiency in a health service setting, the variables chosen as inputs and outputs generally tend to fall into five major categories, as follows.

5.5.2.1 Manpower levels (Inputs)

The number of staff employed in the health service environment under investigation has almost always been included as an input to the DEA model in some form, such as, the number of full-time equivalent (FTE) staff employed, the number of hours worked or the direct salary expenses. These can be divided into categories for staff type or calculated as totals, depending upon circumstances. Hogan *et al.* (1987) used eight categories of labour to measure manpower level for 300 hospitals, whilst Nyman *et al.* (1990) used the number of hours worked to measure manpower and divided this into nursing, social worker, therapist and other worker hours for a study of nursing homes.

5.5.2.2 Resources (Inputs)

Resources can include measures of capacity and availability, patient expenses, supply and building expenses and administrative costs. A frequently used measure of capacity has been 'bed days available', as used by Bowlin *et al.* (1985) for an efficiency study of hospitals in Massachusetts. Sherman (1984) also used this in an examination of the medical services departments of 7 teaching hospitals. However, some studies have used hospital size as a surrogate for some measure of potential availability of beds.

'Routine inpatient expenses' was the factor used by Nunamaker (1985) as the only input for one of his studies. It was identified as a useful indicator of capacity, as it indirectly gave a measure of the number of patients seen. Grosskopf and Valdmanis (1987) simply used the number of admissions into the hospitals as the input measure for capacity.

Supply expense and equipment expense are found in many studies, often combined into an 'other expenses' category, along with staff and patient expenses, for example, by Bitran and Valor-Sabatier (1987).

Examples of other factors used are: Banker *et al.* (1986), who used administrative costs as an input; Hogan *et al.* (1987), where 'capital' and 'energy expenses' were incorporated and Morey *et al.* (1985), who identified 'overhead costs' and the 'value of the inventory' as important inputs when examining pharmacies. Each of these could be aggregated to form an 'allocated costs' category, as included in some of the models in section 5.4 above.

Manpower and resource variables have also been combined to form inputs, for example, in the two cost factors used as inputs in the examples used earlier in section 5.4, 'total direct costs' and 'total allocated costs'.

Also, a factor such as 'average staffed beds' has two elements, one of which is a resource input (the physical existence of a bed) and the other is a manpower input (the availability of the relevant personnel to provide the 'staff'). Such factors take account of the financial aspects of providing health services but do not necessarily represent manpower and resources as distinct factors.

5.5.2.3 Environmental factors (Inputs)

Environmental factors refer to local demographic information, such as, population size and catchment area, or any factor beyond the control of the units. They have not necessarily been factors specifically describing the environment in which a DMU operates - they can be any type of descriptive variable, used to explain any external factors influencing the service operation.

The type of environmental factors used, if any, depends heavily on the context of the research. Most surveys of health services efficiency have not incorporated any environmental factors to date, although they have been more widely used in other areas of DEA application. For example, in a retail environment, the 'number of competitors in the locality' and the 'size of the resident population' could be important factors affecting efficiency and included as inputs to the DEA model.

In an evaluation of local authorities, Smith and Mayston (1987) suggested that the inclusion of environmental factors is critical, as they are crucial determinants of the nature and scale of services needed in each authority. They listed social and economic considerations, topography, climate and demographics as useful indicators of environmental influences, to be included in the DEA model if they can be measured.

Using environmental factors in a DEA model must be undertaken with some amount of caution, however. They may have to be re-scaled, for example inverted, to ensure they are positively correlated with the output factors, that is, output increases with an increase of the environmental factor. A problem can occur if the environmental factors are positively correlated with some outputs but negatively correlated with others.

However, environmental factors have been utilised in some health studies, such as the use of 'service area population' and 'the percentage of users under four years' as inputs by Huang and McLaughlin (1989) in a survey of rural health programmes. 'Service area population' was also used by Morey *et al.* (1985) in an investigation of pharmacies.

Those environmental factors that are exogenously fixed can only be included in a DEA investigation following a small adjustment to the DEA model and are discussed in greater detail in section 5.8.3. They are also generally referred to as non-discretionary or uncontrollable variables, as they cannot be changed at the discretion of management, as many other variables can.

5.5.2.4 Activity levels (Outputs)

A frequently used output measure is that of 'patient days', as noted by Ehreth (1994), who found that it was used in all but one study in the DEA literature as a measure of output. Depending on the focus of the study, this output is often split by type of treatment or by age structure. For example, Banker *et al.* (1986) used geriatric, adult and paediatric inpatient days as the outputs for a survey of overall hospital efficiency in North Carolina. Valdmanis (1990) split patient days into 'acute care days' and 'intensive care unit days' for a similar study of Michigan hospitals. Another well-used measure is the number of discharges, again often divided up by type of treatment (diagnostic group) and age structure, as used by Bitran and Valor-Sabatier (1987), who used the discharges in fifteen major diagnostic categories to examine the medical services departments of hospitals.

Whilst many studies have used 'patient days' as a measure of inpatient activity, the validity of this output variable has been questioned. Hollingsworth and Parkin (1995) utilised a variety of 'patient day' categories, but in a subsequent paper (Parkin and Hollingsworth, 1996), they had replaced this with 'number of discharges'. The implications of using either of these factors are examined more closely in chapter seven, with a full application of the DEA methodology to hospital data.

For studies examining health care provision other than in hospitals, other measures of activity levels have been used. These include the 'number of encounters with medical staff' by type (rural health programmes), the 'volumes of prescriptions dispensed' (pharmacies) and the 'number of residents' by level of care required (nursing homes).

5.5.2.5 Quality measures (Outputs)

The inclusion of quality variables is much more complex, as it is difficult to find any acceptable measures, as noted by Kleinsorge and Karney (1992) in an investigation of nursing homes:

> "Quality is a multidimensional phenomenon, with no measure of quality capable of accounting for all the many facets of quality provided in a nursing home."
>
> KLEINSORGE AND KARNEY (1992)

Studies, such as this one, have therefore had to use surrogate measures for quality. Kleinsorge and Karney (1992) used three outputs to reflect quality of care. These were 'total days available', to account for occupancy rates and patient satisfaction, 'employees per resident', to reflect the amount of time spent providing care, and 'the number of bedsore-free days of care provided', as bedsores were identified as indicators of poor technical quality. Thanassoulis *et al.* (1995), in their examination of perinatal care, defined quality as encompassing two elements: service quality, recorded by patient satisfaction, and the quality of medical outcomes, measured by the number of 'at risk' babies that survived.

In the NHS, the quality of services is measured by surrogate factors such as, 'the number of cancelled operations', 'the number of complaints received' and 'the number of patients waiting longer than one year for treatment'.

In the following chapter, where the value of DEA in analysing the overall efficiency of acute hospitals is examined, the above considerations are used in the selection of all variables to be used in the analysis.

5.6 The Development of Alternative DEA Models

Once the issues surrounding variable selection have been addressed, there are other elements of the DEA model that can be adapted and enhanced to examine efficiency in slightly different ways. The original CCR model can be developed with an input or output orientation, has constant returns to scale, a free allocation of weights and a piecewise linear efficiency frontier. The methodology has been developed so that a variety of other models now exist, for which the crucial differences are 'the envelopment surface and the projection path to the envelopment surface for the inefficient DMUs' (Charnes *et al.*, 1994). These different models are now presented, using data from the gynaecology department as an illustration, and the theoretical concepts relating to the technical development of DEA since 1978 are discussed.

5.6.1 The BCC Model

Banker, Charnes and Cooper (1984), after whom the model is named, developed this extension to the DEA methodology as a means for distinguishing between technical and scale inefficiencies and for investigating the idea of varying returns to scale (VRS) (Charnes *et al.*, 1994). The nature of the returns to scale can be classified, be they increasing, constant or decreasing, and an associated numerical value can also be calculated.

Using the same variables included in model G3 above, the nature of the efficiency frontier using the BCC model with varying returns to scale, would be as represented in Figure 5.7 overleaf. As can be observed by comparison with figure 5.2, it is the shape of the efficiency frontier that has been changed by the VRS methodology.

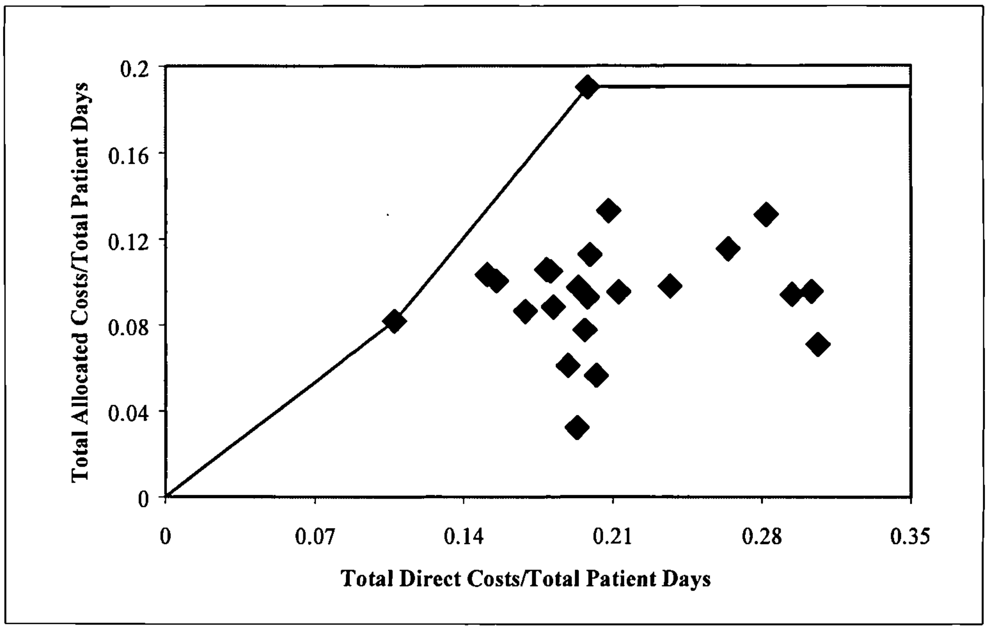


Figure 5.7: Locating the Efficiency Frontier with Varying Returns to Scale

Figure 5.8 below illustrates how the inefficient units are related to the efficiency frontier using the BCC model, with an input minimisation orientation, simplified to a single input and single output model.

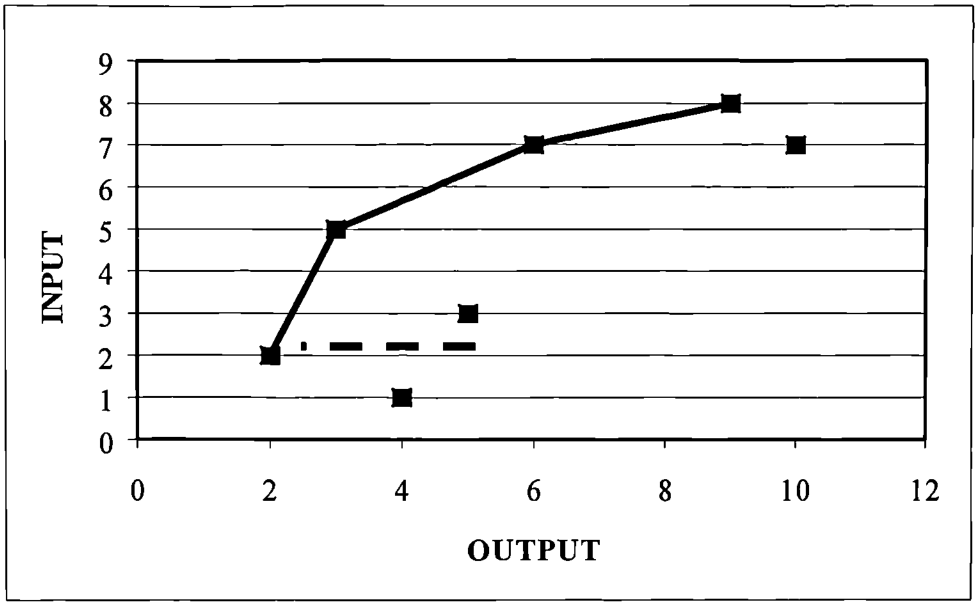


Figure 5.8: Relating Inefficient Units to the Efficiency Frontier using the BCC Model (Input Orientation)

For the output maximisation model, the lines relating the inefficient units to the frontier would be vertical, rather than horizontal. To achieve efficiency, the inefficient DMU would need to increase its outputs to meet the level of the frontier, rather than reduce its inputs.

Including the concept of varying returns to scale in the DEA methodology requires a change in the formulation of the model. The original methodology was again identified in the ratio format, which was formulated as a primal and dual linear programme for solution purposes. The dual form is shown figure 5.9 below, adapted from Charnes *et al.* (1994):

$$\text{MAXIMISE: } e_\alpha = \sum_{r=1}^{s} U_r O_{r\alpha} + U_\alpha ,$$

$$\text{SUBJECT TO: } \sum_{r=1}^{s} U_r O_{rj} - \sum_{i=1}^{m} V_i I_{ij} + U_\alpha \leq 0, \; j = 1....\alpha....n$$

$$\text{AND: } \sum_{i=1}^{m} V_i I_{i\alpha} = 1$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),
U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,
V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,
s = number of outputs and
m = number of inputs.

Figure 5.9: The Input-Oriented BCC Model (Dual)

The main difference between the BCC and CCR model formulation is in the inclusion of an extra term, $U\alpha$, which represents the scale variance in the varying returns to scale model.

Taking the example of the gynaecology departments, with two inputs ('total direct costs' and 'total allocated costs') and one output ('total inpatient discharges') (denoted by G8), the outcome of using the BCC model can be examined. The general output is compared with that of the equivalent CCR model (G3), with full results contained in appendix 2C.

Model	No .of Efficient Units	Efficient Units	Average Efficiency Score	Lowest Efficiency Score	Least Efficient Unit
G3 (CRS)	3	#11, #12, #19	72.33%	44.7%	7
G8(VRS)	5	#1, #11, #12, #15, #19	85.52%	61.2%	18

Table 5.5: The Impact of Varying Returns to Scale on Overall Efficiency

As can be seen in Table 5.5, the impact of including the measure of scale variance is the increase in the number of units on the efficiency frontier, with the average efficiency score also increasing. The efficiency of each individual unit either remains the same, or has increased. Also, the nature and size of the returns to scale can be investigated using the above form of the BCC model. For example, in model G8, 12 of the DMUs exhibit negative returns to scale and 11 demonstrate positive returns to scale. Of the 5 efficient DMUs, the returns to scale coefficient is negative for DMUs #1 and #12, whilst positive for DMUs #11, #15 and #19.

Four of the DMUs, one of which is efficient, have a returns to scale coefficient very close to zero, implying that the assumption of constant returns to scale is probably valid for these DMUs. This can be seen by the relatively minor change to their efficiency scores with the introduction of the VRS model, shown in table 5.6 overleaf, along with further details of the output for these DMUs.

Table 5.6 also includes the output for the DMUs with the largest returns to scale coefficient, both positive and negative, whose efficiency scores have changed dramatically:

Returns to Scale Type	DMU Number	Efficiency Score (%)	Scale Variance Coefficient	Change from CRS Model
Close to Zero	#4	61.76	0.01325	+ 0.03%
	#11	100	0.01976	0%
	#14	87.28	0.02144	+ 0.31%
	#16	72.04	0.01918	+ 0.11%
Large, Positive	#15	100	1	+ 23.98%
Large, Negative	#10	92.35	-0.78375	+ 20.24%

Table 5.6: Nature of Scale Variance

For the DMUs with negative scale variance, the average variance is 0.48158. Conversely, the average positive scale variance is 0.19638. As can be seen from the above analysis, allowing for the influence of varying returns to scale has altered the efficiency scores calculated for each DMU, although the impact does depend upon numerous factors. For example, despite the scale variance for #12 being large and negative (-0.51946), it is actually classified as efficient under both the CRS and VRS models.

5.6.2 Additive Models

The Additive model was developed in Charnes *et al.* (1985) and further extended in Charnes *et al.* (1987). It relates DEA to some of the very early work on efficiency in Charnes and Cooper (1957) and Koopmans (1951) (Charnes *et al*, 1994). It measures variable returns to scale and has a piecewise linear efficiency frontier. The linear programming formulation in the dual form is called the multiplier form, the primal form being denoted by the envelopment form (Charnes *et al.*, 1994).

The shape of the envelopment surface or efficiency frontier for the gynaecology example with two inputs and one output is equivalent to Figure 5.7 above, illustrating the BCC model. The difference between the two models occurs in how the inefficient units are related to the efficiency frontier. In figure 5.8, the way in which units were related to the efficiency frontier was illustrated for the BCC model, with an input minimisation orientation. This can be contrasted with figure 5.10 below, again using a simplified single input and single output model, which shows how inefficient units are related to the efficiency frontier under the additive model (the dotted black lines illustrate this relationship).

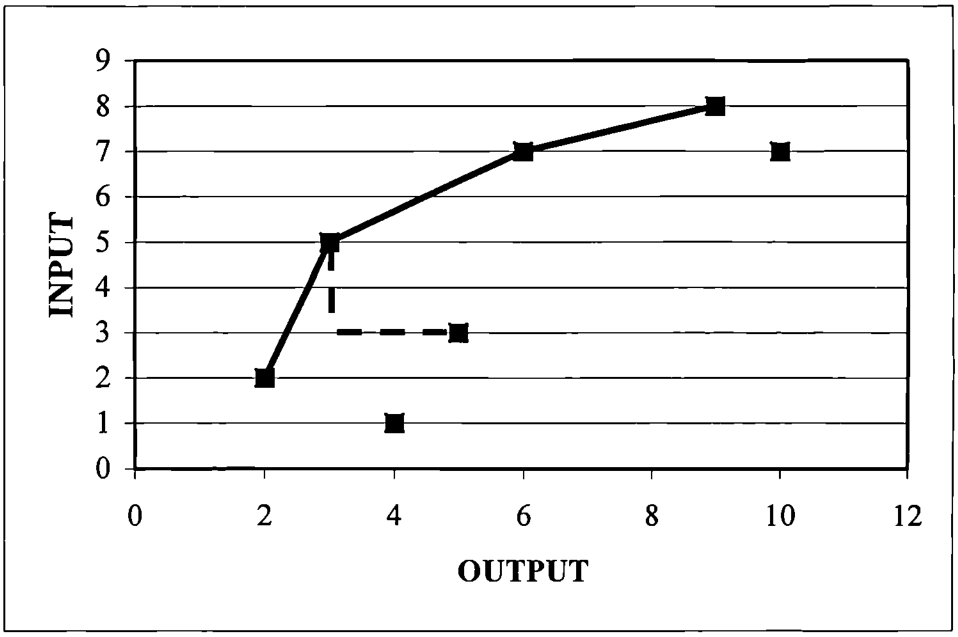


Figure 5.10: Relating Inefficient Units to the Frontier using the Additive Model.

The formulation of the additive model is not included in this chapter, as it is not used in the analysis presented in the following chapters. This is further discussed in section 5.6.4 to follow, which evaluates the different model formulations.

5.6.3 Multiplicative Models

The multiplicative model, found in Charnes *et al.* (1982, 1983), is quite different from the other models in that it has a piecewise log-linear or a piecewise Cobb-Douglas envelopment surface. It is an extension of the additive model, with the model applied 'to the logarithms of the original data values' (Charnes *et al.*, 1994). It has two parts: the variant multiplicative model has constant log-linear returns to scale, whilst the invariant multiplicative model has variable log-linear returns to scale.

The DMUs on the efficiency frontier are the same as would be identified under the previous model. The difference is seen in the shape of the efficiency frontier. Figure 5.10 illustrated a piecewise linear frontier – the lines joining points on the frontier are 'straight'. With the multiplicative model, the lines are 'curved' according to the log-linear or Cobb-Douglas formulation. As with the additive model, the formulation of the multiplicative model is not included.

5.6.4 Evaluation of the Different Model Types

The different model types as mentioned above have all been developed to enhance the theoretical capabilities of the DEA technique. However, the focus of the detailed analysis to come is on developing practical and useful DEA models, which would be acceptable to health care managers and policy makers. The later two developments, the multiplicative and additive models, are more complex and no particular benefits for health care analysis have been established in their development. Therefore, the analysis uses the CCR model primarily, with further discussion of the BCC model also included, focusing on the validity of the constant returns-to-scale assumption.

5.7 Weight Restriction Issues

The modifications to the DEA methodology considered up to this point have been concerned with the shape of the envelopment surface and the way in which inefficient DMUs are related to it. However, over the last five to ten years, the issue of weight flexibility has become of central concern to the development of DEA modifications. The reasons for this development are numerous but essentially it has followed 'as a natural by-product of real-life applications' (Allen *et al.*, 1997), as unrestricted DEA models have been seen as inappropriate outside the academic arena.

This corresponded to the experience of Val Belton of the University of Strathclyde (Belton, 1992), who encountered problems in trying to introduce DEA into large organisations:

> "Amongst the difficulties mentioned are those associated with interpreting the weights and also how setting them to zero on certain inputs and outputs ignores efficiency with respect to those measures." Reported by TOFALIS (1993)

Roll and Golany (1993) presented the same case, in that they found that 'virtually unconstrained factor weights' were usually unacceptable in real world applications. Similarly, it was generally 'deemed inappropriate to accord widely differing weights to the same factor, when assessing different DMUs' (Roll and Golany, 1993).

A similar response was found by Wilkinson (personal communication), when analysing data relating to a major oil company - the findings of the DEA analysis were unacceptable to the company when presented with no restrictions to the weights.

A totally free system of weights, with some of them set at (virtually) zero, made no sense to the company as a means for examining efficiency. Some units known by the company to be highly inefficient were ranked as 100% efficient by the analysis, which exacerbated this confusion. However, the revised models, which incorporated weight restrictions, were acceptable to the company's management.

Ball *et al.* (1997) summarised the reasons why the introduction of weight limitations might be acceptable to the policy-maker, in the light of the experiences referred to above, if DEA applications are to be useful in public sector organisations:

- A DMU that has specialised in a particular area to the neglect of others currently has more chance of being classified as efficient than the good all-rounder;
- The lack of discrimination, with a reasonable number of inputs and outputs, is unsatisfactory. Eliminating factors is conceptually unsound and a very crude form of weight limitation - a variable gets a weight of either zero or one;
- In many problems, not all inputs contribute to the production of every output. This raises the possibility of reaching 100% efficiency on the basis of a meaningless ratio;
- Allowing some inputs and outputs to be more highly weighted than others may be beneficial, where specialist knowledge or intuition suggests this ought to be the case.

Amongst the earliest studies in this area were those by Dyson and Thanassoulis (1988), Charnes *et al.* (1989) and Thompson *et al.* (1990). Prior to this, only passing comment had been given to the problems that can arise with totally unrestricted weights (other than the stipulation of non-negativity), as in the original DEA models.

The supporters of weight restrictions would argue that the relevance and applicability of the technique is extended by the introduction of some method of restricting the weights.

> "Few would argue against reducing weight flexibility in DEA, since doing so would ensure that the subsequent assessment cannot effectively ignore any inputs or outputs but also would assign weights to inputs and outputs more in line with some general view of their perceived importance."
>
> DYSON AND THANASSOULIS (1988)

One of the most widely supported arguments for including weight restrictions is that they are seen to 'clearly enhance the basic DEA approach and enable judgements as to the relative importance of input/output measures to be incorporated' (Beasley, 1990). The argument is further supported below:

> "In choosing the factors to enter the analysis, one has already expressed strong opinion on the importance of these factors versus others which were left out. Imposing bounds on factor weights renders this process more flexible by defining the importance of the different factors." ROLL AND GOLANY (1993)

There are several other reasons for the adoption of some method for restricting the free assignment of weights. In DEA, some variables have weights so small as to effectively be zero and they contribute nothing to the efficiency score obtained for the unit. Imposing factor weight ensures all variables contribute something to the efficiency score for each DMU (Roll *et al.*, 1991). This is illustrated by returning to the example of the gynaecology departments.

Some additional information from models G5 and G6 introduced in section 5.5 is recorded in table 5.7 below, showing the number of input factors contributing to the efficiency score for each DMU. These are defined as those factors with non-zero virtual weights. Particular reference is made to the efficient units.

Model Number	No. of DMUs where efficiency is calculated using:			Number of Inputs used for Efficient Units
	3 Inputs	2 Inputs	1 Input	
G5	13	8	2	#11 – 2, #12 – 3 #15 – 1, #17 – 3
G6	8	9	6	#10 – 2, #11 – 2 #12 – 2, #19 – 1

Table 5.7: Analysis of Factors Contributing to Efficiency Scores

As can be seen, a substantial number of units, including some of the efficient units, have efficiency scores determined with no significant contribution from at least one of the defined inputs.

One of the most widely quoted advantages of DEA is that the free weighting system allows each DMU to be seen in the best possible light, first stated by Charnes *et al.* (1978). However, when this proposition was discussed with those involved with performance assessment in the NHS, it was found to be unacceptable. It was felt to be much more important for hospitals to be assessed in their true light, for the assessment to show up inefficiencies in all areas where they exist. A free system of weights can lead to all the weight being attached to one input, or output, producing an efficiency score which does not reflect the true nature of a unit's performance. Tomkins and Green (1988) found that unrestricted weights allows units to be considered as efficient by placing weights on extremes. They were then defined as relatively efficient only because no one else operated in their niche, which was not necessarily appropriate.

In a similar vein, Dyson and Thanassoulis (1988) identified Liverpool as an efficient Local Authority because it issued the highest number of 'summons for non-payment of rates' amongst all the local authorities under investigation. This would have been prevented by including weight restrictions, as would the relating of one minor aspect of output to a totally unrelated input factor (Wilkinson, 1991).

A free weights approach has also been shown to produce strange results in the health literature. Thanassoulis *et al.* (1995), in their examination of perinatal care, found that for some DMUs 'very satisfied with care' had a lower weighting than 'satisfied with care', when these were used as measures of quality with no restrictions on the weights. Roll and Golany (1993) also found that using alternative weight restrictions could be useful in determining the relative importance of the different variables to each of the DMUs.

The counter view, gradually losing supporters, has been that assigning any restrictions to DEA weights compromises the technique's integrity and 'the objectivity of DEA is lost' (Roll and Golany, 1993). A further problem is then that of selecting the method to restrict the weights, as a number of possibilities have been defined. The method chosen has depended upon the context of the DEA investigation, the units under investigation, the data used, the level of managerial involvement and the focus of the research. This difficulty has often led to bounding techniques being rejected:

> "Different bounding techniques may render significantly different results. The apparent problem, in controlling factor weights, lies, therefore, in the question of how to choose an appropriate set of bounds." ROLL AND GOLANY (1993)

As was discussed above, it has been suggested that anything other than complete flexibility destroys the integrity of the DEA technique (Wilkinson, 1991). The use of weight restrictions automatically necessitates some subjectivity in the selection of weights to be used.

However, despite the many arguments for the necessity to introduce weight restrictions, there is still the further problem of defining how these restrictions should be applied:

> "Deciding on such bounds poses a difficult problem to the efficiency analyst. The relative position of such bounds, the range allowed between upper and lower bounds, etc. may significantly affect the efficiency score." COOK *ET AL.* (1994)

This is now addressed in greater detail.

5.7.1 Choosing the Methods for Restricting the Factor Weights

The dilemma in the use of methods for restricting factor weights is how to strike the right balance between complete flexibility across all DMUs and a fixed set of weights, common to each DMU. Many examples have been found within the DEA literature of the different methods used to restrict the free allocation of weights and the method chosen will reflect the type of information to be included in the model.

Some of these approaches are:

- Dyson and Thanassoulis (1988) proposed one of the earliest approaches to restricting weight flexibility, looking at a single input, multiple output model and imposing upper and lower bounds on the individual variables.

- Thompson *et al.* (1986) amended the basic model by incorporating bounds on ratios of variables.
- Wong and Beasley (1990) appended factor inequalities to the model to restrict the weights
- Charnes *et al.* (1989) introduced the cone ratio approach, requiring all factor weights to belong to given closed cones.

Adapted from CHARNES *ET AL.* (1994)

Thanassoulis *et al.* (1995), Roll and Golany (1993) and Roll *et al.* (1991) also suggested alternative approaches for dealing with factor weights. The DEA computer-based model utilised in this analysis is based on the approach of placing restrictions on the virtual inputs and virtual outputs. This is consistent with the methodology used by Wilkinson (1991), with some degree of success. Five different types of weight restriction are possible using this methodology and these are often also used in combination. These are summarised below.

5.7.1.1 Minimum Contribution

This type of approach constrains the DEA model, such that the virtual percentage for each variable must be above a specified percentage - the restricted factor is required to contribute something to the calculation of the efficiency score. Applying 'minimum' restrictions for all the inputs and outputs would negate the problem of the non-contributing factors, that is, those with a zero weight attached to them. Maximum virtual weights would be implied implicitly, as the sum of the virtual weights can still not exceed 100%. Minimum constraints could be attached such that the virtual weight for each factor must be at least 10%.

5.7.1.2 Maximum Contribution

This approach would be the converse of that discussed in section 5.7.1.2. In this instance, factors are prevented from over-dominating the efficiency score calculations. This was observed with model G2 and illustrated in table 5.2 (see section 5.4.4), where it was seen that a number of the DMUs based their efficiency assessment on the contribution of just one of the inputs ('total direct costs'). As is the case with the introduction of minimum contributions, applying maximum contributions to any of the factors may implicitly introduce minimum values for the other factor weights due to the nature of the efficiency calculations. Maximum constraints could be attached so that the virtual weight for each input, for example, was less than 50% of the total.

5.7.1.3 Range of Values

This is a combination of the two approaches discussed above, with minimum and maximum constraints specified, within which the virtual weight for that factor must lie. This is similar to the 'cone ratio' approach determined by Charnes et al. (1989). Application of such an approach, with equivalent ranges of values used for all of the outputs for example, results in similar virtual weight pattern across all the DMUs. Taken to its extreme, this would lead to the application of a common set of weights, as discussed by Roll and Golany (1993). For example, the virtual weight percentage for each output could be constrained to lie between 25% and 40%.

5.7.1.4 Grouped Restrictions

The fourth alternative for introducing restrictions to the weights is to constrain several factors in combination, with these groups containing either inputs or outputs.

This approach is particularly useful in cases where the data sample contains some missing values, such that applying weight restrictions directly to each factor would be infeasible. In this case, the restriction applied would be that the sum of some combination of the virtual weights for the relevant factors lie in the specified range, using the approach outlined in the previous section. For example, the two inputs 'total direct costs' and 'total allocated costs' used in model G5 in section 5.4.4 could be constrained so that the total virtual weight obtained from both of these factors should be in the range 75% - 85%. Further constraints could be added to control the individual factors if required. In the methodology used for restricting the allocation of weights in this study, it is not possible to include grouped restrictions that relate inputs to outputs, which was an approach adopted by Thanassoulis *et al.* (1995).

5.7.1.5 Specifying Inequalities

The computer-based methodology used for this analysis does not allow for inequalities showing relationships between the virtual weights to be specified directly, although the general DEA models can be adapted for this purpose in a theoretical sense. In the approach used here, inequalities are specified indirectly through the careful application of the previously defined methodologies. For example, it may be required that the virtual weight for 'total direct costs' be larger than the other virtual weights for the other input factors, with the additional proviso that the minimum virtual weight for 'total direct costs' be set at 40%. This would be established by constraining the virtual weight for this factor to have a minimum of 40%, with the other virtual weights using this figure as their maximum value.

5.7.2 Illustrating the Impact of Weight Restrictions

The impact of attaching factor weights in the manner described immediately above can be investigated by returning to the gynaecology inpatient department example. Three models are compared below, each based on the basic CCR model with constant returns to scale and an input orientation. They have been developed with three inputs (total direct costs, total allocated costs and average staffed beds) and one output (total number of discharges), with the only differences in the constraints applied to the weights. Model G6 has no restrictions and was included in table 5.7 above.

The virtual inputs in model G6a are constrained so that each must contribute at least 10% of the efficiency score, that is, the imposition of a lower bound. In model G6b, the restriction imposed is that each virtual input must contribute at least 20% but no more than 40% of the efficiency score, imposing both upper and lower bounds to the virtual inputs. Table 5.8 shows the impact of reducing weight flexibility in general terms.

Model:	Number of Efficient Units	Efficient Units	Average Efficiency Score	Lowest Efficiency Score	Least Efficient Unit
G6	4	#10, #11, #12, #19	81.12%	63.2%	1
G6a	3	#11, #12, #19	79.44%	62.4%	1
G6b	3	#11, #12, #19	76.34%	59.0%	21

Table 5.8: The Impact of Weight Restrictions on Overall Efficiency Scores

In general terms, the impact of reducing weight flexibility has been to reshape the efficiency frontier according to the restrictions imposed, making it more difficult for each DMU to be classified as efficient.

The average efficiency score is reduced if weight restrictions are imposed, as shown above, and tighter restrictions further reduce it. The efficiency score for units classified as inefficient is reduced by the imposition of weight restrictions, and some of the units classified as efficient can become inefficient. For example, DMU #10 is efficient with no weight restrictions, whereas its efficiency score is reduced to 86.8% when the tighter restrictions are introduced in model G6b. In the unrestricted model, its efficiency score was calculated with a zero weighting attached to one of the inputs, 'total allocated costs'. In the weight-restricted models, this is no longer possible and DMU #10 is no longer classified as efficient.

However, three units are classified as efficient in all cases, suggesting that their 100% efficiency score has not been achieved simply because of an unusual input pattern, or due to the over-dominance of one input factor. Such DMUs have demonstrated that their efficiency classification is robust.

Further analysis of the output from the three DEA models introduced in table 5.8 demonstrates the importance of the debate relating to the introduction of weight restrictions, if DEA is to be used successfully in practical applications. In table 5.7, the number of inputs contributing to the efficiency score for each DMU using model G6 was presented, showing that just 8 of the 23 units based their efficiency score on contributions from all three inputs.

However, this does not reflect the size of that contribution, in that the virtual weights for two of the three inputs could be 0.1%, with the third input at 98.8%.

Alternatively, there could be an equal contribution from each input factor. Analysing the patterns of the virtual inputs for model G6 identifies the input factors that have made the dominant contribution to the efficiency evaluation.

For example, in the unbounded model (G6), the pattern of the virtual weights for 'average staffed beds' is interesting. The average value is 49.8%, but there are 8 DMUs with a zero virtual weight and for the remaining DMUs its value is always above 40%. For all but one of the DMUs for which the virtual weight is non-zero, it is the most significant contributor to the efficiency score, or the 'dominant' factor. In fact, the 'average staffed beds' factor is the dominant input, according to the virtual weight percentages, on the most occasions for all three of the models specified in table 5.8. The tight restrictions in model G6b, therefore, have the most effect in terms of reducing efficiency scores for those DMUs heavily reliant on 'average staffed beds'. This is presented in Table 5.9 below:

DMU	Model	Virtual Inputs (%)			Efficiency Score (%)
		DC	AC	ASB	
#2	G6	2.7	0	97.3	**83.22**
	G6a	37.1	10	52.9	**80.54**
	G6b	40	20	40	**77.83**
#4	G6	0	0	100	**89.13**
	G6a	10	10	80	**83.48**
	G6b	40	20	40	**73.00**
#10	G6	38.7	0	61.3	**100**
	G6a	30.2	10	59.8	**97.26**
	G6b	40	20	40	**86.80**

Table 5.9: Analysis of the Virtual Inputs in Relation to the Weight Restrictions and Efficiency Scores for DMUs #2, #4 and #10

The three DMUs that remain efficient across all the three models (#11, #12, #19) do not heavily rely on one input as a dominant contributor to their efficiency scores, which may explain their continued efficiency even under tight weight restrictions.

5.7.3 Weight Restriction Methodology in Practice

Unfortunately, few examples have been found of their use with actual data, in real situations, nor any concrete findings about which methods are the most appropriate for a given situation, a problem also discussed by Roll and Golany (1993). The issue of weight restrictions has often been treated as a problem of formulation, with a variety of statistical and technical additions to the basic DEA models devised. There has been little emphasis on the need to have an understanding of the relative importance of each of the variables before defining the weight restrictions.

Only brief attention has been given to these methods in accordance with the proposition that 'no *mechanical* systemisation of the process of choosing bounds would serve its purpose' (Cook *et al.*, 1994).

The same authors suggest three rules that should govern the process of weight selection, equivalent to the procedures expressed for selecting variables:

> "What is required is a thorough knowledge of the process in which DMUs are engaged, a clear vision of the purposes for which efficiency is being measured, and sound managerial considerations, in order to determine limits within which factor weights may vary."
>
> COOK *ET AL.* (1994)

This view is reflected in the investigation of elderly care services given in Ball and Roberts (1998), where the imposition of weight restrictions, along with the selection of the original variables, was done in conjunction with the policy-makers. It was felt that they had an intrinsic understanding of the services being investigated.

The weight restrictions were imposed in fairly broad bands, with several different models used to investigate the efficiency of the units under a variety of care strategies, utilising the following approach:

1. Minimum weights were applied to all the individual outputs for which no zero values existed (primary weighting);
2. Minimum and Maximum weights were also applied to groups of outputs, according to the three different categories of output defined (secondary weighting);
3. The inputs (of which there were only two) were weighted so that the financial factor would dominate but not over-dominate, allowing the environmental factor to contribute if it would be of benefit.

The impact of introducing weight restrictions in relation to an extensive health care application is investigated in greater detail in chapter eight.

5.8 Other Extensions to the DEA Methodology

Extensions to the basic DEA methodology relating to the formulation of new model types, the inclusion of multiple inputs and outputs and the adoption of weight restrictions have been debated thus far.

However, there are a variety of other approaches for extending the DEA methodology, through alterations to the model formulation or different approaches to the evaluation of data, which are now examined. In most cases, the coverage will be fairly brief and no model formulations are included in the discussion unless the extension is further utilised in later chapters.

5.8.1 Different Types of Efficiency in DEA

The different elements and types of efficiency were described earlier in the dissertation. DEA, in general terms, is a measure of technical efficiency, as opposed to the other types of efficiency defined:

> "... the focus is on the physical levels of outputs achieved, given the various physical level of inputs. This is what we shall mean by technical efficiency, in contrast to allocative efficiency which is concerned with the right mix of inputs, given the prices of the inputs, price efficiency (i.e., the proper mix of outputs to maximise a given objective)." LEWIN AND MOREY (1981)

The CCR model measured the technical efficiency of organisations, whilst specific inputs or outputs can be used to take into account effectiveness, and the BCC model can be used to differentiate between scale and technical efficiencies.

Some recent examples have also attempted to extend the DEA methodology to measure allocative efficiency. For example, Byrnes and Valdmanis (1994) attempted to measure both the allocative and technical efficiency of hospitals in the United States. This paper is comprehensively reviewed in chapter six.

As the DEA methodology is more naturally related to the measurement of technical efficiency, and because this extension to the modelling process is still relatively untested in practical applications, the focus of this analysis is on the measurement of technical efficiency. (See chapter two for a more detailed discussion on the measurement of allocative efficiency.)

5.8.2 The Inclusion of Categorical Data as Variables

The inclusion of categorical (or ordinal) data into the DEA model alongside the continuous data variables adds another dimension to the DEA technique. In the analysis of health service efficiency, this has not been used to any great extent to date and no obvious examples of its value spring readily to mind. Banker and Morey (1986a) developed an approach to incorporate the special characteristics of banks under their investigation, such as whether each bank had a drive-in facility or a cash point, both of which were 0-1 variables. Charnes *et al.* (1994) suggests an approach for dealing with multiple categorical variables, which can be incorporated into all model types and stresses the need for the categories to be comparable. However, as categorical data is not used in the analysis of health service data in later chapters, no further consideration is given to this adaptation of the DEA methodology.

5.8.3 Non-discretionary Variables

The variables included in the analysis up to this point have all been assumed to be controlled by the management of each DMU and varied at its discretion. However, as the managers spoken to within the health service have concurred, there 'exist exogenously fixed or non-discretionary inputs or outputs that are beyond the control of management' (Charnes *et al.*, 1994). These variables are often the environmental variables described in section 5.5.2.3. At the Blackburn, Hynburn and Ribble Valley (BHRV) NHS Trust, these non-discretionary factors include the demographic and socio-economic position of the population, the geography of the locality, the age and condition of the hospital buildings they must use and, to a certain degree, the existing health conditions of the population.

Thanassoulis *et al.* (1995) also identified non-discretionary variables in their analysis of perinatal care in the UK. They classed one of the input variables, 'the number of babies at risk', as exogenously fixed. However, they also categorised the outputs relating to the activity levels as non-discretionary, that is, they assumed that each health authority did not directly control the number of babies delivered or abortions performed.

In most examples, however, it is the input variables that are assumed to be beyond the control of management. Beyond health care settings, examples could be the number of competitors for a fast food company (Banker and Morey, 1986b) or the weather conditions affecting the control of maintenance units. Chalos and Cherian (1995) selected non-discretionary inputs to reflect the socio-economic and demographic make-up of the population of each school district.

The theoretical implications for the DEA modelling process of constraining some variables to be exogenously fixed was first presented by Banker and Morey (1986b) and this approach has been used by most subsequent investigations using non-discretionary variables.

For example, if an expenditure input is assumed to be discretionary, whilst an environmental factor is beyond managerial control, the target for each inefficient DMU is the point on the efficiency frontier which has 'the same output, the same non-discretionary input and less non-discretionary input' (Chalos and Cherian, 1995).

This approach allows for 'the identification of a considerably enhanced opportunity for targeted savings in the controllable inputs and targeted increases in the controllable outputs' (Charnes et. al, 1994).

The various DEA models can all be adapted to accommodate exogenously-fixed factors, by separating the inputs and outputs into subsets of discretionary and non-discretionary variables and making slight adjustments to the model being used. This adaptation to the basic model clearly has some relevance to health care applications of the DEA, which are further discussed in later chapters. The modified model (assuming that the non-discretionary factors are all inputs) has the following formulation, adapted from Norman and Stoker (1991):

$$\text{MINIMISE: } e_\alpha = \sum_{i=1}^{m} V_i I_{i\alpha} ,$$

$$\text{SUBJECT TO: } \sum_{i=1}^{m} V_i I_{ij} + \sum_{k=1}^{t} W_k E_{kj} - \sum_{r=1}^{s} U_r O_{rj} \geq 0,$$

$$j = 1 \ldots\alpha\ldots n$$

$$\text{AND: } \quad \sum_{r=1}^{s} U_r O_{r\alpha} - \sum_{k=1}^{t} W_k E_{k\alpha} = 1$$

WHERE: U_r, V_i and $W_k > \varepsilon$ FOR ALL r, i and k (ε being a very small positive number);

E_{kj} are the positive known environmental inputs of the j^{th} DMU;

W_k are the variable weights applied the environmental inputs of the j^{th} DMU;.

s = number of outputs

m = number of inputs and

t = number of environmental inputs.

Figure 5.11: DEA Model for Uncontrollable Environmental Input Factors

5.8.4 Longitudinal and Window Analysis

In the examples given to date, the samples investigated using DEA comprised a set of distinct organisations. However, as Lewin and Morey (1981) suggested, it is often insightful to examine the performance of a particular DMU with respect only to itself, over a specified period of time. The purpose behind such an examination may be to determine the reasons why the efficiency score of a DMU fluctuates over time – there may have been managerial or structural changes, external factors unique to each DMU or seasonal variations. Such an analysis may also identify 'significant' time periods for further examination.

Lewin and Morey (1981) determined that the efficiency of naval recruitment offices was affected by seasonal influences by carrying out such an analysis. Campbell (1996) investigated the impact of privatisation on British Steel in terms of the effect on the efficiency scores over a twenty-year time span, looking for links with changes in organisational and managerial structure.

The idea of investigating over a particular time period is further extended by the development of window analysis, which is a relatively new addition to the DEA methodology still being investigated:

> "In such a setting, it is possible to perform DEA over time using a moving average analogue, where a DMU in each different time period is treated as if it were a different DMU. Specifically, a DMU's performance is contrasted with its performance in other periods in addition to the performance of other DMUs."
>
> CHARNES *et al.* (1994)

For quarterly data, a window analysis is carried out by choosing a window length and a number of windows. Window length is defined as how many quarters to be used in each DEA run and the number of windows is the number of DEA runs needed to complete the analysis. In the first DEA run, for a window length of three for example, data from the first three time periods (quarters) for each DMU is included in the DEA model, with each treated as distinct and independent DMUs. This is repeated with the data from the second, third and fourth quarters for each DMU and so on until all the quarters have been covered in this way (Charnes *et al.*, 1994a). The efficiency scores for each quarter for each DMU can then be analysed to look for the effects of time lags, seasonal trends and stability, using simple statistics.

Window analysis using DEA has been used to examine both the brewing and soft drinks industries, in which marketing and advertising have a significant role to play and the market can be highly volatile (Day *et al.*, 1994 and Charnes *et al.*, 1994a).

This approach does not require any alterations to the existing models, each of which can be used to carry out a window analysis or analysis over time, if sufficient data can be identified and is consistent over the whole time period. Whilst window analysis does not appear to relate to the analysis of health care to any significant degree, a straight forward longitudinal analysis could be appropriate, in order to assess the effect of structural and managerial changes and their impact on efficiency. However, this would require consistent data from several years, collected and recorded in the same format. This is not always easy in a health service context and, although this approach clearly has some benefits, it is not used in the remainder of this analysis.

5.8.5 Cross-efficiency in DEA

In section 5.4.6, the notion of the reference set and reference set count was established and the suggestion was made that it could be used to further investigate the nature of the efficient units and help to establish points of reference for the inefficient units. Doyle and Green (1995) discussed how analysis of the reference sets and the number of times each efficient unit appears in the reference set for an inefficient DMU can be used as a method for distinguishing between the efficient units, to find which of them is the 'most' efficient. The notion of cross-efficiency evaluation, first proposed by Sexton *et al.* (1986), is another method for examining the efficient DMUs. The technique offers several advantages above those of the original DEA methodology, according to Doyle and Green (1995):

1. It guarantees a unique ordering of the DMUs;
2. The methodology can be used with a small data sample effectively;
3. It produces a measure for all DMUs, not just efficient ones (as is the case with reference set count);
4. Non-efficient DMUs can be rated better than efficient ones under cross-evaluation;
5. It is more robust in the face of variations in input than standard efficiency calculations.

The concept of cross-efficiency has not been widely applied, perhaps because:

> "Users also need an intuitive grasp of what cross-efficiency means, and how it differs from simple efficiency, with at least some idea of how it might be used. Without such an understanding users will not be tempted to invest the effort in finding out."
>
> DOYLE AND GREEN (1994)

5.8.5.1 Definition of Cross-Efficiency

In the DEA model, the weights to be attached to each of the input and output factors have been calculated in order to determine the efficiency score for each individual DMU. In the cross-efficiency methodology, the set of weights for a given DMU is used to weight the inputs and outputs for all the other DMUs. This determines the 'cross-efficiency' for all other DMUs, seen from the perspective of the original DMU. The process is repeated for all DMUs, leading to a cross-efficiency matrix, with the usual efficiency scores found in the leading diagonal (Doyle and Green, 1994). The efficient DMUs in the reference set for each inefficient DMU will have cross-efficiency scores of 100%, as would be expected (see section 5.4.6).

5.8.5.2 Application of Cross-Efficiency

Table 5.9 below gives a snapshot of the cross-efficiency matrix generated from model G3 introduced in section 5.4 above, which has two inputs and one output (the full matrix has 23 by 23 entries):

DMU	#1	#6	#11	#16	#21
#1	57.264	66.826	99.341	71.941	51.573
#6	57.253	66.813	99.323	71.927	51.563
#11	56.609	66.224	100.000	71.448	50.923
#16	57.254	66.814	99.324	71.928	51.564
#21	57.248	66.808	99.312	71.921	51.559

Table 5.10: Cross-Efficiency Matrix for a Simplified Example

From the table it can be seen that, for the small sample of DMUs given, the variance in cross-efficiency scores for each of the DMUs, when looking down each column, does not appear to be too large. This suggests that the efficiency score for each DMU is robust against a slight variation in the weights.

The weights applied to each DMU may not be very different in this case. Table 5.10 below shows the weights applied for the DMUs used in the cross-efficiency matrix above, and as can be seen, the weights for the five DMUs above are fairly consistent, apart from those applied to DMU #1.

DMU\FACTOR	I_2	I_3	I_4
#1	0.0005261	0.0002554	0.0002253
#6	0.0011678	0.0005669	0.0005
#11	0.0011113	0.0005075	0.0004579
#16	0.0014506	0.0007042	0.0006211
#21	0.0012797	0.0006213	0.0005479
Average	**0.0011071**	**0.0005311**	**0.0004704**
St. Deviation	**0.0003497**	**0.0001703**	**0.0001499**

Key: I_2 – Total Allocated Costs, I_3 – Total Direct Costs, I_4 – Average Number of Staffed Beds

Table 5.11: Variations in the Weights Applied

The cross efficiency values can be evaluated in a number of ways, other than the general examination given directly above. For example, averaging the values down the column for DMU #1 can be interpreted as 'an averaged peer appraisal' of DMU #1 and this gives a measure of 'how the unit associated with the column is rated by the rest of the units' (Boussofiane *et al.*, 1991). In this case:

> "A relatively efficient unit with a low such average efficiency is likely to feature in the peer groups of few inefficient units and it is likely to rely on weights dissimilar to those of the rest of the units in order to appear efficient. The converse is likely to be the case with a relatively efficient unit with a high average efficiency."
>
> BOUSSOFIANE *ET AL.* (1991)

There are numerous benefits of the cross-efficiency approach. The efficient DMUs, with high average efficiencies in the relevant column of the cross-efficiency matrix, are more likely to be suitable determinants of best practice (Boussofiane *et al.*, 1991).

Alternatively, averaging along the row for DMU #1 gives an 'averaged appraisal of peers', against which DMU #1 can measure itself (Doyle and Green, 1994).

Clearly, the scope for investigation using the cross-efficiency evaluation is vast, as the cross-efficiency values can be interpreted in a variety of ways (see Doyle and Green, (1995) for examples). However, the practical implications of cross-efficiency have not been investigated to any great degree. For this reason, the methodology is not used in the general analysis of health care institutions in subsequent chapters.

5.8.6 Recent Developments

The technique of data envelopment analysis appears to be in a state of almost continuous evolution:

> "Since the initial article by Charnes et al. (1978) a research industry has blossomed which has led to a richness of theory which the original proponents of the technique can hardly have envisaged."
>
> PEDRAJA-CHAPARRO *ET AL.* (1999)

This has been demonstrated by the numerous adaptations to the original CCR model that have been proposed over the last twenty years and discussed above. The evolutionary process appears set to continue and two of the many areas addressed most recently are briefly discussed here.

5.8.6.1 Super Efficiency

The basis of super-efficiency comes from the observation of the distribution of the efficiency scores for a typical DEA evaluation.

Lovell *et al.* (1994) found that the efficiency scores seemed to be skewed, such that 'a mass of observations achieves the upper bound' and are thus defined to be efficient. In the previous section, the calculation of a cross-efficiency matrix was suggested as a means for distinguishing between the efficient units, as was the reference set count in section 5.4.

However, the notion of super-efficiency has now also been developed to address this issue, so that the differences between the efficient DMUs can be investigated. The theoretical change to the methodology is that each DMU is dropped from its own reference set, with the result that efficiency scores are bounded only by zero. The most efficient DMUs can thus be allocated efficiency scores greater than 100% (Anderson and Petersen, 1993). Including the super-efficiency methodology has an impact on only those DMUs classified as efficient in the basic models.

Model G6, which has three inputs and one output and was introduced in section 5.5, has four efficient DMUs. The effect of super-efficiency on these four DMUs is shown in the table below.

DMU	Super-Efficiency Score	Reference Set Count in CCR Model
#19	133.6	14
#12	123.3	19
#11	142.2	2
#10	103.8	10

Table 5.12: Super Efficiency Scores

As can be seen from table 5.12, the DMU identified as the 'most efficient' using the reference set count (as suggested in section 5.4.6 as a means for distinguishing between the efficient units) does not have the highest super-efficiency score.

The benefit of such a methodological development is in the identification of the most efficient DMUs, as opposed to determining all the efficient DMUs. Lovell *et al.* (1994) made use of super-efficiency to investigate the distribution of efficiency scores for a group of schools in relation to a range of environmental factors using ordinary least squares regression analysis. However, as the impact and validity of super-efficiency has not been fully investigated with health care data, the issue is not debated further or used in the practical applications to follow.

5.8.6.2 The Free Disposal Hull (FDH) Method

This development to the DEA methodology, first proposed by Tulkens (1993), creates an efficiency frontier in the shape of a 'staircase' by removing the convexity constraint from the methodology, as illustrated below:

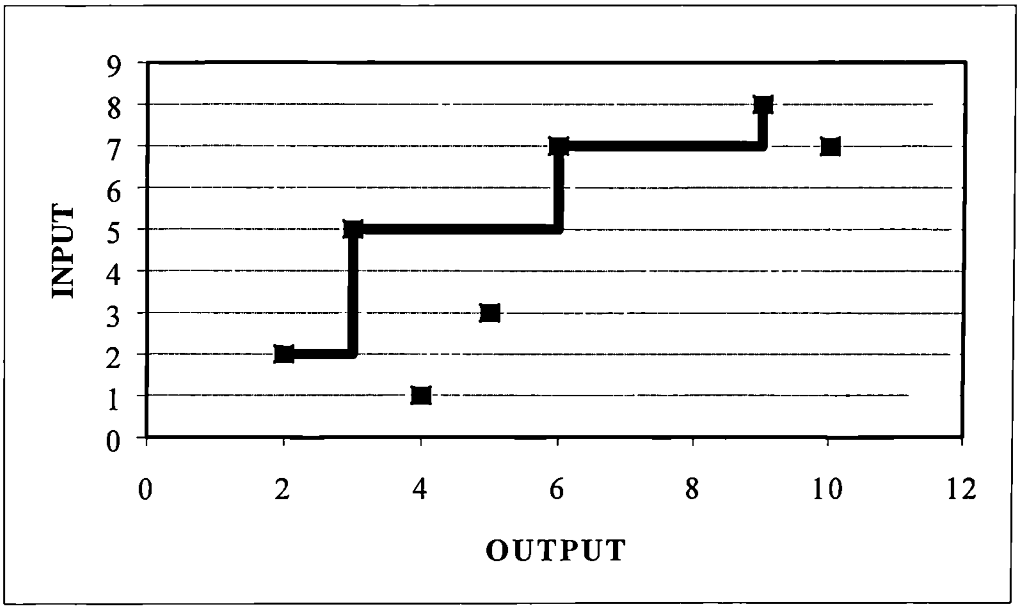


Figure 5.12: The Shape of the Efficiency Frontier in the FDH Model

The impact of the FDH methodology has not been widely debated, particularly in relation to health service applications of DEA. It is not used in the analyses to follow.

5.9 Analysis of the DEA results using Statistical Methods

In the preceding discussion in this chapter introducing the various facets of the DEA technique, several alternative DEA models were developed using the data relating to gynaecology inpatient services. It has been observed that the slightly different efficiency results have been obtained from each of these models for a large number of the DMUs.

For example, table 5.4 showed the impact of using a number of different variable sets and table 5.8 compared the results from models using weight restrictions with the unrestricted CCR model. It would therefore be of value to determine if the observed differences in efficiency score, for example due to the inclusion of an additional factor, are statistically significant.

The traditional statistical analyses, such as the F-test and t-test, are based on an 'underlying assumption of normality' although they 'appear to be fairly insensitive to considerable departures from normality' (Keeping, 1962). Figure 5.13 overleaf shows the distribution of the efficiency scores for model G6. The efficiency scores do not fit the expected 'bell-shaped' pattern of the normal distribution, although the sample size is relatively small.

According to Keeping (1962), in cases such as this where it is not possible to assume that the data can be described by a normal distribution, the alternative is to use 'non-parametric or distribution-free' approaches.

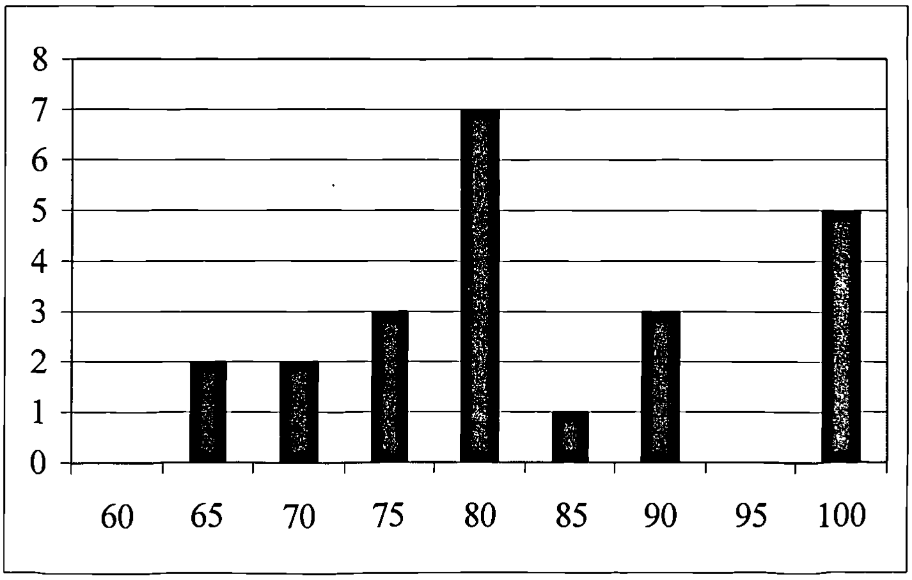


Figure 5.13: Distribution of the Efficiency Scores

Grosskopf and Valdmanis (1987), in an examination of hospital efficiency using DEA-type models, considered the difficulties associated with non-normality:

> "Are these results statistically significant? This is difficult to say using ordinary statistical tests. Since no parametric error structure was included in the model, and we have no reason to assume normal distributions, we use distribution-free tests to analyse these results."
>
> GROSSKOPF AND VALDMANIS (1987)

Similarly, Parkin and Hollingsworth (1996) observed the difficulties in examining the differences in efficiency scores obtained from a number of alternative models, reporting that rank correlation coefficients were the only possible means for comparison as the DEA scores did not fit a known distribution. However, as the approach of Parkin and Hollingsworth (1996) demonstrated, there are a number of possible alternative methods that can be used to investigate the nature of the efficiency results much more closely.

Three alternative methods are introduced below, with examples given of the type of analyses that are conducted in the discussion in later chapters. The traditional techniques, such as the t-test, are also used where appropriate.

5.9.1 Correlation Coefficients

The Spearman Rank Correlation Coefficient 'can be used to analyse the degree of association of two variables' (Black, 1992). Whereas the Pearson correlation statistic uses the raw data, this alternative uses the ranks, adding importance to the ranking assigned to each of the DMUs in a DEA analysis. Therefore, the rankings obtained for two alternative models could be assessed to determine if the inclusion of weight restrictions has an impact on the efficiency ratings assigned to each DMU. The Spearman Rank Correlation Coefficient is calculated as shown below:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

n = number of pairs being correlated

d = the difference in the ranks of each pair

Figure 5.14: The Spearman Rank Correlation Coefficient

Correlation was utilised by Parkin and Hollingsworth (1996) to investigate the impact of various variable sets on the efficiency scores of the DMUs they investgiated (see chapter six for a detailed discussion) and also to examine changes in efficiency score over a number of years.

5.9.2 The Mann-Witney U-test

This is the non-parametric equivalent of the t-test to determine if two independent samples are from the same population. It was utilised by Grosskopf and Valdmanis (1987) to investigate if the distribution of efficiency scores obtained for two different types of hospitals (defined by ownership type) were equivalent.

Data in the two groups under investigation, in this case the efficiency scores, are combined and ranked in increasing order, with the sum of the ranks calculated for the two samples separately (Keeping, 1962). The test statistic is shown in figure 5.16 below, being the smaller of the two quantities calculated:

$$U = N1 + N2 + \frac{N1(N1+1)}{2} - R1$$

$$U' = N1N2 - U$$

N_1, N_2: Sample Sizes for Samples 1, 2

R_1, R_2: Sum of the Ranks for Samples 1, 2

Figure 5.15: The Mann-Witney U-Test

If there are two or more equal values in either or both of the groups, 'the ranks associated with the tie are averaged across the values that tie' (Black, 1992). For large sample sizes, the U statistic can be approximated by the normal distribution, with the significance of the results determined using the normal probabilities. Alternatively, Mann-Witney tables are also available for a selected range of sample sizes. Computation in the following chapters is carried out using appropriate statistical packages, including Microsoft Excel and SPSS.

The Mann-Witney U-test could therefore be used to determine, for example, if the efficiency scores for large and small hospitals (classified using 'average number of staffed beds') are distributed in an equivalent pattern. Additionally, it could be used to determine if the changes in efficiency score found by introducing weight restrictions are the same for the inefficient and efficient hospitals in the unrestricted model.

5.9.3 Two-Tailed Z-test for Proportions

This approach was used by Ozcan *et al.* (1992) to determine if observed differences in efficiency scores for different types of hospitals were significant, with their analysis discussed in greater detail in section 6.2.2. The test statistic is shown in figure 5.16 below:

$$Z = \frac{p_1 - p_2}{\sqrt{(p_1.(1-p_1)/n_1) + (p_2.(1-p_2)/n_2)}}$$

p_1, p_2: *efficiency proportions for samples 1, 2*

n_1, n_2: *frequencies for samples 1,2*

OZCAN *ET AL.* (1992)

Figure 5.16: Two-tailed Z-test for Proportions

The significance of the calculated Z-statistic is tested against the 1% and 5% significance levels obtained from normal distribution tables ($p = \pm 2.58$ and $p = \pm 1.96$ respectively). However, application of the test is restricted, in that it 'should not be used for small samples, for example, fewer than 30 in each group, and proportions outside the range 0.1 and 0.9' (Gardner and Altman, 1989).

5.10 Summary: Perspectives on Data Envelopment Analysis

In section 5.3, the application of DEA to the measurement of health service efficiency was discussed. The intrinsic characteristics of the technique and the multitude of examples showing the use of DEA in the literature were cited as reasons for the inclusion of the technique into the 'health evaluation tool pack' (Nunamaker, 1983).

Throughout the chapter, the characteristics of the original methodology, the developments made over the past twenty years and the issues surrounding its application have been discussed in great detail. In the preceding chapter, the methods currently used in the assessment of performance in a health care environment were introduced. In relation to the limitations of these methods, the potential benefits of the DEA approach, as presented in this chapter, can be summarised in five key points:

5.10.1 Efficiency Measurement

DEA has been defined to measure the efficiency of the units under investigation, rather than any of the other elements of performance or performance as a whole, and therefore can be seen to provide an additional approach to performance evaluation.

5.10.2 Identification of Inefficiencies

The DEA methodology does not simply identify the inefficient units. It records the scale of the inefficiency (by assigning an efficiency score between 0 and 1) and identifies the exact areas where the inefficiencies are occurring, in relation to specific inputs and outputs. The use of the best-practice frontier ensures that changes are related to efficient rather than average DMUs.

5.10.3 Measuring External Influences

DEA can overcome the problems associated with different institutions operating in different environments by adding appropriate input variables, which may relate to the physical environment in which a unit operates or some aspects of its population, by the use of socio-economic indicators. Quality measures can also be specified in the outputs.

5.10.4 Inclusiveness

DEA has the potential to be inclusive because it can handle inputs and outputs of different dimensions. This means it can combine financial inputs and outputs, physical inputs and outputs and non-tangibles such as service quality indicators without requiring each to be measured on a single usually monetary scale of value. Including a comprehensive set of inputs and outputs, upon which the efficiency classification is based, can provide a more complete picture of performance than a single indicator. This is enhanced if weight restrictions are included.

5.10.5 Encompassing Local Policies and Innovations

The DEA model can be enhanced, through the introduction of restrictions to the factor weights or the measurement of other aspects of efficiency, to incorporate this type of issue into the model. A DMU that operates under particular constraints, or has adopted a particular policy in its service provision that is not automatically evidence of efficiency, can still be assessed favourably should it be appropriate.

In the chapters to follow, the theoretical framework presented above is translated into more extensive data samples and health care scenarios. Many of the theoretical developments, such as super efficiency, the measurement of allocative efficiency, additive and multiplicative models and the use of categorical variables, are beyond the scope of the further investigation. These have been presented in order to illustrate the extensive and varied nature of the DEA methodology.

Instead of focusing on these theoretical developments, the emphasis of the following chapters is on practical applications of the DEA methodology, in order to determine if the optimism of the DEA theorists in their perception of the technique can be supported in practical applications:

> "This relatively new weapon in the researcher's arsenal opens up unlimited, heretofore unconsidered opportunities, for examination of production efficiencies at a great many levels and settings for health care delivery." HOLLINGSWORTH AND PARKIN (1995)

5.11 References

1. **Allen, R., Athanassopoulos, A., Dyson, R. G. and Thanassoulis, E.** (1997) Weight Restrictions and Value Judgements in Data Envelopment Analysis: Evolution, Development and Future Directions, *Annals of Operations Research*, Vol. 73, pp. 13 – 34.

2. **Anderson, P. and Petersen, N. C.** (1993) A Procedure for Ranking Efficient Units in Data Envelopment Analysis, *Management Science,* Vol. 39, pp. 1261 – 1264.

3. **Ball, R. and Roberts, E.** (1998) The Relevance of DEA in the Public Sector - A Discussion Paper for Young OR 10, Young OR 10 Keynote Papers, Operational Research Society.

4. **Ball, R., Monaghan, C., Thomas, R. E. and Wagner, R.** (1997) Data Envelopment Analysis: A Practical Tool for Policy Makers?, University of Stirling Departmental Working Paper.

5. **Banker, R. D. and Morey R. C.** (1986a) The use of categorical variable in Data Envelopment Analysis, *Management Science*, Vol. 32, No. 12, pp. 1613 – 1627.

6. **Banker, R. D. and Morey, R. C.** (1986b) Efficiency Analysis for Exogenously Fixed Inputs and Outputs, *Operations Research*, Vol. 34, No. 4, pp. 513 – 521.

7. **Banker, R. D., Charnes, A. and Cooper, W. W.** (1984) Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis, *Management Science*, Vol. 30, No. 9, pp. 1078 –1092.

8. **Banker, R. D., Conrad, R. F. and Strauss, R. P.** (1986) A Comparative Application of Data Envelopment Analysis and Translog Methods: an Illustrative Study of Hospital Production, *Management Science*, Vol. 32, pp. 30 – 44.

9. **Banker, R. D., Charnes, A., Cooper, W. W., Swarts, J. and Thomas, D. A.** (1989) An Introduction to Data Envelopment Analysis With Some of its Models and Their Uses, *Research in Governmental and Nonprofit Accounting*, Vol. 5, pp. 125 – 163.

10. **Beasley, J.** (1990) Comparing University Departments, *Omega*, Vol. 18, pp. 171 – 183.

11. **Belton, V.** (1992) An IDEA – Integrating Data Envelopment Analysis with Multiple Criteria Decision Analysis, In Proceedings of the Ninth International Conference on Multiple Criteria Decision Making: Theory and Applications in Business, Industry and Commerce, **Goicoechchea, A., Duckstein, L. and Zionts, S.** (eds.), pp. 71 – 79, Springer-Verlag, Berlin.

12. **Belton, V. and Vickers, S. P.** (1993) Demystifying DEA – A Visual Interactive Approach Based on Multiple Criteria Analysis, *Journal of the Operational Research Society,* Vol. 44, No. 9, pp. 883 – 896.

13. **Bitran, G. R. and Valor-Sabatier, J.** (1987) Some Mathematical Programming Based Measures of Efficiency in Health Care Institutions, *Advances in Mathematical Programming and Financial Planning*, Vol. 1, pp. 61 – 84.

14. **Black, K.** (1992) Business Statistics An Introductory Course, West Publishng Co.

15. **Boussofione, A., Dyon, R. G. and Thanassoulis, E.** (1991) Applied Data Envelopment Analysis, *European Journal of Operational Research,* Vol.

16. **Bowlin, W. F., Charnes, A., Cooper, W. W. and Sherman, H. D.** (1985) Data Envelopment Analysis and Regression Approaches to Efficiency Assessment and Evaluation, *Annals of Operations Research,* Vol. 2(1), pp. 113 – 138.

17. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

18. **Campbell, E.** (1996) British Steel: Changes in Technical Efficiency and in Decision Making Behaviour, Paper Presented at the Young OR Conference, York.

19. **Capettini, R., Dittman, D. A. and Morey, R. C.** (1985) Reimbursement Rate Setting for Medicaid Prescription Drugs Based on relative Efficiencies, *Journal of Accounting and Public Policy*, Vol. 4, No. 2, pp. 83 – 110.

20. **Chalos, P. and Cherian, J.** (1995) An Application of Data Envelopment Analysis to Public Sector Performance Measurement and Accountability, *Journal of Accounting and Public Policy*, Vol. 14, No. 2, pp. 143 – 160.

21. **Charnes, A. and Cooper, W. W.** (1957) Management Models and Industrial Applications of Linear Programming, *Management Science,* Vol. 4, No. 1, pp. 38 – 91.

22. **Charnes, A. and Cooper, W. W.** (1962) Programming with Linear Fractional Functionals, *Naval Research Logistics Quarterly*, Vol. 9(3/4) pp. 181 – 185.

23. **Charnes, A. and Cooper, W. W.** (1985) Preface to Topics in Data Envelopment Analysis, *Annals of Operations Research*, Vol. 2(1), pp. 59 – 94.

24. **Charnes, A., Cooper, W. W. and Rhodes, E.** (1978) Measuring the Efficiency of Decision-Making Units, *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429 – 444.

25. **Charnes, A., Cooper, W. W., Golany, B., Learner, D. B., Phillips, F. Y. and Rousseau, J.** (1994a) A Multiperiod Analysis of Market Segments and Brand Efficiency in the Competitive Carbonated Beverage Industry, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (eds.) (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

26. **Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M. and Stutz, J.** (1985) Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions, *Journal of Econometrics*, Vol. 30(1/2), pp. 91 – 107.

27. **Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (eds.)** (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

28. **Charnes, A., Cooper, W. W., Rousseau, J. and Semple, J.** (1987) Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets, Research Report CCS 558, Centre for Cybernetic Studies, The University of Texas at Austin.

29. **Charnes, A., Cooper, W. W., Seiford, L. M. and Stutz, J.** (1982) A Multiplicative Model for Efficiency Analysis, *Socio-Economic Planning Sciences*, Vol. 16, No. 5, pp. 223 – 224.

30. **Charnes, A., Cooper, W. W., Seiford, L. M. and Stutz, J.** (1983) Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopment, *Operations Research Letters*, Vol. 2, No. 3, pp. 101 – 103.

31. **Charnes, A., Cooper, W. W., Wei, Q. L. and Huang, Z. M.** (1989) Cone-ratio Data Envelopment Analysis and Multi-objective Programming, *International Journal of Systems Science*, Vol. 20, pp. 1099 – 1118.

32. **Cook, W. D., Kazakov, A and Roll, Y.** (1994) On the Measurement and Monitoring of Relative Efficiency of Highway Maintenance Patrols, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

33. **Day, D., Lewin, A. Y., Li, H. and Salazar, R.** (1994) Strategic Leaders in the U.S. Brewing Industry: A Longitudinal Analysis of Outliers, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (eds.) (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

34. **Department of Health** (1997) 'The New NHS: Modern, Dependable', HMSO.

35. **Doyle, J. R. and Green, R. H.** (1994) Efficiency and Cross-Efficency in DEA: Derivations, Meanings and Uses, *Journal of the Operational Research Society*, Vol. 45, No. 5, pp. 567 – 578.

36. **Doyle, J. R. and Green, R. H.** (1995) Cross-Evaluation in DEA: Improving Discrimination Among DMUs, *INFOR*, Vol. 33, Pt. 3, pp. 205 – 222.

37. **Dyson, R. G. and Thanassoulis, E.** (1988) Reducing Weight Flexibility in DEA, *Journal of the Operational Research Society*, Vol. 39, No. 6, pp. 563 – 576.

38. **Dyson, R. G., Thanassoulis, E. and Boussofiane, A.** (1990), Data Envelopment Analysis, in "Operational Research Tutorial Papers", L.C. Hendry and R. Eglese (eds), The Operational Research Society, pp. 13 – 28.

39. **Ehreth, J. L.** (1994) The Development and Evaluation of Hospital Performance Measures for Policy Analysis, *Medical Care*, Vol. 32, No. 6, pp. 568 – 587.

40. **Farrell, M. J.** (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A (General)*, Part III, Vol. 120, pp. 253 – 279.

41. **Farrell, M. J. and Fieldhouse, M.** (1962) Estimating Efficient Production Functions Under Increasing Returns to Scale, *Journal of the Royal Statistical Society Series A (General)*, Part II, pp. 252 – 267.

42. **Feldstein, M. S.** (1967) Economic Analysis for Health Service Efficiency, North Holland.

43. **Finkler, M. D. and Wirtschafter, D. D.** (1993) Cost-Effectiveness and Data Envelopment Analysis, *Health Care Management Review*, Vol. 18, No. 3, pp. 81 – 88.

44. **Gardner, M. J. and Altman, D. G.** (1989) Statistics with Confidence – Confidence Intervals and Statistical Guidelines, British Medical Journal.

45. **Grosskopf, S. and Valdmanis, V.** (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, Vol. 6, No. 2, pp. 89 – 107.

46. **Hogan, A. J., Chesney, J., Wroblewski, R. and Fleming, S.** (1987) The Impact of the Medicare Prospective Payment System on Hospital Efficiency: a Data Envelopment Analysis, *Paper Presented to the American Economic Association Annual Meeting*, Chicago.

47. **Hollingsworth, B. and Parkin, D.** (1995) The Efficiency of Scottish Acute Hospitals: An Application of Data Envelopment Analysis, *IMA Journal of Mathematics Applied in Medicine and Biology*, Vol. 12, pp. 161 – 173.

48. **Huang, Y. L. and McLaughlin, C. P.** (1989) Relative Efficiency in Rural Primary Health Care: an Application of Data Envelopment Analysis, *Health Services Research*, Vol. 24, No. 2, pp. 143 – 158.

49. **Information and Statistics Division (ISD)** (1996) Scottish Health Service Costs 1995 – 96, ISD, NHS in Scotland.

50. **Keeping, E. S.** (1962) Introduction to Statistical Inference, Van Nostrand.

51. **Kleinsorge, I. and Karney, D.** (1992) Management of Nursing Homes Using Data Envelopment Analysis, *Socio-Economic Planning Science*, Vol. 26, pp. 57 – 71.

52. **Koopmans, T. C.** (1951) Analysis of Production as an Efficient Combination of Activities, In Koopmans, T. C. (ed.) Activity Analysis of Production and Allocation, Wiley.

53. **Lewin, A. and Morey, R. C.** (1981) Measuring the Relative Efficiency and Output Potential of Public Sector Organizations: An Application of Data Envelopment Analysis, *Journal of Policy Analysis and Information Systems*, Vol. 5, No. 4 (December 1981), pp. 267 – 285.

54. **Lovell, C. A. K., Walters, L. C. and Wood, L. L.** (1994) Stratified Models of Education Production Using Modified DEA and Regression Analysis, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

55. **MacMillan, W. D.** (1987) The Measurement of Efficiency in Multiunit Public Services, *Environment and Planning*, Vol. 19, pp. 1511 – 1524.

56. **Morey, R. C., Capettini, R. and Dittman, D. A.** (1985) Pareto Rate Setting Strategies: an Application to Medicaid Drug Reimbursement, Policy Sciences, Vol. 18, No. 2, pp. 169 – 200.

57. **Norman, M. and Stoker, B.** (1991) DEA: The Assessment of Performance, Wiley.

58. **Nunamaker, T.** (1983) Measuring Routine Nursing Service Efficiency: A Comparison of Cost Per Patient Day and Data Envelopment Analysis Models, *Health Services Research*, Vol. 18, No. 2, pp. 183 – 205.

59. **Nunamaker, T.** (1985) Using Data Envelopment Analysis to Measure the Efficiency of Non-profit Organizations: a Critical Evaluation, *Managerial and Business Economics*, Vol. 6, No. 1, pp. 50 – 58.

60. **Nyman, J., Brick, D. and Link, D.** (1990) Technical Efficiency in Nursing Homes, *Medical Care*, Vol. 28, No. 6, pp. 541 – 551.

61. **Ozcan, Y. A., Luke, R. and Haksever, C.** (1992) Ownership and Organizational Performance: a Comparison of Technical Efficiency Across Hospital Types, *Medical Care*, Vol. 30, pp. 781 – 794.

62. **Parkin, D. and Hollingsworth, B.** (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991 – 1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

63. **Pedraja-Chaparro, F., Salinas-Jimenez, J. and Smith, P.** (1999) On the Quality of the Data Envelopment Analysis Model, Journal of the Operational Research Society, Vol. 50, No. 6, pp. 636 – 644.

64. **Roll, Y. and Golany, B.** (1993) Alternate Methods for Treating Factor Weights in DEA, *Omega*, Vol. 21, pp. 99 – 109.

65. **Roll, Y., Cook, W., and Golany, B.** (1991) Controlling Factor Weights in DEA, *IIE Transactions*, No. 23, pp. 2 – 9.

66. **Sczcepura, A., Davies, C., Fletcher, J. and Boussofiane, A.** (1993) Efficiency and Effectiveness in General Practice, *Journal of Management in Medicine*, Vol. 7, no. 5, pp. 36 – 47.

67. **Sexton, T. R., Silkman, R. H. and Hogan, A.** (1986) Data Envelopment Analysis: Critique and Extensions in **Silkman, R. H.** (ed), Measuring Efficiency: An Assessment of Data Envelopment Analysis, Publication No. 32 in the Series New Directions for Program Evaluation, A Publication of the American Evaluation Association, Jossey Bass, Inc.

68. **Sherman, H. D.** (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927 – 938.

69. **Smith, P. and Mayston, D.** (1987) Measuring Efficiency in the Public Sector, *Omega*, Vol. 15, No. 3, pp. 181 – 189.

70. **Thanassoulis, E., Boussofiane, A. and Dyson, R. G.** (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588 – 608.

71. **Thanassoulis, E., Dyson, R. G. and Foster, M. J.** (1987) Relative Efficiency Assessments Using Data Envelopment Analysis: an Application to Data on Rates Departments, *Journal of Operational Research Society*, Vol. 38, No. 5, pp. 397 – 411.

72. **Thompson, R. G., Langemeier, L. N., Lee, C. T., Lee, E. and Thrall, R. M.** (1990) The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas Farming, *Journal of Econometrica*, Vol. 46, pp. 93 – 108.

73. **Thompson, R. G., Singleton, F. D. Jnr., Thrall, R. M. and Smith, B. A.** (1986) Comparative Site Evaluations for Locating a High-Energy Physics Lab in Texas, Interfaces, Vol. 16, No. 6 (December), pp. 35 – 49.

74. **Tofalis, C.** (1993) Estimating the Efficiency with which Inputs are Utilised, University of Hertfordshire Business School Working Paper Series, UHBS 1993:5.

75. **Tomkins, C. and Green, R.** (1988) An Experiment in the Use of Data Envelopment Analysis for Evaluating the Efficiency of UK University Departments of Accounting, *Financial Accountability and Management*, Vol. 3, No. 4, pp. 3331 – 342.

76. **Tulkens, H.** (1993) On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit, *Journal of Productivity Analysis*, Vol. 4(1 – 2), pp. 183 – 210.

77. **Valdmanis, V.** (1990) Ownership and Technical Efficiency of Hospitals, *Medical Care*, Vol. 28, No. 6, pp. 552 – 561.

78. **Wilkinson, R.** (1991) The Measurement of University Performance: an Application of Data Envelopment Analysis, PhD. Thesis.

79. **Wong, Y-H. B. and Beasley, J.** (1990) Restricting Weight Flexibility in DEA, *Journal of the Operational Research Society*, Vol. 41, No. 9, pp. 829 – 835.

CHAPTER SIX

APPLYING DEA TO HEALTH CARE ORGANISATIONS: INFORMATION FROM THE LITERATURE AND HEALTH CARE MANAGERS

6.1 Introduction

The information presented in this chapter has been obtained from two distinct sources and is intended to present the differing perspectives on the measurement of hospital efficiency. The first section provides a detailed summary of some of the investigations into hospital efficiency, as located in the literature, in order to illustrate the nature and type of analyses undertaken. This follows on from the brief discussion in chapter five, confirming the applicability of DEA to performance evaluation in a health service context.

The second main element of this chapter focuses on the perspective of the health care manager, highlighting the aspects of the DEA methodology that are most problematic to its general acceptance. These observations are intended to contrast with the evaluation of the DEA methodology given in the conclusion to chapter five. They are also used to guide the application of DEA in the following chapters.

An extensive bibliography of health care applications of the DEA methodology is also included (see section 6.5), illustrating the variety of investigations carried out and the scope of the DEA approach. Information from a number of these papers is also used to guide the application process in later chapters.

6.2 Information from the Literature

Seven key papers have been identified for inclusion in this section, with the focus on identifying the important stages involved in an analysis of hospital efficiency. The key themes obtained from the discussion are summarised in section 6.2.8.

6.2.1 Ehreth (1994)

The analysis in this paper was focused on the development and evaluation of 'hospital performance measures to include aspects of hospital behaviour beyond the traditional use of hospital profit margins for policy analysis' (Ehreth, 1994). The use of DEA models formed some part of the analysis presented, with all hospitals in the USA for which Medicare cost reports and balance sheet data were available included in the sample. A method developed by federal policy analysts was used to break down the large sample of hospitals into smaller, more homogeneous groupings, according to 'size, geographic region, proximity to an urban centre, and teaching status' (Ehreth, 1994). As a starting point, a number of DEA models were defined using the CCR methodology, with the impact of the different variable mixes to be investigated. The factors included were selected from the following lists, the data for which was all available in the annually published Medicare Cost Report:

> INPUTS - Full-time Equivalent Employees, Total Other Direct Costs, Direct Salary Costs, Number of Beds (Number of Coronary Care Unit (CCU) Beds, Number of Intensive Care Unit (ICU) Beds, Number of Other Beds), Fixed Assets (CCU Fixed Assets, ICU Fixed Assets, Other Fixed Assets)
>
> OUTPUTS - Inpatient Total Days (ICU Total Days, CCU Total Days, Other Total Days) or (Medicare Inpatient Days, Other Inpatient Days), Total Discharges (Medicare Discharges, Medicaid Discharges, Other Discharges), Total Patient Revenue, Adjusted Total Discharges (Adjusted Medicare Discharges, Adjusted Medicaid Discharges, Adjusted Other Discharges), Outpatient Discharge Equivalent.

Ehreth (1994) originally developed seventeen different DEA models, using different combinations of the inputs and outputs presented above, as well as numerous other performance measures based on simple ratios.

More than fifty indicators of this type had been identified from the literature, although several potential measures had to be excluded prior to the analysis due to a lack of available data for all Decision-Making Units (DMUs).

Only one of the efficiency measures identified, using both DEA and ratio analysis, was acceptable to the policy analysts with whom the alternative models were discussed. This was based on its applicability to policy analysis and link with hospital efficiency measurement. The DEA model was found to be superior to the ratio measures, which were seen to be less applicable for the measurement of efficiency, as well as having other limitations.

The variables included in the acceptable DEA model, denoted by FTE_MDA, were:

> *Inputs* - Fixed Assets, Full-Time Equivalent Employees.
> *Outputs* - Adjusted Medicare Discharges, Medicaid Discharges, and Other Discharges.

However, the policy analysts also felt that this one measure had several limitations and proposed that two further DEA models should be developed. They suggested that an acceptable model would be one that incorporated some form of adjustment for case-mix and/or a measure of outpatient activity in the outputs and two measures of inputs to reflect hospital size and structure. Two new models were constructed, therefore, which had the same inputs as used in FTE_MDA defined above (Fixed Assets, Full-time Equivalent Employees). There was a slight variation in the output variables selected, as shown below, with the second of the two models referred to as a 'technical efficiency measure' (TEM):

FTE_DASO: Adjusted Medicare Discharges, Adjusted Medicaid Discharges, Adjusted Other Discharges, Outpatient Discharge Equivalent.

TEM: Adjusted Medicare Discharges, Medicaid Discharges, and Other Discharges, Outpatient Discharge Equivalent.

The analysis of the efficiency scores, calculated for each hospital using the different DEA models, took several forms. Initially, the different DEA models were evaluated by looking at the impact of changes in the variable set on the average and individual efficiency scores. For example, disaggregating the discharges factor into three categories of patient based on payment category (Medicare, Medicaid or Other) had a significant effect on a sample of large urban hospitals, with one of the DMUs increasing its efficiency score by over 40%.

As the above example illustrates, a further assessment of the overall sample and the DEA models was made according to various hospital groupings. The overall sample of hospitals was broken down into smaller clusters, defined by several factors, such as size, rural/urban and teaching status. The number of DMUs in each of the smaller samples was very varied, according to the level of clustering applied. Average efficiency scores were calculated and compared for the different types of hospital within each region. For example, for the 568 hospitals in the Middle Atlantic Region, the medium-sized, non-teaching urban hospitals had the highest average efficiency scores.

Finally, the efficiency of the hospitals over a three-year time span was investigated. Differences were observed, in that some groups increased their average score from one time period to the next, whilst others showed a significant decrease.

For example, the average efficiency of large rural hospitals in the Middle Atlantic region was seen to have decreased, whereas the equivalent hospitals in the South Atlantic region showed an increase in their average efficiency score.

The conclusions of the study in relation to the position of DEA found that the TEM model defined above was the most robust of all the DEA and non-DEA measures evaluated. The sensitivity of the DEA measure to errors and omissions in the data was seen as a disadvantage of the approach. The difficulty in comparing efficiency across the hospital groupings using this type of methodology was also noted, as were the complications involved in explaining the DEA theory to hospital administrators.

However, the benefits of using the DEA model in conjunction with other financial measures was seen as a major advantage, as 'efficiency problems could be identified using data available at organisational level' and 'more efficient ways to spend our health care dollars' could be determined (Ehreth, 1994).

6.2.2 Ozcan, Luke and Haksever (1992)

As with the previous study, the focus of this paper was on the impact of hospital type, in terms of ownership, on the technical efficiency of a sample of urban hospitals. 3,000 acute care general hospitals were identified in the sample taken from the American Hospital Association Annual Survey, which could be divided up into groupings based on the standard metropolitan statistical areas defined in the USA, of which there are 317. In the instances where the groupings contained less than 13 hospitals, with that number selected arbitrarily as leading to too few hospitals in a grouping, two or more groupings were combined.

The groupings chosen for combining were those in a similar population category and geographical region. The necessity for pooling some of the groupings led to the formulation of 82 distinct groups within the sample of 3,000 hospitals.

The investigation was based on the original DEA methodology, using the unrestricted CCR model with constant returns to scale. Their primary reason for selecting the DEA methodology, thought to be its greatest advantage over the other methods used previously in the measurement of efficiency in health services, was its ability to directly incorporate multiple inputs and outputs.

The purpose of the study was to carry out a general evaluation of efficiency using DEA, investigating the usefulness and diversity of the results produced by the DEA models. There was little attention given to the inclusion of managerial perspectives in the modelling processes or on the views of health care managers as the value of a DEA approach to health care evaluation.

Inputs were selected in four categories, to reflect service complexity, hospital size, personnel requirements and supply expenses. The factors recording hospital size and service complexity were used as proxy measures of capital assets, for which the data was not available for all hospitals.

The selection of the output variables was based on the premise that the hospitals in the sample produced three distinct types of output: inpatient discharges, outpatient activity and medical trainees. The inpatient output factor was adjusted for case-mix complexity, using the Medicare case-mix index, available for each of the hospitals.

The index was seen to have some limitations as an adequate measure of the impact of case-mix, although no alternative indicator was available to the researchers. Due to a lack of available data, it was not possible to include other aspects of hospital output, such as their research and community activities.

The specific input and output factors were thus:

> INPUTS: Total Number of Inpatient and Outpatient Diagnostic and Special Services, Number of Operational Hospital Beds, Number of Non-Physician Full-time and part-time Personnel, Amount of Operational Expenses (Excluding Payroll, Capital and Depreciation Expenses).
>
> OUTPUTS: Case-Mix Adjusted Inpatient Discharges, Outpatient and Emergency Visits, Weighted Sum of Medical, Dental and Other Professional Trainees.

The main focus of the analysis was on the assessment of efficiency according to hospital ownership, with the three categories of ownership being government, non-profit and for-profit. This was to determine if the different types of hospitals 'differ significantly in their relative technical efficiencies' (Ozcan *et al.*, 1992).

The sample of 3,000 DMUs was divided into peer groups, in order to control for 'local environmental variations' (Ozcan *et al.,* 1992), which was defined as a 'restricted peer group strategy'. The control groups were defined according to market size, geographical region, system membership and hospital size. Thus, the DEA score for each DMU was calculated four times, with the sample size and structure varying in each case.

Analysis of the efficiency scores across the peer groups, however, was thought to be problematic, as it emphasised that DEA is a measure of relative efficiency and not absolute efficiency:

> ".. we compare efficiency scores that are not absolute, but relative to the levels of performance achieved within each local area. Thus, hospital found to be inefficient relative to its own peers may not necessarily be inefficient relative to hospitals in other areas that receive within their own areas." OZCAN *ET AL.* (1992)

Despite this reservation, this approach was utilised and analysis concentrated on the percentage of the hospitals classified as either efficient or highly inefficient, compared across 'ownership types and within control variable categories' (Ozcan *et al.,* 1992). The significance of the results was investigated statistically, based on 'pair-wise comparisons of these percentages using the two-tailed z-test for proportions' (Ozcan *et al.*, 1992).

In relation to ownership type, the overall percentage of DMUs being classified as efficient ranged from 36.5% (for-profit hospitals) to 57.1% (government hospitals). This was found to be statistically significant ($p < 0.01$). However, it was found that 'the observed differences between ownership categories are fairly consistent across the control variables' (Ozcan *et al.*, 1992). For example, the percentage of government hospitals found to be efficient was significantly higher than for the other hospital types across almost all of the other control variables.

The percentage of hospitals classified as efficient was also investigated within the control groups. Looking at the government hospitals as an example, 65.2% of the small hospitals (less than 100 beds) were found to be efficient, although only 42% of the medium-sized hospitals (100 – 349 beds) were similarly classified.

Further analysis on the contribution made by each variable to the calculation of the efficiency scores was also included. It was suggested, for example, that the inefficiencies identified in government hospitals could be related to an underproduction of discharges relative to the other hospital types, linked to the longer lengths of stay traditionally associated with government hospitals.

The authors believed that the DEA approach enabled them to draw conclusions on the differences in efficiency amongst the hospital types, with a higher percentage of government hospitals found to be efficient, with the proviso that the efficiency scores calculated were relative and not absolute. The impact of ownership on efficiency is a major source of debate, simply due to the diversity of ownership types within the US health care system. The lack of an indicator of quality was identified as problematic, but the technique of DEA was perceived to have much to offer in this type of analysis.

Whilst not directly related to the NHS in the UK, where ownership is not an issue at present, the various investigations in this field can provide useful guidelines for using DEA with data from other health care systems. The survey discussed how DEA could be used to examine different types of managerial practice and also the impact of regional and environmental influences on efficiency evaluations.

6.2.3 Sherman (1984)

This is one of the earliest investigations using DEA in relation to health care data, evaluating a small sample of just seven teaching hospitals in Massachusetts and with the focus on the medical-surgical (MS) areas rather than the performance of the whole hospital.

Data Envelopment Analysis was used to investigate elements of hospital performance in this case, as it was seen to address 'the limitations associated with ratio analysis and regression techniques' (Sherman, 1984). As with the previous discussion, one of the primary reasons for the use of DEA was the inclusion of multiple inputs and outputs, particularly in relation to case-mix complexity.

The DEA model selected was again the CCR model as defined in section 5.4, with the free allocation of the factor weights seen as a major strength of the methodology. Sherman (1984) argued that hospitals identified as inefficient could be considered to be truly inefficient, as the pattern of weights applied had been selected to maximise their efficiency score.

The sample used in the analysis was again selected using some form of clustering, using the comparable groups defined by the state rate-setting commission. The variable selection was restricted by the limited availability of some data. The major discussion in the selection of outputs focused on how the number of inpatient days should be disaggregated to accurately represented case-mix complexity. Age was selected as the defining factor with some reservations, as the DEA results could be skewed because the impact of case mix complexity was not reflected. As the hospitals in the sample were defined to be comparable using the clustering technique, it was hoped that this effect would not be too apparent. The factors were identified and accepted in conjunction with a number of health-care experts, with the final selection being as follows:

> INPUTS: Full-time Nonphysician Equivalents Employed in MS Area, Total Dollar Value of Supplies, Number of Bed Days Available in MS Area.

OUTPUTS: Number of Patient Days of Care in MS Area (Patients Aged 65 and Over, Patients Aged Under 65), Number of Nursing Students Trained, Number of Interns and Residents.

In the small sample, five of the seven DMUs were in fact identified as being efficient. Thus, the focus of the analysis was on a specific hospital that had been shown to be inefficient. The value of the analysis was perceived to be in the fact that the two hospitals identified as inefficient, with the nature of the inefficiencies also characterised, would not have been so identified using the more traditional ratio analysis.

One of the hospitals shown to be inefficient in its provision of medical-surgical services was found to have an unusually high number of staff, in comparison with the rest of the sample, which was the primary cause of its inefficiency rating.

As with all stages of the research, the input from health care experts was used to analyse the results from the DEA model. In the case of this inefficient DMU, two senior managers identified the cause of its apparent inefficiency as a deliberate hospital policy to provide 'more personalised patient care' (Sherman, 1984).

6.2.4 Grosskopf and Valdmanis (1987)

The majority of the other papers discussed in this section have relied on the traditional DEA measurements of technical efficiency, calculated using the CCR model. However, this investigation (and the one to follow by Byrnes and Valdmanis (1994)) considered other elements of efficiency and alternative formulations of the DEA methodology. Data from a set of urban hospitals within the state of California was used to illustrate the facets of the efficiency measures being investigated.

These were developed from the theories proposed by Farrell (1957) and adapted by Fare *et al.* (1985). Two types of model were investigated, to investigate the influence of constant or variable returns to scale (CRS or VRS). This was equivalent to using either the CCR or BCC (defined in section 5.6.1) form of the DEA methodology.

The sample investigated contained 82 DMUs, restricted to hospitals with 'at least 200 beds in metropolitan areas with population 500,000 or more' (Grosskopf and Valdmanis, 1987). 60 of the DMUs in the sample were private, not-for-profit hospitals, with the remaining 22 units being public hospitals. The sample was restricted to hospitals in just one state in order to avoid differences in regulation, unlike the samples selected by Ehreth (1994) and Ozcan *et al.* (1992) that included hospitals from across the USA.

As with many other surveys using samples from the USA, the source of the data was the American Hospital Association Survey for Hospitals. The choice of outputs reflected the authors' attempts to incorporate 'an array of outputs which are assumed to be related to improved health status' (Grosskopf and Valdmanis, 1987). The data for inpatient activity was measured in 'days' rather than 'discharges', although an additional output relating to intensive care activity was included to reflect the differences in resource use in acute care.

In order to measure capital costs, 'net plant assets' was used as an input, whilst staff numbers in various categories were used to measure labour. 'Number of admissions' was included as an input in the additional analysis, although its removal was shown to have no significant impact on the efficiency results obtained.

The variables used for the analysis were as follows:

> INPUTS: Number of Physicians, Full-time Equivalent Employment of Non-Physician Labour, Number of Admissions, Net Plant Assets.
> OUTPUTS: Acute Inpatient Days, Intensive Care Inpatient Days, Inpatient Surgeries, Outpatient Surgeries, Ambulatory and Emergency Care Visits.

The variation in activity levels at each of the hospitals was investigated, by identifying the mean, standard deviation, minimum and maximum values for each of the factors included in the model. These statistics were calculated for the sample as a whole, then for the two types of hospitals individually. The range between minimum and maximum was quite considerable for the majority of the variables, suggesting that the hospitals varied in size quite extensively. The sample was also partitioned for the calculation of the efficiency scores, similar to the restricted peer group approach used to Ozcan *et al.* (1992).

In this case, efficiency scores were calculated for the sample as a whole and for the two types of hospital individually. In the case with the sample treated as a whole, it was found that over 90% of the hospitals were located on the efficiency frontier.

Additionally, the average technical efficiency score (most closely related to the CCR model with constant returns to scale) was found to be 94.2% for the public hospitals and 90.1% for the not-for-profit hospitals. In the varying returns to scale model, the average efficiency scores followed the same pattern as above, with the value for public hospitals approximately 3% higher than the not-for-profit mean value.

Investigating the two types of hospital separately resulted in the mean value for both efficiency measures increasing, with the average at 99% for the public hospitals, using the varying returns to scale model. The variation in efficiency scores for the two types of hospitals led to three conclusions being drawn by Grosskopf and Valdmanis (1987):

1. The performance of public hospitals differed from not-for-profit hospitals when it was assumed that they faced the same 'best practice' frontier;
2. Public hospitals were closer, on average, to their own separate frontier, that is, more efficient on average;
3. The best of the public hospitals were generally different, in terms of their efficiency, from the best of the not-for-profit hospitals even when they had their own frontiers.

The differences observed in average efficiency scores were relatively small in percentage terms and it was necessary to employ statistical measures to investigate the validity of their conclusions, as was discussed in section 5.9. Using the non-parametric Mann-Witney test, they additionally investigated whether the efficiency scores obtained for the two types of hospital were significantly different, under a variety of hypotheses.

The first null hypothesis, which stated that the efficiency distributions were the same for the two types of hospital investigated as a whole, was rejected with a 10% significance level. This was valid for both types of efficiency measure (with constant and varying returns to scale), confirming the first of the conclusions drawn above. Similar applications of the Mann-Witney test resulted in the other two points being verified, again using a 10% level of significance.

The observations on hospital ownership cannot be directly translated to the provision of health services in the UK, as was referred to in the discussion on Ozcan *et al.* (1992). However, the type of analysis has value, in that hospitals in Scotland fall into several categories that can be based on the services provided by them, their relative size and the region in which they are situated.

6.2.5 Byrnes and Valdmanis (1994)

This paper used a sample of 123 community (nonteaching) not-for-profit hospitals in California, using a combination of DEA and other efficiency measures, some of which were based on the Farellian definitions of efficiency, found in Farrell (1957), and all of them were based on a cost-minimising approach. The analysis was an extension to the more basic approaches to hospital efficiency evaluation, as the models were developed to go beyond the traditional measurement of technical efficiency. Therefore, the 'overall cost-minimizing efficiency measure is decomposed into the allocative-efficiency component and three technical-efficiency components' (Byrnes and Valdmanis, 1994).

The element relating to allocative efficiency was incorporated to measure if the hospitals were selecting the 'right' mix of inputs. It was suggested that focusing solely on technical efficiency could result in invalid conclusions being drawn on the appropriate steps for the individual hospitals to achieve overall efficiency. The various elements of efficiency were discussed in section 2.3, with its implications on the efficiency assessment in a health service context highlighted in section 2.5.4. The decomposition of the model, measuring allocative and technical efficiency, also allowed for an investigation into the nature of the returns to scale, as would be the case if the BCC model introduced in section 5.6 was used.

As referred to above, the sample investigated contained only not for profit hospitals, as the authors were mindful of the different operating structures for the various types of hospital in the USA, as discussed in the previous section. The sample was further constrained to exclude teaching hospitals, as an alternative to including an index to measure case-mix complexity. Additionally, they elected to focus on 'the production of inpatient care, because it composed the largest component of hospital costs' and also because the data was in an appropriate format (Byrnes and Valdmanis, 1994).

Unlike the previous paper (Grosskopf and Valdmanis, 1987), the inpatient activity was specified as 'discharges' rather than 'days' in order to avoid confusing 'differences in efficiency with occupancy rates' (Byrnes and Valdmanis, 1994). They did use the distinction between acute and intensive care activity, also specified in Grosskopf and Valdmanis (1987), to take into account differences in the use of resources.

Analysis of the output data (the calculation of mean, standard deviation, maximum and minimum) also revealed that there were some zero values in the output data, as not all the hospitals provided maternity and ICU services. However, these factors were still incorporated into the model.

The numbers of staff in five categories were used to measure personnel, rather than using financial measures. This was a much larger degree of disaggregation than utilised in the studies discussed in the preceding sections, although it corresponds quite closely to the input selection of Hollingsworth and Parkin (1995), debated in the next section. In order to incorporate a measure of capital, the number of staffed beds was also included. The factors included in the efficiency models are listed below:

INPUTS: Number of full-time Equivalence Hours for Registered Nurses (RN), Management and Administrative Personnel (MGT), Technical Services Personnel (TECH), Aides and Orderlies (AIDE) and Licensed Practical Nurses (LPN); Average Number of Staffed Beds (BEDS).

OUTPUTS: Medical-Surgical Acute Discharges (ACU); Medical-Surgical Intensive Care Discharges (ICU); Maternity Discharges (MAT).

Analysis of the overall efficiency scores, which measured both technical and allocative efficiency, found that just 6 of the 123 hospitals in the sample were identified as efficient, a much smaller percentage of efficient DMUs than in any other of the samples investigated in this section. The average overall efficiency score was 61%, suggesting that 'on average, inefficient hospitals would have needed to lower operating costs by 39% in order to perform as well as other similar, best practice hospitals in the sample' (Byrnes and Valdmanis, 1994).

The level of allocative efficiency was identified separately, with an average efficiency score in this model of 73% and just six allocatively efficient DMUs. However, 49 of the hospitals were found to be efficient in the technical efficiency model, which was the closest model in its formulation to the basic CCR model introduced in section 5.4. In this case, the average technical efficiency score was 84%, with the minimum individual score being 10%. Thus, whilst many of the hospitals were transforming inputs into outputs efficiently, they still demonstrated allocative inefficiency, in that they employed an inappropriate mix of inputs in order to minimise overall costs.

Further analysis of the results from the efficiency models focused on the nature of the returns to scale. This determined that the nature of the returns to scale was linked to hospital size (using 'average number of discharges' as the distinguishing factor), suggesting that 'the sample hospitals follow the traditional U-shaped average cost curve' (Byrnes and Valdmanis, 1994). Thus, the small hospitals demonstrated increasing returns to scale on average. The average efficiency score in the scale model was found to be 94%, noticeably higher than for any of the other models. Attempts to validate the results by comparing them with traditional measures of hospital performance were included, primarily to assure hospital managers that the results calculated were accurate guides to the technical inefficiency of the hospitals. It was found, for example, that the hospitals classified as technically and allocatively inefficient had a much higher average value for their average cost per discharge.

In conclusion, the authors proposed three areas where they felt that the results obtained from this type of efficiency analysis would be most beneficial to hospital managers:

1. To identify if their hospital was using too many or the wrong mix of inputs;
2. To determine how well their hospital was competing for patients;
3. To calculate how much to charge for the services they offered in the future.

The investigation of allocative efficiency is beyond the scope of this research, as was discussed in sections 2.3.3 and 5.8.1. However, some of the approaches discussed above still have some relevance to the discussion to follow, particularly in terms of the variable selection process, the implication of returns to scale and the observations on overall efficiency.

6.2.6 Hollingsworth and Parkin (1995) and Parkin and Hollingsworth (1996)

These two papers were among the first to directly focus upon the application of DEA to hospital evaluation using data from the UK. (Other authors had applied to DEA to other aspects of the health service, such as Norman and Stoker (1991), Thanassoulis *et al.* (1995), Salinas-Jiménez, J. and Smith, P. (1996) and Scezepura *et al.* (1993).) Both these papers were primarily concerned with an investigation into the applicability of the DEA technique, using information about acute hospitals in Scotland primarily as a source of data. The sample comprised data on the 75 hospitals classified as acute by the Information and Statistics Division (ISD) of the Scottish Office, with the Scottish Health Service Costs used as the source of the data. The focus of the discussion in relation to the type of DEA methodology to be utilised was on the measurement and importance of the nature of the returns to scale for each of the DMUs.

Therefore, initial applications of the model were restricted to the constant returns to scale assumption of the CCR model, using the methodology of Charnes, Cooper and Rhodes (1978). The inclusion of the varying returns to scale option (Parkin and Hollingsworth, 1996), using the BCC model defined in chapter five, was seen to be beneficial to the individual hospitals in that it allowed them to identify scale and technical inefficiencies.

Corresponding to the view expressed by Grosskopf and Valdmanis (1987), the range of outputs included in the model used by Hollingsworth and Parkin (1995) were seen to be a proxy for the 'ideal' output measuring a change in health status. They did, however, choose to measure inpatient activity in terms of 'days' rather than discharges. In terms of inputs, the factors to measure personnel were recorded as the number of staff in various categories, not as financial measures.

The inputs were similar to those used by Byrnes and Valdmanis (1994). The only measure of financial input was 'capital charge'. The twelve factors selected, which were combined in a variety of ways to form a major part of the analysis, are shown below:

> INPUTS: Average Number of Staffed Beds; Total Number of Trained, Learning and Other Nurses; Total Number of Professional, Technical, Clerical and Administrative Staff; Total Number of Junior and Senior Non-Nursing Medical and Dental Staff; Total Cost of Drug Supply; Capital Charge.
>
> OUTPUTS: Total Number of Medical Inpatient Days; Total Number of Surgical Inpatient Days; Total Number of Obstetrics and Gynaecology Inpatient Days; Total Number of Other Inpatient Days; Total Number of Outpatient Attendances; Total Number of A&E Attendances.

The later paper, Parkin and Hollingsworth (1996), revised the outputs so that 'discharges' were used as a measure of inpatient activity, as opposed to 'days' used in the original models. The change was made in accordance with the argument presented by Feldstein (1967) that 'cases treated are the correct unit of output, and that measures which incorporate length of stay penalise possible substitution between length of stay and intensity of resource use' (Parkin and Hollingsworth, 1996).

The investigations in both papers focused on the impact of different variable mixes. Thus, the variables reflecting staffing levels or inpatient activity were aggregated in additional formulations of the model. This approach was undertaken to 'allow for hospitals that do not deal with all specialties' or 'may not employ staff in some of the disaggregated categories' (Parkin and Hollingsworth, 1996).

In relation to the data set being investigated in the earlier paper, they found more justification for the aggregation of the outputs rather than the inputs, in terms of the impact on the level of discrimination in the efficiency scores and the reduced amount of information obtained.

In the later paper (Parkin and Hollingsworth, 1996), the average efficiency score in the model with all 12 variables included was 91.5%, with 38 DMUs out of 75 rated as efficient. Aggregating the inpatient outputs, with the resultant model having just 9 factors, led to an average efficiency score of 79.1% and 17 efficient DMUs in the sample. A third model, with six outputs but four inputs (the variables measuring staff levels were aggregated) gave an average efficiency score of 84.0%, with 27 efficient DMUs.

The impact of removing one of a pair of factors found to highly correlated, such as 'average number of staffed beds' and 'capital charge', was also investigated. Removing 'average number of staffed beds', with the 11 variables in the resultant model, reduced the average efficiency score by just 1% in comparison with the model with 12 variables (Parkin and Hollingsworth, 1996). The number of efficient DMUs was reduced from 38 to 34. The analysis was on the general impact of the changes in the variable set, in terms of the average efficiency score and the number of efficient units, rather than their influence of the efficiency evaluations for individual DMUs.

The overall impact of changing the variable set was investigated through the calculation of the Spearman Rank Correlation Coefficient (defined in section 5.9.1) in Parkin and Hollingsworth (1996).

This was obtained for each pair of the various models in turn, with the value calculated to be significantly different from zero in each case. However, it was felt that the calculated values were in fact quite low 'if the hypothesis is that the efficiency scores are essentially the same thing' (Parkin and Hollingsworth, 1996). This was particularly apparent if the comparisons were repeated, excluding the efficient DMUs, suggesting that the impact was more apparent for the inefficient DMUs.

Overall, the analysis of the various models was based on 'pragmatic' criteria: the identification of 'a reasonable number of efficient hospitals – small enough to discriminate but large enough to provide potential peers – whilst retaining a good degree of disaggregation in inputs and outputs' (Parkin and Hollingsworth, 1996). Some consideration was also given to the impact of allowing for varying returns to scale on the efficiency scores and the number of efficient units. The 'divergence between the results obtained under conditions of CRS and VRS' was discussed with some disquiet, as 'there is no theoretical or empirical evidence about which is the correct assumption' (Parkin and Hollingsworth, 1996).

Additionally, they observed that the discrimination of the DEA methodology was greatly reduced by removing the constant returns to scale assumption. For example, using the final specification of variables identified above, the varying returns to scale model resulted in over 75% of the DMUs being classified as efficient, as opposed to the 44% classified as efficient in the CRS model. This corresponds to the observation in section 5.6, such that the BCC methodology usually results in more DMUs being classified as efficient as was an expected outcome.

The validity of the DEA model was tested by comparing the efficiency results obtained from data from three year period, as it was assumed that changes in efficiency over time 'would not be so large between two years that the results would be radically affected' (Parkin and Hollingsworth, 1996).

Efficiency scores were calculated for each of the DMUs, using a CCR model with 11 variables ('average number of staffed beds' was excluded from the list above), for data from three consecutive years. Calculating the Spearman Rank correlation coefficient, it was found that the coefficients seemed to be low if the assumption was made that the same thing was being measured, year on year. Analysis of the results from the three years also showed that the targets identified for improving efficiency were not consistent over the time frame investigated.

As the above discussion has illustrated, this paper questioned the applicability of DEA and determined that 'we need better means of assessing validity and specification issues than those currently available before DEA can be widely accepted' (Parkin and Hollingsworth, 1996).

6.2.7 Other Sources

As was observed in chapter five, the DEA methodology is continuously evolving, with new amendments to the model being developed. There are also new applications of the methodology being researched, some of which relate specifically to the measurement of hospital efficiency. There are several more examples of the use of DEA, which can be found in the literature, some of which are used to provide additional information on the DEA modelling process in chapters seven and eight.

A comprehensive bibliography on DEA applications to health care applications is given at the end of this chapter and some of this research will also be referred to in the following chapters.

6.2.8 Evaluation of DEA Applications in the Literature

The extensive discussion presented above shows the extent to which the DEA methodology has been applied to the evaluation of hospital performance. Drawing together the information obtained from this survey of the literature, a number of observations can be made on the use of DEA for assessment of hospital efficiency. These cover five main areas: modelling strategy, DEA model formulation, variable selection, including individual perspectives and evaluating and presenting results.

6.2.8.1 Modelling Strategy

A uniform approach for the application of the DEA methodology, which could be transferred to any similar health care environment, was not observed. This corresponded to the observation that there is an 'absence of a convincing model-building methodology for the user of DEA' by Pedrajo-Chapparo *et al.* (1999).

Although the same stages of research were carried out (related to the following four sections in this discussion), the emphasis placed on each of them seemed to be dictated by the perceptions of the authors or the pre-determined focus of their analysis. They were not always connected to the specific requirements of the health care environment.

6.2.8.2 DEA Model Formulation

The majority of applications used the original form of the DEA methodology, with constant returns to scale and the free allocation of the factor weights. In the cases where extensions to model have been utilised, this was for the measurement of other types of efficiency and varying returns to scale, rather than, for example, the inclusion of weight restrictions. In most cases, the appropriateness of the model formulation, having been defined, was not questioned.

6.2.8.3 Variable Selection

The importance attached to the selection of the data set was heavily stressed, with a variety of measures employed to determine the most appropriate factors. Lack of appropriate data was often influential in the final selection of variables. The number and nature of the variables included in the models varied considerably. The maximum number of variables employed was twelve (Parkin and Hollingsworth, 1996). There was no consensus amongst the authors as to the ideal factors for the measurement of efficiency using DEA. This lack of consensus was also reflected in the sample sizes chosen for analysis, which ranged from seven to several thousand.

6.2.8.4 Including Individual Perspectives

Incorporating the views of health care managers into the DEA modelling process, from the selection of variables to the analysis of the results, was clearly stated as an important element in the successful application of DEA. However, in the papers presented, any involvement tended to be in the evaluation of the results, rather than in the model-building process as a whole.

6.2.8.5 Evaluating and Presenting Results

The analysis of results from the models tended to on the development of a number of general statistics (average efficiency score, number of efficient units, range of efficiency scores). These were often compared across a number of models or hospitals of different types within the data sample, using a selection of statistical or observational methods. The analysis of individual DMUs was generally restricted to one or two, with the inefficient units receiving more attention.

The information obtained from the analysis of the literature is related to the application of DEA to the evaluation of acute hospitals in Scotland in the following chapters. The key points identified above are also used to guide the evaluation of the two-stage application procedure, discussed in chapter nine. The literature observations on DEA are now contrasted with opinions on the technique from the perspective of the health care manager.

6.3 Evaluation of Data Envelopment Analysis from the Perspective of the Health Care Manager

Techniques such as Data Envelopment Analysis have not been widely utilised within the health sector in the UK (as is demonstrated by the DEA bibliography at the end of this chapter, where the majority of papers relate to data sources in the USA). In section 6.2.8.4, however, it was suggested that incorporating the perspectives of health care managers into the modelling process was a key element of a successful application of DEA. This section, therefore, focuses on the use of the DEA methodology from the perspective of the health service manager.

In the discussion in chapter four, it was suggested that the various types of health service manager had differing perceptions of performance indicators and used them in different ways (see section 4.3.3). For example, a difference was observed between financial mangers at hospital level and the District General Managers by Jenkins *et al.* (1987).

Therefore, the information included in this discussion comes from a cross-section of professionals involved in health service management and has been collected from a number of sources, using a variety of methodologies. It was intended that this would provide a 'snapshot' of perceptions on performance assessment, rather than be a comprehensive survey bringing to light all possible opinions, which was beyond the scope of this research.

Semi-structured interviews were held with four members of the management team at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust, these being the directors of Finance and Corporate Development, an Audit Manager and an information specialist. It was intended that this would give the perspective on performance assessment at local level, to be contrasted with views from a national or regional perspective. Semi-structured interviews were therefore also held with two members of the Performance Directorate of the NHS Executive, based in Leeds. Their interest in performance assessment is at a regional level, looking at the performance of the Regional Health Authorities in England. Feedback and further correspondence from these interviews has also been filtered into the following discussion.

The other views incorporated into this section were collected on a more informal basis, through discussions at health care, operational research and DEA conferences and conversations with health care mangers and information specialists.

Many of the health care managers with whom performance assessment issues were discussed were completely unfamiliar with DEA. Having been introduced to the technique, they proposed a number of suggestions about its use, some of which correspond to the general thoughts on performance assessment given in chapter four. There were also contradictory opinions from the few people who were already aware of the DEA methodology. For example, the response by a member of the Audit Commission was 'not that old thing again' (personal communication). Others, however, particularly those at the NHS Performance Directorate in Leeds, were potentially more receptive.

The general observations on DEA and its acceptability are presented below, summarised into five areas, with specific issues introduced into the DEA application procedure in chapters seven and eight.

6.3.1 Selection of Samples and Variables

The importance of including a homogeneous sample of DMUs was considered to be vitally important (as was stressed in chapter four in relation to the use of performance indicators). The types of sample thought to be appropriate varied amongst those interviewed, that is, using DEA to compare hospitals, hospital departments, trusts (rejected by the Financial Director of BHRV NHS Trust) or health authorities (favoured by a member of the NHS Performance Directorate).

The importance of understanding the context and environment within which the units were operating was determined to be highly significant. The same issue was also raised with performance indicators – a major criticism was that they did not take into account the particular situation of each hospital or health authority. A number of relevant variables were identified, particularly in relation to measures of hospital activity, environmental influences and quality.

6.3.2 Providing Useful Information

The introduction of new methods for performance assessment was considered to be a worthwhile exercise, only if the information provided by them could be shown to have actual benefits and be more than just a 'number'. This was in response to the limitation identified for performance indicators, such that, knowing that the 'cost per patient discharge' was below average for a particular output did not really help to improve performance. Additional information was considered to be essential. It was felt that new approaches should provide better information than the currently used techniques, not just be different for the sake of it.

6.3.3 Avoiding Complicated Theory

The DEA technique, especially in comparison with the simplicity of the performance indicator, was seen to be very complex. It was observed (at a number of conferences where DEA presentations were made to non-practitioners, particularly Health Care and Public Sector Managers) that the benefits of the DEA methodology were lost in the mathematical equations and theoretical intricacies. It was determined that it was important to develop a clear and straightforward approach to the presentation of the results in the modelling strategy.

6.3.4 Presenting a True Reflection of Performance

Performance Indicators have often seen to present a distorted picture of a hospital's performance, especially when a small number have been looked at in isolation (as is the case with the indicators used in the league tables). Producing a technique that could present a true picture of performance (or efficiency) was seen as an important development. It was critical that the results would be trustworthy.

6.3.5 Relevant at a Local Level

Performance Indicators and Clinical Audit (as discussed in chapter four) were developed at a national level. It was seen as significant, in the development of the DEA methodology, that local policies and initiatives could be reflected in, and evaluated by, the measurement of efficiency.

These observations have been incorporated into the DEA modelling strategy developed in the following chapters, with specific points raised in the specific sections.

6.4 Summary

In this chapter, the technique of Data Envelopment Analysis, as applied to the measurement of hospital efficiency in particular, has been evaluated from two perspectives: the DEA literature and the Health Care Manager. This had led to the identification of a number of key issues, which are addressed in the following two chapters through the development of a two-stage application procedure for the measurement of hospital efficiency, using data from the NHS in Scotland.

6.5 DEA Health Services Bibliography

The technique of data envelopment analysis has been presented in the previous chapter as an effective method, which can be used to measure the efficiency of a sample of health care organisations. In the analysis in chapter five, the focus has been on the performance of hospitals. However, as was referred to in section 5.3, the DEA methodology has also been utilised in the assessment of a wide range of health service organisations. This section contains a bibliography of much of this research, some of which will be referenced again in chapter seven.

6.5.1 Health Care Summary

1. **Rosko, M. D.** (1990) Measuring Technical Efficiency in Health Care Organizations, *Journal of Medical Systems*, Vol. 14, No. 5 pp. 307-322.

6.5.2 Regional/district Health Authorities or Health Districts

1. **Norman, M. and Stoker, B.** (1991) DEA: The Assessment of Performance (Chapter 8.2), Wiley.

2. **Thanassoulis, E., Boussofiane, A. and Dyson, R. G.** (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588-608.

6.5.3 Hospitals

1. **Banker, R. D., Conrad, R. F. and Strauss, R. P.** (1986) A Comparative Application of Data Envelopment Analysis and Translog Methods: an Illustrative Study of Hospital Production, *Management Science*, Vol. 32, pp. 30-44.

2. **Burgess, J. F. and Wilson, P. W.** (1998) Variation in Efficiency Among US Hospitals, *INFOR*, Vol, 36, No. 3, pp. 84 – 102.

3. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

4. **Ehreth, J. L.** (1994) The Development and Evaluation of Hospital Performance Measures for Policy Analysis, *Medical Care*, Vol. 32, No. 6, pp. 568-587.

5. **Ersoy, K., Kavuncubasi, S., Ozcan, Y. A. and Harris, J. M.** (1997) Technical Efficiencies of Turkish Hospitals: DEA Approach, *Journal of Medical Systems,* Vol. 21, No. 2, pp. 67 – 74.

6. **Grosskopf, S. and Valdmanis, V.** (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, Vol. 6, No. 2, pp. 89-107.

7. **Hollingsworth, B. and Parkin, D.** (1995) The Efficiency of Scottish Acute Hospitals: An Application of Data Envelopment Analysis, *IMA Journal of Mathematics Applied in Medicine and Biology*, Vol. 12, pp. 161-173.

8. **Ozcan, Y. A., Luke, R. and Haksever, C.** (1992) Ownership and Organisational Performance: a Comparison of Technical Efficiency Across Hospital Types, *Medical Care*, Vol. 30, pp. 781-794.

9. **Ozcan, Y. A. and McCue, M. J.** (1996) Development of a Financial Performance Index for Hospitals: DEA Approach, *Journal of the Operational Research Society*, Vol. 47, No. 1, pp. 18 – 26.

10. **Parkin, D. and Hollingsworth, B.** (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991-1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

11. **Valdmanis, V.** (1990) Ownership and Technical Efficiency of Hospitals, *Medical Care*, Vol. 28, No. 6, pp. 552-561.

6.5.4 Hospital departments

1. **Bitran, G. R. and Valor-Sabatier, J.** (1987) Some Mathematical Programming Based Measures of Efficiency in Health Care Institutions, *Advances in Mathematical Programming and Financial Planning*, Vol. 1, pp. 61-84. [Focus on medical services department.]

2. **Sherman, H. D.** (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927-938. [Focus on medical-surgical departments.]

6.5.5 Individuals (physicians and/or managers)

1. **Chilingerian, J. A. and Sherman, H. D.** (1990) Managing Physician Efficiency and Effectiveness in Providing Hospital Services, *Health Services Management Research*, Vol. 3, No. 1, pp. 3-15.

2. **Chilingerian, J. A.** (1989) Investigating Non-medical Factors Associated with the Technical Efficiency of Physicians in the Provision of Hospital Services: A Pilot Study, *Annual Best Paper Proceedings of the Academy of Management*, pp. 85-89.

3. **Chilingerian, J. A.** (1994) Exploring Why Some Physicians' Hospital Practices Are More Efficient: Taking DEA Inside the Hospital, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

4. **Chilingerian, J. A. and Sherman, H. D.** (1996) Benchmarking Physician Practice Patterns with DEA: A Multi-stage Approach for Cost Containment, Annals of Operations Research, Vol. 67, pp. 83 – 116.

6.5.6 General Practice/Primary Care

1. **Salinas-Jiménez, J. and Smith, P.** (1996) Data Envelopment Analysis Applied to Quality in Primary Health Care, *Annals of Operations Research*, Vol. 67, pp. 141 – 161.

2. **Sczcepura, A., Davies, C., Fletcher, J. and Boussofiane, A.** (1993) Efficiency and Effectiveness in General Practice, *Journal of Management in Medicine*, Vol. 7, no. 5, pp. 36-47.

6.5.7 Nursing homes

1. **Kleinsorge, I. and Karney, D.** (1992) Management of Nursing Homes Using Data Envelopment Analysis, *Socio-Economic Planning Science*, Vol. 26, pp. 57-71.

2. **Nyman, J., Brick, D. and Link, D.** (1990) Technical Efficiency in Nursing Homes, *Medical Care*, Vol. 28, No. 6, pp. 541-551.

6.5.8 Pharmacies

1. **Capettini, R., Dittman, D. A. and Morey, R. C.** (1985) Reimbursement Rate Setting for Medicaid Perscription Drugs Based on relative Efficiencies, *Journal of Accounting and Public Policy*, Vol. 4, No. 2, pp. 83-110.

2. **Lothgren, M. and Tambour, M.** (1999) Productivity and Customer Satisfaction in Swedish Pharmacies: a DEA Network Model, *European Journal of Operational Research*, Vol. 115, No. 3, pp. 449 – 458.

3. **Morey, R. C., Capettini, R. and Dittman, D. A.** (1985) Pareto Rate Setting Strategies: an Application to Medicaid Drug Reimbursement, *Policy Sciences*, Vol. 18, No. 2, pp. 169-200.

6.5.9 Health Care Programmes

1. **Huang, Y. L. and McLaughlin, C. P.** (1989) Relative Efficiency in Rural Primary Health Care: an Application of Data Envelopment Analysis, *Health Services Research*, Vol. 24, No. 2, pp. 143-158.

6.6 References

1. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

2. **Charnes, A., Cooper, W. W. and Rhodes, E.** (1978) Measuring the Efficiency of Decision- Making Units, *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429 - 444.

3. **Ehreth, J. L.** (1994) The Development and Evaluation of Hospital Performance Measures for Policy Analysis, *Medical Care*, Vol. 32, No. 6, pp. 568-587.

4. **Fare, R., Grosskopf, S. and Lovell, C. A. K.** (1985) The Measurement of Efficiency of Production, Kluwer-Nijhoff Publishing.

5. **Farrell, M. J.** (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A (General)*, Part III, Vol. 120, pp. 253-279.

6. **Feldstein, M. S.** (1967) Economic Analysis for Health Service Efficiency, North Holland.

7. **Grosskopf, S. and Valdmanis, V.** (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, Vol. 6, No. 2, pp. 89-107.

8. **Hollingsworth, B. and Parkin, D.** (1995) The Efficiency of Scottish Acute Hospitals: An Application of Data Envelopment Analysis, *IMA Journal of Mathematics Applied in Medicine and Biology*, Vol. 12, pp. 161-173.

9. **Jenkins, L., Bardsley, M., Coles, J., Wickings, I. and Leow, H.** (1987) Use and Validity of NHS Performance Indicators - A National Survey, CASPE.

10. **Norman, M. and Stoker, B.** (1991) DEA: The Assessment of Performance, Wiley.

11. **Ozcan, Y. A., Luke, R. and Haksever, C.** (1992) Ownership and Organizational Performance: a Comparison of Technical Efficiency Across Hospital Types, *Medical Care*, Vol. 30, pp. 781-794.

12. **Pedraja-Chaparro, F., Salinas-Jimenez, J. and Smith, P.** (1999) On the Quality of the Data Envelopment Analysis Model, Journal of the Operational Research Society, Vol. 50, No. 6, pp. 636-644.

13. **Parkin, D. and Hollingsworth, B.** (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991-1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

14. **Salinas-Jiménez, J. and Smith, P.** (1996) Quality in Primary Health Care, *Annals of Operations Research*, Vol. 67, pp. 141 – 161.

15. **Sczcepura, A., Davies, C., Fletcher, J. and Boussofiane, A.** (1993) Efficiency and Effectiveness in General Practice, *Journal of Management in Medicine*, Vol. 7, no. 5, pp. 36-47.

16. **Sherman, H. D.** (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927-938.

17. **Thanassoulis, E., Boussofiane, A. and Dyson, R. G.** (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588-608.

6.7 Additional Sources

1. Interviews at the NHS Executive Performance Directorate in Leeds (conducted in October, 1996).

2. Interviews at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust (conducted in July, 1996).

3. Discussions in the DEA and Health Streams at Young OR 10, Guildford, March 31st – April 2nd, 1998.

4. Informal Discussions at ORGS Meetings (Operational Research Group, Scotland).

5. Informal Interview with a member of the Audit Commission (December, 1997).

CHAPTER SEVEN

ANALYSING HEALTH CARE ORGANISATIONS USING DATA ENVELOPMENT ANALYSIS (DEA)

7.1 Introduction

The technique of data envelopment analysis (DEA) and the theoretical issues surrounding its application were discussed in depth in chapter five. It was also noted that, despite the fact that DEA appears to be ideally suited to the measurement of efficiency in health care organisations, there are few examples to be found where DEA has been used as a practical tool for efficiency assessment at a managerial level. In many cases, the DEA literature relating to health care organisations used the health data to illustrate facets of the DEA technique, rather than using the DEA technique to investigate health care data, such as Nunamaker (1985).

However, there are an increasing number of examples showing the benefits of a DEA approach to the assessment of health services, illustrating how the technique can be applied and the observations that can be made about be made about efficiency in health services. These have been discussed extensively in chapter six, although there are new examples appearing regularly. One of the most recent of these is Ersoy *et al.* (1997), which used DEA to make an assessment of the overall efficiency of hospitals in Turkey.

A DEA analysis is presented in this chapter, which uses data from the NHS in Scotland. This illustrates the issues and problems to be addressed when applying DEA to a health service environment, specifically related to the selected data set and the issues surrounding the health care system in Scotland. This chapter addresses the first phase of a two-stage application of DEA. In chapter eight, the second stage of the approach is presented, which is the application of restrictions to the factor weights.

In chapter five, a brief summary of the types of investigations carried out using DEA with health service data was presented. As this summary demonstrated, virtually all elements of health service provision have been investigated using DEA, such as general practice, health education programmes and health districts.

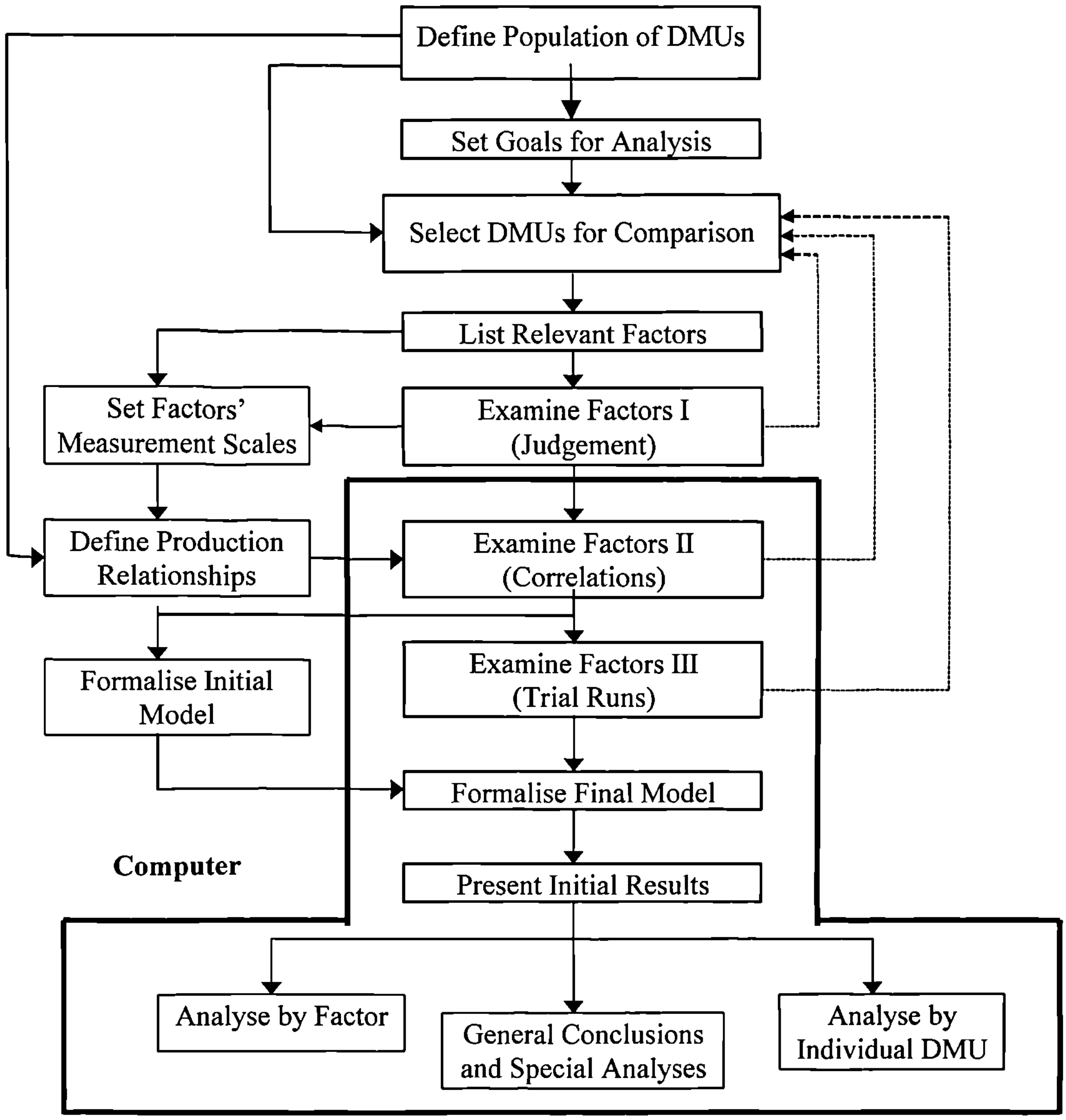
Whilst maintaining a broad interest in health services across the UK, the majority of the analysis presented here is, however, focused on hospital activity in Scotland, with data taken from a variety of sources published by the Scottish Office and the NHS in Scotland. The decision to focus on Scottish hospitals and hospital trusts was taken for numerous and important reasons:

1. The availability and accessibility of data sources and other information;
2. Reflecting the priorities of both the Government and the NHS on improving the performance and efficiency of individual hospitals (Department of Health, 1997 and the Scottish Office, 1997).
3. The availability of papers and academic research relating to the performance of hospitals, as discussed in chapter six.
4. Pre-existing knowledge, contacts and interests.

This chapter begins with a brief summary of the managerial and organisational structure of the NHS in Scotland, to provide a context for the application of the DEA technique. It was observed in chapter two, and referred to again in the investigation of performance indicators in chapter four, that this can have significant impact on the type of techniques employed and certainly on the success, or otherwise, of this application. The remainder of the chapter is devoted to the detailed analysis of acute hospitals in Scotland using data envelopment analysis.

7.2 The DEA Modelling Process

If DEA is to be used to successfully examine the efficiency of a group of hospitals, or other health care organisations, it can be seen that there are several key stages to the modelling process. These have been highlighted in the discussion in chapters five and six. Figure 7.1 below, adapted from Golany and Roll (1989), has been proposed as a guide to the key stages of the modelling process. As can be seen, the process is cyclical, with some elements to be repeated if suggested by the results and analysis.



GOLANY AND ROLL (1989)

Figure 7.1: An Application of Data Envelopment Analysis to Health Care Data

However, whilst figure 7.1 can be seen as a useful indicator of the complexity of the DEA modelling process, it is not followed directly in this application. As was stated in section 7.1, the analysis in chapters seven and eight presents a two-stage approach to the application of the DEA technique. The elements comprising these two stages are shown below, with the first stage covered in this chapter, beginning with the definition of the sample of DMUs in the next section. The second stage is discussed in chapter eight.

STAGE ONE:	1.	Definition of the Sample;
(Chapter Seven)	2.	Formulation of DEA Model Type and Orientation;
	3.	Selection of Possible Inputs and Outputs
	4.	Results of the Preliminary Analysis;
	5.	Revisions to the Model – Sample Size, Variable Set.
STAGE TWO:	1.	Results from the Revised Model;
(Chapter Eight)	2.	Evaluation of Factor Weights;
	3.	Application of Weight Restrictions;
	4.	Development of Efficiency Strategies;
	5.	Evaluation of Model – Sensitivity and Robustness.

7.3 Definition of an Appropriate Sample

Earlier in the chapter, the decision to focus on hospital-based activity as the area for assessment using DEA was specified, as was the use of data relating to health services in Scotland. The National Health Service in Scotland is financed through the Scottish Office and there are 15 health boards covering the main geographical regions, some 50 hospital and healthcare trusts and over 200 hospitals.

The recent reforms to the NHS proposed by the Labour government in 1997 have been based on the following premise, intending to put an end to the practice of changing the service in Scotland simply because things are changing in the rest of the UK:

> "The Government's vision is a National Health Service for the people of Scotland that offers them the treatment they need, where they want it, and when: a modern, 'designed' health service putting patients first. We want a seamless health service centred on primary care, designed to ensure that patients receive care quickly and with certainty." THE SCOTTISH OFFICE (1997)

The following diagram illustrates the complex organisational structure of the NHS in Scotland:

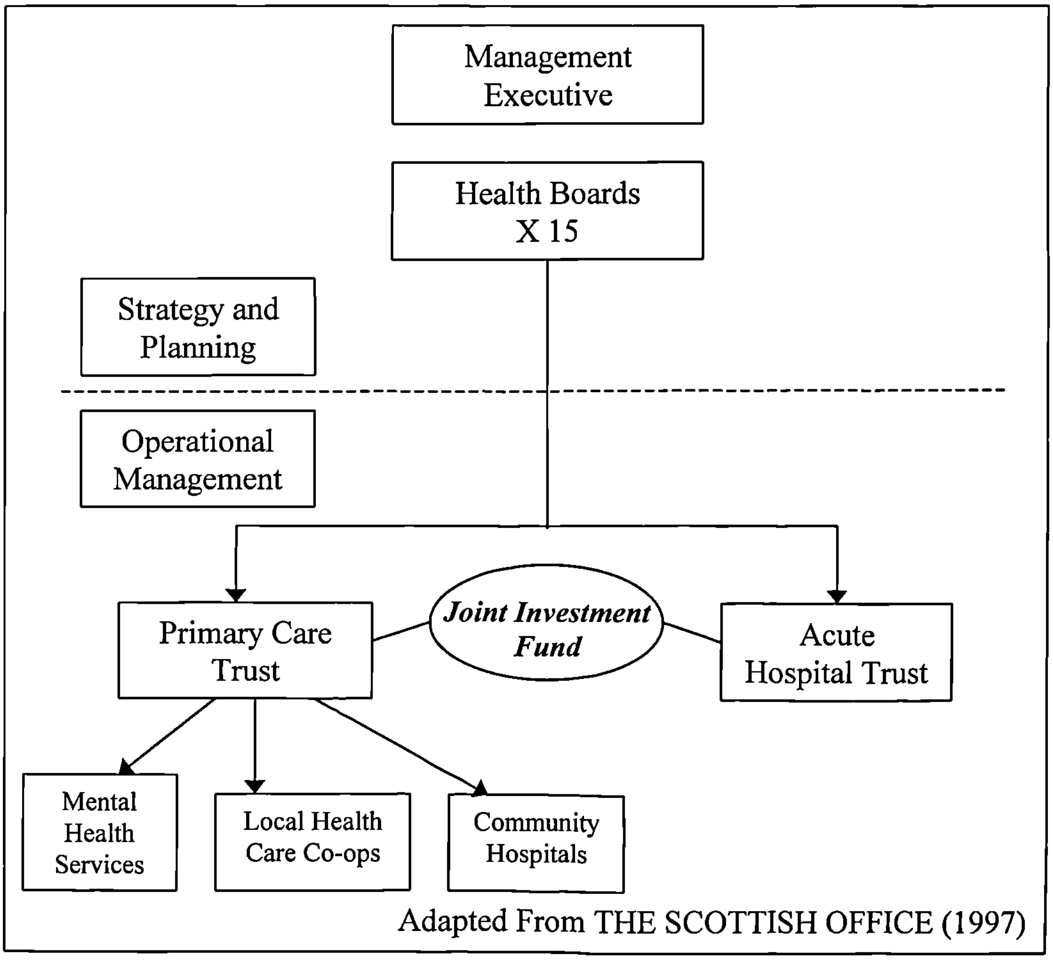


Figure 7.2: Structure of the NHS in Scotland in 1998

Each of the trusts, as defined in chapter three, is a provider of health care services, with some of them concentrating on specific areas such as community, mental health and dental services. Others provide only one type of service such as outpatient facilities, although many provide a cross section of services at the hospital sites under their control. The size and scope of the trusts vary quite considerably.

For example, using ISD (1996) as the source of information, the Royal Alexandra Trust provided a cross-section of services across a small mix of specialties at just one hospital site. West Glasgow Trust has nine hospitals, providing a large cross-section of services in almost all specialties. Examples of the different types of Trust are presented in Table 7.1:

Examples of the Different Types of Provider Trusts in Scotland			
Hospital	Aberdeen Royal	Stirling Royal Infirmary	Law
Teaching Hospital	Dundee		
Children's Hospital	Edinburgh	Yorkhill	
Healthcare	Central	Dundee	Grampian
Community	Highland	Dumfries and Galloway	Borders
Mental Health	Greater Glasgow		
Dental	Glasgow		

Table 7.1: Hospital Trusts in Scotland

Clearly, the very different types of trust, and the fact that many of the trusts provide service at a number of hospital sites, mean that to use all the trusts in a single DEA sample would not be appropriate or useful. The Finance Director of the Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust believed that trusts were inherently difficult for across-the-board comparisons. This was discussed in chapter four in reference to performance indicators and medical audit.

Therefore, the level of service provision that is used for comparison in the following analysis is the hospital level. However, it is still necessary to reduce the sample still further, as there are over 200 hospitals in Scotland, which vary in size and in the services they offer quite considerably. The sample of all hospitals can be reduced according to some specified characteristics, which is consistent with many of the approaches in the literature. For example, Byrnes and Valdmanis (1994) focused on community hospitals, Sherman (1984) used a sample of teaching hospitals and Grosskopf and Valdmanis (1987) used large, urban hospitals. These types of study, however, developed their sample from very large groupings of hospitals. With Scottish data, refining the sample according to very strict characteristics, such as location, size and type, would result in a large number of very small samples.

There are only a handful of hospitals in each area, which are often all of different types, such as community, acute, teaching, children's, maternity and geriatric. The hospitals in Scotland can actually be divided into some 49 categories, called functional classifications, and defined by the ISD (1996). Many of these categories are quite distinct and the services they provide do not overlap. Therefore, it has been necessary to make some generalisations in order to obtain an appropriate sample with an adequate number of DMUs, focusing on just one type of hospital.

The sample evaluated initially consisted of the 74 hospitals in Scotland that were classified as 'acute', which was consistent with the approaches in Parkin and Hollingsworth (1996) and Ozcan *et al.* (1992).

These 'acute' hospitals cover 15 functional classifications, with the full definition of each classification given in appendix one, and the hospitals in these classes can be divided into four broad categories:

1. General Hospitals - may have a teaching element and a maternity department, will cover the majority of surgical functions but not necessarily all of them and may vary in size quite considerably;
2. GP Practitioner Cottage Hospitals - will have a limited surgical function, may have a maternity department but no teaching element;
3. Mixed Specialist Hospital - may have a maternity department, will cover the basic surgical functions but not highly specialised treatment;
4. Specialist Children's Hospitals - may have teaching element, will cover complete range of paediatric services.

In the following analysis, the hospitals are referred to by a numerical reference (#1 - #74) and are listed alphabetically in appendix three. (No cross-referencing is given in order to maintain anonymity.)

7.4 Formulation of the DEA Model

In the majority of investigations into health service organisations, the original DEA model, developed by and Charnes, Cooper and Rhodes (1978) and denoted by CCR, is used for the analysis of efficiency. (See Hollingsworth and Parkin (1995), Sherman (1984) and Ehreth (1994) above.) Thus the majority of models have been developed with the assumption of constant returns to scale (CRS) and the free allocation of weights.

Parkin and Hollingsworth (1996) introduced the idea of examining the efficiency of hospitals, whilst also taking account of varying returns to scale (VRS), thus utilising the BCC model as defined in section 5.6.1. Also, Byrnes and Valdmanis (1994) presented the approach for examining non-technical types of efficiency. The more technical developments to the basic DEA model, such as the multiplicative and additive models discussed in sections 5.6.2 and 5.6.3, have not been translated into any practical applications using health care data.

In relation to the evaluation of hospital efficiency, the impact of weight restrictions has not been widely investigated, although Thanassoulis *et al.* (1995) included restrictions in their analysis of perinatal care at a health authority level. The analysis presented in this chapter, and in subsequent chapters, will therefore use the CCR methodology for the majority of the DEA investigations. This is appropriate in this case for several reasons:

1. The majority of the models used in the analysis of hospital efficiency have relied on the CCR model for their initial investigations, as presented in the discussion above of several of the most significant papers in this field.
2. The data sample contains a set of DMUs, each of which have been classified as 'acute' using definitions determined by the Scottish Office and can thus be said to be 'homogeneous', meaning the constant returns to scale argument is still valid (Smith and Mayston, 1987).
3. There are no 'non-discretionary' or 'categorical' variables used in the analysis.
4. Weight restrictions are not introduced until the second stage of the analysis in chapter eight.

The validity of this approach is discussed in chapter nine, where the overall DEA application procedure is evaluated, focusing on each of the points made above in the decision to use the basic CCR methodology.

In addition to the choice of model, it is necessary to determine the orientation of that model, be it input minimisation or output maximisation, as defined in section 5.4. The orientation of input minimisation has been selected in this case, although this does not alter the efficiency scores calculated.

This has been done according to the premise that the type of thinking reflected by an input minimisation orientation, that is, minimising costs without reducing services, has motivated much of the recent debate relating to health service performance assessment. The efficiency problem has generally been investigated within the framework of 'minimising the ratio of inputs to outputs' (Ball and Roberts, 1998), as was also discussed in chapter two. It is consistent with the approach of Byrnes and Valdmanis (1994), who also used an input-oriented model.

Therefore the DEA model to be solved, using the DMUs presented above and variables to be defined in the following section, is shown in figure 7.3 overleaf. Following the definition of the model type and the sample of DMUs, the next stage of the analysis has been to define the variable set to be used in the solution of the CCR (Charnes, Cooper and Rhodes) model shown overleaf.

$$\text{MAXIMISE: } e_\alpha = \sum_{r=1}^{s} U_r O_{r\alpha},$$

$$\text{SUBJECT TO: } \sum_{r=1}^{s} U_r O_{rj} - \sum_{i=1}^{m} V_i I_{i\alpha} \leq 0,\ j = 1....\alpha....n$$

$$\text{AND: } \sum_{i=1}^{m} V_i I_{i\alpha} = 1$$

WHERE: U_r and $V_i > \varepsilon$ FOR ALL r AND i (ε being a very small positive number),
U_r are the weights applied to the outputs O_{rj} for the j^{th} DMU,
V_i are the weights applied to the inputs I_{ij} for the j^{th} DMU,
s = number of outputs and
m = number of inputs.

Figure 7.3: The CCR Model with Input Minimisation Orientation

7.5 Selecting the Variables to be used in the DEA Model

The selection of the variables is as important in a DEA investigation as with all other methods of performance assessment. Without the 'right' factors, the analysis will have little or no benefit. In simple organisations or groups of DMUs, it may be a relatively straightforward exercise to define the variables to be used in an efficiency analysis. However, in the complex world of health care organisations, there are many and varied factors, which could be used as inputs and outputs in a DEA model of efficiency, as the discussions in sections 5.5.2 and chapter six have highlighted.

In some cases, the most appropriate variables can be defined by 'intuition'. For example, the output variables for a hospital evaluation will generally be some 'measure of the number of cases dealt with by the hospital' (Parkin and Hollingsworth, 1996). Ozcan *et al.* (1992) followed a similar process, by assuming that hospital output was limited to three main areas, to be included as the output factors in their DEA model.

In terms of the data relating to Scottish Acute Hospitals, the cases dealt with by each hospital fall into six main categories: inpatients, consultant outpatients, PAM (Professions Allied to Medicine) outpatients, day cases, day patients and A&E attendances. (Full definitions for each of these patient categories can be found in appendix one). However, there were often several potential variables for each area of activity that could be used as outputs, such as 'surgical inpatients' or, disaggregating still further, 'orthopaedic surgical inpatients'.

The same situation is also apparent for the inputs, where numerous possibilities and assumptions have to be made to identify the significant inputs 'to represent the major factors used in producing the... outputs' (Ozcan *et al.*, 1992). Also, as was the case with the outputs, there were often several possibilities, for example, measuring staffing levels in financial terms or as the number of full-time equivalents (FTEs). Additionally, deciding upon the level of disaggregation is an important part of the variable selection process, as discussed by Parkin and Hollingsworth (1996). For example, should 'total staff costs' be used, or is dividing it into 'nursing costs', 'PAM costs' and so on, through all the staff categories more appropriate?

Having decided upon the areas to be included as factors, the specific data to represent them must also be identified, which is not always straightforward. There may be a number of alternatives or a lack of data. In this analysis, a potential input factor would be the 'Total Number of Trained, Learning and Other Nurses', as used by Hollingworth and Parkin (1995). However, this data was not included in the information to which open access was available in this case and thus could not be incorporated.

Additionally, the figures for inpatient activity have not been scaled by an index of case-mix complexity, as proposed by Ehreth (1994) and Ozcan *et al.* (1992). In both cases, they were able to make use of the Medicare case-mix index, a statistic produced for each hospital in the USA. However, an equivalent statistic is not easily identified for hospitals in the UK. Parkin and Hollingsworth (1996) did not consider employing a case-mix index in their analysis of acute hospitals. This is clearly significant, particularly if the inpatient figures were aggregated into a single output factor, treating discharges in all categories as equivalent.

The figures for inpatient activity have not been broken down into age categories, as proposed by Sherman (1984), for a similar lack of the appropriate data in this form or any demonstration of its importance in this case.

Excluding variables due to a lack of appropriate data is clearly problematic and raises questions on the validity of the models produced and the results obtained, particularly if the excluded factors were considered to be important. This is an issue considered in many of the studies discussed in chapter six and it was felt that the factors they identified, whilst not necessarily ideal, were acceptable measures of hospital activity.

The list of factors that could be used in some combination in a DEA model for the evaluation of acute hospitals in Scotland, and for which the data is available, is given in Table 7.2 to follow. The scale of each of the factors is not given, as this has no bearing on the DEA modelling process. For example, the cost factors have been recorded in £000s, whereas the output variables were given as the actual numbers.

Inputs (5)	'Average Number of Staffed Beds' 'Capital Charge' 'Total Costs' 'Total Direct Costs', 'Total Allocated Costs'
Outputs (30)	'Total Inpatient Discharges' 'Inpatient Discharges, Surgical' 'Inpatient Discharges, ITU' 'Inpatient Discharges, Medical' 'Inpatient Discharges, Obstetrics and Gynaecology' 'Inpatient Discharges, Other' 'Total Inpatient Days' 'Inpatient Days, Surgical', 'Inpatient Days, Medical' 'Inpatient Days, Obstetrics and Gynaecology' 'Inpatient Days, Other' 'Inpatient Days, ITU' 'Consultant Outpatients Attendances, Other including ITU' 'Consultant Outpatients Attendances, Surgical' 'Consultant Outpatients Attendances, Medical' 'Total Consultant Outpatient Attendances' 'Total New Consultant Outpatient Attendances' 'Total A&E Attendances' 'Total New A&E Attendances' 'Total Day case Attendances' 'Day case Attendances, Medical' 'Day case Attendances, Surgical' 'Day case Attendances, Other' 'Total Day Patient Attendances' 'Day Patient Attendances, Medical' 'Day Patient Attendances, Other including Mental Health' 'Total PAM Outpatient Attendances' 'Total New PAM Outpatient Attendances' 'Total Outpatient Attendances, A&E, Consultant and PAM'

Key: PAM – Professions Allied to Medicine ITU – Intensive Therapy Unit A&E – Accident and Emergency

(See appendix one for a full definition of each of the factors).

Table 7.2: Potential Variables

The inclusion of each of the factors can be supported for several reasons:

1. Some evidence exists in the literature for the factor, or one of an equivalent nature (see the discussions in section 5.5 and chapter six);
2. They intuitively describe some element of the performance of each hospital, in relation to the idea of the conceptual input and output proposed by Beasley (1990);
3. They were acceptable to 'health care experts' in that they are used in internal evaluations, published by the Scottish Office and supported by the Department of Health.

However, the list above, containing some 38 factors, is extensive and to include all of these variables in some form would have been inappropriate and unnecessary, in accordance with the discussion in section 5.5. The resulting DEA modelling process would also have concluded with a large majority of DMUs being classified as efficient, which may not necessarily be appropriate or particularly useful.

Therefore, it was necessary to reduce the above list of factors to more manageable and appropriate proportions, and determine the most applicable variable mix, particularly as the models discussed in the literature had a maximum of twelve variables. The reduction process was guided by some of the observations found in the literature examined in chapter six and other information gained from discussions with health care managers and professionals. Several key points were identified to assist the process of obtaining the variable sets to used in the DEA models, which also reflects the approaches used in the literature, as follows:

1. Should inpatient activity be measured as days or discharges?
2. What level of disaggregation should be used to reflect inpatient activity?
3. What other measures should be used to record output: outpatients, day cases, day patients, A&E admissions, and PAM appointments? Should there be further disaggregation?
4. Are there any other variables that should have been identified?
5. Are the financial measures of input comprehensive enough?
6. Is some of measure of hospital size, such as 'Average Number of Staffed Beds', necessary?
7. How should capital costs be reflected?

Clearly, it would have been possible to focus on other points in the selection of variables, depending on the nature of the data to hand and the motivation for the study. However, in this case, the above points were considered to represent a broad spectrum of issues relating to variable selection in an acute hospital context.

Five approaches were identified, again guided by the literature, to reduce the number of potential variables and to address the seven points listed above. These were:

(i) taking the advice of experts;
(ii) looking at previous use in the literature;
(iii) evaluating the data;
(iv) manipulating the data using statistical methods;
(v) heuristic approaches.

These were utilised in turn, as reported in the following sections, although some overlap does occur.

7.5.1 The Opinion of Experts

This has been widely proposed as an essential element of the model building process. Some of the views included in this section were first presented in the general discussion in section 6.3. A member of the NHS Executive Performance Directorate suggested that 'inpatient discharges' should always be used as an output in place of 'inpatient days'. The use of 'inpatient days' was thought to be highly flawed and open to distortion if it was used to investigate efficiency in a health care environment. The disaggregation of the inpatient activity factors was also supported, as 'inpatients' was regarded as too broad a category of hospital activity.

The importance of including some measure of day case activity was also stressed, although this was not reflected in the literature discussed in chapter six. This may be due to the relatively recent emergence of the day case approach to treatment, which was discussed in chapter three.

In terms of inputs, it was proposed by the member of the NHS Executive in Leeds referred to above that financial measures should be used as inputs whenever possible, as all factors can be translated back into some actual 'cost' or 'financial value'. As such factors such 'number of FTEs' should always be replaced by 'staffing costs', in his opinion. This relates back to the very basic model of public sector performance given in figure 2.1, which illustrated the translation from costs to benefits. The redundancy of the 'average staffed beds' factor was also suggested, reflecting the observation by Parkin and Hollingswoth (1996) that 'capital charge' actually takes some account of this.

7.5.2 Previous Use in Literature

The main focus in terms of outputs was on the selection of factors to reflect inpatient activity, presumably as this is related to the largest area of expenditure for the majority of hospitals. The use of 'inpatient discharges' in place of 'inpatient days', as proposed by the 'experts', is supported by the majority of the literature. Ehreth (1994) rejected the use of 'inpatient days' despite finding it as a factor in the majority of previous research approaches investigated. Parkin and Hollingsworth (1996) also stressed the importance of using 'inpatient discharges', having included 'inpatient days' in a previous investigation (Hollingsworth and Parkin, 1995) and been dissuaded against its continued inclusion in the modelling process.

The validity and value of this focus on the use of 'discharges' rather than 'days' is examined in the next section, by using both of these output alternatives in two separate models and comparing the results.

The pattern followed by many of the previous studies, including Ehreth (1994) and Sherman (1984), was to adjust the inpatient figures to reflect either age or case-mix severity, whichever type of inpatient statistic was included. Unfortunately, the data relating to acute hospitals in Scotland was not directly available in an appropriate format in the published statistics being used for this analysis, despite the apparent value of this approach, as highlighted above.

The different levels of disaggregation were also of interest in the final selection of variables. Grosskopf and Valdmanis (1987) split inpatient activity into three areas (acute, ICU and surgical) whilst a four-way split was the approach suggested by Hollingsworth and Parkin (1995), using 'medical', 'surgical', 'obstetrics and gynaecology' and 'other' as the four inpatient factors. The importance of the level of disaggregation is investigated in the next sections, where a number of models with several different mixes of outputs were investigated. This utilises the final approach to variable selection, that is heuristics, to be discussed in section 7.5.5.

In terms of outputs, outpatient discharges and A&E attendances have frequently been used as output factors, whilst there has been no reference to the other types of patient activity as mentioned above. For the input factors, it appears from the literature that these generally reflect three characteristics, measured in a variety of ways:

1. Hospital size - number of staffed beds, net plant assets, number of admissions, fixed assets and bed days available.
2. Staffing levels - number of full-time employees, number of staff in various categories, payroll expenditure and direct salary costs.
3. Supply expenses - amount of operational costs, total value of supplies and cost of the drug supply.

There was a lack of extensive data in all these areas, particularly for the staffing levels in the different categories, which were only available at Trust level or for individual departments within each hospital. Additionally, they were almost always measured in financial ('medical costs') rather than numerical ('number of medical staff') terms (ISD, 1996).

Therefore, the input factors in this initial analysis were limited to the small number shown in table 7.2. The final mix of inputs was obtained from: 'average number of staffed beds', 'capital charge', 'total expenditure', 'total direct costs' and 'total allocated costs'. The two cost categories have been defined, according to the ISD (1996), to incorporate the following areas:

> DIRECT COSTS: Medical, dental and nursing staff costs; pharmacy, PAM, theatre and laboratory staff and supply costs; other direct care staff and supply costs.
> ALLOCATED COSTS: The costs of Administration, Catering, Laundry, Cleaning, Transport, Travel, Portering, Heating, Nurse Teaching, Maintenance, Furniture, inter alia.

Capital charge has frequently been used as an input factor and has been defined in the context of hospital services in the UK as follows:

> "This is a charge levied by the NHS on hospitals for use of capital assets and new capital investment. It reflects the requirement that hospitals make a specified return on asset value including an allowance for depreciation."
>
> PARKIN AND HOLLINGSWORTH (1996)

Whilst not directly equivalent, these inputs have been determined to generally reflect the three characteristics described above and, as such, were appropriate input factors, according to the information from the literature and the availability of data. Restricting the number of inputs has also reduced the overall number of factors included in the models, which was discussed in chapter five as being significant to the number of units classified as efficient.

7.5.3 Evaluation of the Data

Some degree of disaggregation of the inpatient factors was appropriate in the view of the 'experts' (also reflected in the literature). Additionally, several other categories of output were also suggested. However, analysis of the output variables, in terms of the number of missing or zero values, showed that it would not necessarily be possible to disaggregate to the level that was required. Table 7.3 overleaf records the number of DMUs for which the particular output had a zero or missing value and also as a percentage of the total number of DMUs, which equalled 74 in this case.

It can be seen that in some categories of output, over 25% of the 74 DMUs had a zero value and there were no categories, apart from the aggregated inpatient factor, for which all the DMUs had a non-zero value.

Output Factor	Number of DMUs with Zero Values	Percentage of the Total No. of DMUs
'Inpatient Discharges, Surgical'	20	27.0%
'Inpatient Discharges, Medical'	21	28.4%
'Inpatient Discharges, Obs and Gynae'	20	27.0%
'Inpatient Discharges, Other'	2	2.7%
'Total Inpatient Discharges'	0	0%
'Total Consultant Outpatient Attendances'	3	4.1%
'Total A&E Attendances'	9	12.2%
'Total Day case Attendances'	15	20.3%
'Total Day Patient Attendances'	36	48.6%
'Total PAM Outpatient Attendances'	4	5.4%

Table 7.3: Zero Values in the Output Factors

Continuing to examine the influence of zero values, table 7.4 gives the breakdown of all the DMUs, with the number of factors for which the data for each DMU was non-zero. The nine factors assessed were those in table 7.3 above, excluding 'total inpatient discharges'. The table also records the number of DMUs as a percentage of the total number in the sample.

Number of Output Factors With Non-Zero Values	Number of DMUs	Percentage of DMUs
1	1	1.35%
2	1	1.35%
3	0	0%
4	2	2.70%
5	11	14.86%
6	4	5.41%
7	15	20.27%
8	19	25.68%
9	21	28.38%

Table 7.4: Non-zero Data in the Output Factors

From the table, the number of DMUs for which there were no zero values in any of the output factors was actually only 21, that is, only 28% of the DMUs showed activity in each of the patient categories. Also, the table shows that there were 2 DMUs for which there were just one or two output factors with non-zero values. The information from tables 7.3 and 7.4 was important as it showed that there were some combinations of the patient activity categories for which at least one of the DMUs had no non-zero outputs, thus making efficiency calculations impossible for that DMU under the DEA methodology.

Combining the two tables showed with further investigation that, for example, if 'total inpatient discharges' was excluded from any of the DEA models, there would be one DMU for which no output values were available if 'total inpatient discharges, other' was also excluded.

Therefore, the impact of the disaggregation of the inpatient factors or the exclusion of some of the other output factors, in particular 'number of day patients', would have a significant effect on efficiency ratings simply due to the high number of zero values. It would also be highly significant in the development of weight restrictions. The method used in later chapters for the restriction of weights, whereby limits are attached to the virtual inputs and outputs as described in section 5.8, is limited to factors for which there are no zero values if minimum weight restrictions are specified. The impact of these particularities of the data was investigated by using several DEA models, with a mixture of output factors, leading to the refinement of the sample of DMUs if required.

Of particular concern were those DMUs for which there were just one or two output factors with non-zero values. It would be necessary to remove these from the sample due to their 'unique' nature in the later analysis using the weight restriction methodology in the second stage of the analysis in chapter eight.

It was also noted that there were some inconsistencies in the data, in that the summation of the different figures did not match the published totals. For example, the summed total expenditure figure obtained for DMU #1 was £3 million lower than the figure given by the ISD (1996). For DMU #2, the summed figure was £3 million above the given total whereas the figures obtained for DMUs #4 and #7 were equal to the published totals.

Despite recourse to the ISD and the Scottish Office, no explanation could be found for the discrepancies and it was felt that this was a reflection of the nature of the NHS, where there have often been concerns over the quality of the data. In the cases where the differences occurred, the calculated totals were used, rather than the published totals, to provide continuity with the rest of the output data, which also used calculated figures in many cases.

Analysis of the data for the other inputs did not reveal anything as startling, although the data on 'capital charge' was not complete as the figures for two of the hospitals were missing or excluded from the published sources of information. There have been some reservations expressed regarding the use of 'capital charge' as an input variable, as the nature of the calculations from which it was obtained can appear quite arbitrary.

It was particularly difficult to ascertain if the same calculations had been employed in the calculation of capital charge, in relation to each of the 74 DMUs, as was noted with some of the performance indicators discussed in chapter four.

By focusing so closely on the nature of the data available, there is the possibility of this type of analysis becoming data driven – measuring hospital efficiency using the data that is available, whether it is appropriate or not. This has been a charge levelled against many types of performance assessment techniques, as was debated in chapter four in relation to performance indicators. Due to the nature of the data available relating to hospital performance, it is difficult to avoid this to a certain degree. However, a precedent was identified for each of the variables selected for the DEA modelling process. Additionally, the overall analysis of hospital performance using some or all of the factors included in this case is widespread within the NHS and generally acceptable, as well as being strongly promoted by the Department of Health and the Scottish Office (Department of Health, 1997 and Scottish Office, 1997).

7.5.4 Statistical Manipulation

A number of statistical approaches have been used to distinguish between similar factors, usually using correlation, as discussed by Parkin and Hollingsworth (1996) and more theoretically by Nunamaker (1985). The premise that it would be unnecessary to include two highly correlated factors has been widely supported, as would be the case in a regression analysis. Parkin and Hollingsworth (1996) used this to suggest the redundancy of 'average number of staffed beds', as they found it to be highly correlated with 'capital charge' in their analysis of a sample of acute hospitals in Scotland.

This was not found to be the case with the data sample used in this investigation, as the Pearson Correlation Coefficient for these two variables was found to be ***0.039, p = 0.741***. However, 'average number of staffed beds' was found to be highly correlated with some of the other potential input factors, in particular, 'total direct costs' (***0.9517, p = 0.00***) and 'total allocated costs' (***0.9607, p = 0.00***). As to be expected, 'total allocated costs' and 'total direct costs' were also found to be highly correlated (***0.9667, p = 0.00***). As a consequence of this, following on from past examples, it would have been appropriate to remove at least one of the input factors, if not two, for example 'average number of staffed beds' and 'total allocated costs' or to use the aggregated cost input, 'total costs'.

The same exercise was also carried out using the output factors, with the results from the correlation analysis illustrated in the table below.

	COA	**AEA**	**IDS**	**IDM**	**IDG**	**IDO**	**TID**	**DCA**	**DPA**
AEA	0.88								
IDS	0.93	0.86							
IDM	0.96	0.81	0.94						
IDG	0.63	0.53	0.56	0.62					
IDO	0.52	0.55	0.51	0.48	0.59				
TID	0.96	0.85	0.97	0.98	0.70	0.57			
DCA	0.89	0.80	0.88	0.86	0.64	0.61	0.91		
DPA	0.76	0.54	0.70	0.80	0.47	0.47	0.76	0.63	
POA	0.69	0.71	0.68	0.64	0.65	0.65	0.71	0.76	0.34

KEY: IDS - 'Inpatient Discharges, Surgical'
IDM - 'Inpatient Discharges, Medical'
IDG - 'Inpatient Discharges, Obs and Gynae'
IDO - 'Inpatient Discharges Other'
AEA - 'Total A&E Attendances'
COA - 'Total Consultant Outpatient Attendances'
TID - 'Total Inpatient Discharges'
DCA - 'Total Day case Attendances'
DPA - 'Total Day Patient Attendances'
POA - 'Total PAM Outpatient Attendances'

Table 7.5: Pairwise Correlation Coefficients for Output Factors

As table 7.5 shows, several pairs of output factors were very highly correlated, which suggesting that they should have been excluded from the modelling process. In fact, all the pairwise correlations except one (DPA and POA) were found to be significant at the 95% level, with many with a much greater significance value. None of the factors were perfectly correlated, with a correlation coefficient of 1.

COA, which measured outpatient activity, was highly correlated (with a Pearson Correlation Coefficient of greater than 0.95) with IDM and TID, two factors relating to inpatient discharges. This may be expected, as the number of discharges in any given category is likely to be related to the size of each hospital. However, as these factors measured two totally different aspects of the services provided by each hospital, the exclusion of one of them simply on the basis of these results would be difficult to justify.

The suggestion of removing one of each pair of highly or perfectly correlated inputs or outputs has not been universally accepted or followed as a matter of course in the DEA studies reported in the literature sources discussed in chapter six. With techniques such as regression, with its centralising tendency, this approach has been found to be valid. However, with techniques based on frontier methodologies, the influence of correlation has not quite been so straightforward.

Nunamaker (1985), most notably, stated that adding in a new variable that was perfectly correlated to an existing factor could not result in an inefficient unit achieving efficiency.

The converse was also true, that is, the impact of removing one of a pair of perfectly correlated variables, would be the same. However, if 'the correlation was less than perfect' it would be possible for the DMU to become efficient with the addition of the new variable (Wilkinson, 1991).

Additionally, if one of each pair of highly correlated factors was removed, some approach would be required to determine which of the factors should be excluded, particularly as the impact on individual DMUs could be significantly different using a frontier methodology such as DEA.

For example, based on the above correlation analysis, it might have been appropriate to remove 'inpatient discharges, surgical' or 'inpatient discharges, medical', as the correlation coefficient was 0.94. However, for several of the DMUs, one of these factors had a zero value but not both of them, most notably DMU #71, for which there were 3286 surgical inpatient discharges and none in the medical inpatient category. DMU #74 had a similar pattern for its inpatient activity in these two areas. Conversely, DMU #28 had no surgical inpatient activity but 981 medical discharges, with DMU #39 also following this pattern. Therefore, removing either of these factors would have an impact on these types of DMUs if they used it the calculation of the efficiency scores.

Hence, whilst analysing correlation coefficients has provided some interesting observations, it has not been used to a great extent as the defining factor for the inclusion or exclusion of particular variables. Instead, a number of factors have been used in conjunction, which will be further discussed in the next section.

A further statistical approach was the use of regression as a means for establishing that the factors to be included were related to the technical efficiency that is being measured and not simply arbitrary measures of input or output, supported by the following:

> ".. the inputs and outputs included in the model should be somewhat related experimentally, statistically, and/or conceptually, and it is also important to have information on the direction of the relationship, be it positive or negative."
>
> CHARNES *ET AL.* (1994)

A multiple regression procedure was applied, with 'total expenditure' used as the single independent variable initially, to determine if there was a link between this input and the main output factors identified from the list in table 7.2. The most significant factors were identified from a list containing all those included in the output factors correlation table excluding 'total inpatient discharges' using the stepwise regression function in MINITAB.

Stepwise regression is the process of adding in the independent variables in turn, beginning with the factor that has the most significant contribution to the R-Square value and continuing until the addition of any more factors does not significantly improve the R-Square value. This procedure is frequently used to determine the significant factors affecting sales in a marketing environment. The results obtained from a standard stepwise procedure are presented in table 7.6 overleaf.

Step	1	2	3	4
Constant	3444	2568	1764	1443
IDM	3.96	2.05	2.15	1.89
T-Value	*36.24*	***6.74***	***7.61***	***6.56***
COA		0. 189	0.153	0.142
T-Value		***6.56***	***5.37***	***5.15***
POA			0.049	0.064
T-Value			***3.56***	***4.46***
DPA				0.57
T-Value				***2.67***
S	5830	4634	4294	4118
R-Sq.	**94.80**	**96.76**	**97.26**	**97.51**

KEY: See Table 7.5 for an explanation of the codes for the output factors.

Table 7.6: Regression of Output Factors

The above regression exercise suggested, therefore, that the most significant factors affecting total costs (used as a surrogate performance measure) were 'inpatient discharges, medical', 'consultant outpatient attendances', 'PAM outpatient attendances' and 'day patient attendances', explaining 97.5 % of the variation, using the r-squared statistic. In statistical terms, such a value of R-Square suggests 'a very strong linear relationship' (Swift, 1997) between 'total costs' and the factors included and would certainly be a very acceptable result in an assessment using regression analysis.

Alternatively, the reverse of the above procedure can also be used to identify the most significant factors, whereby all the output factors were included in a regression equation, with those not classified as significant to be removed one by one, with the significance level determined to be $p < 0.05$. The completion of such a procedure also resulted in the identification of the above factors.

Allowing for a slightly lower level of significance ($p < 0.10$), six output factors were identified as significant giving a R-Square value of 97.4%, such that the regression equation would be:

TC = 1718 + 0.146*COA + 0.529*IDS + 0.628*IDG + 1.24*IDM + 0.553*DCA + 0.470*DPA

Predictor	Coefficient	St. Dev.	T Value	P Value
Constant	1718.1	663.9	2.59	0.012
COA	0.146	0.029	5.09	0.000
IDS	0.529	0.280	1.89	0.064
IDG	0.628	0.371	1.69	0.095
IDM	1.24	0.368	3.37	0.001
DCA	0.553	0.236	2.35	0.021
DPA	0.470	0.217	2.17	0.034

Figure 7.4: The Regression Equation for Total Costs

One interesting result of the above investigations using the regression process was the exclusion of the output factors for 'A&E attendances' and 'Inpatient Discharges, Other'. All the other factors appeared in one or both of the regression equations identified, whereas neither of these two was found to be significant.

The factor chosen to be the dependent variable could have been important in terms of the output factors that were found to be significant. In the simple example presented in section 5.4, with two inputs and one output analysed in a DEA model, examination of the virtual weights showed that 'total direct costs' was the more dominant of the input factors, contributing more to the efficiency score on the majority of occasions. This suggested that the two elements of cost were not related equally in the efficiency analysis and aggregating them in the regression equation could be misleading.

Repeating the stepwise regression analysis using 'total direct costs', the significant factors were identified as 'consultant outpatient attendances' and 'inpatient discharges, medical', explaining 97.4% of the variation based on the r-square value. These factors were also significant in the 'total allocated costs' regression model, along with 'PAM outpatient attendances' and 'day patient attendances', although the r-square value was 95.5% in this case. Therefore, investigating the two parts of the 'total costs' factor individually identified a similar group of factors were significant, at least in the context of a regression analysis.

However, as DEA is a frontier approach to efficiency measurement, whereas regression uses a centralising tendency, the above results provided guidance but did not give absolute answers on the factors to be included in a DEA model. One of the questions to be addressed with the heuristic analysis was whether the factors identified through regression were consistent with those included in the DEA models.

7.5.5 A Heuristic Approach

By following the above approaches, the most important factors (dependent on the available data) were identified, with their probable combinations determined. However, there were still several appropriate alternatives, which could have been used in the DEA models, and the best mix of input and output factors had not been finally determined. The final approach in the identification of the most appropriate DEA model was to run several alternative models. The results were then analysed to determine which groupings of inputs and outputs were the most appropriate measures of hospital efficiency in this case, according to some pre-determined factors to be discussed below.

This corresponded with the approach used by Ehreth (1994) that resulted in the identification of three useful DEA models from the nineteen developed originally, as was discussed in chapter six.

7.6 Analysing the Output from DEA Models

In the previous section, the different issues surrounding the selection of the variables were debated and the extensive list of potential factors reduced to more manageable proportions. In this section, the results obtained from applying the DEA methodology to the hospital data set are presented. Based on the discussion above, the alternative DEA models were selected in order to investigate the following issues:

1. The importance of using 'discharges' not 'days';
2. The effect of removing highly correlated factors;
3. The most appropriate measures of activity to be used as outputs;
4. The impact of disaggregating both input and output factors.

By focusing on these four points and using the discussion above as a guide, ten different variable combinations were eventually identified and evaluated. The variables included in each of the models are presented in table 7.7 overleaf. The table also summarises the efficiency analyses for each of the models using five general statistics. The coding system to be used in the remainder of the discussion is explained in the key below the table. The figures in brackets for each model correspond to the number of variables (number of inputs, number of outputs). The data for each of these factors is included in appendix three with information on each of the hospitals included in the sample. A comprehensive breakdown for the results from one of the models is also included in appendix three.

Model Name	Inputs	Outputs	Average Efficiency Score	Standard Deviation of Efficiency Scores	No. of Efficient DMUs	Minimum Efficiency Score	Least Efficient DMU
7.1 (4, 6)	AVE, CAP, TDC, TAC	IDS, IDM, IDG, IDO, COA, AEA	87.5%	17.44	31	11.2%	#72
7.2 (4, 6)	AVE, CAP, TDC, TAC	IYS, IYM, IYG, IYO, COA, AEA	96.1%	7.99	47	57.6%	#43
7.3 (3, 6)	CAP, TDC, TAC	IDS, IDM, IDG, IDO, COA, AEA	86.6%	17.41	28	11.2%	#72
7.4 (3, 6)	AVE, CAP, TDC	IDS, IDM, IDG, IDO, COA, AEA	86.3%	18.22	29	11.2%	#72
7.5 (3, 7)	CAP, TDC, TAC	IDS, IDM, IDG, IDO, COA, AEA, DCA	88.9%	16.62	35	11.2%	#72
7.6 (3, 6)	CAP, TDC, TAC	IDS, IDM, IDG, IDO, COA, DCA	87.7%	17.78	33	11.2%	#72
7.7 (4, 7)	AVE, CAP, TDC, TAC	IDS, IDM, IDG, IDO, COA, AEA, DCA	89.6%	16.54	38	11.2%	#72
7.8 (3, 6)	AVE, CAP, TC	IDS, IDM, IDG, IDO, COA, AEA	84.3%	18.74	28	9.7%	#72
7.9 (4, 4)	AVE, CAP, TDC, TAC	TID, COA, AEA, DCA	82.6%	19.65	24	9%	#72
7.10 (4, 6)	AVE, CAP, TDC, TAC	TID, COA, AEA, DCA, POA, DPA	88.9%	17.39	39	9%	#72

Key to Factor Codes: AVE - ‘Average Number of Staffed Beds’ CAP - ‘Capital Charge’ TDC - ‘Total Direct Costs’ TAC - ‘Total Allocated Costs’ TC - ‘Total Costs’

IDS - ‘Inpatient Discharges, Surgical’ IDM - ‘Inpatient Discharges, Medical’ IDG - ‘Inpatient Discharges, Obstetrics and Gynaecology ’ IDO - ‘Inpatient Discharges Other’

IYS - ‘Inpatient Days, Surgical’ IYM - ‘Inpatient Days, Medical’ IYG - ‘Inpatient Days, Obstetrics and Gynaecology’ IYO - ‘Inpatient Days, Other’

AEA - ‘Total A&E Attendances’ COA - ‘Total Consultant Outpatient Attendances’ TID - ‘Total Inpatient Discharges’ DCA - ‘Total Day case Attendances’

DPA - ‘Total Day Patient Attendances’ POA - ‘Total PAM Outpatient Attendances’

Table 7.7: Definitions and General Results for the DEA Models under Investigation

The models were developed to investigate four key points, as given in section 7.5, relating to identifying the most appropriate mix of variables for this analysis. Model 7.1, which had four inputs and six outputs, corresponded most closely to the models developed by Parkin and Hollingsworth (1996) to investigate a similar data set. The main differences were caused by data availability, particularly in the selection of the input factors. This model was used as the starting point for the comparative analysis to follow, with the results from model 7.1 assessed against those obtained from the other models where appropriate. Model 7.1 is referred to as the base model.

Prior to addressing the key points from this investigation, the results from the base model are presented and evaluated, by looking at the results of the model in general terms and by focusing on several DMUs in particular.

Model 7.1 had four inputs (‘Average Number of Staffed Beds’, ‘Capital Charge’ ‘Total Direct Costs’ and ‘Total Allocated Costs’) and six outputs (‘Inpatient Discharges, Surgical’, ‘Inpatient Discharges, Medical’, ‘Inpatient Discharges, Obs and Gynae’, ‘Inpatient Discharges Other’, ‘Total Consultant Outpatient Attendances’ and ‘Total A&E Attendances’).

Table 7.7 above showed that the DEA analysis of the 74 acute hospitals in Scotland, using model 7.1, resulted in 31 hospitals being classified as efficient, with the average efficiency score being 87.5%. The results from this model were investigated further, using similar approaches to those found in the DEA papers evaluated in chapter six.

The efficiency scores for each of the DMUs have been dissected according to several categories, in order to investigate the nature of the efficient hospitals. Unlike the 'restricted peer group approach' used by Ozcan *et al.* (1992), all the DEA efficiency scores have been calculated using the sample as a whole. Each hospital was assumed to operate in relation to the same efficiency frontier.

In figure 7.2 above, illustrating the structure of the NHS in Scotland, it was shown that there were fifteen health boards, each of which covers a particular geographical region. The health boards have responsibility for local strategy and implementation and each of the NHS Trusts within their boundary is accountable to them for the services they provide (The Scottish Office, 1997). Hospitals in each category should have similar strategies and environmental influences. Table 7.8 shows the average efficiency score for each health board region, based on the results obtained from model 7.1.

As was discussed in section 7.3, there are some regions, notably B, R and Z, with just one acute hospital within them and several more of the health boards are responsible for just two acute hospitals. The largest health board is N, which has thirteen units and over half of them were classified as efficient. Excluding the very small health boards with less than five acute hospitals, board S had the highest average efficiency score at 93.1% in comparison with the sample average of 87.5% and half of its DMUs were found to be efficient. (The codes used are equivalent to those used by the Scottish Office and are defined in full in appendix one.)

Health Board	Number of DMUs	Average Efficiency Score	Standard Deviation	No. of Efficient DMUs
A	4	93.3	8.7	1
B	1	94.9	-	0
C	9	86.2	15.9	3
F	5	78.4	27.7	2
G	7	92.7	8.1	3
H	6	85.4	10.3	1
L	4	95.5	5.5	2
N	13	81.5	28.1	7
R	1	83.8	-	0
S	8	93.1	8.0	4
T	8	84.4	19.1	3
V	2	93.7	8.9	1
W	2	80.4	27.8	1
Y	2	100	0	2
Z	1	78.3	-	0

Table 7.8: Analysis of Efficiency Scores by Health Board

Efficiency scores were also evaluated according to hospital type, using the Scottish Office functional classifications discussed in section 7.3 and defined in appendix one. The sample contained hospitals from twelve of the acute hospital categories and some of the classes contain a very small number of hospitals, most notably 04, 07, 13 and 14, which all have three or less hospitals allocated to them. (A similar situation can be observed with the health board categories as shown in table 7.8.)

Table 7.9 overleaf shows the average efficiency scores for each of them, as well as a number of other summary statistics. (Definitions of the functional classes can be found in appendix one.)

Functional Classification	Number of DMUs	Average Efficiency Score	Standard Deviation	No. of Efficient DMUs
01	7	96.8	6.79	5
02	13	92.6	12.5	6
04	2	91.5	12.1	1
05	10	86.8	9.3	2
07	3	95.4	8.0	2
08	4	72.0	23.43	1
09	4	86.7	26.7	3
10	12	89.1	16.0	7
11	7	87.3	14.2	2
12	8	84.2	20.4	2
13	3	60.7	44.1	0
14	1	89.8	-	0

Table 7.9: Efficiency Score by Functional Classification

For the larger classes, 01 had the highest average efficiency score (96.8%) and five of its seven hospitals were classified as efficient. Class 12 had the lowest average efficiency score (84.2%) but two of the eight hospitals were still classified as efficient. In terms of the spread in efficiency scores for a particular class, 01 and 05 had two of the lowest values for their standard deviations, although they had very different average efficiency scores.

In order to determine if the observed differences were significant, some statistical analysis was required. Tests such as ANOVA (Analysis of Variance) were discarded (see section 5.9) as their assumption of normality was found to be unacceptable. However, other statistical methods were available, such as the Mann-Witney test and the approach used by Ozcan *et al.* (1992), whereby they examined the percentage of efficient DMUs in various categories using the two-tailed z-test for proportions.

Examining the efficiency scores in relation to functional classification or health board using such approaches was inappropriate, as there were large number of categories containing a very small number of DMUs. Therefore, broad categories with a greater number of DMUs in each were developed, with further statistical analysis to follow.

Table 7.10 shows the equivalent analysis to those illustrated in tables 7.8 and 7.9, with the DMUs divided up by size rather than functional classification or health board region. Six categories of size were defined based on the factor 'average number of staffed beds', with between 9 and 18 DMUs in each of them.

Hospital Size (By No. of Beds)	No. of DMUs	Average Efficiency Score	Standard Deviation	No. of Efficient DMUs
< 50	18	90.9	13.2	10
50 – 100	13	75.7	21.9	4
100 – 200	9	80.3	27.6	2
200 – 400	16	90.1	10.8	4
400 – 600	9	92.3	14.7	5
> 600	9	95.9	7.3	6

Table 7.10: Efficiency Scores and Hospital Size

The highest average efficiency score was found to be for the hospitals with the greatest number of staffed beds, although the category for small hospitals also demonstrated a high value for its average efficiency score.

The observations based on table 7.10 suggested that the classification of efficiency in the sample of DMUs could be connected to hospital size, defined according to the number of beds.

There were other potential options for comparisons of the efficiency scores across the DMUs to reflect hospital size, as this approach measures size according to inpatient activity whereas some of the hospitals in the sample focus more heavily on other areas, such as outpatient attendance.

Grouping efficiency scores according to two equally sized categories of total expenditure (less than £10million and greater than £10million) resulted in the larger hospitals having an average score of 92.7% compared with 82.3% for the smaller hospitals. However, application of the two-tailed z-test for proportions using the approach defined in section 5.9.3 and discussed in chapter six, did not identify this observed difference to be significant ($p < 0.05$). The spread of the efficiency scores, based on standard deviation, was much wider in the 'small' category but this is not measured in the z-test applied. Further analysis of this type is discussed in later chapters.

As an alternative, the efficiency calculations were also analysed according to the output factors, in order to determine if there were any particular characteristics of the efficient units. This was carried out by examining the dominant output factors, that is, those that contributed most to the efficiency calculations according to the virtual output percentages. The virtual output was defined in chapter five to be the 'product of that output and its corresponding weight' (Thanassoulis *et al.*, 1987).

In this instance, 'consultant outpatient attendances' and 'inpatient discharges, medical' were the two factors contributing most to the efficiency score on the majority of occasions (18 out of 31) for the efficient units.

Table 7.11 below shows the frequency distribution for each the dominant factors (for the output variables) for all the DMUs, not just the efficient units. As was discussed in chapter five, the weights identified for each DMU were not necessarily unique. The analysis that follows is based upon one possible set of weights for each DMU.

FACTOR	IDS	IDM	IDG	IDO	COA	AEA
FREQUENCY	18	20	5	4	19	8

KEY: See table 7.7 for a definition of the output codes

Table 7.11: Distribution of Dominant Outputs

As can be seen from the above table, although three factors clearly dominated the efficiency calculations, all of the factors were dominant for at least four DMUs. However, for the inputs, the 'total direct costs' factor was clearly the most significant factor in the efficiency calculations as it was the dominant input factor for 53 of the 74 DMUs. Further discussion of the importance of the virtual weights is given in chapter eight in the second stage of the DEA application procedure.

A further phase of the analysis at this stage was to examine the reference groups for the non-efficient DMUs. This illustrated which of the efficient DMUs were the most influential in the determination of the efficiency calculations for the inefficient DMUs and thus the efficiency frontier. In this case, six of the efficient DMUs (#35, #50, #22, #51, #13 and #46) formed part of the reference group on more than ten occasions, with #35 cited in the reference group for 29 of the inefficient DMUs. A comprehensive breakdown of the reference set information for model 7.1 is found in appendix three.

Investigating the DMUs individually, particularly those with a very low efficiency score, provided useful information on inefficient practice. In this model, there were six DMUs with efficiency scores of less than 60%, referenced in table 7.12. Each of these had no patient activity in at least two of the output categories.

DMU	Efficiency Score	Non-zero Output Categories	Dominant Input Factor (Virtual)	Dominant Output Factor (Virtual)
#14	56.6%	COA, IDS, IDM, IDO	TDC	IDM
#36	46.5%	COA, AEA, IDO	TAC	AEA
#41	46.7%	COA, AEA, IDG, IDO	TDC	IDO
#52	55.3%	COA, AEA, IDG, IDO	AVE	IDO
#70	38.4%	COA, IDS, IDM, IDO	TDC	COA
#72	11.2%	IDO	TAC	IDO

KEY - See Table 7.7 for a definition of the factor codes

Table 7.12: Analysis of the Least Efficient DMUs

The least efficient unit, DMU #72, had only one non-zero output category ('inpatient discharges, other'), which may have had a distorting effect on its efficiency assessment. This will be further discussed in the next section, as using inpatient days in the DEA model (model 7.2) resulted in DMU #72 being classified as efficient.

There did not seem to be any compelling similarities between these DMUs, which could be used to explain their very low efficiency rating. They were all among the smaller DMUs, with four of the six having a total expenditure of less than £3million. They also generally tended to have a very small amount of inpatient activity in comparison with their outpatient activity levels (excluding DMU #72 as discussed above).

The impact of changing the variable set on this particular group of DMUs is debated in the following sections, along with reference to other DMUs that were particularly affected by the changes. Additionally, other options available for analysing the output from the DEA model were used to determine the importance of the particular variables included in each model, beginning with a discussion on the nature of the variables used to measure inpatient activity.

7.6.1 Using 'Discharges' or 'Days' to Measure Inpatient Activity

As can be seen from table 7.7 above, where just simple summary information was given for the output from each of the DEA models, there were extensive differences between the efficiency calculations for models 7.1 and 7.2. Model 7.2 was equivalent to model 7.1 except that inpatient activity was measured according to the number of days (length of stay) rather than the number of patients treated.

The average efficiency score was lower in model 7.1 by 8.6% (87.5% compared to 96.1%). There were also fewer DMUs classified as efficient (31 compared to 47). The changes in comparison with model 7.1 were as follows: 10 efficiency scores decreased, 37 increased and 27 remained the same. Of the 31 efficient units under model 7.1, all but four of these were efficient in model 7.2, and were thus unaffected by a change in the approach used to measure inpatient activity.

As the majority of DMUs either increased their efficiency score in the second model, or remained the same, the DMUs that were less efficient in model 7.2 were of particular interest.

Whilst the average decrease in efficiency score over all the ten DMUs was small (less than 10%), two DMUs have very large decreases (more than 20%) in their efficiency score (#30: 88.0% -> 66.0% and #43: 100% -> 57.6%). Of the four DMUs that were no longer classified as efficient, the reduction in the efficiency score was less than 10%, apart from DMU #43, which was highlighted above.

The output patterns for these DMUs, when compared to those with increased efficiency scores were investigated, with the relationship between days and discharges appearing to be of particular importance. For DMU #43, which showed a dramatic reduction in its efficiency rating under model 7.2, the calculated figure for 'days per discharge' was 5.19. The average 'days per discharge' figure for all the DMUs was 15.34 and for DMU #44, which increased its efficiency score from 75.57% to 100%, the equivalent value was 49.97. The significance of this relationship was further investigated by an examination of the group of DMUs identified in table 7.12 above as the least efficient units in the sample under the 'discharges' model (model 7.1). This is shown in table 7.13.

DMU	Model 7.1	Model 7.2	Difference	Days/Discharge
#14	56.6	100	43.4	22.17
#36	46.5	100	53.5	52.43
#41	46.6	74.5	27.9	18.75
#52	55.3	89.6	34.3	23.58
#70	38.4	90.4	52.0	62.47
#72	11.2	100	88.8	215.75
Average	**42.4%**	**92.4%**	**50.0%**	**65.86**

Table 7.13: Changes in Efficiency Scores

Three of the least efficient DMUs were re-classified as efficient in model 7.2 and the average efficiency score for this small group showed an increase of 50%, from 42.4% to 92.4%.

The relationship between 'days per discharge' and the changes in efficiency score between the two models was further investigated by looking at the whole sample of DMUs. The analysis is shown in table 7.14 below. The DMUs with the highest average value for 'days per discharge' were identified as those whose efficiency score had increased when 'days' was used to measure inpatient activity rather than 'discharges'.

Nature of Change In Efficiency Scores	**Average no. of days per discharge**
Decreased	6.29
Stayed the Same	9.78
Increased	21.85
Overall	**15.34**

Table 7.14: Evaluation of Changes in Efficiency Scores

Using a one-way ANOVA test on the 'days per discharge' data, the relationship between the change in efficiency scores and the number of 'days per discharge' was found to be slightly significant ($F = 2.41$, $p = 0.097$), that is, acceptable at the 90% level but not at the 95% level. This suggested that, whilst 'length of stay' was a significant contributing factor, it was not the only influence affecting the differences in efficiency scores. The inclusion of a case-mix index could negate this influence – hospitals with above average values for the length of stay indicator may have been dealing with more severe cases, requiring lengthy hospital treatment.

The analysis carried out comparing the results from these two models confirmed the observations made in the majority of the research discussed in chapter six, where the use of 'patient days' was generally rejected. A model that suggested that hospitals with above average lengths of stay should be regarded as more efficient than those with lower values was likely to be unacceptable, as it was contradictory to the view expressed below:

> "… a hospital that takes more days of care to cure a patient is less efficient than one curing a similar patient with less days of care."
>
> EHRETH (1994)

Aggregation of the inpatient variables or the removal of particular categories of inpatient activity may have negated the immediate impact of the inclusion of 'inpatient days', particularly in relation to the small number of DMUs with a number of zero values amongst their outputs factors. DMU #72, for example, which demonstrated the most dramatic change in its efficiency score in the second model, had zero values in all but one of the inpatient categories, which was 'inpatient discharges (days), other'. However, as was to be expected from the earlier discussion, the greater acceptability of the 'discharges' models was clear and all the subsequent models used 'discharges' to measure inpatient activity.

7.6.2 Removing Highly Correlated Factors

Following the decision to concentrate on models that used 'discharges' to measure inpatient activity, the next phase of the investigation into variable selection was to identify any redundant factors, based on the statistical analysis discussed in section 7.5.4.

Amongst the inputs, it was observed that there was a high correlation between 'average number of staffed beds' and the two expenditure factors, 'total direct costs' and 'total allocated costs'. Also, there was a high correlation between the two cost variables themselves.

Analysis of the output factors using correlation, with the results detailed in table 7.6, also identified that certain factors may be redundant, using a high value for the correlation coefficient as the indicator for redundancy. The impact of removing some of the highly correlated outputs is reported in the following sections, where the investigation focuses upon selecting the most appropriate mix of outputs and the impact of aggregation. Links with the correlation analysis are also given.

The impact of removing or replacing some of the correlated inputs was investigated by comparing the output from models 7.1, 7.3, and 7.4, using the definitions found in table 7.7 above. In model 7.3, the 'average number of staffed beds' input was excluded, whereas in model 7.4 it was the 'total allocated costs' factor. As to be expected in DEA analyses, the number of efficient units was reduced by the removal of one of the input factors when compared to the base model (model 7.1 in this case). The changes were quite slight, in that model 7.3 had 28 efficient DMUs and in model 7.4 there were 29 efficient units, in comparison with the 31 efficient units in model 7.1.

Also, differences in the average efficiency score were negligible – the average efficiency score for the base model was 87.5% whereas for model 7.3 it was calculated to be 86.6% and for model 7.4 it was 86.3%.

Further information on the impact of the variable set was observed by investigating the distribution of the efficiency scores. Figure 7.5 below illustrates this pattern for the three models (7.1, 7.3 and 7.4), showing that the distribution of the efficiency scores appeared to be fairly constant throughout, despite the changes to the variable set.

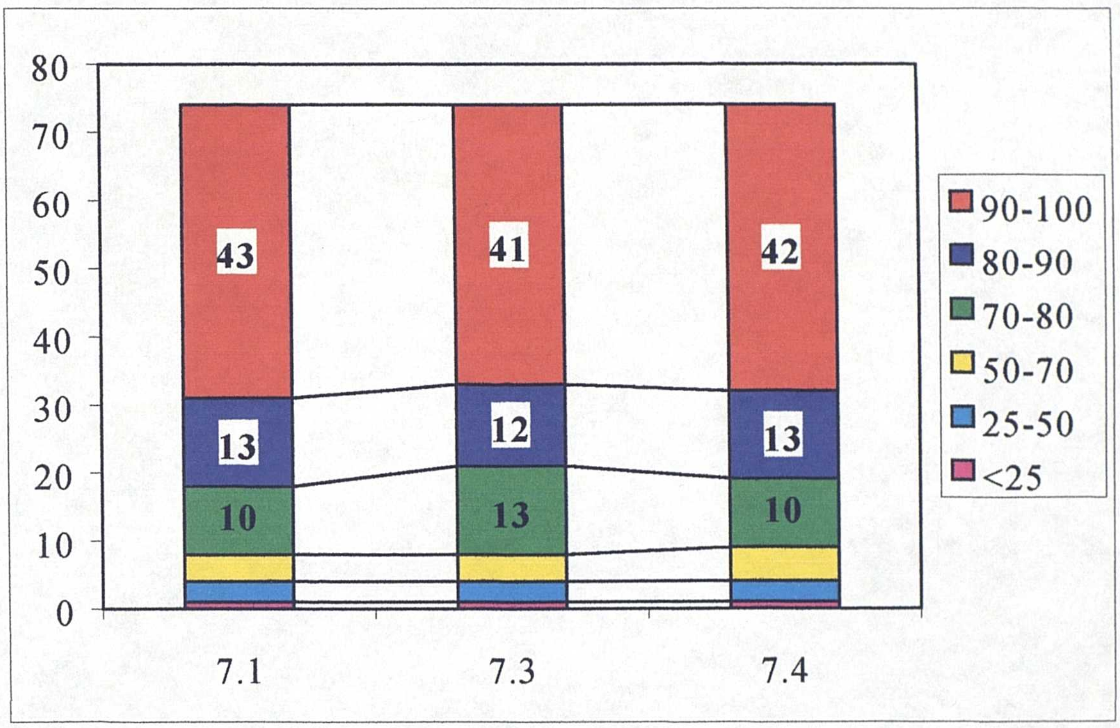


Figure 7.5: The Distribution of Efficiency Scores for Models 7.1, 7.3 and 7.4

The above discussion shows that the changes in the variable set did not appear to have a significant effect on the efficiency scores obtained in general terms. However, as was observed above, the number of efficient units had decreased along with the average efficiency score.

Therefore, it was important to identify if there were particular characteristics displayed by those DMUs whose efficiency ratings did change, which distinguished them from the group of DMUs relatively unaffected by the changes to the variable set.

The differences between the efficiency score in the base model and the alternative models were calculated for each of the DMUs in order to identify those most or least affected by the exclusion of either of the two inputs. In those cases where an amendment to the variable set produced a large change in efficiency score, the DMU was investigated in greater detail. This approach is used throughout the thesis to assess the impact of changes to the variable set or amendments to the model formulation.

The two models discussed in this section in comparison with the base model both have one less variable included. As such, it was to be expected that the impact of this would be negative, in terms of reductions to the efficiency scores, if there was any change at all.

Looking first at the differences between model 7.1 and model 7.3, where the 'average number of staffed beds' factor was excluded, there were just eleven DMUs with a reduction in their efficiency score, with this greater than 5% for six of them. The average change in the efficiency score was 0.9%, although the average size of the change (with all the zeros excluded from the calculations) was 6.1%.

This was an expected result, as further analysis of model 7.1 shows that only 13 of the DMUs actually had any contribution from this factor in the calculation of their efficiency score. Two of these DMUs were able to obtain an equivalent efficiency to that in model 7.1 by reallocating the factor weights applied. (This result demonstrated that the factor weights identified were not necessarily unique, as an equivalent efficiency could be achieved if they were assigned in a slightly different pattern.)

The remaining 11 DMUs were unable find an alternative allocation of their factor weights to maintain their efficiency score, as including a contribution from 'average number of staffed beds' ensured they were represented in 'the best possible light'.

The largest single change between the two models was for DMU #54, with a reduction of 21.1% in its efficiency rating. As was identified above, three DMUs (#4, #34 and #38) were no longer classified as efficient in model 7.3. Investigation of the virtual weights shows that all of the DMUs with a large reduction in their efficiency scores based their assessment in model 7.1 on a significant contribution from the 'average number of staffed beds' factor.

For example, for DMU #38, its efficiency score in model 7.1 was calculated with a virtual weight of 100% attached to 'average number of staffed beds'. In model 7.3, this 100% virtual weight was transferred to the 'total direct costs' factor, with the result that its efficiency score was reduced by just 7.8%. In general, the DMUs with a contribution from the 'average number of staffed beds' transferred the attached weight to the 'total direct costs' factor in model 7.3.

The impact of removing the 'total allocated costs' factor in model 7.4 was equally insignificant, as only fourteen of the DMUs showed any reduction in their efficiency score, with a decrease of greater than 5% for just seven DMUs. The average change across all 74 DMUs was 1.2%, slightly higher than the figure for the differences between models 7.1 and 7.3. The impact of the change on the DMUs affected was a reduction in their efficiency score of 6.4% on average.

DMU #30 was most affected by the change, with its efficiency score reduced by 16.8% from 88% to 71.2%. The virtual weight attached to 'total allocated costs' in model 7.1 was 100%; that is, the other inputs all have a zero weight attached to them. In model 7.4, this has transferred to the 'average number of staffed beds' factor, although the level of inpatient activity at this hospital was relatively small, with less than 1,000 discharges in total. Two DMUs rated as efficient in model 7.1 were no longer efficient in the amended model, these being #48 and #57, with a corresponding reduction in their efficiency score of 9% and 12.4% respectively.

The impact of removing the 'total allocated costs' factor was more tangible, as a greater number of DMUs had some contribution from this factor in their efficiency calculation in model 7.1. However, the number of DMUs potentially affected was still only 25 (the number of DMUs with a contribution from the 'total allocated costs' factors in the efficiency calculations). Only 14 of these 25 DMUs actually showed a reduction in their efficiency score in model 7.4, with the remaining 11 DMUs able to reallocate their factor weights and maintain their efficiency rating at the same level.

As has been observed, the impact of removing either 'total allocated costs' or 'average number of staffed beds' was quite negligible in general terms and for the majority of the DMUs in particular. This seems to be primarily due to the fact that very few of the DMUs based their efficiency assessment on a significant contribution from either of these factors. However, as differences were observed between models 7.3 and 7.4, the impact of removing one of a pair of highly correlated inputs was clearly dependent on the factor chosen for removal.

In the original analysis, 'average number of staffed beds' was found to be highly correlated with both 'total allocated costs' and 'total direct costs'. Therefore, the 'total direct costs' factor could be excluded from the model, with the other two factors remaining. The 'total direct costs' factor was a significant contributor to the efficiency evaluations of the majority of the DMUs.

The removal of this factor would have been much more influential, particularly in the patterns for the factor weights. Also, this factor was the most significant cost input into the hospital system, representing the personnel involved in providing all types of treatment. It would be difficult to justify its exclusion at a 'policy' level, not just in terms of the results from the DEA analysis.

In the two models investigated in this section, the 'total allocated costs' factor appeared to be of greater importance – a higher number of DMUs relied on this factor and the impact of removing the variable was slightly more apparent in the efficiency assessments. Whilst 'average number of staffed beds' was related primarily to the provision of inpatient services, 'total allocated costs' was linked to all the outputs.

The removal of 'average number of staffed beds' from the variable set may be justified, although the decision should be based on other factors than just the results from the correlation and DEA analyses. This is further discussed in chapter nine, in relation to the amount of managerial involvement required for a successful application of the DEA methodology. The further importance of this factor is assessed in some of the subsequent models, as it was still included in a number of them.

This factor ('average number of staffed beds') also illustrates one of the key issues involving the restriction of factor weights, as was discussed in chapter five. It was selected for inclusion in the models, suggesting its importance to efficiency evaluation, but was then excluded from the efficiency assessments for the majority of the DMUs. This is further discussed in chapter eight, where weight restrictions are introduced in the second stage of the DEA model-building process.

7.6.3 Investigating Alternative Measures of Output

It was seen in the previous section that the removal of certain highly correlated input factors had little or no effect on the efficiency assessment. It appeared as though this was primarily due to the fact that the factors removed in each case were not significant contributors to the efficiency evaluation. It was also suggested that the decision on which factors to exclude should not have been taken based only on the results of the statistical analysis, but also on an understanding of the situation under investigation.

Using a combination of models with three or four inputs (equivalent to the inputs used in models 7.1 and 7.3), a number of additional models were identified in order to investigate the impact of alternative output patterns. In section 7.5, eight potential measures of output were identified, excluding the 'total inpatient discharges' factor and focusing on measures that reflected inpatient activity by the number of discharges. Before assessing the impact of aggregating both the input and output factors, some alternative combinations of output factors were examined, which also reflected on the results obtained from the correlation analysis. Three additional models are included in the discussion in this section, these being 7.5, 7.6 and 7.7, as defined in table 7.7.

Model 7.5 included an additional output measuring day case activity, whereas model 7.6 assessed the impact of replacing 'A&E attendances' with 'day case attendances'. Both of these models have just three inputs and as such were related back to model 7.3, which had the same set of input factors. This ensured that the differences observed were related only to the change in the outputs. Model 7.7 examined the impact of an additional output factor ('day case attendances') in relation to the original model with four input factors. The general results for these models, in terms of the average efficiency score and the number of efficient units, was given in table 7.7.

The general impact of changing the variable set was investigated by looking at the number of units identified as efficient, the average efficiency score and the distribution of the efficiency scores. Referring to the information contained in table 7.7, the addition of extra variables (with model 7.5 compared with model 7.3 and model 7.7 compared with model 7.1) had the effect of increasing both the average efficiency score and the number of efficient variables. There were an extra 7 efficient DMUs in model 7.5 (35 compared to the 28 for model 7.3) and the average efficiency score increased by 2.3%. In model 7.6, the number of efficient units (33) and the average efficiency score (87.7%) had both increased, although the impact of replacing the 'A&E attendances' variable with 'day case attendances' did not necessarily have a positive effect on all DMUs. This was investigated with reference to the individual DMUs.

It was observed in the previous section that the changes in the variable set did not have a dramatic effect on the distribution of the efficiency scores and the pattern remained fairly constant across all the models investigated.

This pattern was repeated with the changes made to the output factors, as can be seen in figure 7.6 below, comparing the efficiency distributions for models 7.5 and 7.6 with that from model 7.3, which had the same number of inputs.

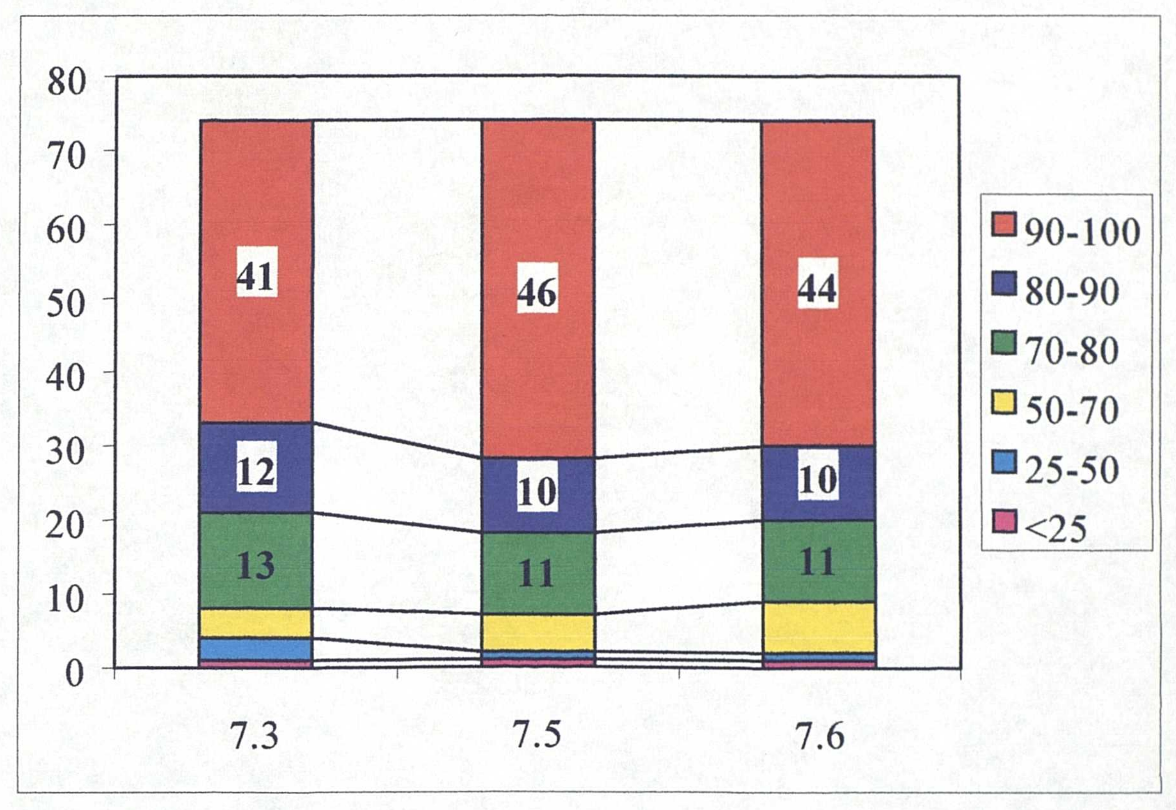


Figure 7.6: Distribution of Efficiency Scores for Models 7.3, 7.5 and 7.6

A similar pattern was observed in comparisons between model 7.1 and 7.7, where the number of DMUs with an efficiency score of above 90% was 43 in the base model and 48 in the amended model with the additional 'day case attendances' factor included. The number of efficient DMUs increased from 31 to 38, with a corresponding increase of 2.1% in the average efficiency score.

Looking at the individual DMUs, comparisons between model 7.3 and 7.5 demonstrated that 17 of the DMUs showed an increase in their efficiency score following the addition of 'day case attendances'.

The range of the increases was interesting, in that DMU #8 showed an improvement in its efficiency score of just 0.2% whereas the efficiency score for DMU #37 increased by 39.6%. This DMU achieved an efficiency score of 100% in the model with the extra variable included and the virtual weight for the 'day case attendances' was set at 100% in model 7.5. In model 7.3, the dominant output factor for DMU #37 was 'A&E attendances' with a virtual weight of 65.8%. (The two other factors with non-zero virtual weights were 'consultant outpatient attendances' and 'inpatient discharges, other'.)

Two other DMUs (#6 and #68) showed an increase of greater than 20% in their efficiency scores, which also led to them being classified as efficient. For the seven DMUs that were reclassified as efficient (#6, #31, #37, #38, #66, #68 and #71), the average increase in efficiency score was 15.6%, with 5 of them having 'day case attendances' as the dominant output factor in the amended efficiency assessment. Thirty-four of the DMUs in the sample included some contribution from the new factor in the efficiency analysis, that is, 'day case attendances' had a non-zero virtual weight for almost half of the DMUs.

Model 7.6 was developed in order to assess the impact of replacing 'A&E attendances' with 'day case attendances', such that there were still six outputs included in the model. The overall effects of the changes were highlighted previously, such that there was an overall increase in the number of efficient DMUs and the average efficiency score, when compared with the results from model 7.3. However, closer inspection of the results showed that the effect on the individual DMUs was not consistent across the whole sample, as was suggested above.

Sixteen of the DMUs showed an increase in their efficiency score, with the largest increase being 39.6% for DMU #37 (an identical increase to that observed for model 7.5). With the exception of DMU #29, the DMUs that showed an increase in their efficiency score with the inclusion on 'day case attendances' in model 7.5 demonstrated an equivalent improvement in model 7.6.

The exclusion of 'A&E attendances' had no effect on these DMUs, primarily because they had no significant contribution from this factor in their efficiency assessment or the inclusion of 'day case attendances' overshadowed its influence.

DMU #29 was one of eleven DMUs that had a reduced efficiency score in model 7.6, in comparison with the results obtained from model 7.3. For DMU #29, the reduction was relatively small (1.2% less than its efficiency score in model 7.3 and 1.5% lower than in model 7.5). However, three DMUs (#36, #40 and #45) had a reduction of more than 10% in their efficiency score, with two of these (#40 and #45) no longer classified as efficient following the exclusion of 'A&E attendances'. Each of these DMUs had this factor ('A&E attendances') as its dominant output in model 7.3, with the weight transferred to a variety of other factors in model 7.6.

Model 7.7 was identified in order to assess the impact of including 'day case attendances' whilst still incorporating the same four inputs as were used in model 7.1. The impact of this additional output factor was consistent with the observations made regarding the differences between models 7.3 and 7.5. There were seven extra units classified as efficient and the average efficiency score increased by 2.1%.

Sixteen DMUs increased their efficiency score and the largest change was again for DMU #37 (39.6%), which was identical to the improvement shown in both of the other two models investigated in this section. Each of the DMUs that showed an increase in their efficiency score under model 7.5 exhibited an equivalent increase in model 7.6, when compared with the efficiency scores for model 7.1, excluding DMU #38.

DMU #38 maintained its efficiency score (100%) across models 7.1 and 7.7, suggesting that the combination of 'average number of staffed beds' and 'day case attendances' was influential in its efficiency analysis. It was rated as 100% efficient in the model that included both of these factors (7.7) and also in the models that incorporated either of them (7.1 and 7.5). However, its efficiency score was reduced to 92.2% in model 7.3, which excluded both 'average number of staffed beds' and 'day case attendances'.

The three models discussed in this section (7.5, 7.6 and 7.7) were identified in order to investigate the impact of including an extra output, considered to be a significant measure of hospital activity. The increase in the number of surgical procedures carried out as day cases was identified as a major operational change for the future of the NHS, as discussed in chapter three.

It was found that several DMUs were able to significantly improve their efficiency score when this factor was included, although the exclusion of 'A&E attendances' was found to decrease the efficiency rating for a small number of the DMUs.

In the regression analysis discussed in section 7.5.4, 'day case attendances' was identified as a significant factor, whereas 'A&E attendances' was not included in any of the regression equations. The analysis of the DEA models in this section, and the extra information on hospital activity,. would confirm this result in relation to the efficiency analyses. Including 'day case attendances' alongside, or in place of, 'A&E attendances' would seem to be an appropriate step in the development of future DEA models of hospital performance.

The importance of including alternative measures of output is discussed in the following section, where the implications of aggregating various factors in the DEA models is also discussed. It was not possible to investigate every combination of outputs, particularly the exclusion of some of the inpatient factors, as there were several DMUs with zero entries in these variables, as explained in section 7.4.3. For example, as the 'inpatient discharges, other' was distorted by the number of long stay patients at some of the hospitals, it might be sensible to remove the factor. However, this was impossible without first refining the data sample, which is discussed in section 7.7. Additionally, the number of combinations was limitless and therefore some restrictions were necessary. However, in this case, most of the subsequent models have included both 'day case attendances' and 'A&E attendances'.

7.6.4 The Impact of Variable Aggregation

Unlike the audit approach considered in chapter four, where each case is evaluated individually, the application of the DEA technique to health care data automatically requires some degree of aggregation amongst the variables.

For example, patients in specific categories such as 'medical inpatients' have been grouped together in the models used in the preceding analysis, although there are actually ten different specialties within this grouping. All 'day case attendances' have been included in one output, although this covers many different specialties.

The issue discussed in this section is whether further aggregation of some of the inputs and outputs is viable and appropriate. Aggregating specific factors has the effect of reducing the number of variables required, whilst still representing the required areas of input and output. However, information produced in the modelling process is generalised, so that observations on efficiency can only be related to 'total costs' or 'total inpatient discharges', rather than the individual categories. The level of aggregation selected depends, primarily, on the degree of information required from the DEA model.

In the models used in this analysis, there were two groups of factors for which aggregation could have been appropriate: the cost factors ('total direct costs' and 'total allocated costs') and the inpatient discharges ('inpatient discharges, medical', 'inpatient discharges, surgical', 'inpatient discharges, obstetrics and gynaecology' and 'inpatient, discharges, other'). To investigate these factors, three further models were developed.

Model 7.8 was developed to investigate the impact of aggregating the cost factors ('total costs') and contained the same output factors as were used in model 7.1. The aggregated inpatient factor ('total inpatient discharges') was used in two models (7.9 and 7.10) with a number of other output factors also included.

7.6.4.1 Aggregating the Cost Factors (Inputs)

The general impact of aggregating the cost factor (comparing models 7.1 and 7.8) again appeared to be quite negligible, with the general statistics presented in table 7.7. The average efficiency score decreased by 3.2% (87.5% -> 84.3%) and the number of efficient DMUs was reduced from 31 to 28. This change to the model appeared to be slightly more influential in general terms than the removal of the 'total allocated costs' factor (model 7.4, addressed in section 7.6.3), especially in terms of the number of DMUs whose efficiency scores changed.

The number of DMUs affected by the change was considerable, in that 44 of the 74 DMUs in the sample showed a reduction in their efficiency score, comparing model 7.8 with model 7.1. Additionally, the size of the decrease was greater than 5% for exactly half of these DMUs.

In model 7.1, all of the DMUs (except #4 and #54) included contributions from either 'total direct costs' or 'total allocated costs' (or both of them) in the calculation of their efficiency score. Therefore, it was not unexpected that such a large number of DMUs were affected by the aggregation of the expenditure factors.

The DMU with the largest reduction in its efficiency score was #30, as was the case for model 7.4, when the 'total allocated costs' factor was removed. Its efficiency score decreased by 16.8% (88.0% -> 71.2%), with the factor weight transferred from 'total allocated costs' to 'average staffed beds' (repeating its assessment in model 7.4).

DMU #23 showed a very similar decrease (91.2% -> 75.0%), although its allocation of the input weights was very different. In model 7.1, 100% of its virtual input weight was attached to 'total direct costs', whereas in the aggregated costs model, almost all of this was transferred to 'total costs', with 0.7% being allocated to 'capital charge'. Three other DMUs with large decreases (#24, #58 and #63) had 'total direct costs' as their dominant input factor in model 7.1, replaced by 'total costs' in model 7.8.

The three DMUs that were no longer classified as efficient (#42, #57 and #64) showed reductions in their efficiency scores of 6.2%, 8.2% and 1.5% respectively. DMUs #42 and #64 were both efficient in model 7.4, where 'total allocated costs' was excluded, and attached no weight to this factor in model 7.1. In the aggregated input model (7.8), their dominant input factor was 'total costs'.

DMU #57, on the other hand, showed a much larger decrease in its efficiency in model 7.4 (100% -> 87.6%) and it had 'total allocated costs' as its dominant input factor in model 7.1. DMU #57 appeared to be unusual, as all the other DMUs rated as 100% efficient in model 7.1 with 'total allocated costs' as their dominant input factor, were still efficient in model 7.8, which had an aggregated cost factor.

Although the impact of the aggregation appeared to be negligible in general terms, it reduced the amount of information produced by the DEA modelling process. This was particularly significant in relation to the 'total allocated costs' factor, which was an indirect measure of the amount each hospital spent on 'bureaucracy'.

The Labour Government, as was discussed in chapter three, has emphasised that hospitals should reduce the amount of their budget spent in this area, focusing instead on direct spending on patient care, measured in the DEA models by 'total direct costs' (Department of Health, 1997). Also, as the DEA model was formulated with an input-minimisation orientation, it would seem appropriate to provide as much information about the significant inputs as possible.

Therefore, aggregating the cost factors did not seem to be appropriate in this study, as it reduced the amount of information provided and did not comply with recent policy objectives in health care. This was consistent with the results produced by Parkin and Hollingsworth (1996), who found little justification for further aggregation of the cost inputs used in their analysis of hospital efficiency.

7.6.4.2 Aggregating Inpatient Discharges (Outputs)

Two new models are discussed in this section, each of which includes the aggregated output factor 'total inpatient discharges' and a selection of other outputs. Model 7.9 included the same inputs as were used in model 7.1, but has only four outputs: 'total inpatient discharges', 'day case attendances', 'A&E attendances' and 'consultant outpatient attendances'. Model 7.10 had the same inputs and two additional outputs ('PAM attendances' and 'day patient attendances'). In order to assess only the impact of aggregation, however, the results from model 7.9 were compared with those obtained from model 7.7, which has an equivalent set of variables apart from the disaggregated inpatients factors. The effect of the additional outputs included in model 7.10 was investigated by comparing its results with those from model 7.9.

Figure 7.7 below shows the distribution of the efficiency scores for models 7.7, 7.9 and 7.10. As can be seen, the general impact of the aggregation of the inpatient factors (comparing the distribution of the efficiency scores between model 7.7 and 7.9) was a reduction in the number of DMUs with efficiency scores at the top end of the scale, that is, above 90%. Adding in extra variable (comparing models 7.9 and 7.10) had the effect of increasing the number of highly efficient units.

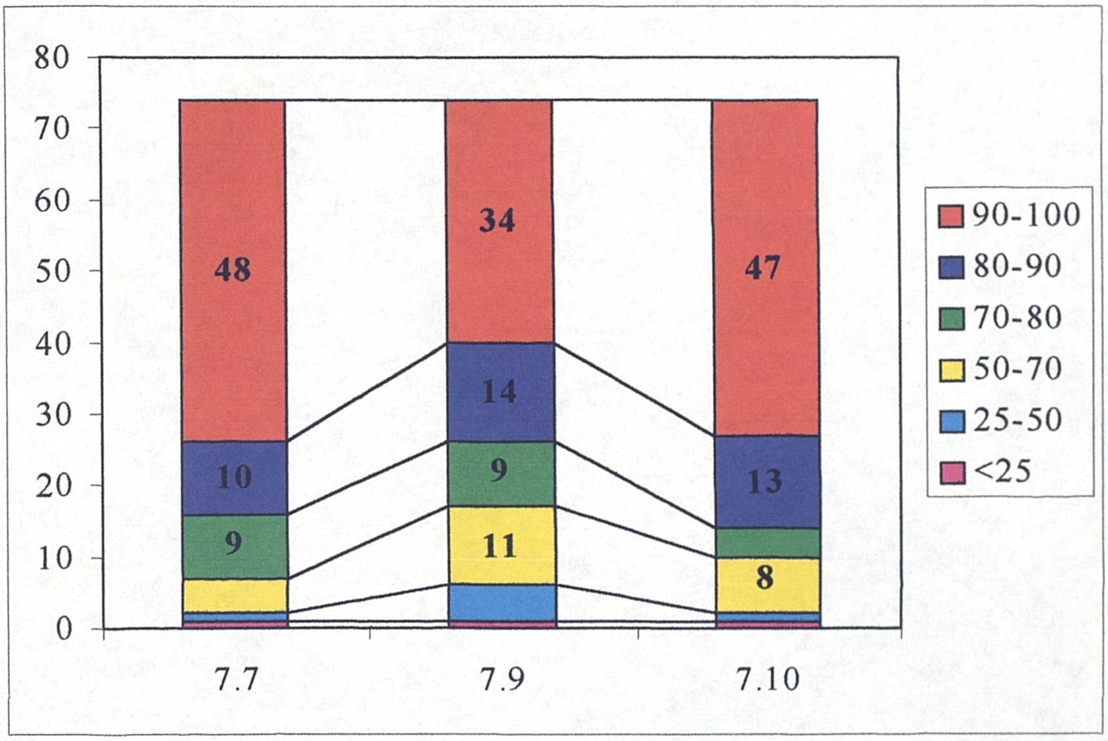


Figure 7.7: Distribution of Efficiency Scores for Models 7.7, 7.9 and 7.10

Looking first at the impact of aggregating the inpatient factors, there were just twenty-four efficient units in model 7.9, all of which were efficient in model 7.7. This is the lowest number of efficient units across all the models investigated in section 7.6. These were the only DMUs whose efficiency scores remained unchanged following the aggregation of the inpatient discharges factors. The average efficiency score decreased by 7% and there were fourteen DMUs that were no longer rated as efficient.

Every DMU that was classified as inefficient in model 7.7 showed a reduction in its efficiency score in model 7.9, following the aggregation of the inpatient factors. Overall, twenty DMUs showed a reduction in their efficiency score of greater than 10%, with six of these rated as efficient in model 7.7. The scale of the reduction ranged from 0.3% for DMU #64 to 36.5% for DMU #30. (DMU #30 was found to demonstrate a similarly large decrease in its efficiency score in several of the models discussed previously, particularly for model 7.2, which used 'number of days' rather than 'number of discharges' to measure inpatient activity.) Table 7.15 below lists the DMUs that show the greatest reduction in their efficiency score in this case. The level of inpatient activity for each of these DMUs was considerably less than the average value for all the DMUs (10977 inpatient discharges), except for DMU #19. In most cases, these DMUs also had a much higher number of 'consultant outpatient attendances' in comparison with their inpatient activity level.

DMU	Efficiency Score (%)		Reduction (%)	Inpatient Activity Level	Dominant Output Factor	
	Model 7.7	Model 7.9			Model 7.7	Model 7.9
#30	88.0	51.5	36.5	908	IDS (98.1%)	TID (100%)
#44	75.6	41.2	34.4	443	IDM (100%)	AEA (64.4%)
#69	78.7	52.8	25.9	1704	IDM (78.3%)	TID (73.8%)
#74	89.8	66.7	23.1	2869	IDS (70.5%)	COA (100%)
#42	100.0	78.3	21.7	383	COA (63.9%)	COA (89.1%)
#48	100.0	78.3	21.7	401	COA (53.4%)	COA (100%)
#25	71.7	50.4	21.3	1982	IDS (63.5%)	TID (92.2%)
#19	90.0	69.2	20.8	9312	IDS (90.9%)	TID (100%)
#43	100.0	82.3	17.7	719	IDS (100%)	TID (100%)
#53	74.0	57.7	16.3	222	IDG (45.7%)	TID (100%)

Key: See Table 7.7 for a full definition of the output codes

Inpatient Activity Level = Total Number of Inpatient Discharges

Table 7.15: Comparisons Between Models 7.7 and 7.9

The further influence of the 'total inpatient discharges' factor can be demonstrated by an analysis of the factor weights. Fifty-four of the DMUs have a non-zero weight for this factor and it is the dominant output factor (with the largest virtual percentage) for thirty-four of the DMUs. Table 7.16 below shows the distribution of the factor weights in model 7.9 based on one application of the DEA methodology. (The factor patterns recorded were not necessarily unique, particularly for the efficient units.)

Factor	Contribution to Efficiency Score		Dominant Factor	
	All DMUs	Efficient DMUs	All DMUs	Efficient DMUs
COA	47	15	25	8
AEA	23	10	9	5
DCA	32	14	6	4
TID	54	13	34	7

Key: See Table 7.7 for a Definition of the Factor Codes
Contribution to Efficiency Score: Number of DMUs with Non-Zero Factor Weight.
Dominant Factor: Number of DMUs with Largest Virtual Percentage.

Table 7.16: Distribution of the Factor Weights for Model 7.9

As was expected, aggregating the inpatient factors into 'total inpatient discharges' has reduced the number of efficient DMUs and the average efficiency score. However, it has become the dominant output factor for almost half of the DMUs in the sample and 73% (54 out of 74) have some contribution from this factor in their efficiency assessment. It was suggested above that aggregating the expenditure factors reduced the amount of useful information produced by the DEA model and was not necessarily appropriate in this analysis, particularly in the light of recent government proposals. In the case of the inpatient factors, the decision on aggregation was not as obvious and much more dependent on the focus of the research as a whole.

Clearly, if the focus were specifically on the identification of inefficiencies in inpatient services, aggregating the inpatient factors would not be appropriate. In the case of this research, the analysis was a general investigation into the efficiency of acute hospitals using DEA. Thus, the aggregation could be justified.

The final model to be investigated is model 7.10, which included two additional outputs ('PAM attendances' and ' day patient attendances') alongside the four outputs used in model 7.9 ('total inpatient discharges', 'day case attendances', 'A&E attendances' and 'consultant outpatient attendances'). The inputs used in both models were identical. As could be seen in figure 7.7, the number of DMUs with efficiency scores above 90% increased with the inclusion of the two extra factors. This is consistent with the DEA methodology, where increasing the number of factors provides more opportunities for each DMU to be classified as efficient.

The number of efficient units in model 7.10 was thirty-nine and the average efficiency score was 88.9%. Therefore, in comparison with model 7.9, there were fifteen additional DMUs classified as efficient and the average efficiency score had increased by 6.3% (82.6% -> 88.9%). Overall, there were thirty-five DMUs with no change to their efficiency score, twenty-four of which were efficient in both models. Of the eleven inefficient DMUs that showed no change in their rating, only two of these had no activity for the additional variables (the factor value was zero). The remaining eight DMUs did treat patients in these categories but this was not reflected in their efficiency assessment, that is, the factor weights for 'PAM outpatient attendances' and 'day patient attendances' were zero.

The range of increases for the fifteen newly efficient DMUs was interesting, in that DMU #64 increased its efficiency score by 0.3% (99.7% -> 100%) whereas for DMU #36, the increase was 57.1% (42.9% -> 100%). Table 7.17 below shows the DMUs with increases over 15%, eight of which became efficient following the inclusion of 'PAM outpatient attendances' and 'day patient attendances'. It can be seen that in each case, one of these two factors became the dominant output for this selection of DMUs.

DMU	Efficiency Score (%)		Increase	Dominant Output Factor	
	Model 7.9	Model 7.10	(%)	Model 7.9	Model 7.10
#36	42.9	100	57.1	TID (100%)	POA (71.2%)
#41	48.5	97.2	48.7	AEA (64.4%)	DPA (100%)
#70	54.8	100	45.2	TID (73.8%)	DPA (100%)
#74	66.7	100	33.3	COA (100%)	POA (75.8%)
#47	69.6	100	30.4	COA (89.1%)	POA (69.1%)
#26	71.2	100	28.8	COA (100%)	POA (51.9%)
#44	41.2	65.1	23.9	TID (92.2%)	POA (82.2%)
#42	78.3	100	21.7	TID (100%)	POA (61.4%)
#19	69.2	89.9	20.7	TID (100%)	DPA (63.8%)
#59	80.1	100	19.9	TID (100%)	POA (51.1%)
#21	84.7	100	15.3	COA (56.7%)	POA (33.7%)
#9	78.6	93.8	15.2	COA (49.6%)	DPA (65.9%)

Key: See Table 7.7 for a full definition of the output codes

Table 7.17: Comparisons Between Models 7.9 and 7.10

The impact of the additional factors on this set of DMUs was clear – they were able to show a significant improvement in their efficiency assessment. However, the additional factors also impacted on many of the other DMUs, in that they were given a non-zero factor weight on a number of occasions, as shown in table 7.17 below.

Factor	Contribution to Efficiency Score		Dominant Factor	
	All DMUs	Efficient DMUs	All DMUs	Efficient DMUs
COA	39	24	18	10
AEA	13	9	3	3
DCA	27	14	29	6
TID	46	20	6	10
DPA	23	12	5	1
POA	40	21	13	9

Key: See Table 7.7 for a Definition of the Factor Codes

Contribution to Efficiency Score: Number of DMUs with Non-Zero Factor Weight.

Dominant Factor: Number of DMUs with Largest Virtual Percentage.

Table 7.18: Distribution of the Factor Weights for Model 7.10

The table shows that over half the DMUs had a non-zero factor weight for 'PAM outpatient attendances'. This factor was also the dominant output for thirteen of the DMUs, nine of which were efficient. However, this result was not surprising as all but four of the DMUs treated patients in this category (see table 7.3). The impact of the 'day patient attendances' factor was slightly less influential, although only thirty-eight of the DMUs carried out this type of service (see table 7.3).

The two extra factors included in model 7.10 dramatically changed the efficiency assessment for a number of the DMUs (as was shown in table 7.17 above), although it was the 'PAM outpatient attendances' factor that appeared to be the most important. If it was important to consider the wide range of services provided by each acute hospital in Scotland, these variables could be included in the DEA model. However, they could be excluded if the efficiency assessment was concerned with the core activities of the hospitals.

7.6.5 Evaluation of the Models (Variable Set)

In the previous sections, ten different models have been considered to evaluate the efficiency of a sample of acute hospitals in Scotland. The use of 'inpatient days' (model 7.2) was rejected by looking at the results produced by the DEA model using this factor. However, the acceptability or otherwise of the other models could not be determined as easily, as it was found to be dependent on a number of other considerations, such as the focus of the research and the policy objectives of central government. Table 7.19 summarises the general results for these models, reproduced from table 7.7 above.

Model Description	Model Name	Average Efficiency Score	Standard Deviation of Efficiency Scores	No. of Efficient DMUs
Disaggregated Inputs and Outputs	7.1 (4, 6)	87.5%	17.44	31
	7.2 (3, 6)	86.6%	17.41	28
	7.4 (3, 6)	86.3%	18.22	29
	7.5 (3, 7)	88.9%	16.62	35
	7.6 (3,6)	87.7%	17.78	33
	7.7 (4, 7)	**89.6%**	**16.54**	**38**
Disaggregated Outputs and Aggregated Inputs	7.8 (3, 6)	84.3%	18.74	28
Aggregated Outputs and Disaggregated Inputs	7.9 (4, 4)	82.6%	19.65	24
	7.10 (4, 6)	**88.9%**	**17.39**	**39**

Key: The model codes are defined in Table 7.7

Table 7.19: Evaluation of the Models

The average efficiency score for the models that measured inpatient activity using 'number of discharges' varied between 82.6% (7.9) and 89.6% (7.7), a difference of 8.2%. The standard deviations appeared to be fairly consistent (16.54 -> 19.65).

The number of efficient units ranged from 24 (7.9) to 39 (7.10), which was 32% - 52% of the total number of DMUs. The lowest numbers of efficient units (24) corresponded to the model with the fewest factors (8 were used in model 7.9), as to be expected according to the DEA methodology. In comparison, 40% of the DMUs in the sample investigated by Byrnes and Valdmanis (1994) were found to be technically efficient using DEA-type methodologies. The investigation by Parkin and Hollingsworth (1996) resulted in up to 38 of the 75 DMUs being labelled as efficient, using various groupings of variables. This study used a sample of acute hospitals in Scotland, which was similar to those used in this analysis.

For some of the DMUs, the changes to the variable set had little or no impact on their efficiency score. For example, DMU #9 achieved the following four efficiency scores for models 7.1, 7.3, 7.4 and 7.8 respectively, with its ranking given in brackets: 86.2% (49), 86.2% (48), 86.1% (43) and 85.8% (47). Despite various changes to the variable mix, including the aggregation of the input factors in model 7.8, the efficiency rating of the DMU stayed fairly robust throughout. Under the same four models, DMUs #2, #3 and #5 were identified as 100% efficient throughout. Other similar examples were identified, including the three least efficient DMUs (#31, #70 and #72), which all remained as inefficient under each of the four models identified above.

The analysis also showed that there were a number of DMUs whose efficiency scores changed dramatically following minor changes to the variable set. For example, DMU #30 benefited from the inclusion of the 'day case attendances' (model 7.5), as did DMU #36 when 'PAM outpatient attendances' was included in model 7.10.

The development of the ten models using the DEA methodology has led to the following observations on the variables required in the evaluation of hospital efficiency:

- 'Number of discharges' was found to be the most appropriate measure of inpatient activity;
- 'Inpatient discharges, other' had a distorting effect on efficiency scores for several of the DMUs, particularly when 'number of days' was used in model 7.2. It should be excluded or aggregated into other inpatient factors to negate its influence;
- The effect of excluding 'average number of staffed beds' was negligible on the efficiency assessments for most of the DMUs. It is also only related to the provision of inpatient services, rather than to all the outputs in general;
- Various combinations of output factors were possible, with efficiency scores fairly robust to the exclusion/inclusion of the factors for which only a small number of the DMUs treated patients, such as 'day patient attendances'. The factors to be included should be related to the objectives of the research, which could be a wide-ranging efficiency assessment or an analysis of the core activities.
- The aggregation of the cost inputs appeared to be significant for the efficiency assessments and not entirely appropriate. This was not the case with the inpatient discharge factors used as the outputs, which could be aggregated when the study is looking for general observations on efficiency.

In the case of this research, which is an investigation into the applicability of the DEA technique for the general evaluation of acute hospitals in Scotland, the factors to be included in the next stage of the analysis are as follows (using the points made above as a guideline):

	FACTOR NAME	CODE
INPUTS	Capital Charge	CAP
	Total Direct Costs	TDC
	Total Allocated Costs	TAC
OUTPUTS	Consultant Outpatient Attendances	COA
	Accident and Emergency Attendances	AEA
	Day Case Attendances	DCA
	Total Inpatient Discharges	TID

Table 7.20: Factors to be Used in Stage Two of the Application Process

The formulation of the DEA model (defined as 7.11), with the same seventy-four DMUs as were included in the models discussed above, resulted in twenty-two DMUs classified as efficient and the average efficiency score was 81.4%. However, as was stated in section 7.2, the evaluation of the DEA application process was related not just to the composition of the variable set but also to the sample of DMUs.

7.6.6 Evaluation of the Models (Sample Definition)

In chapter six and also in the discussion on performance indicators in chapter four, the importance of comparing 'like with like' in a health care environment was firmly stressed. This was particularly related to the range of services offered by the hospitals being compared. In past applications of the DEA technique to the evaluation of hospital efficiency (see section 6.2), restricting the samples analysed to acute hospitals has tended to satisfy this requirement. In the framework of this analysis, which has incorporated the perspectives of health service personnel into the modelling process, the restriction to acute hospitals was no longer deemed to be a sufficient.

Seven variables have been selected to be used in the second stage of the analysis (see table 7.20 above), which involves the application of weight restrictions. However, as has been discussed previously, not all of the DMUs included in the original sample have treated patients in each of these output categories (see table 7.3). Prior to the second stage, therefore, hospitals that did not have activity for each of the four output categories were excluded from the data sample. This was to ensure that all the DMUs treated patients in each of the four categories chosen to reflect their output. It was intended that this would produce a homogeneous sample, such that each of the DMUs was operating under a similar strategy for their patient care and hospital organisation.

The changes to the sample resulted in some twenty-one DMUs being excluded initially. The excluded DMUs included each of those highlighted in table 7.7 for their unusual data patterns (apart from #41). In general terms, the DMUs removed from the data sample were those from six of the functional classes defined in appendix one and referred to in section 7.3, with these being classes 04, 05, 08, 09, 10, 13 and 14. A list of the excluded DMUs, with the justification for their exclusion, is given in appendix three.

Further analysis of the data showed that there were some DMUs with a very small number of patients treated in one of the categories. For example, the figure for 'A&E attendances' for DMU #21 was found to be just four, that is, only four patients were recorded in this category. These DMUs were also removed to ensure that the DMUs in the refined sample could all be seen to follow the same pattern in their patient activity – each of them treated patients in the four main categories.

Unfortunately, in a small number of cases, there was also some missing data for one of the input factors ('capital charge'). Thus, two additional DMUs were excluded for data reasons (#22 and #56) in order to successfully apply the weight restriction methodology used in this analysis.

Appendix three includes the efficiency score for each of the excluded DMUs in model 7.11, with the seven variables shown in table 7.20 and all seventy-four acute hospitals included. Overall, six of the DMUs excluded from stage two of the application process were identified as efficient in the model that included all of the acute hospitals in Scotland (#22, #45, #49, #50, #51 and #56).

Following the exclusion of these additional DMUs, the refined sample contained forty-seven DMUs, each of which had non-zero values for all of the inputs or outputs. Clearly, the choice of the sample was dependent on the variables included in the analysis, which were chosen to investigate the efficiency of hospitals on a very general basis. The methodology to be used for the application of weight restrictions was of secondary importance in this process, as only two DMUs were excluded for 'data' reasons.

In chapter eight, the second stage of the application process is presented, based on the development of a new model (defined as model 8.1). The results obtained from this model will be discussed in detail in section 8.2.

7.7 Summary of the First Stage of the Application Process

In this chapter, the first stage of the DEA application process has been presented. This comprised five key elements: the definition of the sample, the formulation of the DEA model type and orientation, selection of possible inputs and outputs, the results of the preliminary analysis and the revisions to the model.

The sample selected for analysis was the seventy-four acute hospitals in Scotland, with this determined to be the most appropriate level in the organisational structure of the NHS in Scotland for the application of the DEA technique. In relation to the formulation of the DEA technique, the CCR model defined in chapter five was selected, which was based on an assumption of constant returns to scale. An input orientation was chosen to reflect an ideology of 'minimising costs without reducing services'.

A large number of potential variables were identified and ten different models were defined, using different combinations of these factors. The important factors, and the most appropriate combinations of them, were obtained, based on the nature of the study and its objectives. This resulted in a revision to the data sample and the selection of seven variables, which is discussed in the second stage of the application in chapter eight.

7.8 References

1. **Ball, R. and Roberts, E.** (1998) The Relevance of DEA in the Public Sector - A Discussion Paper for Young OR 10, Young OR 10 Keynote Papers, Operational Research Society.

2. **Beasley, J.** (1990) Comparing University Departments, *Omega*, Vol. 18, pp. 171-183.

3. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

4. **Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.)** (1994) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

5. **Charnes, A., Cooper, W. W. and Rhodes, E.** (1978) Measuring the Efficiency of Decision- Making Units, *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429 - 444.

6. **Department of Health** (1997) The New NHS, Modern, Dependable, HMSO.

7. **Ehreth, J. L.** (1994) The Development and Evaluation of Hospital Performance Measures for Policy Analysis, *Medical Care*, Vol. 32, No. 6, pp. 568-587.

8. **Ersoy, K., Kavuncubasi, S., Ozcan, Y. A. and Harris, J. M.** (1997) Technical Efficiencies of Turkish Hospitals: DEA Approach, *Journal of Medical Systems,* Vol. 21, No. 2, pp. 67 – 74.

9. **Golany, B. and Roll, Y.** (1989) An Application Procedure for DEA, *Omega*, Vol. 17, No. 3, pp. 237-250.

10. **Grosskopf, S. and Valdmanis, V.** (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, Vol. 6, No. 2, pp. 89-107.

11. **Hollingsworth, B. and Parkin, D.** (1995) The Efficiency of Scottish Acute Hospitals: An Application of Data Envelopment Analysis, *IMA Journal of Mathematics Applied in Medicine and Biology*, Vol. 12, pp. 161-173.

12. **Information and Statistics Division** (1996) Scottish Health Service Costs 1995-96, ISD, NHS in Scotland.

13. **Nunamaker, T.** (1985) Using Data Envelopment Analysis to Measure the Efficiency of Non-profit Organizations: a Critical Evaluation, *Managerial and Business Economics*, Vol. 6, No. 1, pp. 50-58.

14. **Ozcan, Y. A., Luke, R. and Haksever, C.** (1992) Ownership and Organizational Performance: a Comparison of Technical Efficiency Across Hospital Types, *Medical Care*, Vol. 30, pp. 781-794.

15. **Parkin, D. and Hollingsworth, B.** (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991-1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

16. **Sherman, H. D.** (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927-938.

17. **Smith, P. and Mayston, D.** (1987) Measuring Efficiency in the Public Sector, *Omega*, Vol. 15, No. 3, pp. 181-189.

18. **Swift, L.** (1997) Mathematics and Statistics for Business, Management and Finance, MacMillan Press Ltd.

19. **Thanassoulis, E., Boussofiane, A. and Dyson, R. G.** (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588-608.

20. **Thanassoulis, E., Dyson, R. G. and Foster, M. J.** (1987) Relative Efficiency Assessments Using Data Envelopment Analysis: an Application to Data on Rates Departments, *Journal of Operational Research Society*, Vol. 38, No. 5, pp. 397-411.

21. **The Scottish Office** (1997) Designed to Care - Renewing the National Health Services in Scotland, HMSO.

22. **Wilkinson, R.** (1991) The Measurement of University Performance: an Application of Data Envelopment Analysis, PhD. Thesis.

7.9 Additional Sources

1. Interviews at the NHS Executive Performance Directorate in Leeds.

2. Interviews at Blackburn, Hyndburn and Ribble Valley NHS Trust (conducted in July, 1996).

CHAPTER EIGHT

INTRODUCING RESTRICTIONS TO THE FACTOR WEIGHTS IN THE DEA METHODOLOGY

8.1 Introduction

In the previous chapter, the first stage of the two-phase application process for the DEA technique was presented. The key elements of this were the definition of the sample, the formulation of the DEA model type and orientation, the selection of possible inputs and outputs, the results of the preliminary analysis and the revisions to the model. Following the selection of the variable set, it was determined that the sample of DMUs should be amended accordingly. This reflected the view expressed within the NHS that any sample of hospitals being compared should be 'homogeneous'. Each DMU in the refined sample, therefore, was an acute hospital that treated patients in each of the four categories of output: inpatient, day case, consultant outpatient and accident and emergency.

The discussion in this chapter is concerned with the second phase of the DEA application process, as defined in section 7.2. This comprised five elements, the first of which is the analysis of the results from the amended model. Therefore, in the following section, the key results from the model are presented, using a similar type of analysis to that used in section 7.7. This is extended in section 8.3, with a comprehensive analysis of the factor weights used in the efficiency assessment for each of the DMUs.

The third stage of the application process was the application of weight restrictions, using the methodology defined in chapter five. The introduction of weight restrictions is related to the potential benefits of this approach, enhancing the applicability of the DEA technique to 'real-world' applications.

Weight restrictions to the basic DEA model (model 8.1) were introduced using a number of alternative scenarios, each with a different pattern of restrictions. Each of the scenarios was developed to relate to a specific policy or managerial objective or observation within the health service. The main emphasis in the analysis was on the development of a practical approach to efficiency assessment, focused on the inclusion of weight restrictions.

The application of weight restrictions, using the different scenarios related to policy objectives, has led to the identification of efficiency strategies for each of the DMUs, which is also discussed in this chapter. The two-stage approach to the application of the DEA technique was completed by an evaluation of the models, in terms of their sensitivity and robustness. The importance of stability in the results obtained is specifically related to the health service environment.

8.2 Analysis of the Results from Model 8.1

Following the analysis in section 7.6, a revised model for the evaluation of efficiency was developed, which included seven variables, and the sample contained forty-seven DMUs. The revisions to the data sample and variable set were based on the objective of the analysis, which was defined to be a general evaluation of hospital efficiency in this case. The DMUs included in the revised sample, numbered 1 to 47 and prefixed by a '*', are listed in appendix four, with the data used in the modelling process. (A cross-referencing system with the DMUs used in the original analysis is also included in appendix four.)

Table 8.1 below summarises the results obtained from model 8.1 and also defines the characteristics of the model and the variables that were included in the analysis.

Model Name	**8.1**
Model Type	CCR
Orientation	Input Minimisation
Returns to Scale	Constant
Inputs	'capital charge' (CAP), 'total direct costs' (TDC), 'total allocated costs' (TAC)
Outputs	'total inpatient discharges' (TID), 'consultant outpatient attendances' (COA), 'day case attendances' (DCA), 'A&E attendances' (AEA).
Sample Size	47
Number of Efficient DMUs	20
Percentage of Efficient DMUs	42.6%
Mean Efficiency Score	87.98%
Standard Deviation	15.08
Minimum Efficiency Score	45.2%
Least Efficient DMU	*25
Reference Set Citations (>10)	*30 (18), *12 (13), *13 (12), *14 (10)

Table 8.1: Summary Description of Model 8.1

In chapter seven, reference was made to model 7.11, which included the same seven variables as model 8.1 but investigated the efficiency of all seventy-four acute hospitals in the original sample. There were twenty-two DMUs classified as efficient in model 7.11 and the average efficiency score was 81.4%. Six of these DMUs were excluded from the second stage of the analysis (see section 7.6.6). The remaining sixteen efficient DMUs were also classified as efficient in the corresponding model with fewer DMUs (model 8.1 detailed in table 8.1).

In addition to these sixteen DMUs, an extra four were classified as efficient in model 8.1, each of which was classified as inefficient using model 7.11 that included all seventy-four acute hospitals. This group of efficient units is shown in table 8.2 below.

DMU (Model 7.11)	Efficiency Score (Model 7.11)	DMU (Model 8.1)
#7	88.5%	*7
#10	93.9%	*10
#11	97.2%	*11
#64	99.7%	*43

Table 8.2: Newly Efficient DMUs

The exclusion of a number of efficient units altered the efficiency frontier, such that these four DMUs were placed upon it in the subsequent DEA efficiency assessment. In the case of the set of four DMUs highlighted above, it was the exclusion of DMU #22 from the sample of DMUs that was significant. This DMU was included in the reference set for each of them in model 7.11, although a number of other DMUs were also included. Changes to the efficiency frontier, and the efficiency assessment for the individual DMUs, illustrated that the efficiency scores calculated using the DEA technique were relative and not absolute.

In the analysis of the results presented in section 7.7, the efficiency scores for the DMUs were dissected according to a number of criteria, such as grouping by functional classification, health board or hospital size. A similar approach is used in the later sections of this chapter, following the inclusion of weight restrictions in the modelling process. Prior to this, the factor weights obtained in the unbounded model (model 8.1) are analysed.

8.3 Analysis of the Factor Weights for Model 8.1

In chapter five, it was observed by way of a introduction to the debate on restrictions to the factor weights that this development to the methodology has stemmed from attempts to transform DEA from a theoretical to a practical approach to efficiency assessment. As Allen *et al.* (1996) commented, weight restrictions have followed 'as a by-product of real-life applications.'

In other words, when DEA has been used in conjunction with a real data set and input from the managers who understand the minutiae of the organisations being evaluated, it has been found to be flawed. Results obtained from the unrestricted models in these types of analyses have often been unacceptable and unsupported. For example, Roll and Golany (1993) reported that in real world situations, 'virtually unconstrained factor weights are usually unacceptable.' As reported in chapter five, this was the situation observed by Wilkinson (personal communication) in applying DEA to data from a major oil company. Pedraja-Chaparro *et al.* (1997) also addressed the total flexibility of the factor weights and the introduction of weight restrictions.

In reference to the results from model 8.1, table 8.3 overleaf summarises the factor weights applied to each of the DMUs in the calculation of their efficiency score. As has been observed throughout the analysis, the factor weights used by each DMU were not necessarily unique (see section 5.4.3). This is further investigated in later sections, following the inclusion of weight restrictions, particularly in relation to the efficient DMUs. The information presented in table 8.3 has been obtained from one possible set of factor weights.

	Inputs			Outputs			
	CAP	TDC	TAC	COA	AEA	TID	DCA
Mean	9.88	89.71	35.6	46.86	28.86	59.646	21.97
St. Deviation	6.89	11.48	37.14	24.89	32.96	33.28	23.57
Minimum	1.1	57.7	2.4	12.7	0.3	0.2	0.7
Maximum	32.5	100	100	99.3	92.1	100	100
Count	24	43	17	25	18	39	31
Dominance	0	43	4	11	5	25	6

KEY: Factor Codes are defined in table 8.1.

Mean/Standard Deviation/Maximum/Minimum: calculated by excluding zero values for virtual weights (Mean does not sum to 100% as zeros have been excluded).

Count: Number of times input/output factor is used in efficiency calculations.

Dominance: Count of the number of times that factors contributes most in terms of virtual percentage to the efficiency score.

Table 8.3: Analysis of the Virtual Weights by Mean, Standard Deviation, Maximum, Minimum, Count and Dominance

Further analysis of the factor weights, in conjunction with the observations made in chapter five on the reasons for adopting weight restrictions in the DEA methodology, is included in the following discussion. This analysis is used as a justification for the two-phase approach to the application of the DEA technique, such that the methodology is incomplete without weight restrictions being included.

8.3.1 Widely Varying Factor Weights

One of the most widely stated advantages of the DEA technique (see discussion in chapter five) as a means of efficiency assessment was that it allowed for the free allocation of the factor weights (Dyson *et al.*, 1990). The efficiency score for each DMU was calculated such that the contribution of each variable in terms of the factor weight was determined to ensure the efficiency score for each DMU was maximised.

Therefore, virtually any pattern for the factor weights was potentially a possibility, as can be observed in table 8.4 below, showing the virtual weight percentage for each variable for a selection of DMUs using model 8.1, which had seven variables and included forty-seven DMUs overall.

DMU Code	Eff. Score	Virtual Factor Weights (%)						
		INPUTS			OUTPUTS			
		CAP	TDC	TAC	COA	AEA	TID	DCA
*1	95.2	7.8	92.2	0	31.3	0	68.7	0
*7	100	7.7	0	92.3	13.4	85.0	0	1.6
*12	100	0	100	0	12.7	0	43.1	44.3
*16	100	0	64.9	35.1	69.5	0	12.0	18.5
*26	90.1	0	100	0	33.8	0	37.7	28.5
*31	100	0	100	0	0	0	0	100
*36	92.0	0	0	100	0	0	100	0
*42	90.3	11.3	88.7	0	20.9	1.9	73.0	4.2

Key: CAP – 'capital charge' TDC – 'total direct costs' TAC – 'total allocated costs'
COA – 'consultant outpatient attendances' AEA – 'A&E attendances',
TID – 'total inpatient discharges' DCA – 'day case attendances'

Table 8.4: Virtual Factor Weights for Selected DMUs

As the table 8.4 shows, the range of values used by this small number of DMUs is extremely wide. This was further illustrated in table 8.3 above, which showed the maximum and minimum contributions from each of the factor weights for all the DMUs in the sample. This information is extended in table 8.5 overleaf, showing that for all the outputs, the DMUs selected a wide range of weights. For the inputs, 'capital charge' (CAP) had a maximum virtual weight of 32.5%, whereas 'total direct costs' (TDC) had a minimum virtual weight of 57.7%. Four of the seven factors (TDC, TAC, TID and DCA) had a 100% virtual weight for at least one of the DMUs.

	Virtual Factor Weights (%)						
	INPUTS			OUTPUTS			
	CAP	TDC	TAC	COA	AEA	TID	DCA
Min	1.1	57.7	2.4	12.7	0.3	0.2	0.7
Max	32.5	100	100	99.3	92.1	100	100
Range	31.4	42.3	97.6	86.6	91.8	99.8	99.3

KEY: Max/Min: calculated by excluding zero values for virtual weights.
See table 8.4 for a definition of the factor codes.

Table 8.5: Analysis of the Virtual Weights

The types of variation shown above may have been considered advantageous in a theoretical sense, in that DEA can be applied without having to know in advance what weights should be used. However, in relation to the specific investigation into the efficiency of a sample of acute hospitals, such widely varying patterns were found to be unacceptable for several reasons.

In the unrestricted DEA model, the weights were calculated in such a way that it appeared as if each DMU had its efficiency evaluated under its own distinct set of rules, as each DMU determined its own set of weights. This was clearly advantageous at the level of the individual DMUs, as each DMU chose the set of weights most beneficial to its efficiency rating. However, if DEA is being used as a comparative technique, the varying patterns of the factor weights become more problematic. For example, with reference to table 8.4 above, both DMUs *7 and *12 achieved an efficiency rating of 100% with very different patterns of factor weights. The efficiency of DMU *7 was based upon a relationship between 'total allocated costs' as the significant input and 'A&E attendances' as the dominant output factor.

Alternatively, for DMU *12, 'total direct costs', 'total inpatient discharges' and 'consultant outpatient attendances' were considered to be the major contributors to the efficiency measure. For these two DMUs, efficiency has been defined in two different ways.

This varying pattern of factor weights was also repeated for the inefficient DMUs in relation to their peer group or reference set (see section 5.4.6). Each DMU in the reference set 'is efficient with the inefficient unit's weights' (Dyson *et al.*, 1990). However, the DMUs in the peer group for each inefficient DMU were often found to have widely varied factor weight patterns, based on their virtual weight percentages. In such cases, it does not appear as though the efficiency of each of the DMUs has been calculated using the same guidelines. For example, DMU 4 had an efficiency score of 85.0% and there were four DMUs in its reference set (*2, *3, *14 and *16). The virtual weight percentages are shown in table 8.6 below.

DMU	Virtual Factor Weights (%)						
	INPUTS			OUTPUTS			
Code	CAP	TDC	TAC	COA	AEA	TID	DCA
*4	16.5	83.5	0	81.5	4.4	14.1	0
*2	1.5	79.0	19.5	0	0	100	0
*3	1.1	0	98.9	95.9	4.1	0	0
*14	1.9	57.7	40.4	32.0	30.7	37.3	0
*16	0	64.9	35.1	69.5	0	12.0	18.5

Key: See table 8.4 for a definition of the factor codes.

Table 8.6: Factor Weight Percentages for an Inefficient DMU and its Peer Group

In relation to performance indicators, the means of assessment with which most health care evaluators are familiar, the efficiency index is calculated as: a ratio of weighted sum of costs to the weighted sum of activity. This is close to the DEA approach to efficiency measurement. However, with the efficiency index, the weights used in the calculation are pre-determined, with the same set of weights used for each hospital being assessed. Therefore, the DEA approach, with the internally generated set of weights, distinct for each hospital, appears to be strange to the non-academic. This can be related to the discussion to follow on the difficulties involved in the interpretation of the factor weights.

Introducing a pattern of weight restrictions that controls this variation to a certain degree, whilst still allowing some freedom in the selection process, would reduce the problems highlighted above. Hospitals would then be compared according to their efficiency in relation to externally defined patterns of service delivery.

8.3.2 Non-Contribution of Factors

The information contained in table 8.4 in the previous section illustrated that amongst the vastly different allocations of factor weights for each DMU, there were a number of zeros in every case illustrated. For example, DMUs *1 and *12 both had three factors with a zero value for the virtual weight. Also, analysis of the results obtained for all DMUs showed this pattern to be repeated across the whole data set, as illustrated in the table overleaf. This shows the number of times each variable had a zero value for the factor weight and also the number of times it was the dominant factor, that is, had the highest virtual percentage.

	INPUTS			OUTPUTS			
	CAP	TDC	TAC	COA	AEA	TID	DCA
Count	23	4	30	22	29	8	16
Dominant	0	43	4	11	5	25	6

KEY: See table 8.4 for a definition of the factor codes.

Count: Number of times no contribution from this factor (zero virtual weight)

Dom: Count of the number of times that factors contributes most to the efficiency score calculation.

Table 8.7: Contribution of Factors

Additionally, it was observed that for each of the DMUs, there tended to be several variables with a zero factor weight, notably *25, *31 and *36, whose efficiency ratings were based on contributions from just two variables. The table below confirms these observations, showing how many variables contributed to the efficiency score of each DMU by counting the non-zero contributions. It can be seen that none of the DMUs included all seven of the variables in their efficiency assessment.

No. of Factors	2	3	4	5	6	7
Count	3	8	17	15	4	0

Table 8.8: Distribution of Factors Used to Calculate Efficiency Scores.

These observations were significant in relation to the data sample utilised here for several key reasons, primarily that relating to the selection of the variables included in the model. A great deal of attention was given to the selection of the variables to be included in the model 8.1, as shown by the detailed analysis in chapter seven. The variables, and the sample of DMUs, were selected with a specific objective in mind (see sections 7.6.5 and 7.6.6). It would seem inappropriate to then ignore the influence of several of these variables in the subsequent calculation of the efficiency scores.

The question most likely to be asked by observers of this approach to the evaluation process would be, if the variables were important enough to be selected in the first place, how can their exclusion from the efficiency calculation be justified?

Looking at a particular DMU emphasises this contradiction more clearly. DMU *7 was rated as 100% efficient using model 8.1, with contributions in terms of the factor weights from five out of the seven variables. However, the two variables from which this DMU had no significant contribution to its efficiency score were 'total direct costs' (TDC) and 'total inpatient discharges' (TID). These were the variables with the highest scores when taking an average of the virtual weight percentages and contributed to the efficiency scores of more DMUs than any other factors. In other words, DMU *7 is rated as 100% efficient with no contribution at all from the two variables considered most important by the majority of other DMUs and also the most intuitively important factors. Inpatient treatment consumed the largest proportion of the resources for most hospitals. 'Total direct costs' measured the expenditure on direct patient care and was, generally, the greater of the two expenditure factors (TDC and TAC).

The non-contribution of certain variables to the calculation of efficiency scores was significant, as it potentially allowed the DMUs to exclude areas of weakness or low levels of efficiency, as was noted with the previous discussion on the variations observed among the patterns of the factor weights. The methodology of the unrestricted DEA model was that the efficiency score calculated was that which showed each DMU in the best possible light.

Contrary to this, however, was the viewpoint of some of the health care managers interviewed (as discussed in chapter six). They considered it to be of much greater importance to look at the performance of each hospital in its true light. A high number of non-contributing factors made this unlikely in the example of the data sample used in model 8.1.

To counteract the problem of non-contributing factors, restrictions could be introduced such that each factor must be included in the calculation of the efficiency score. This has been achieved by specifying a minimum value for the virtual weight for each of the seven factors, not necessarily with each of them having the same minimum value (as discussed in section 5.7.1). This ensured that all variables were taken into account in the efficiency assessment, whilst still allowing some to be more significant than others, especially if maximum values for the factor weights were not specified.

8.3.3 Interpretation of the Factor Weights

For those newly acquainted with the DEA technique, the nature and development of the factor weights was a particularly difficult concept to grasp and relate to the actual performance of each hospital. It was difficult to visualise what each weight signified, simply by looking at their numerical values. This was improved by consideration of the virtual inputs and outputs, although the difficulties have not been completely nullified. Introducing restrictions to the factor weights in the form of inequalities would potentially make the concept more comprehensible.

This approach, as used by Beasley (1990) and Thanassoulis *et al.* (1995), involved stating the relative importance of the various factors in terms of their impact on the efficiency of the DMU.

In the case of the hospital investigation, this could be achieved by specifying that 'inpatient' activity was more important in relation to efficiency than 'outpatient services'. The virtual weight for the TID factor would be forced to be greater than that for the COA variable for every DMU.

Translating this back to the interpretation of factor weights in general, it can be understood more clearly that the factor with the largest virtual weight for both inputs and outputs was the biggest influence on the efficiency of that DMU. For the majority of the DMUs in the sample, therefore, the dominant input factor was 'total direct costs', as this variable has the largest value for its virtual weight for forty-three of the DMUs. In terms of outputs, the dominant factor was 'total inpatient discharges', with this having the largest virtual weight for twenty-five of the DMUs.

8.3.4 Inappropriate Relationships between Inputs and Outputs

Taking in combination some of the points already discussed, the situation might have arisen in the unrestricted DEA model whereby a particular DMU has its efficiency score calculated using a small minority of the possible number of variables. This might have resulted in a meaningless or inconsequential relationship being defined, especially if there were particular inputs and outputs that were not directly related, without the influence of other factors.

This phenomenon has been observed in the DEA literature, most notably in the application by Thanassoulis *et al.* (1987), investigating the efficiency of local authorities in the UK, as referenced in chapter five. In this investigation, a local authority was classified as efficient because it had issued the largest number of summons for the non-payment of rates. The application of weight restrictions to the same data set was discussed in Dyson and Thanassoulis (1988).

Using model 8.1, three of the DMUs had their efficiency scores calculated on the basis of contributions from just two factors, as can be observed by reference to table 8.8. Looking at DMUs *25 and *36 in particular, both of these attached all the weighting in their efficiency calculation to the same two factors, 'total allocated costs' as the input and 'total inpatient discharges' as the output. In other words, the efficiency of these two DMUs was defined as: the ratio of TAC:TID. This suggested that hospital efficiency was a measure of the level of inpatient activity in terms of the amount spent on administration, laundry services, maintenance costs and utilities, using the definition of 'total allocated costs' found in chapter seven.

This would not be considered to be an appropriate definition of efficiency, as many of the other influential factors have been excluded, over-simplifying the calculation into a simple ratio. In this analysis, it also highlighted that not all factors should be considered to be equally important in the evaluation of efficiency. As was observed in the previous section, the exclusion of several factors from the efficiency assessment (by giving them a zero weight) not only over-simplifies the efficiency calculations but can also potentially make them meaningless.

The findings from the evaluation of the factor weights for model 8.1 correspond to the following assessment on total weight flexibility:

> "Factors of secondary importance may dominate a DMU's efficiency assessment. If the inputs and outputs included in the analysis are not equally important, it is not sensible to claim that a DMU is relatively efficient if the weights assigned to the important inputs and outputs are zero. The total flexibility of the unbounded model may lead to an unfounded emphasis on efficient use of relatively unimportant factors or the production of relatively unimportant outputs, concealing inefficiencies in the most important activities undertaken by the unit."
>
> PEDRAJA-CHAPARRO *ET AL.* (1997)

8.4 Further Justification for the Introduction of Weight Restrictions

In the previous section, the factor weights obtained for each of the DMUs using model 8.1 have been investigated. Four areas of difficulty were identified: widely varied factor weights, the non-contribution of factors, interpretation and inappropriate relationships between the factors. Each of these areas was seen as a reason for the inclusion of weight restrictions in the DEA methodology. However, in addition to the treatment of the factor weights, there were a number of other areas within the DEA methodology that have led to the adoption of weight restrictions.

8.4.1 The Number of Units Classified as Efficient

In the application of model 8.1, twenty of the forty-seven DMUs in the sample were classified as 100% efficient, using the basic DEA model with no restrictions to the factor weights.

The average efficiency score was 87.98%, with the lowest ranked DMU receiving an efficiency rating of 45.2%. Additionally, in the unrestricted model (model 8.1), 60% of the DMUs were given an efficiency score of at least 90%. In a sense, these results have suggested that the hospitals were generally relatively efficient, with only a very small number of them demonstrating extensive inefficiencies (five of the DMUs had an efficiency score of less than 70%).

Many of the applications of DEA found in the literature have produced similar results. As was observed in chapter seven, the number of efficient units identified in the initial analysis was fairly consistent with two examples taken from the literature, Byrnes and Valdmanis (1994) and Parkin and Hollingsworth (1996).

Boussofione *et al.* (1991) observed that the number of efficient units in an unrestricted DEA model was likely to be approximately equal to the product of the number of inputs and the number of outputs. Therefore, especially as the data samples to be used in this analysis were relatively small in comparison to some of those discussed in the review of the literature in chapter six, a large proportion of the DMUs have been classified as efficient.

However, improving the efficiency with which health services are provided has been a priority in recent reforms to the National Health Service (Department of Health, 1989 and 1997), as was discussed in chapters two and three. In the assessment of the relative efficiency of a sample of hospitals, it would be highly unlikely that each of them would be actually be as efficient as the 'best practice' DMUs.

As was discussed in chapter five, the impact of introducing weight restrictions on actual efficiency scores was to increase discrimination amongst the DMUs. Fewer DMUs were classified as efficient and the average efficiency score was reduced (see section 5.7.2). The reduced flexibility made it more difficult for each DMU to reach the efficiency frontier. Overall, efficiency scores tended to decrease, as the gap between 'most efficient' and 'least efficient' widened.

In the analysis of the gynaecology inpatient services, introducing weight restrictions meant that the hospital departments had to 'work harder' to achieve a 100% efficiency score, encouraging improved performance overall. It also resulted in the identification of a central core of efficient DMUs, able to maintain their efficiency under a strong degree of restriction to the factor weights, which could be used as 'role models' for the inefficient DMUs.

8.4.2 Discrimination Against the All-Rounder

As has been observed previously, it was possible for a DMU to be classified as efficient in the unrestricted DEA model by operating at the extremes of the efficiency frontier. A DMU with an unusually high ratio between two potentially unrelated variables (one input and one output) was likely to receive a high efficiency score simply because there are no other similarly operating DMUs with which to accurately compare it. In terms of the efficiency frontier, such a DMU was operating at the extremes. The DMUs operating away from the extremes had a greater number of competitors for a rating as 100% efficient – it was more difficult to achieve.

Introducing weight restrictions would make it more difficult for the DMUs at the extremes to be rated as 100% efficient unless they really were efficient – all the DMUs are assessed under a similar operating strategy.

8.4.3 Expert Opinion

In the unrestricted DEA model, there was no scope for including some pre-determined information on the relationship between a given set of factors and their contribution to the efficiency of the DMU. This restricted experts in the efficiency and management structures of each organisation from incorporating their knowledge into the model building process, outwith the selection of the set of variables. However, the inclusion of some set of weight restrictions allowed for their knowledge to be built into the DEA model. For example, it could be apparent that some of the variables, whilst significant influences on efficiency, are not as influential as other factors in the data set. Weight restrictions could be added to ensure that the virtual weights conform to this required pattern. In the analysis of hospital efficiency, this involved the identification of primary and secondary outputs, with a different level of importance attached to these two groups.

8.5 Specification of Weight Restrictions

In the previous sections, the various reasons for introducing some level of restriction on the factor weights were debated in relation to the results obtained from model 8.1. Additionally, some suggestions were given as to how best to introduce factor weights and what form these restrictions should take, relating to the theoretical investigation in chapter five.

As was noted there, restrictions could be introduced in a variety of forms, most usually dependent on the modelling approach and the computer-based package being utilised for the analysis. In this case, weight restrictions have been introduced by attaching constraints to the virtual inputs and outputs, following the approach presented in Ball and Roberts (1998) and used to introduce the theory of weight restrictions in chapter five.

For the seven variables used in the analysis, the restrictions were specified in relation to the virtual inputs and outputs. In section 5.7.1, the five different approaches for including weight restrictions of this type were discussed. In the investigation of model 8.1, three of these possibilities have been utilised, summarised below:

1. Minimum Contribution: the virtual percentage for the factor was greater than a specified percentage - the restricted factor was required to contribute something to the calculation of the efficiency score.
2. Maximum Contribution: factors are prevented from over-dominating the efficiency score calculations, such that no factor was able to have a virtual weight percentage of 100%.
3. Range of Values: combining minimum and maximum constraints results in a range of values being specified, within which the virtual weight for that factor must lie.

These approaches can be used to develop any required set of weight restrictions, using careful application, particularly in relation to zero values in the data. The following section continues the process of applying weight restrictions, focusing on the key issue of determining how the restrictions are to be selected, using the idea of ‘scenarios’.

Each of the scenarios reflected a particular policy objective or approach to efficiency assessment. Some care was needed, as the imposition of weight restrictions can impact on the feasibility of the DEA model – the dilemma is in finding the right balance. If the restrictions were too loose then their adoption would be pointless – an unbounded model would have produced equivalent results. On the other hand, very tight restrictions leave no room for flexibility and the resulting model could be infeasible, transforming the DEA model into a weighted ratio.

8.6 Selection of Weight Restrictions

In the previous section, the different approaches to weight restriction have been discussed in relation to the analysis presented in the second phase of the DEA application process. This section is concerned with how the actual weight restrictions were selected, referring to the sample of forty-seven DMUs being investigated in this chapter. It was determined in chapter seven that the objective of this analysis was to investigate efficiency at a general level, looking at the hospital as a whole. The weight restrictions that were included have also been selected with this objective in mind. As was discussed in chapter five, weight restrictions need to be developed in relation to the objectives of the research and within a framework of consultation (Cook *et al.*, 1994 and Ball and Roberts, 1998).

Therefore, several different scenarios were developed, with each of these having a slightly different pattern of restrictions, to determine the impact of weight restrictions on the overall sample and to assess alternative operating strategies on the efficiency of the individual DMUs.

Each scenario could be related to a specific aspect of hospital efficiency measurement or requirement of the analysis. In this case, five alternative scenarios were developed, illustrating several different possibilities in terms of restrictions to the factor weights. The restriction requirements for the inputs and outputs were considered separately before being combined into the five scenarios.

8.6.1 Restrictions Attached to the Outputs

The different weighting scenarios examine three main options to represent different approaches to health care provision in hospitals, in terms of the priorities of patient care and its relation to efficiency assessment. These were as follows:

8.6.1.1 Equal Contributions from All Outputs

Each output should contribute equally to the efficiency score; that is, a DMU could not be classified as efficient if there was a weakness in any of its outputs. In relation to the specification of the restrictions, a narrow band would be defined for the virtual weight percentage, within which a small amount of variation would be acceptable.

A DMU could not neglect any area of service provision in the pursuit of efficiency. However, this would not allow a hospital to deliberately over-resource a specific output, even if a justification for this was proposed, and still be classified as efficient.

In terms of policy-making, restrictions of this type would be appropriate if it was assumed that each of the outputs were equally important in the definition of acute hospital efficiency and should be weighted equally.

8.6.1.2 Minimum Contributions from All Outputs

Each output should contribute something to the efficiency score; that is, none of the outputs could be completely ignored in the assessment of efficiency. In relation to the specification of the weight restrictions, a minimum virtual weight percentage would be defined for each of the outputs. None of the outputs could have virtual weight percentage of 100%.

This was further extended by also attaching maximum percentages for the outputs. In the cases where only minimum restrictions would apply, one or two of the outputs would still be allowed to dominate, with a virtual weight percentage in excess of 50%. In the cases where maximum percentages were also specified, it would be impossible for a single factor to have a virtual weight percentage of greater than 50%. However, the choice of the dominant factor would still be determined by the DMU itself, rather than externally fixed.

A minor weakness in terms of efficiency in one of the outputs would not prevent a DMU being classified as efficient if it had other areas of strength. In terms of policy objectives, this approach would allow the hospital management to decide where its resources should be concentrated, whilst still pursuing all-round efficiency. It would also not specifically penalise a hospital, which has chosen to concentrate its service provision in one or two areas. This approach recognised that not all DMUs operate under the same objectives or in the same environment. The specification of minimum and maximum values for the virtual weights was used in conjunction with the final approach, which defined some of the outputs as more significant than others.

8.6.1.3 Externally Defining the Important Outputs and their Contribution

In this final case, it was recognised that it would not be appropriate to assume that all of the outputs included in this analysis were equally important in the evaluation of hospital efficiency (as was referred to in section 8.3 above). This decision was based upon an understanding of the hospital environment, as well as an evaluation of the results from the DEA models applied in the preceding analysis.

It was suggested in the discussion in chapter seven that the four outputs included in model 8.1 ('total inpatient discharges', 'day case attendances', 'consultant outpatient attendances' and 'A&E attendances') were the central features of hospital activity, which had to be included in a general evaluation of efficiency. At that stage, two other potential measures of output ('PAM outpatient attendances' and 'day patient attendances') were excluded.

In model 8.1, the importance attached to each of the individual outputs was determined by each DMU, as no weight restrictions were included. Analysis of the factor weights for each of the output factors (see table 8.3) suggested that the most important factors (in terms of the number of times they have contributed to the efficiency assessment for each DMU) were 'total inpatient discharges' and 'day case attendances'. 'Total inpatient discharges' was included by thirty-nine of the DMUs, with thirty-one of the DMUs having some contribution from 'day case attendances'. 'Consultant outpatient attendances' was the dominant output on a greater number of occasions than 'day case attendances' but it had a zero weight for twenty-two of the DMUs.

Within the DEA modelling process, therefore, it seemed that two of the factors have been recognised as the most important. Focusing on these two factors, it was also observed that inpatient and day case treatment were the most expensive service elements to provide, requiring a greater amount of expenditure in terms of 'cost per case'. Table 8.9 below shows the average cost per case for several groups of hospitals for each of the four output categories used here, where clear differences can be seen. The main functional class averages representing the forty-seven hospitals included in model 8.1 have been used.

Functional Class	Output Categories – Average Cost per Case (£)			
	Inpatients	Day Cases	Outpatients	A&E
1	1450	396	56	41
2	1344	263	46	39
5	1786	476	41	41
11	1366	305	50	55
12	1360	210	34	44

Table 8.9: Differences in Average Cost per Case by Functional Class

Health care policy over the last five years has promoted an increase in the amount of general surgeries to be carried out using the 'day case' approach to treatment (as was discussed in chapter three). This suggested that 'day case attendances' was an important factor and should be included in the efficiency assessments for the DMUs. Therefore, in this analysis, 'total inpatient discharges' and 'day case attendances' were defined as the primary factors, as they represent the 'cornerstone' of the treatments offered by acute hospitals. The two remaining outputs ('consultant outpatient attendances' and 'A&E attendances') were defined as the secondary factors.

The weight specifications for the two types of outputs were then specified independently (using the approach discussed in the previous section) and minimum and maximum virtual weight percentages were defined. This approach to weight restriction ensured that, to be classified as efficient by the DEA, a DMU would have to be providing its primary services efficiently. The DMUs that had not included the primary outputs in their efficiency assessment in the unbounded model were less likely to be defined as efficient in the restricted models.

It was also possible to only specify weight restrictions for the primary factors, such that each DMU could chose to ignore the secondary factors in the calculation of its efficiency score if this was beneficial. This would help the DMUs that were efficient in the provision of primary services, but demonstrated inefficiencies in their secondary outputs.

8.6.2 Restrictions Attached to the Inputs

As was the case with the outputs, the strategies for specifying the input restrictions were developed in relation to the situation being investigated and following an analysis of the factor weights in the unbounded model, which identified 'total direct costs' as the dominant input factor. This reflected the fact that this was intuitively the most significant input, as it measures the physical cost of treating patients. Thus, as was the case with the outputs, this was defined as the primary input factor, with 'total allocated costs' and 'capital charge' classified as the secondary inputs. In one of the scenarios, the restrictions for the two types of factors were specified independently to reflect this difference in their influence.

In this case, it would not be possible for a DMU to be efficient by attaching a 100% weighting to 'total allocated costs', resulting in efficiency being defined by an inappropriate ratio (as was discussed in section 8.3.4).

In other scenarios, the same restrictions were applied to each of the inputs, either as a minimum virtual percentage or as a range, within which the virtual weight could be varied. Specifying minimum restrictions for each of the inputs ensured that, for example, it was not possible for a DMU to be classified as efficient by ignoring 'total allocated costs' completely. This factor, as was discussed in chapter seven, was a measure of 'bureaucracy'. Reducing the amount spent by hospitals in this area has been an important objective of recent health care policy (Department of Health, 1997).

It was also possible to remove the restrictions attached to the inputs, whilst still enforcing output restrictions. The models used in this analysis have all been specified with an input minimisation orientation. Removing the restrictions from the input factors allowed the DMUs to identify their own targets for improved efficiency, rather than forcing each DMU to follow a given pattern.

8.6.3 Definition of the Weight Restriction Scenarios

In the previous sections, the rationale behind the different weight restriction patterns has been discussed. Thus, five alternative scenarios were developed, each of which could be related to a specific policy objective in terms of the restrictions attached to the inputs and the outputs.

It would have been possible to develop many other scenarios, using slightly different restrictions, as was the case with the possible variable combinations used in the analysis in chapter seven. However, the scenarios included in the following discussion represent the main options in relation to this analysis, and they were defined as follows:

8.6.3.1 Scenario One (All Round Efficiency)

The output weights were constrained such that the virtual weight for each of the four factors must lie between 20% and 35%, examining all round efficiency. Such a pattern of weight restrictions suggested that each of the factors was equally important in the evaluation of efficiency. Under this model, the pattern of the factor weights should have been fairly standard across all the DMUs (as discussed in section 8.3.1). The notion of all-round efficiency was extended to the inputs, as each was constrained to contribute to the efficiency assessment. However, the importance of the 'total direct costs' factor was recognised by specifying the range for its virtual percentage as 50% - 100%. The other two input factors were constrained such that their virtual weights were between 5% and 25%.

8.6.3.2 Scenario Two

The minimum virtual weight for the secondary factors was 5%, with a 25% minimum for the primary factors. In this case, some contribution was required from all factors, but the restrictions were not as tight as with scenario one. It was still possible for one factor to be dominant, with a virtual weight of up to 65%.

This related to the policy objective discussed above, such that, a DMU could not be classified as efficient without a contribution from each of the output factors in its efficiency assessment. However, it was still possible to focus in one or two areas (including the secondary factors)..

A minimum virtual weight of 10% was set for all the three input factors, such that each DMU had considerable freedom in the allocation of its input weights, as was the case with the outputs. It would be possible for one of the inputs to continue to dominate, with a virtual percentage of up to 80%. However, none of the inputs could be excluded, so it was impossible for each DMU to completely ignore any areas of weakness.

8.6.3.3 Scenario Three

In scenario three, the range of values for the virtual weights for the secondary output factors was 5% to 50%, increasing to 25% to 50% for the primary factors. None of the factors could dominate the efficiency calculation, due to the inclusion of additional maximum values for the weights. This type of weight restriction pattern allowed for judgements on the relative importance of the factors to be incorporated into the model. It also allowed for DMUs strong in the areas of the secondary factors to gain an efficiency score relative to their strengths. A secondary factor was still able to contribute up to 45% of the virtual weight if the minimum end of the range for the other factors was used. A similar additional maximum restriction was imposed on the inputs to prevent one of them being completely dominant, such that the virtual weight percentage for each them was constrained to lie between 10% and 50%.

8.6.3.4 Scenario Four

This used the same pattern of output restrictions as were defined for scenario three, such that minimum and maximum constraints were specified for each of the factors, although the primary and secondary factors had a different set of restrictions applied. The input restrictions were completely removed. The contributions from each of the outputs were restricted but the choice of input weights was left to each DMU.

8.6.3.5 Scenario Five

In scenario five, the constraints attached to the secondary outputs and all the inputs were removed, with restrictions attached only to the primary output factors, using the same ranges as with scenarios three and four (a range of 25% – 50%). The important outputs had to be included in the evaluation of efficiency, with the position of the secondary factors left to each DMU. Significant weights could still be attached to the secondary outputs if this was beneficial.

Table 8.10 below summarises the restrictions required for each of the scenarios, with the results from their application discussed in the following section.

	Weighting on Virtual Weight for Each Factor (Min – Max %)						
Scenario	**CAP**	**TDC**	**TAC**	**COA**	**AEA**	**TID**	**DCA**
One	5 – 25	50 - 100	5 – 25	20 – 35	20 – 35	20 – 35	20 - 35
Two	10 – 100	10 - 100	10 – 100	5 – 100	5 – 100	25 – 100	25 - 100
Three	10 – 50	10 - 50	10 – 50	5 – 50	5 – 50	25 – 50	25 - 50
Four	/	/	/	5 – 50	5 – 50	25 – 50	25 - 50
Five	/	/	/	/	/	25 – 50	25 - 50

Key: See table 8.4 for a definition of the factor codes.

Table 8.10: Weight Restriction Scenarios

8.7 Application of Weight Restrictions

Following the application of the scenarios defined in the previous section, their impact on the nature of the efficiency calculations and their effect on the efficiency assessment can be investigated. As with the analysis in previous chapters, this is presented in a number of different ways. Firstly, the impact of weight restrictions is discussed in general terms, looking at the number of efficient units for each of the scenarios, average efficiency scores, and so on. Attention is also be given to the impact the alternative scenarios have on the different types of hospital, according to functional class and size and also the effect of the different weight restrictions selected in general terms.

The second part of the analysis is focused on a DMU specific evaluation, looking closely at individual DMUs under each of the alternative scenarios. Addressing the efficient and inefficient DMUs separately, the analysis of the efficient DMUs is concentrated on identifying the characteristics of efficiency. In relation to the inefficient units, the analysis is focused on the impact of weight restrictions to the reference groups and target setting. This stage of the analysis is related to the identification of efficiency strategies for the inefficient DMUs, which was one of the elements of the application process discussed in chapter seven.

Finally, there is an evaluation of the various factors and factor weights, looking at the patterns of the factor weights in the unbounded model (model 8.1) in relation to changes in efficiency scores in the scenarios.

8.7.1 General Analysis of the Alternative Weights Scenarios

In section 8.3, some of the reasons for including factor weights were given in relation to the results obtained from model 8.1. Imposing such restrictions by way of the five scenarios shown in table 8.10, it can be seen that some of these areas of difficulty have been overcome. Table 8.11 gives a summary of the results obtained from the scenario models, illustrating their impact on a number of key characteristics.

	Model or Weighting Scenario					
	Model 8.1	One	Two	Three	Four	Five
Mean Score	87.98	67.74	70.49	65.07	72.70	74.53
Standard Deviation	15.08	22.61	21.97	22.48	21.84	22.11
Number of Efficient DMUs	20	4	4	4	6	7
Percentage of Efficient DMUs	42.6	8.5	8.5	8.5	12.8	14.9
Minimum Score	45.2	9.7	7.7	7.4	8.0	8.0

Table 8.11: Impact of Weight Restrictions

In comparison with the unrestricted model, where almost half of the DMUs were evaluated as being 100% efficient, it can be seen that the number of efficient units was substantially reduced. There were just four efficient DMUs in scenarios one, two and three, which are the scenarios with the tightest restrictions applied. In scenario five, where restrictions are only applied to the primary output factors, there are seven efficient units in all, with three of the DMUs identified as efficient in the model 8.1 regaining their efficiency rating.

A further expected impact of weight restrictions (see section 8.4.1) was the overall reduction in efficiency scores, as can be seen by observing the average values in table 8.11 above.

The difference between the average efficiency score for each of the scenarios and model 8.1 appeared to be quite dramatic. The inclusion of factor weights, therefore, appeared to have an impact on the efficiency evaluation of the sample of DMUs, in relation to the average efficiency scores and the number of efficient units. Conversely, when assessing the results from the alternative scenarios, the observable differences did not appear to be quite as dramatic. As shown in table 8.11, the number of efficient units is fairly static throughout and the average efficiency scores did not vary to the same extent, particularly in the analysis of scenarios one to three where the level of restrictions applied was similar.

As was discussed in chapter five, the non-parametric nature of the DEA results restricted the extent to which the significance of the differences discussed above can be verified using traditional statistical approaches, such as a test for the comparison of means based on the normal distribution. However, the relationship between the efficiency scores for every DMU in the model 8.1 (the unbounded model) and each of the scenarios was investigated using correlation analysis, as suggested by Parkin and Hollingsworth (1996). Thus, the Spearman Rank Correlation Coefficient was calculated for each of the scenarios, with both the basic model and each other, to further investigate the impact of including weight restrictions. The correlation between efficiency scores for each scenario and the unbounded model was relatively low, in that it was approximately equal to 0.5 in each case. This implied that there was some relationship between the efficiency score calculated for each DMU in the unbounded model and its equivalent score in the weighting scenario, as the correlation coefficient was significantly different from zero in each case.

However, as the coefficient was not close to one, this suggested that including weight restrictions did have some effect on the ranking for each DMU.

As was observed in the analysis of the average efficiency scores above, analysing between the scenarios produced very different conclusions. The correlation coefficients calculated between the scenarios were relatively high (much closer to one) in every case. This suggested that a DMU with a high ranking (efficiency score) in scenario one, for example, was likely to have a high ranking (efficiency score) in scenario two. These correlation figures are shown in table 8.12 below:

	Model 8.1	**ONE**	**TWO**	**THREE**	**FOUR**	**FIVE**
Model 8.1	1					
ONE	0.49	1				
TWO	0.54	0.89	1			
THREE	0.51	0.92	0.98	1		
FOUR	0.53	0.88	0.99	0.95	1	
FIVE	0.55	0.78	0.95	0.90	0.96	1

Table 8.12: Correlation Statistics for Efficiency Rankings in Model 8.1 and the Weightings Scenarios

The preceding analysis, therefore, suggested that applying weight restrictions has significantly affected the overall efficiency rating of the DMUs. However, the actual choice of weight restrictions was of less importance – the rankings of the DMUs under each of the scenarios were very highly correlated (using the Spearman Rank Correlation Statistic). Also, significant differences between the mean efficiency scores for each of the scenarios were not generally observed.

The impact of the alternative scenarios was also be investigated by looking at the changes to the efficiency score for the individual DMUs, under each of the weight restriction patterns. As was discussed above in analysing the average efficiency scores, some considerable differences were observed. Table 8.13 references the five largest variations in efficiency score between the unbounded model and each of the scenarios. As can be seen, three particular DMUs (*23, *33 and *35) are included in the list for each of the scenarios, suggesting that these DMUs were most severely affected by the introduction of weight restrictions.

Weight Restriction Scenario				
ONE	**TWO**	**THREE**	**FOUR**	**FIVE**
*35 (90.3)	*35 (92.3)	*35 (92.6)	*35 (92.0)	*35 (92.0)
*5 (74.5)	*23 (74.0)	*23 (74.5)	*23 (73.9)	*23 (73.8)
*23 (68.7)	*33 (53.6)	*33 (55.8)	*33 (53.1)	*33 (52.8)
*33 (46.3)	*21 (45.5)	*21 (50.7)	*21 (44.0)	*21 (43.4)
*47 (42.9)	*5 (42.7)	*31 (46.8)	*5 (41.8)	*5 (38.1)

Table 8.13: The Five Largest Differences in Efficiency Scores between Model 8.1 and Each Scenario

In regard to DMU *35, the output value for 'day case attendances' was particularly small in relation to its other outputs and the data sample as a whole. Therefore, attaching any form of restriction to this factor that enforced a minimum contribution had a particularly noticeable affect on its efficiency rating. Other DMUs with low values for one or more of the primary factors were affected, in terms of their efficiency rating, in a similar way. Further discussion of the inefficient DMUs will be given later in this section, in order to understand some of the other reasons why these DMUs in particular were most affected by the introduction of the weight restrictions.

Alternatively, table 8.14 summarises the impact of weight restrictions by listing the average difference between the efficiency score under model 8.1 and that obtained for each of the scenarios and additionally the differences between the scenarios. In relation to the unbounded model, efficiency scores were reduced by between 15% and 20% on average when weight restrictions were introduced into the DEA models. As was the case with the analysis of actual efficiency scores above, the differences between model 8.1 and each of the scenarios appeared to be much greater than the variations across the various scenarios. This was most notable for comparisons between the results from scenarios four and five, where the difference between efficiency scores was less than 2% on average.

	Model 8.1	**ONE**	**TWO**	**THREE**	**FOUR**
ONE	20.23	-			
TWO	17.49	5.80	-		
THREE	22.91	6.81	5.42	-	
FOUR	15.28	6.52	2.49	7.63	-
FIVE	13.44	8.24	4.05	9.47	1.84

Table 8.14: Average Differences In Efficiency Scores

Additionally, the spread of the efficiency scores was investigated by reference to the calculated standard deviation values as shown in table 8.12. The standard deviation of the efficiency scores for model 8.1 was 15.08, whereas for the scenarios it has been calculated as between 21 and 23. This suggested that the spread of efficiency scores in the scenarios had become much wider, demonstrated by the increased number of DMUs receiving an efficiency score of less than 50%. This is illustrated in figure 8.1 overleaf.

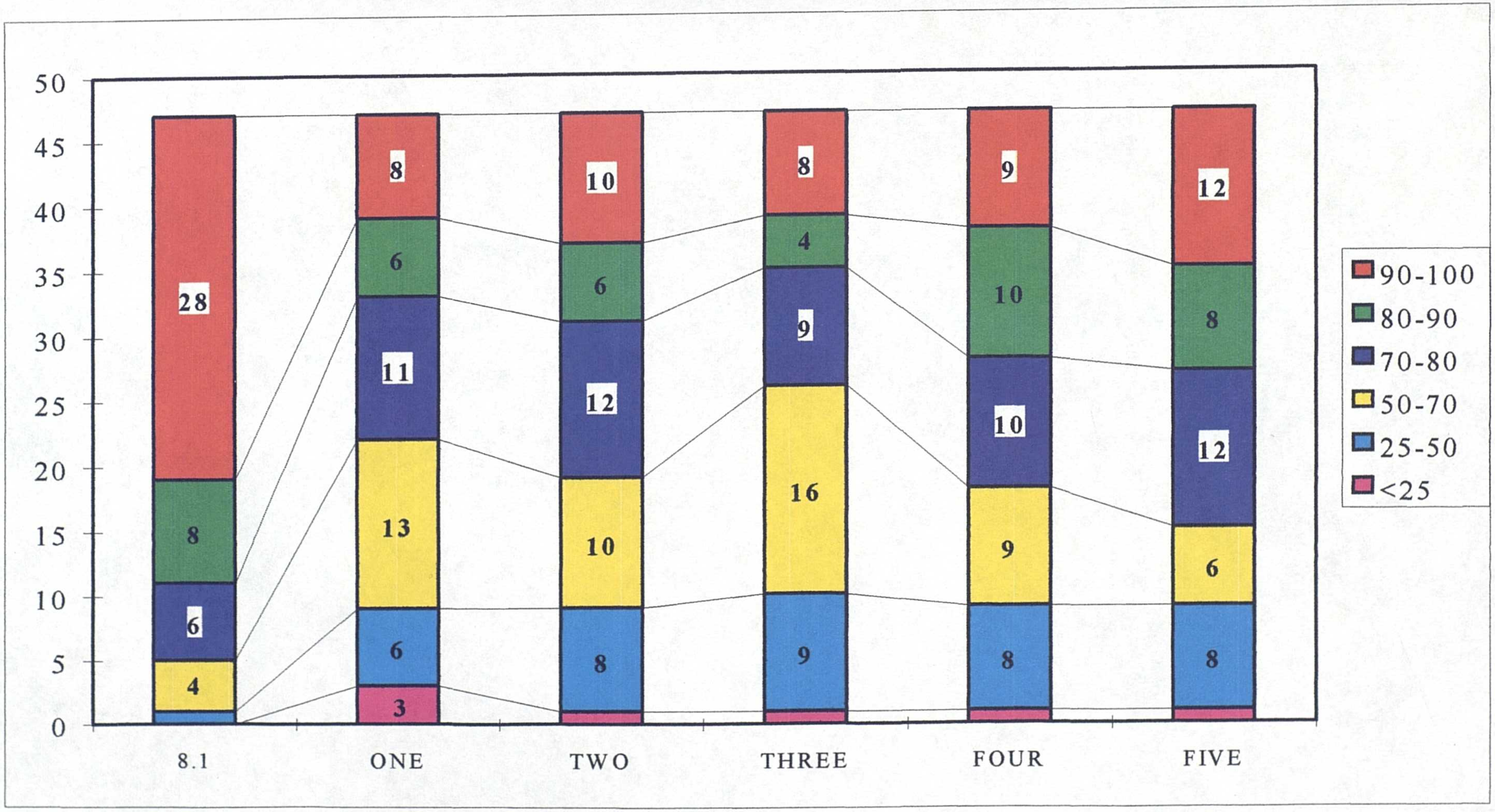


Figure 8.1: Distribution of the Efficiency Scores for Model 8.1 and the Scenarios

As figure 8.1 demonstrates, the redistribution of the efficiency scores was most apparent at the highest end of the value ranges, that is, efficiency scores between 90% and 100%. In the unrestricted basic model, 28 of the DMUs had efficiency scores of greater than 90%, with just one DMU rated as less than 50% relatively efficient.

However, under all the alternative scenarios, which demonstrated fairly similar patterns of efficiency scores in terms of their distribution, the number of DMUs in the highest range was greatly reduced, repeating the pattern found for the number of efficient units.

In the unbounded model, there was little incentive for the DMUs to improve their operating practices, as the majority of them appeared to be operating at a very high level of efficiency. The introduction of weight restrictions, however, has emphasised the existing inefficiencies much more clearly, demonstrating that additional effort should be given to improving the efficiency with which services are provided.

There was a much more cogent focus for addressing these inefficiencies, as a further impact of weight restrictions was the identification of a core group of efficient DMUs. Under all scenarios, there were just four efficient DMUs (*2, *3, *13 and *43) and they have shown, by the fact that they remain efficient under all five scenarios, that their efficiency rating was robust. These hospitals could be used as 'role models' for the inefficient DMUs and are discussed again in greater detail later in this chapter.

The impact of the weight restriction scenarios on several DMUs in particular was noted in table 8.13, where large differences between the efficiency score in the unbounded model and that obtained under each scenario were observed. Of the seven different DMUs referenced in table 8.13, five of these were rated as efficient in the basic model.

Analysing the overall results for each scenario, the sample of DMUs was partitioned into two distinct groups: (i) efficient in model 8.1 and inefficient in the weight restriction scenarios and (ii) inefficient in model 8.1 and inefficient in the weight restriction scenarios. Marked differences were observed in the way the weight restrictions impacted on these two groups, particularly in the average change to their efficiency score. Table 8.15 summarises this information for each of the five scenarios, including only the data on units inefficient in the weighting scenarios.

Scenarios	Units Efficient in Model 8.1			Units Inefficient in Model 8.1		
	Average Difference	Standard Deviation	Number of Units	Average Difference	Standard Deviation	Number of Units
ONE	28.9	24.8	16	18.1	15.4	27
TWO	26.4	27.3	16	14.8	10.7	27
THREE	32.0	26.1	16	20.9	11.5	27
FOUR	**28.6**	**27.6**	**14**	**11.8**	**11.6**	**27**
FIVE	28.7	29.2	13	9.6	10.0	27

Table 8.15: The Impact of Weight Restrictions on Efficient and Inefficient DMUs

The Mann-Witney U-test for non-parametric data was used to analyse the information presented in table 8.15, testing whether the observed differences were significant. This proved to be case for scenario four only, using a 95% significance level.

The implication of this analysis was that it appeared as though the inclusion of weight restrictions affected the efficiency assessments of the individual DMUs in different ways in this example, depending on their efficiency evaluation in the unbounded model and also on other factors. The DMUs rated as efficient in the unbounded model and inefficient using the weighting scenarios appeared to be much more dramatically affected by the inclusion of weights on some occasions. However, the efficient DMUs also had more to lose, in terms of their efficiency score being reduced, than the very inefficient DMUs.

The efficiency rating for each of the efficient DMUs in the unbounded model was certainly not robust, as it could not be sustained following restrictions to the factor weights (apart from the core group of efficient DMUs identified above). Their rating as efficient was essentially a feature of one of the properties of the DEA methodology, such that each DMU is allowed to select the weights that show their efficiency 'in the most favourable light' (Dyson *et al.*, 1990). The contribution from a particular factor was also suggested earlier in this section as a primary cause for several of the DMUs, including *35, having greatly reduced efficiency scores in the weighting scenarios.

As was discussed earlier in this chapter, the health care managers interviewed in relation to this research would prefer to use a method of assessment that showed each DMU in its true light. Introducing weight restrictions has clearly identified those DMUs whose efficiency rating in the unbounded model was suspect. The inclusion of weighting scenarios would help the health care managers to be much more aware of any apparent inefficiency that exists in the provision of services in their hospital.

Clearly, it would have been possible to investigate the general impact of weight restrictions in a number of other ways. However, the analysis above has shown that the general implications of weight restrictions were:

1. The average efficiency score was reduced but the spread (measured by standard deviation and distribution) of the efficiency scores increased;
2. A core group of efficient DMUs was identified, which were efficient in the unbounded model and also in each of the scenarios;
3. The impact of weight restrictions on the efficient DMUs in the unbounded model appeared to be greater than on the inefficient units;
4. The difference between the efficiency scores in the various scenarios was not as great as between the scenarios and the unbounded model (model 8.1).

8.7.2 Analysis of the Individual DMUs

As was stated in the previous section, introducing weight restrictions had the effect of reducing the number of efficient DMUs, with the result that a core group of efficient units was identified. These units could be thought of as role models for the inefficient DMUs and investigated in order to determine patterns of efficient practice.

Additionally, the impact of the different weighting scenarios on the individual inefficient units was examined in order to ascertain how these units could best improve their efficiency. Therefore, these two groups of DMUs will now be investigated.

8.7.2.1 The Efficient DMUs

The core group of efficient DMUs, identified as efficient under all scenarios and in the model 8.1, numbered just four: *2, *3, *13, *43. Table 8.17 below lists the input and output factors for these DMUs.

Factors	Inputs			Outputs			
DMU	CAP	TDC	TAC	COA	AEA	TID	DCA
*2	21.7	**71314**	**33121**	257572	62177	**62654**	**20091**
*3	14.9	69489	24140	**265709**	**91626**	38247	12431
*13	79.4	29744	10600	128834	66457	27913	15228
*43	51.0	28332	14370	121895	59293	26085	11208
Sample Average	61.9	20112	10249	74463	28614	15709	6179

Table 8.16: Inputs and Outputs for the Core Group of Efficient Units

As the table shows, the core group of efficient units had above-average values for each of the variables, excluding 'capital charge'. In the case of the other factors, the values for the core group were up to four times above average. The figures highlighted in bold were the maximum values for all the DMUs in the sample. Using the simplest form of the DEA model without weight restrictions would automatically lead to the DMUs with the highest ratio for any given pair of input and output factors being labelled as efficient or very close to efficient (Dyson *et al*, 1990).

This applied to DMU *2 in this sample and in the unbounded model, this DMU achieved its 100% rating with all the weight attached to just one of its outputs, 'total inpatient discharges'.

As this DMU operated at the extremes (it is one of the largest hospitals in the data set according to expenditure and the number of patients treated), few of the other inefficient units in the sample related to it directly in their operating practices. This was reflected in the reference set count to be discussed below. However, as the DMU was rated as efficient in each of the scenarios (with a different set of factor weights applied in each case), its rating was certainly more robust than many of the other DMUs in the sample.

In order to distinguish between the efficient DMUs in any sample, several methods have been suggested, including the use of cross-efficiencies and the super efficiency model (see sections 5.8.5 and 5.8.6). The most frequently used method in the literature, however, was to look at the reference set count, that is, the number of occasions that each of the efficient DMUs was found in the reference groups of the inefficient units. Looking at the core group of efficient units, the same DMU was the most frequently referenced unit for each of the scenarios, that, is, DMU *13. This implied that the inefficient DMUs associated most closely with the efficiency evaluation of DMU *13. DMU *2 was the next most referenced DMU, with DMU *43 included in the reference set for the inefficient DMUs to a much lesser extent.

The remaining DMU (*3) in the core group of efficient units, which was identified as operating at the extremes of the efficiency frontier, was included in the reference set for very few, if any, of the inefficient DMUs under all of the scenarios and also in the unbounded model. This confirms the suggestion made above that few of the inefficient DMUs focused on DMU *3 in their efforts to improve their efficiency.

Looking at the patterns of the virtual weights for the core group of efficient DMUs also provided additional information on efficient practice in this data sample. Table 8.17 below shows the dominant output factor for each of the DMUs, under the basic model and the alternative weighting scenarios.

Model DMU	8.1	ONE	TWO	THREE	FOUR	FIVE
*2	TID	TID	DCA	DCA	TID	TID
*3	COA	AEA	AEA	AEA	COA	COA
*13	DCA	DCA	DCA	DCA	DCA	DCA
*43	COA	COA	COA	AEA	COA	COA

Key: TID – 'total inpatient discharges' AEA – 'A&E attendances'
COA – 'consultant outpatient attendances' DCA – 'day case attendances'

Table 8.17: Dominant Output Factors for the Core Group of Efficient Units

As can be observed, the DMU identified above as providing the most important data on efficient practice using the idea of the reference set count (*13), based its efficiency assessment on one of the primary output factors in all of the scenarios and the unbounded model. DMUs *3 and *43, which were identified as being less useful in providing targets for the inefficient units in relation to the reference set count, based their efficiency contributions on significant contributions from the secondary output factors.

It was interesting to note, however, that only one of the core group of efficient DMUs had the same dominant output factor across all the scenarios, this being the DMU repeatedly referenced as the most influential of the efficient units (DMU *13).

In relation to the factor weights applied to the inputs, the pattern for each of the core group of efficient DMUs was quite varied across the scenarios and in comparison with the model 8.1. Under scenarios one to three, the dominant input for each of the DMUs was constrained by the weight restrictions applied so that 'total direct costs' would be the dominant factor. However, in the scenarios where the input constraints were removed, only three of the DMUs had this as the dominant factor. DMU *3, previously recognised as operating at the extremes of the efficiency frontier, reverts to using 'total allocated costs' as the significant contributor to its efficiency score on the input side, as was the case in the basic model.

Focusing on DMU *13 in order to identify aspects of efficient practice for the inefficient DMUs, the following table shows the patterns in which the virtual weights for all of the factors were applied in the weighting scenarios and the unbounded model. As can be seen, this DMU has a fairly constant pattern of factor weights across all scenarios, despite the changing weight restrictions.

	Inputs			**Outputs**			
	CAP	**TDC**	**TAC**	**COA**	**AEA**	**TID**	**DCA**
Model 8.1	0	97.6	2.4	30.6	4.6	31.3	33.5
One	5	90	5	20	20	25	35
Two	10	80	10	5	5	25	65
Three	10	50	40	20	5	25	50
Four	0	97.6	2.4	30.2	5	31.5	33.3
Five	0	97.6	2.4	30.6	4.6	31.3	33.5

Key: Figures given are the virtual weight percentages for each factor.
See table 8.17 for a definition of the factor codes.

Table 8.18: The Factor Weight Patterns for DMU *13

The pattern of factor weights for DMU *3 showed that its efficiency rating was based on significant and equal contributions from three of the four factors in the majority of the scenarios and in unbounded model ('total inpatient discharges', 'day case attendances' and 'consultant outpatient attendances'). In each of the scenarios, the relative unimportance of the factor recording 'A&E attendances' is highlighted (excluding scenario one where the outputs were constrained to be of roughly equal importance).

In scenarios four and five, several other DMUs were also identified as being 100% relatively efficient, having been rated similarly in the unbounded model but less efficient in the scenarios with the very tight weight restrictions.

For example, DMU *45 was efficient in model 8.1 and in scenarios four and five, losing its efficiency rating only when constraints were attached to the inputs. Analysis of its factor weights in the basic model show that its efficiency rating was achieved with a virtual input of 100% for the 'total direct costs' factor. A similar pattern was observed for DMU *12. Both of these DMUs, however, had very similar patterns for their output weights as DMU *13. DMU *13 was in their respective reference sets in each of the scenarios for which they were inefficient. DMU *47 again followed a similar pattern, although it was only efficient in model 8.1 and scenario five. The removal of the input constraints was not enough to allow it to regain its efficiency rating, as it also attached a significant weight to one of the secondary output factors. In terms of targets for improved efficiency for these DMUs, the necessary reductions in input required were all related to the 'total allocated costs' factor.

8.7.2.2 The Inefficient DMUs

In the previous section, observations were made on the impact of weight restrictions in relation to the core group of efficient DMUs. The discussion is now directed towards the DMUs evaluated as inefficient in the different weighting scenarios, as it was observed in tables 8.14 and 8.15 that the inclusion of weight restrictions caused a considerable change in the efficiency scores for many of the DMUs.

It was noted, for example, that DMU *35 was heavily affected by the introduction of weight restrictions, in that its efficiency score decreased by over 90%. An observation on the DMU's inputs and outputs were included in order to explain the causes for this change.

Therefore, in order to assess the impact of the various scenarios on individual DMUs more comprehensively, several other inefficient DMUs were selected for further analysis. These were used to investigate some of the issues raised in the previous section in the general analysis, such as determining the characteristics of the DMUs most affected by weight restrictions.

The weighting scenarios have affected the DMUs included in the selection in a variety of ways, with some achieving an efficiency rating in the unbounded model but not in any of the scenarios whilst others have remained inefficient throughout. The DMUs that achieved an efficiency score of 100% in some but not all of the scenarios were discussed in the previous section.

Of particular interest in this evaluation of the inefficient DMUs were the units included in the reference sets under each of the scenarios and whether these remained constant, taking into account the reduced number of efficient DMUs. Also, it was necessary to determine if the targets set for the inefficient units were similar in each of the scenarios in comparison with those identified under the unbounded model. This analysis is also concerned with a further stage of the DEA application process, that is, the development of efficiency strategies.

Looking first at the targets set for the inefficient DMUs, four were selected to illustrate the impact of weight restrictions. Table 8.19 below shows the efficiency scores for these DMUs (*4, *11, *28 and *42) in the unbounded model and in each of the scenarios. As can be observed, each had a slightly different profile in terms of their efficiency score. For example, the efficiency scores for DMU *4 were fairly constant throughout each of the scenarios. The scores for DMU *42 fluctuated to a much greater extent. The figure in parentheses for each DMU refers to its functional classification (defined in section 7.3).

Model DMU	8.1	One	Two	Three	Four	Five
*4 (1)	85	72.2	70.8	70.7	70.8	71.2
*11 (2)	100	96.5	97.1	96.6	97.6	98.1
*28 (7)	73.8	57.3	56.6	52.6	57.5	57.7
*42 (12)	90.3	80.5	78.6	74.1	82.3	82.4

Table 8.19: Efficiency Scores for a Selection of Inefficient DMUs

The DEA models used in the two-stage application process have been developed with an input minimisation orientation, reflecting the focus of efficiency evaluations in public sector services in the UK (see section 7.4). Therefore, the targets set for the inefficient DMUs were primarily focused on reducing inputs and particularly the cost inputs.

The impact of the weighting scenarios on the target for 'total direct costs' (the core input factor) for the selected group of DMUs can be seen in table 8.20 below. This shows the expected percentage reduction in the input factor required for each DMU to meet the target, which was set for each DMU by its reference set. The targets obtained from the unbounded model and each of the scenarios are included, along with the input value for 'total direct costs' (TDC).

DMU Model	Percentage Reduction for 'Total Direct Costs'						Input Value for TDC (£000)
	8.1	One	Two	Three	Four	Five	
*4 (1)	15.0	35.4	29.2	29.9	29.2	28.8	63496
*11 (2)	0.0	8	2.3	0.3	2.4	1.9	28820
*28 (7)	25.3	40.8	40.3	40.3	41.7	41.6	26594
*42 (12)	9.7	17.1	16.5	9.1	17.7	17.6	21178

Table 8.20: Targets for Improving Efficiency for a Selection of Inefficient DMUs

As can be observed, the percentage decrease that was required for the 'total direct costs' factor was considerably larger for each of the DMUs in the scenarios with weight restrictions applied. Following the targets would substantially reduce the expenditure levels for each of the DMUs.

DMU *11, which was in fact identified as efficient in the unbounded model, was required to reduce the specified input by up to 8%, according to the targets from the five scenarios. Thus, a DMU identified as efficient has been shown to have an area where cost reductions could be made, when the efficiency analyses were amended to include some form of weight restrictions. The proposed target for this DMU does vary considerably across the scenarios, as its efficiency ranges from 96.1% to 100%. If DMU *11 aimed for the maximum reduction in its 'total direct costs' (in scenario one), its expenditure could be reduced by over £2 million. Analysis of DMU *28, which also has a fairly consistent efficiency score across each of the scenarios, rated at around 55% throughout the scenarios, was consistently shown to have a potential reduction in the 'total direct costs' factor of around 40%.

Unusually, the reduction in 'total direct costs' for DMU *42 was lower in scenario three than in the unbounded model, which restricted the contribution from this factor to a maximum of 50%. However, looking at the targets set for the other input factors, the reduction in the 'total allocated costs' factor was considerably higher in scenario three, when compared with the unbounded model.

Figure 8.2 overleaf illustrates this information visually, showing the difference in the target levels for 'total allocated costs' and 'total direct costs' for DMU *4, whose efficiency score in the scenarios ranged from 70.7% to 72.8%. In the unbounded model, its efficiency score was 85%. The chart shows the DMU's current input levels for the two expenditure factors, the target set by the unbounded model and the minimum target identified in the five scenarios (relating to scenario one for DMU *4).

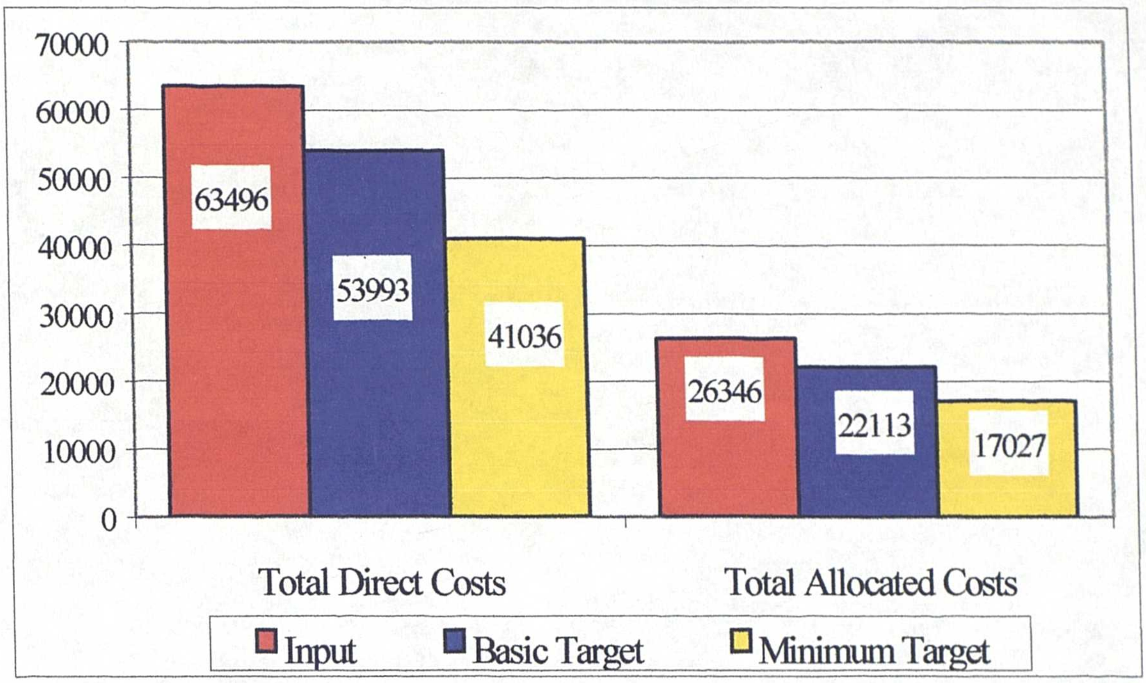


Figure 8.2: Input Reduction Targets for DMU *4 (Costs in £000s)

As the figure and the above discussion demonstrated, the targets for these two factors were very different for the unbounded model and the various scenarios. From the perspective of the hospitals, the target set by model 8.1 could be seen as a first step towards improved efficiency. The lower targets set by the scenarios could be seen as long-term goals. In the case of DMU *4, the first target would produce an overall reduction in expenditure of more than £13 million (reducing its budget by 15%). The long-term goal, following the targets set by scenario one, would be a decrease in its expenditure of over £30 million. The results from the analysis have been used to produce an efficiency strategy for DMU *4.

A similar efficiency strategy could be developed for each of the inefficient DMUs. However, as each scenario produced a slightly different set of targets for the inefficient DMUs, it would be necessary for each DMU to identify the scenarios that were related to their operational structure or their long-term performance strategy.

For example, the DMUs attempting to provide an equally efficient service in all areas could use the results from the scenario that assumes all of the outputs should be treated equally in the efficiency calculations (scenario one). If they wanted to continue to focus on one or two output areas in particular, the results from scenarios four and five would be more appropriate.

For every DMU, a long-term efficiency strategy could be determined following the application of weight restrictions to the basic DEA methodology. The approaches required to reach the targets were not directly identified by the DEA methodology (as was the case with performance indicators). However, the identification of a core group of efficient units provided each inefficient DMU with a role model, from whom 'best-practice' could be observed and adopted.

Looking at the efficiency debate from the alternative perspective, targets could be related to increasing output levels, using the output maximisation orientation of the DEA methodology. Outputs would be increased relative to the current inputs, with the targets being increases rather than decreases in each factor. The long-term goal would be to reach the maximum output levels set for each factor across the scenarios. This approach to the efficiency strategy is discussed in chapter ten.

The final part of this investigation into the inefficient DMUs was to examine the units included in their reference sets, from which the targets discussed above and their whole efficiency analysis were developed.

Introducing weight restrictions was shown to reduce the number of efficient DMUs and this will automatically affect the nature of the reference sets for each inefficient DMU, primarily because there are a greater number of inefficient DMUs.

The major concern in the investigation of reference sets was the significance or influence of each DMU having a slightly different set in each of the scenarios. The four DMUs included in the discussion above, one of which was efficient in the unbounded model, actually had a varied range of DMUs included in their respective reference sets, as can be seen in table 8.21 below.

Model DMU	8.1	One	Two	Three	Four	Five
*4 (1)	*2, *3, *14, *16	*2, *13, *43	*2, *3, *13	*2, *13	*2, *3, *13	*2, *3, *13
*11 (2)	*11	*2, *43	*2, *3, *13	*2, *43	*2, *43	*2, *43
*28 (7)	*11, *13, *20	*13	*13	*2, *13	*12, *13, *45	*12, *13, *45, *47
*42 (12)	*11, *12, *14, *30, *43	*13	*13	*2, *13, *43	*12, *13, *45	*12, *13, *45, *47

Table 8.21: References Sets for a Selection of Inefficient DMUs

The reference sets for the selection of inefficient DMUs was seen to vary quite considerably, although in each case there were one or two of the efficient DMUs that were included across all the scenarios. Interestingly, all of the DMUs that were only efficient in some of the scenarios appear in the reference sets for two of these DMUs (*28 and *42).

However, this range of reference sets did not seem to impact to any great degree on the targets set for each of the inefficient DMUs, which were fairly consistent across all the scenarios.

For example, for DMU *42, five different DMUs appeared its reference set, looking at all of the scenarios, in four different combinations. The target for the 'total direct costs' factor ranged from £19.2 million (scenario three) to £17.4 million (scenario four). For four of the scenarios (one, two, four and five), the target range was £17.4 million to £17.7 million.

Clearly, analysing each of the inefficient DMUs would highlight numerous other implications of the inclusion of weight restrictions in the DEA methodology. However, the cross-section of DMUs included in the discussion here have shown that introducing weight restrictions has reduced the targets set for the inputs in comparison with the basic model. Additionally, the reference set for each of the DMUs has been shown to vary quite widely across the scenarios, although the nature of the target set for each of the input factors was seen to be quite consistent. The identification of a number of scenarios was used to develop an efficiency strategy for one DMU in particular (*4), although the same approach was applicable for all of the inefficient DMUs.

In the analysis that concludes the discussion on the application of weight restrictions, the patterns of factor weights in the unbounded model are related to the changes in efficiency scores observed.

8.7.3 Analysis of the Factor Weights

In the preceding analysis, a selection of inefficient DMUs were identified, each affected by the introduction of weight restrictions in a number of different ways. Table 8.13, for example, listed the DMUs most affected by the weight restrictions, in terms of the largest reduction in their efficiency scores. Additionally, four inefficient DMUs were highlighted in section 8.7.2.2, with the effect of the weight restrictions on their efficiency scores shown in table 8.15.

Looking first at the group of DMUs referenced above as being most affected by the inclusion of factor weights, it was observed that the factor weights applied in model 8.1 all followed a very similar pattern. The efficiency assessment for each DMU was generally dominated by a single input, usually 'total direct costs'. Additionally, there were large contributions from the secondary factors for the outputs or a 100% contribution from just one output. The secondary output factors were defined to be 'consultant outpatient attendances' and 'A&E attendances' and they were restricted from dominating the efficiency rating for each of the DMUs in the weighting scenarios. Additionally, the output values for the primary factors were also found to be small in comparison with the other factors and with the average levels across all the DMUs, as was observed for DMU *35 in the previous section.

The scenarios introduced above enforced contributions from the two primary output factors in every case and also from each of the inputs in scenarios one to three. Thus, it could be expected that the DMUs that did not include such contributions were likely to be heavily penalised in their efficiency rating.

Table 8.22 below lists the virtual factor weights for the specific group of DMUs identified in table 8.13 in the model 8.1 (defined in section 8.2). DMU *35 is excluded from the analysis, as the efficiency assessment for this hospital has been discussed in detail in section 8.7.1.

DMU	Virtual Inputs (%)			Virtual Outputs (%)			
	CAP	TDC	TAC	COA	AEA	TID	DCA
*5	15.4	84.6	0	86.5	0	13.5	0
*21	32.5	67.5	0	0	69.9	30.1	0
*23	0	94.9	5.1	48.4	51.6	0	0
*31	0	100	0	0	0	0	100
*33	0	100	0	0	92.1	0	7.9
*47	0	100	0	40.3	0	8.5	51.2

Key: See table 8.17 for a definition of the factor codes.

Table 8.22: Virtual Weights for a Selection of Inefficient DMUs in Model 8.1

As the table shows, each of these DMUs had 'total direct costs' as the dominant input and four out of the six has one of the secondary factors as the dominant output (the largest contribution in terms of the virtual output percentage).

DMU *47 was unusual amongst the selection, in that it was one of three DMUs that were efficient in the basic model, inefficient in some of the scenarios and then regained its 100% efficiency score in scenario five. Scenario five has restrictions attached only to the primary output factors, such that DMU *47 could include a significant contribution from 'consultant outpatient attendances' in this case and also attached a virtual weight of 100% to 'total direct costs'.

DMUs *5 and *47 were also unusual amongst the selection of DMUs identified above, in that their lowest efficiency score was for scenario one, which enforced equivalent contributions from the outputs. The other four DMUs recorded their minimum efficiency score in scenario three, which restricted the input from 'total direct costs' (a maximum virtual weight of 50%) and prevented any of the secondary outputs from contributing more than 45%, in terms of their virtual weights.

The DMUs discussed above were seen to have their efficiency rating dramatically reduced in one or more of the scenarios. Not all the DMUs were similarly affected, as was observed in analysis of DMU *11. Its efficiency scores ranged between 96.1% and 100% across the scenarios and model 8.1. Several other DMUs followed this pattern. Their efficiency score in model 8.1 and the five scenarios are listed below:

Model DMU	8.1	One	Two	Three	Four	Five
*8	77.5	72.6	75.4	72.6	76.6	77.3
*9	79	74.9	75.7	73.7	76.4	77.1
*14	100	97.4	97.9	97.9	97.9	98.6
*40	66	62.3	60.6	54.8	64.5	64.5

Table 8.23: Efficiency Scores for DMUs Robust against Weight Restrictions

The DMUs referenced above all have a very similar pattern of factor weights, in that their efficiency scores are based on significant contributions from the primary factors or the virtual weights are fairly evenly distributed. They were not all 100% efficient in model 8.1 but their efficiency rating was robust against the type of restrictions included in this investigation, due to their apparent strength in the important areas.

8.8 Summary of the Second Stage of the Application Process

In this chapter, the key elements of the second stage of the DEA application process have been discussed. The results from the revised model were investigated, focusing specifically on the patterns of the factor weights and the virtual weight percentages. This led to weight restrictions being included in the DEA model, with the development of five alternative scenarios, each related to a specific policy objective or operational strategy.

The impact of weight restrictions was investigated in general terms and by looking at a number of individual DMUs. It was found that the efficiency scores for many of the DMUs had decreased and there were changes in the units included in the reference sets for the inefficient DMUs. It was shown that the efficiency scores for some of the DMUs were robust against changes in the factor weights, whilst others were particularly sensitive. This depended on the weight restrictions applied and the particular strengths and weaknesses of each DMU.

It was observed that the alternative scenarios could be used to develop an efficiency strategy for each DMU, giving short-term and long-term goals for reducing its inputs in relation to the targets identified by the modelling process.

In the following chapter, the two-stage application process for the DEA technique is evaluated in relation to the requirements of the health care manager, as discussed in chapter six. However, the key benefits of the second stage, as discussed in this chapter, can be summarised in three points:

1. The identification of a core group of efficient hospitals: the small group of efficient hospitals provide a more cogent focus for the inefficient DMUs, in terms of identifying efficient practice and treating the core group as role models.
2. The improvements in overall standards: applying weight restrictions has shown up apparent inefficiencies in many of the hospitals identified as efficient in the basic model, showing them areas where there is potential for improving efficiency. Additionally, the levels of inefficiency in the majority of the inefficient hospitals have been shown to be much higher than would be the case without restrictions.
3. The identification of efficiency strategies: applying weight restrictions in the form of a series of scenarios, consistent with current operating practices and management strategies, could assist all the DMUs in improving their efficiency. The tightest weight restrictions produce lower targets for input reduction (or higher targets for output maximisation) and these could be seen as long-term goals. A series of scenarios, with the restrictions becoming tighter in the direction of some predetermined plan, would highlight the steps necessary to reach the long-term aims.

8.9 References

1. **Allen, R., Athanassopoulos, A., Dyson, R. G. and Thanassoulis, E.** (1997) Weight Restrictions and Value Judgements in Data Envelopment Analysis: Evolution, Development and Future Directions, *Annals of Operations Research*, Vol. 73, pp. 13 – 34.

2. **Ball, R. and Roberts, E.** (1998) The Relevance of DEA in the Public Sector - A Discussion Paper for Young OR 10, Young OR 10 Keynote Papers, Operational Research Society.

3. **Beasley, J.** (1990) Comparing University Departments, *Omega*, Vol. 18, pp. 171 – 183.

4. **Boussofione, A., Dyon, R. G. and Thanassoulis, E.** (1991) Applied Data Envelopment Analysis, *European Journal of Operational Research,* Vol.

5. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

6. **Cook, W. D., Kazakov, A and Roll, Y.** (1994) On the Measurement and Monitoring of Relative Efficiency of Highway Maintenance Patrols, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

7. **Department of Health** (1989) Working for Patients, HMSO Cm 555.

8. **Department of Health** (1997) The New NHS, Modern, Dependable, HMSO.

9. **Dyson, R. G. and Thanassoulis, E.** (1988) Reducing Weight Flexibility in DEA, *Journal of the Operational Research Society*, Vol. 39, No. 6, pp. 563 – 576.

10. **Dyson, R. G., Thanassoulis, E. and Boussofiane, A.** (1990), Data Envelopment Analysis, in "Operational Research Tutorial Papers", L.C. Hendry and R. Eglese (eds), The Operational Research Society, pp. 13 – 28.

11. **Parkin, D. and Hollingsworth, B.** (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991 – 1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

12. **Pedraja-Chaparrro, F., Salinas-Jiménez, J. and Smith, P.** (1997) On the Role of Weight Restrictions in Data Envelopment Analysis, *Journal of Productivity Analysis,* Vol. 8, Pt. 2, pp. 215 – 230.

13. **Roll, Y. and Golany, B.** (1993) Alternate Methods for Treating Factor Weights in DEA, *Omega*, Vol. 21, pp. 99 – 109.

14. **Thanassoulis, E., Boussofiane, A. and Dyson, R. G.** (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588 – 608.

15. **Thanassoulis, E., Dyson, R. G. and Foster, M. J.** (1987) Relative Efficiency Assessments Using Data Envelopment Analysis: an Application to Data on Rates Departments, *Journal of Operational Research Society*, Vol. 38, No. 5, pp. 397 – 411.

8.10 Additional Sources

1. Interviews at the NHS Executive Performance Directorate in Leeds.

2. Interviews at Blackburn, Hyndburn and Ribble Valley NHS Trust.

3. Personal Communication with Dr. R. Wilkinson (former research student at the University of Stirling) through Professor R. Ball, Department of Management and Organization, University of Stirling.

CHAPTER NINE

EVALUATING THE APPLICATION PROCEDURE FOR DATA ENVELOPMENT ANALYSIS IN THE ASSESSMENT OF HOSPITAL EFFICIENCY

9.1 Introduction

Evaluation was an essential part of the discussion into the approaches currently used for performance assessment within the NHS, with the main techniques being performance indicators and clinical audit (see chapter four). They were each found to have numerous benefits as well as serious limitations and deficiencies. This chapter is concerned with an evaluation of data envelopment analysis, focusing not just on the procedures presented for its application to the measurement of hospital efficiency in the preceding two chapters, but also on the nature of the methodology itself.

The theoretical concepts behind DEA and the framework of the methodology were discussed in chapter five. It was suggested that the technique was an adequate and appropriate method for the evaluation of efficiency in the health sector. The DEA model can handle multiple incommensurate inputs and outputs and it is not necessary to have a pre-defined production function, or input-output relationship (Sherman, 1984).

This discussion was extended in chapter six to focus on the measurement of hospital efficiency, which included an extensive review of previous applications of the DEA technique to this area taken from sources in the academic literature. It was noted that DEA has frequently been used to make observations on the nature of health services (see the DEA bibliography at the end of chapter six for a comprehensive list of papers in this area). However, the literature sources showed little evidence to suggest that DEA has been used as a tool for improving services, changing procedures or policy-making exercises, confirming the following observation:

> "A number of attempts to apply DEA have not met with the resounding success one might expect from an approach enjoying its current level of popularity." BELTON (1992)

Chapter six also included the views of a number of managers from within the National Health Service, emphasising the importance of understanding the context and culture of the organisations being evaluated. The key areas of concern in the use of techniques such as DEA were also highlighted. It was noted that DEA has not been widely utilised in the UK and was unfamiliar to a number of those interviewed.

In chapters seven and eight, a two-stage application process was presented, for the evaluation of health service efficiency using data envelopment analysis, guided by some of the ideas presented in chapters five and six. The key elements of the process were emphasised, with particular attention given to the definition of an appropriate sample, the selection of the variable set, the inclusion of weight restrictions, the setting of targets for the inefficient DMUs and the development of efficiency strategies. The purpose of the application was to investigate overall hospital efficiency on a general basis, which was seen to directly influence each stage of the application process. It was suggested that different policy objectives and management strategies could be incorporated into, or measured by, the DEA modelling process.

The evaluation of the DEA technique and its application is focused on three areas: highlighting the positive aspects, discussing the limitations and suggesting potential improvements. Many of the issues raised in this chapter can be seen as signposts to further study, to be carried out in the future to extend the scope of this research.

The discussion is looked at from two perspectives, the DEA theorist and the health service manager, with the perspective of the health service manager considered to be the most relevant. This study has been focused on developing the practical application of DEA rather than on particularly extending the theoretical debate. Therefore, the following sections correspond to the key issues raised in section 6.3.

Due to the nature of the application process, some evaluation has already occurred, particularly in relation to the selection of the variables and the sample of DMUs. These issues are also included in the discussion in this chapter, although with slightly less detail, in order to make the evaluation as comprehensive as possible.

The evaluation of the modelling strategy as a whole is considered initially, prior to a discussion of the specific issues raised in chapter six.

In the following sections, DMUs prefixed with '#' relate to the seventy-four hospitals included at the beginning of the application process. DMUs prefixed with a '*' are taken from the forty-seven DMUs used in model 8.1 and the subsequent weight restriction scenarios in the second stage of the application process in chapter eight.

9.2 The Modelling Strategy

The two-stage application procedure for DEA, illustrated in chapters seven and eight, was developed in response to the evaluation of the literature (see section 6.2.8.1), where the 'absence of a convincing model-building methodology' (Pedraja-Chapparo *et al*, 1999) was noted.

It was intended that this strategy could be applied not only to the sample of acute hospitals discussed in this research, but also to any other similar application of the DEA technique for the measurement of health service efficiency. The ability of the modelling strategy to be translated to other health care environments is considered in this section.

The ten elements of the modelling strategy are shown in table 9.1 below (repeated from chapter seven):

STAGE ONE:	1. Definition of the Sample; 2. Formulation of DEA Model Type and Orientation; 3. Selection of Possible Inputs and Outputs; 4. Results of the Preliminary Analysis; 5. Revisions to the Model – Sample Size, Variable Set.
STAGE TWO:	1. Results from the Revised Model; 2. Evaluation of Factor Weights; 3. Application of Weight Restrictions; 4. Development of Efficiency Strategies; 5. Evaluation of Model – Sensitivity and Robustness.

Table 9.1: The Two-Stage Application Procedure for Data Envelopment Analysis

The general nature of the modelling strategy should mean that it would be easily applicable to other health service situations, for example, hospital departments, health boards and primary care facilities or programmes.

The analysis presented in chapter five, for example, used data relating to the provision of inpatient services in the specialty of gynaecology and covered similar areas to those specified by the modelling strategy.

However, for some of the elements in the evaluation of acute hospitals in Scotland, it was necessary to apply a number of restrictions to the assessment process. The application process was not straightforward, in that, the same principles or strategies would not be appropriate in every application of the DEA technique. The modelling strategy required additional information, such as a clear definition of the objectives for the evaluation, which was particularly important in the selection of the variables and the development of the weight restriction scenarios.

Managerial perspectives also impinged most significantly on the formulation of the DEA model (chosen to be input-minimisation to reflect recent government policy) and the reduction of the sample (to ensure that the DMUs being investigated were homogeneous).

Therefore, whilst the modelling strategy could be considered to be a useful guide, there are clearly some aspects of it that require further investigation if it is to be applicable to other types of efficiency assessment. The key areas of concern, identified from the analysis of acute hospitals in Scotland, are in the definition of the sample, the selection of the variables and the development of the weight restriction scenarios. These are considered further in the following sections, where alternative applications of the modelling strategy are also considered.

9.3 Selection of Samples

In the analysis presented in chapters seven and eight, a sample of acute hospitals in Scotland was evaluated using the DEA methodology. The suitability of this sample is evaluated in this section.

Discussions with health care managers, as referred to in chapter four in the discussion on performance indicators and medical audit, emphasised that methods of comparative performance assessment such as data envelopment analysis and performance indicators have been regarded with a certain amount of mistrust. The widely held belief within many hospitals, stressed by two managers at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust, was that each hospital was unique in its operation, facilities and environment.

It was suggested that the impact of this view was that a hospital should not be directly compared with any other hospital as a whole. It was also implied that this argument could be used to avoid participating in any level of performance assessment or as a justification for rejecting results obtained from them, particularly if they were unfavourable.

However, in the DEA papers evaluated in chapter six, samples of hospitals were compared using the DEA methodology, with a number of criteria used to determine exactly which hospitals could be included in the analysis.

Therefore, in chapter seven, the original sample of 74 DMUs to be used in the analysis was identified as those hospitals classified as acute using the guidelines defined by the NHS in Scotland (ISD, 1996). The use of hospital trusts as the DMUs was rejected, as many of the trusts in Scotland consist of hospitals at more than one location and of many different types.

It was perceived at the initial stage of the analysis that restricting the sample to just those hospitals defined to be acute would provide the homogeneous sample preferred by the DEA theorists and the health care managers. This was consistent with the applications by Parkin and Hollingsworth (1996) and Ozcan, Luke and Haksever (1992) found in the DEA research detailed in chapter six.

However, many differences in output and input patterns were identified in the subsequent analysis of the data and the determination of the most appropriate set of input and output factors (see section 7.5).

Certainly, such a large variation in the activity levels of the hospitals being compared would be unacceptable to the Financial Director of the Blackburn, Hyndburn and Ribble Valley NHS Trust, who was particularly concerned by such methods of comparative analysis. Taking two of the DMUs in the original data sample of 74 acute hospitals illustrates this area of concern. The largest value for the 'total inpatient discharges' factor was for hospital #2, with 62861 cases undertaken and a total expenditure of £73,336,000. In comparison, DMU #53 dealt with just 222 inpatient cases, at a total cost of £473,000.

As was discussed in section 7.7, a number of DMUs were excluded from the second stage of the application, as they did not treat patients in each of the output categories defined to be the most significant in this case. This was carried out in order to satisfy the requirements for homogeneity, such that each of the DMUs should have had a similar organisational strategy.

However, this approach to the efficiency assessment could appear to be somewhat arbitrary. It is possible that the number of the DMUs in the sample could vary from year to year in an ongoing analysis of efficiency, as hospitals made changes to their operating practices. A number of DMUs were also excluded as they treated a very small number of patients in one of the given categories of output. However, the definition of 'very small' was clearly a matter of judgement, based on an analysis of the data and the nature of the hospitals. Additionally, the efficiency of the excluded DMUs, in relation to the sample as a whole and also to each other, was not assessed to any great degree. This would be significant if DEA were to be adopted as a means for efficiency assessment in the NHS, as all DMUs would need to be incorporated into a sample, in order to have their efficiency assessed.

In relation to the assessment of acute hospitals in Scotland, further investigation of this sample of forty-seven DMUs demonstrated that even the reduced sample still contains DMUs with extensive differences in their input and output patterns of activity. Table 9.2 below contains the minimum and maximum values for each of the factors (with the number of the DMU from which this has been obtained also shown).

Considerable variation can be noted in each case, with the average value also included to illustrate this still further. A definition of the factor codes is given in the key below the table.

	CAP	TDC	TAC	COA	AEA	TID	DCA
Minimum	14.9 (*3)	668 (*26)	470 (*27)	2080 (*34)	1093 (*25)	326 (*31)	25 (*35)
Maximum	208.16 (*41)	71314 (*2)	33121 (*2)	265709 (*3)	91626 (*3)	62654 (*2)	20091 (*2)
Average	61.87	20112	10249	74463	28614	15709	6179

KEY: Figures in brackets represent the DMU from which the minimum or maximum value was obtained using the coding system defined in section 6.8.

CAP: 'capital charge' TDC: 'total direct costs' TAC: 'total allocated costs'

COA: 'consultant outpatient attendances' AEA: 'A&E attendances'

TID: 'total inpatient discharges' DCA: 'day case attendances

Table 9.2: Minimum, Maximum and Average Values for Factors used in Model 8.1

As the table shows, the maximum expenditure on direct patient care (TDC) is more than one hundred times greater than the minimum, similarly for the allocated expenditure (TAC). In terms of output, similar large differences can be identified between the figures for minimum and maximum values. Each of the minimum values has been obtained from a different DMU, whilst the maximum values belong to either DMU *2 or *3 for all the factors apart from 'capital charge'.

Extending the analysis of the data, table 9.3 overleaf gives the average values for each of the four output categories used in model 8.1, and compared across smaller groupings of the hospitals within the data sample, developed from the functional classifications (defined in appendix one).

It can be seen that each different output category has a particular characteristic for each of the functional classifications. For example, for inpatient activity (TID), the values for codes 7+, 11 and 12 are fairly consistent and close to the average activity value. However, for outpatient activity, the values for codes 11 and 12 are again very similar but are considerably higher than the value for code 7+.

Output/ Functional Class	COA	AEA	TID	DCA
1	191938	53722	36836	13376
2	113755	46711	24174	10115
5	7829	5950	1875	528
7+	24876	29946	13354	3254
11	59233	23828	13592	5748
12	64324	29835	15706	7270
Average	**74463**	**28614**	**15709**	**6179**

KEY: The values calculated were for hospitals within a single functional classification apart from 7+, which is a combination of several functional classification with only a small number of hospitals used in model 8.1. The average is calculated from all 47 DMUs. The output codes are those first defined in section 7.5 and represent outpatient, inpatient, day case and A&E activity.

Table 9.3: Examining Differences in Output Activity

Figure 9.1 overleaf represents this information graphically for two of these outputs (AEA and DCA) and shows up some of the differences quite clearly. Most apparent in the figure was the difference between functional class five and the averages for the other classes. As with the values for 'total inpatient discharges' and 'consultant outpatient attendances', these values were considerably smaller.

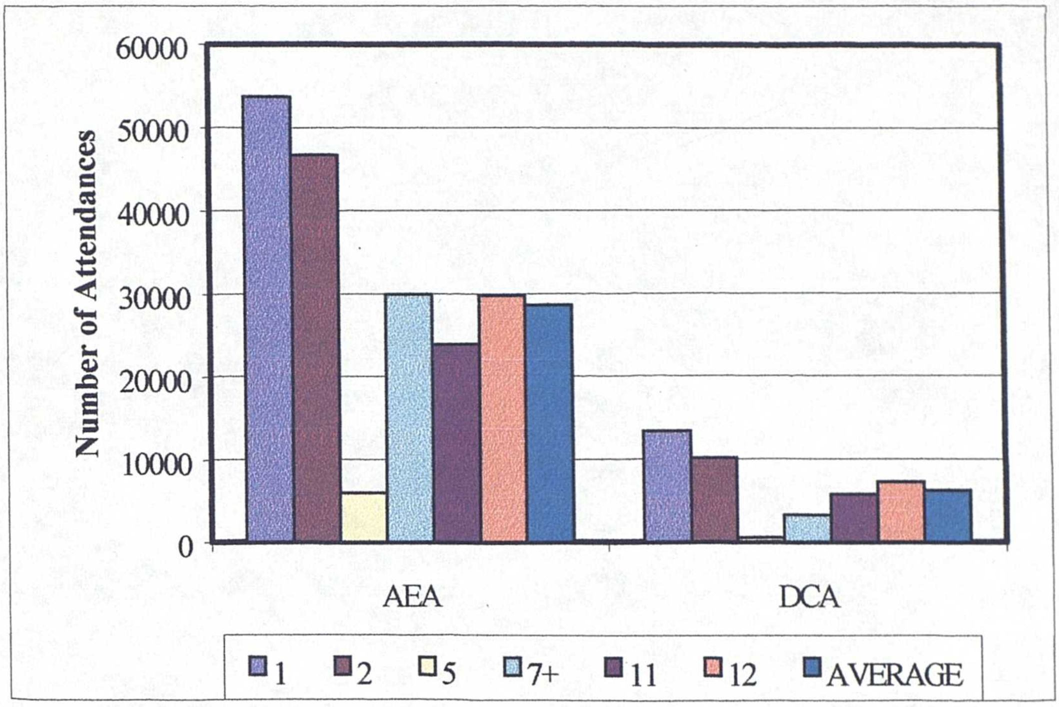


Figure 9.1:Comparison of Output Patterns According to Functional Classification

Examination of the financial factors (TAC and TDC) in a similar graphical form (figure 9.2) highlights the variations that occur for the inputs, illustrating the differences between the DMUs that exists in model 8.1. As with the outputs, it was functional class five with the smallest average value and class one with the largest.

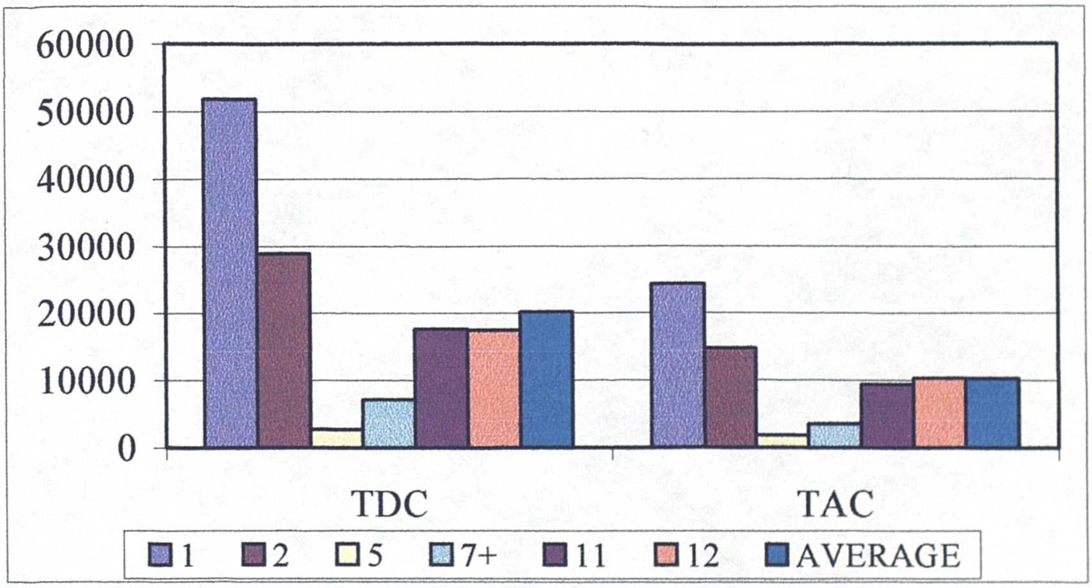


Figure 9.2: Comparison of Input Patterns According to Functional Classification

Evaluation of the results from the unbounded model (model 8.1) and the weight restriction scenarios showed that this might have impacted on the efficiency assessments of a number of DMUs. The efficiency scores from the unbounded model were analysed according to the functional codes identified in table 7.3 above. Although the variation cannot be confirmed by conventional statistical means (as was discussed in chapter five), some variation can be clearly observed in the average efficiency scores, particularly for classes 5 and 11, which appear to be considerably lower than the other classes and the overall sample average.

Functional Codes	No. of DMUs	Average Efficiency Score	Standard Deviation	No. of Efficient DMUs
1	7	96.6	5.53	4
2	10	94.1	8.95	6
5	10	77.7	18.44	2
7+	8	88.9	17.35	5
11	6	77.0	15.38	0
12	6	94.6	7.21	3
Overall	**47**	88.0	15.08	**20**

Table 9.4: Analysis of Efficiency Scores by Functional Classification

In chapter eight, a number of weight restriction scenarios were introduced, each of which was related to a specific aspect of hospital efficiency measurement (see section 8.6). The weight restrictions for scenario three, for example, were applied in such a way that 'total inpatient discharges' and 'day case attendances' were required to be the significant contributors to the efficiency calculations. The other output factors ('A&E attendances' and 'consultant outpatient attendances') were able to contribute to the efficiency calculations, although were not able to dominate completely.

This type of restriction, therefore, was likely to have the most significant impact, in terms of reducing efficiency scores in relation to the unbounded model, on those hospitals with a weakness in either or both of the areas covered by the primary output factors. As can be seen in table 9.5 below, the efficiency scores for each of the functional classes were reduced by the imposition of weight restrictions. However, in this case, functional classes 5 and 7+ were most affected by the restrictions (in terms of the average reduction in their efficiency score).

Functional Class	Number of DMUs	Model 8.1		Scenario Three		Average Change
		Average Efficiency Score	No. of Efficient DMUs	Average Efficiency Score	No. of Efficient DMUs	
1	7	96.6	4	81.2	2	21.49
2	10	94.1	6	85.2	1	8.15
5	10	77.7	2	44.8	0	25.59
7+	8	88.9	5	47.2	0	32.70
11	6	77.0	0	60.7	0	13.78
12	6	94.6	3	74.7	1	19.82
Overall	**47**	**88.0**	**20**	**65.1**	**4**	**22.91**

Table 9.5: Analysis of Scenario Three by Functional Classification

Statistical testing of the data, using Mann-Witney U-tests for example, showed that the observed differences between the DMUs in functional classes 5 and 7+ and those in class 2 were in fact significant for scenario three. (The Mann-Witney tested for changes in the ranking of the DMUs rather than the actual efficiency scores.) Similar significant differences were observed between the DMUs in class 2 and class 5 in the other four scenarios and between the DMUs in classes 2 and 7+ in scenarios one and two.

Additionally, the observable differences between the DMUs in classes 7+ and 11 were found to be significant for all but one of the scenarios. In these cases, applying weight restrictions did not appear to have an equal effect on all DMUs, such that the DMUs in some of the functional classes appeared to be more dramatically affected than those in others.

Looking more closely at these DMUs suggests that there might be underlying reasons for the observable differences, highlighted even more strongly once weight restrictions have been introduced. The DMUs with the lower than average efficiency scores are those belonging generally to the categories of hospital with a very limited surgical function, such that they tend not to concentrate to the same degree as the larger hospitals on inpatient activity. Additionally, these hospitals (excluding the children's hospitals that make up functional class 7) are also based in rural areas, with potentially a different set of operating conditions.

Therefore, the scenarios that enforced a strong contribution from 'total inpatient discharges' and 'day case attendances' were likely to have a strong effect on these hospitals, as they are particularly related to surgical functions and location. In chapter eight, DMU *35 was highlighted as the DMU with the largest reduction in its efficiency score following the inclusion of weight restrictions. In terms of its outputs, it can be said to fit into this category. It is in functional class 7+ (its actual class being 10) and located in a rural part of Scotland, that is, Argyll and Clyde.

The analysis of the results from model 8.1 and the weight restriction scenarios carried out above has suggested that some of the concerns of the health care managers in relation to comparative performance assessment could be justified. The managers of hospitals in classes 5 and 11 would be more likely to reject the results and the DEA technique itself, as they were unfavourable to their hospitals. It would be difficult to determine without further investigation whether these types of hospitals were generally less efficient than the remainder of the sample, as the results have implied. The DEA model applied may not accurately reflect their level of efficiency, due to their size, structure or the particular nature of the services they provided.

In the initial analysis only those hospitals classified by the NHS in Scotland as providing acute services were included, and then the data sample was further refined to remove some of the smaller or more specialised hospitals. However, strong differences in the input and output factors, which can impact on the efficiency calculations, have been identified in the discussion in this chapter. The inclusion of weight restrictions has heightened these differences.

The traditional approach to sample selection (identified in the literature sources in chapter six), whereby hospitals defined as acute were included in the DEA model, has been shown to problematic, particularly in relation to the acute hospitals in Scotland. In order to produce DEA results on efficiency assessment likely to be more acceptable in a practical situation and to address the concerns of hospital managers, some further refinement of the data sample, or an alternative approach to the efficiency analysis, could be required.

This change of approach corresponds with the nature of research into the investigation of university performance (See Beasley (1990) and Wilkinson (1991), *inter alia*). Comparing entire universities was difficult, if not impossible, as the potential range of subjects offered was vast. Similarly with hospitals, where clear differences in 'output' can be observed, an alternative approach was necessary.

For example, in this analysis 'total inpatient discharges' was used as a measure of output. This can cover an extensive range of specialties, which need not be consistent across all acute hospitals. There were in fact thirty-six different acute specialty groups using inpatient care facilities, incorporating the general categories of surgical, medical, obstetrics and gynaecology and dental, as referenced in detail in appendix one. 47 out of the original 74 hospitals used as the data source have inpatient discharges in the general surgical category, whilst just 21 of the sample have inpatient facilities for the specialty of ophthalmology. Appendix one also lists the detailed breakdown of the number of hospitals with patients treated in each of the main inpatient categories to illustrate this point further.

Discussions with health service managers and observations from the literature have suggested several possible extensions or alternatives to counteract the previously mentioned problems and improve the application of DEA in a health care environment. In each case, the sample size considered is relatively small, in comparison with those identified in the literature in chapter six and the analysis presented in chapters seven and eight. Therefore, a discussion on the use of DEA with small samples is also included.

Changes to the sample are also likely to lead to corresponding changes to the variable set and other aspects of the modelling strategy, also considered in the following sections.

9.3.1 Small Sample Analysis using DEA

The DEA literature on health care applications has tended to concentrate on using DEA to analyse fairly large samples of hospitals or health care institutions. This includes many of the key papers mentioned in chapter six, such as Ehreth (1994), Ozcan, Luke and Haksever (1992) and Grosskopf and Valdmanis (1987). However, Sherman (1984), also discussed in chapter six, used a sample of just seven teaching hospitals within Massachusetts to investigate the DEA modelling process. The hospitals in the sample were selected to conform to some definition of homogeneity, using comparable groups defined by the state rate-setting commission.

Other approaches to the selection of homogeneous data samples have been found in the literature, including a variety of clustering techniques, based on location, hospital size, teaching status and categories of patients treated. Ehreth (1994), in particular, used a number of these approaches, as the hospitals in the sample were grouped according to 'size, geographic region, proximity to an urban centre, and teaching status'.

However, in relation to hospital services in Scotland, it is probable that the resulting 'clusters' could be found to have a very small number of DMUs contained within them. This has been observed to be problematic in the DEA methodology and has attracted a great deal of attention in the DEA literature (see chapter five).

It was suggested that if a large number of variables are included in the DEA model, in relation to the number of DMUs in the data sample, the majority of DMUs could be classified as efficient. For example, it was observed in Sherman (1984), as discussed in chapter six, that five of the seven hospitals included in the data set were found to be efficient.

To counteract such a problem, a number of suggestions have been put forward as to the minimum number of DMUs required in relation to the number of variables, such as the rule of thumb proposed by Banker *et al.* (1989). They suggested that in order to have a satisfactory level of discrimination between efficient and inefficient DMUs, the number of DMUs in any sample should equal the number of variables, multiplied by three. Therefore, unless just two or three variables were included in the model, the sample size would need to be considerably greater than ten.

However, including such a small number of variables negates one of the major benefits of the DEA approach to efficiency assessment in comparison with the more traditional approach such as performance indicators, in that it allows for multiple inputs and outputs to be directly incorporated. Therefore, particularly in relation to the health care literature, the problem has been overcome through the use of fairly large samples of DMUs. In relation to small samples of hospitals or health care organisations, other approaches for overcoming this difficulty have not been addressed in any great detail in the DEA literature.

However, as observed by Ball *et al.* (1997) in their investigation into local authority elderly care services in Scotland, the lack of discrimination often observed in a small sample application of DEA can be overcome by the introduction of weight restrictions. In the sample of twelve local authorities, there was just one DMU classified as inefficient in the unweighted model. Following the introduction of carefully selected weight restrictions, the number of inefficient DMUs increased to six, half of the sample.

As was shown in chapter eight, introducing weight restrictions in the sample of 47 DMUs had the effect of greatly reducing the number of efficient DMUs in each of the five scenarios. Therefore, whilst there are limitations to the DEA models with such a small number of DMUs, these can be overcome to a certain degree through the introduction of weight restrictions, as was highlighted in chapter eight. In the following sections, different approaches to generating a more homogeneous sample of DMUs are considered. The first section, on the evaluation of hospitals as a whole, is looked at in the greatest detail, as this has been the focus of the preceding analysis.

9.3.2 Small Samples of Hospitals

Selecting a sample of hospitals that would be generally accepted as 'homogeneous' within the Scottish data sample is made difficult due to the relatively small number of hospitals in any given cluster. This is especially apparent given the number of different hospital types and sizes, the diverse geography and the range of population demographics in the various regions.

Therefore, it would be virtually impossible to use the approach of Ehreth (1994) discussed above, as there would be only one or two hospitals in each category if such a comprehensive clustering method were utilised. This would involve hospital evaluation using much smaller numbers of DMUs in the samples, which must be similar in activity level, size or location. In the case of Scottish hospitals, this would be through restricting the samples of DMUs to those with the same functional classification, located in the same geographical region or of an equivalent size, for example. The potential for each of these is now addressed.

9.3.2.1 Geographical Region

Small samples of hospitals could be collated according to geographical region in order to assess the impact of location, for example, or population demographics. For example, the hospitals in the Highlands and Islands Region were likely to be subject to a different range of environmental and managerial concerns than the urban hospitals in the Greater Glasgow region. This will be most notable in terms of patient transportation and length of stay for inpatient care, due to the greater travel distances involved.

The use of geographical region with this data set, relating to acute hospitals in Scotland, does not appear to be a realistic approach, as there are so few acute hospitals in any given region that could be used for a comparative analysis. For example, the largest number of acute hospitals in any of the fifteen regions is thirteen, for the Grampian region that includes the city of Aberdeen. Over half of these hospitals, however, were excluded from the second stage of the analysis in chapter eight (see appendix three).

In fact, the Grampian Region is the only one of the fifteen health board regions in Scotland with more than 10 DMUs located within its boundaries. Table 9.6 below lists the number of acute hospitals within each geographical region in Scotland.

Health Board	No. of DMUs	Health Board	No. of DMUs
Ayrshire and Arran	4	Orkney	1
Borders	1	Lothian	8
Argyll and Clyde	9	Tayside	8
Fife	5	Forth Valley	2
Greater Glasgow	7	Western Isles	2
Highland	6	Dumfries and Galloway	2
Lanarkshire	4	Shetland	1
Grampian	13		

Table 9.6: Number of Acute Hospitals in Each Health Board Region

As the table shows, several of the health board regions have just one or two acute hospitals located within them, particularly the island regions of Orkney and Shetland, and would therefore not provide appropriate samples for a DEA analysis. It would be possible to group the regions, by population density or some other factor, but the same problems as identified for the Grampian region above would still be in evidence, in that the assortment of hospitals in each grouping would still be very varied.

In relation to acute hospitals in Scotland, using geographical region to generate smaller samples of hospitals does not seem to provide a more useful approach for the evaluation of hospital efficiency. However, the approach could be used for an alternative sample of DMUs, where appropriate. Using the university analogy, this approach would be equivalent to evaluating universities in the same geographical region or urban centre.

9.3.2.2 Hospital Size

Whilst hospitals may not have the same functional classification or be in a different geographical region, they may well be of a similar size, in terms of the number of patients treated in a given category, the average number of staffed beds available or the total expenditure.

However, focusing primarily on 'size' does not negate the requirement for overall homogeneity. This could be overcome in the sample of Scottish hospitals, by restricting the hospitals included to only those classified as acute. Additionally, many of the similar sized hospitals fall into the same broad categories of hospital type, as were described in chapter seven, combining several of the functional classes together. For example, the hospitals in functional classes 1 and 2 could all be categorised as 'large', as the total expenditure of all the hospitals (excluding one) is in excess of £20million.

However, using hospital size as a defining factor in itself is still problematic, in that the factors used to divide the overall sample would need to be defined, according to some predetermined level or arbitrarily. Also, hospitals with a similar level of expenditure or number of beds could still be from totally different functional classes, with different strategies or patterns of service provision.

For example, the hospitals in functional classes 2 and 11 are of a similar size in terms of their total expenditure, although functional class 2 includes teaching hospitals covering a wide range of services, whereas hospitals in class 11 do not have teaching status and have no special units.

Certainly, dividing up the acute hospitals by size would produce a number of smaller samples, with a similar level of services in terms of volume. However, as was the case with using geographical region, there are still a number of problems associated with this approach, in terms of obtaining a sample with a homogeneous group of hospitals acceptable to the health care professional.

Referring to the university example, using hospital size would compare with examining universities according to the number of students enrolled, the size of the academic budget or the annual budget.

9.3.2.3 Functional Classification

By focusing on the type of hospital, rather than size or geographical considerations, the hospitals within Scotland can be grouped according to functional classification, thereby providing a small sample of hospitals that carry out virtually the same functions and processes. Also, hospitals in the same functional class tend to have a similar level of overall expenditure and are of a similar size. In many cases, they also appear to be located in equivalent locations in Scotland. For example, all the hospitals in functional class one are located in the main population centres in Scotland.

The sample used originally in chapters seven and eight contained all those hospitals classified as providing acute services, and covered some fifteen functional classifications. Each functional classification represents a slightly different type of hospital, in terms of the services they offer. For example, those in functional class one are large general teaching hospitals.

Hospitals in classes eight, nine and ten are all general practitioner cottage hospitals, with a very limited amount of surgery undertaken, if any at all. Appendix one contains the complete definitions for each of the functional classes.

A continued difficulty still occurs with this approach in relation to the sample of Scottish hospitals, in that for some of the functional classes, there are only a handful of hospitals falling into that category, as was observed if the hospitals were divided up by health board region. Table 9.7 below lists the number of hospitals within each functional classification utilised, for the 74 acute hospitals in Scotland. (Note: There are three classes of acute hospital into which no hospitals have been categorised at the present time.)

Functional Classification	Number of Hospitals	Functional Classification	Number of Hospitals
01	7	09	4
02	13	10	12
04	2	11	7
05	10	12	8
07	3	13	3
08	4	14	1

Table 9.7: Number of Hospitals in Each Functional Classification

As can be observed from the table, only three of the functional classes have ten or more hospitals falling into that category, that is 02, 05 and 10 and six classes have fewer than seven hospitals (the sample size used by Sherman, 1984). It would be possible to investigate the efficiency of those functional classes with a large enough sample of DMUs, such as the six classes with seven or more DMUs.

For the remaining categories, it would be necessary to determine which, if any, of the functional classifications could be combined. For example, the hospitals in classes 08, 09 and 10 are all cottage hospitals, which it might be possible to combine. Input from the relevant health care managers would be essential at this stage, in order to determine that the combined categories were reasonable.

This approach to comparative performance assessment was favoured by one of the managers interviewed at the Blackburn Hyndburn and Ribble Valley (BHRV) NHS Trust. However, it would be necessary to demonstrate that the samples could be seen to contain hospitals of a similar nature, with equivalent operational structures. BHRV, for example, has hospitals on a number of sites and also includes a mental health unit under its control. Therefore, all the trusts to which it was to be compared would be required to conform to the same structure in order that the comparisons would be totally acceptable to the manager concerned.

In relation to the sample of acute hospitals in Scotland, the functional classification would seem the most appropriate measure for further dividing up the sample. However, as has been observed above, this would not be entirely straightforward, particularly as some of the classes contain such a small number of hospitals.

This would correspond to comparing small groups of universities, each offering a similar range of subjects to a similar group of students in a comparable environment. It would probably be very difficult to determine such groupings in the higher education field, whereas in the health sector, functional classification can be used.

9.3.3 Comparisons by Patient Type (Inpatients, Outpatients etc.)

In the analyses presented to date, the data samples have been focused on measuring the efficiency of each hospital as a whole. However, it would also be possible to investigate the efficiency with which particular types of patient care are carried out, such as inpatient attendances, day case treatments and outpatient activity. This has frequently been the case with performance indicators and more acceptable to the health managers with whom the issue was discussed. A degree of disaggregation of the output variables was possible without needing to incorporate a large number of factors into the model. For example, the day case figures could be broken down into surgical, medical, obstetrics and gynaecology and other, as was suggested in the whole hospital examples in chapter seven for the inpatient data.

The sample size for each level of the analysis would be variable. For example, of the seventy-four acute hospitals, fifty-nine of them provide day case treatments and fifty-four have a surgical inpatient department. As was the case with the whole-hospital analysis outlined previously, there would be some considerable variance again in patterns of activity and refining the sample would almost certainly be appropriate. As some of the sample sizes may be quite small, some modifications to the traditional DEA methodology may be required as discussed in section 9.3.1.

Returning to the university analogy, this is equivalent to investigating undergraduate courses or research activity separately. This would be a valid approach, as not all universities have a large number of research students - comparing just undergraduate programmes would provide a more homogeneous sample.

9.3.4 Comparisons by Patient Category (Ophthalmology, Medical, Gynaecology)

This would be an alternative approach to the previous section - instead of dividing up the hospital according the different types of treatment being offered, comparisons are made according to the category of patient, the medical specialty or the department providing the treatment. Therefore, all types of ophthalmology service could be evaluated across all the hospitals that have an ophthalmology department. Again problems may arise in that not all hospitals offered all types of service delivery, but this approach would present yet another perspective on the efficiency assessment of health services. It would be equivalent to using DEA to investigate the efficiency of undergraduate and taught postgraduate courses, postgraduate research and staff research activity in physics departments at all universities in the sample.

9.3.5 Comparisons by Patient Category and Type (Gynaecology Inpatients, Ophthalmology Day Cases)

This final type of analysis, presented as an alternative to whole-hospital comparisons, would be to investigate activity on a small scale, the DEA equivalent of the medical audit approach discussed in chapter four. The major difficulty in this type of approach is that the amount of output data available is generally quite limited. Specific performance indicators could be introduced, such as the quality indicators published by the government annually. For example, in the investigation of A&E activity, the indicator 'time to immediate assessment' (highlighted in chapter four) could be used as a quality factor in a DEA model. The inclusion of quality variables will be discussed more thoroughly in the next section, where a number of amendments to the variable set are considered.

Also, there would be scope to further disaggregate the input factors, as the cost factors can be broken down into several categories, such as medical staffing costs and supply expenses. The sample used to illustrate the application procedure for DEA in chapter five, containing data on inpatient services in the specialty of gynaecology, was an example of this type of approach.

A major advantage of using such an approach would be that it allowed for very specific models to be developed, focusing on particular policy issues and managerial concerns, by the selection of additional variables and the adoption of a variety of weight restriction scenarios. Taking the specialty of ophthalmology as an example, the success or otherwise of government policies to increase the number of patients treated as day cases rather than through inpatient treatment could be investigated. In relation to the university example, this corresponds to evaluating undergraduate history courses or postgraduate research degrees in psychology.

The alternative methods of investigating hospital data, in order to analyse efficiency, as described above, could be used to extend the scope of the DEA methodology. They would also make it possible for the efficiency investigations to be tailored to specific scenarios (such as the investigation of day case attendances as highlighted above and in chapter three). However, it was noted that there were potential problems with applying Data Envelopment Analysis to these types of samples, particularly in terms of differing sample sizes and the varied nature of the acute hospitals in Scotland. In other health care systems, the potential difficulties may be less apparent, making the application of the DEA methodology more straightforward.

9.4 Selection of Variables

In the first stage of the DEA application process, considerable importance was attached in the model-building process to the selection of the right variables. Statistical methods, data analysis, judgement and trial and error were all employed to identify the variable set employed in the final model (see sections 7.5 and 7.6). Also, as is emphasised throughout this chapter, the people who manage hospitals and other health care institutions have perceived them to be unique or individual in their operating practices and environmental influences. Therefore, the combination of methods used was important - it was not enough to select a particular variable on the basis that it had been used in other cases in the literature.

The process of variable selection (stage one) resulted in the rejection of certain potential variables, such as the factors measuring inpatient activity in terms of days rather than discharges. Statistical analysis was useful in determining some of the most appropriate combinations of outputs by identifying significant relationships within the data. In the final analysis, there were still several combinations of variables that could have been used to determine the efficiency of the DMUs in the sample of acute hospitals.

In the use of DEA, reservations have been expressed that the results are very heavily dependent on the choice of variables and in particular the number of variables included. This was covered in chapter five quite extensively. Therefore, the fact that slightly different DEA models can produce different efficiency scores and efficiency profiles for each DMU is a major concern.

In the analysis in section 7.6, it was seen that different combinations of variables produced slightly different results from the DEA model, although this was with the original sample of seventy-four DMUs. For example, the average efficiency score for the ten models varied between 82.6% (7.9) and 89.6% (7.7), a difference of 8.2%. The number of efficient units ranges from 24 (7.9) to 39 (7.10), which was 30% - 52% of the total number of DMUs. Each of these models could have been further developed, by including weight restrictions and producing efficiency strategies for the inefficient DMUs.

However, an additional model was produced, which had a disaggregated cost factor and an aggregated inpatient discharges variable, to be used in the second stage of the analysis. It was stressed that the variables included in the second stage of the analysis (defined in table 7.20 in section 7.6.5) were chosen to investigate a particular aspect of hospital efficiency, that is, on a very general level, looking for potential reductions in expenditure. Changing the objective of the analysis would have resulted in changes to the variable set and also to the DEA model applied, particularly if non-discretionary environmental variables were required. Information from the literature sources and a number of 'experts' was used to guide the process of variable selection. Potential changes or additions to this variable set are considered in the following sections.

9.4.1 The Inclusion of a Case-Mix Index for Inpatient Activity

In the evaluation of the DEA papers in chapter six, it was observed that a number of the evaluations of hospital efficiency had made use of a case-mix index in their measure of inpatient activity, including Ozcan *et al.* (1992).

However, Parkin and Hollingsworth (1996) did not employ a similar index in their evaluation of acute hospitals in Scotland. Additionally, Byrnes and Valdmanis (1994) restricted their sample of DMUs to focus on teaching hospitals in order to avoid the need for a case-mix indicator.

The models developed in chapter eight used an aggregated statistic for inpatient activity ('total inpatient days'). It would therefore have been useful to employ a case-mix index or its equivalent, if one could have been identified. As was discussed in chapter seven, such a measure was not available.

Employing further restrictions to the sample of DMUs, such as examining acute hospitals by functional classification (discussed in section 9.3 above), would reduce the need for a case-mix index. Another option would be to calculate a weighted sum of the inpatient discharge categories, if potential weights could be defined (using average cost per case, for example). If the inpatient figures were of primary importance in the analysis, it could be more appropriate to use them in a disaggregated form ('surgical inpatient discharges', 'medical inpatient discharges' and so forth).

In relation to a number of the other output factors employed, such as 'day case attendances' or 'consultant outpatient attendances', case-mix indices or a weighted sum of each of the relevant categories might also be appropriate in particular circumstances.

9.4.2 Alternative Measures of Input

As was the case with the outputs, particularly the inpatient statistic, the level of aggregation for the cost factors was clearly significant. In the analysis presented in chapters seven and eight, costs were divided into 'direct' (measuring the actual cost of patient care provided by doctors, nurses *inter alia*) and 'allocated' (measuring bureaucracy). The specific elements included in each of these factors were defined in chapter seven. Having identified potential savings in each of these general areas, as was the case with the efficiency strategies developed in chapter eight, additional models using disaggregated cost factors could be defined. These could be used to pinpoint exactly where the inefficiencies exist, in relation to various staffing categories, for example. In this case, the changes to the variable set would correspond to a specific objective for the efficiency evaluation. This is further discussed in section 9.5, which focuses on the information produced by the DEA modelling process.

9.4.3 The Influence of the Environment

Following the discussions with health care managers, it has been observed that they tend to be wary of the comparative performance techniques frequently used in the health service (as was discussed in chapters four and six). One of the most commonly stated concerns was that these techniques often do not take into account the individual nature of the hospitals being investigated. Almost as important, however, was that most of the techniques have not taken into account the environmental concerns, under which each individual DMU is operating. In the original analysis, presented in chapters seven and eight, the sample was restricted to acute hospitals, but no additional measures were included to reflect specific environmental concerns.

In the analysis, differences in efficiency evaluations were observed, particularly in relation to hospital size or functional classification. In section 9.3, it was proposed that changes to the sample of DMUs could ensure that environmental differences (be they physical or organisational) did not impact upon the efficiency assessments for the DMUs.

This could also be addressed through the inclusion of environmental variables. Therefore, factors to reflect population density and many other types of demographic information, such as ethnicity, age structure or the prevalence of certain illnesses, could be included. These could be particularly relevant if regional health service activity was being investigated or definitive environmental concerns existed, such as the known prevalence of certain categories of illness. Environmental factors could also reflect the internal environment of each DMU, for example, the amount of space available within the hospital or the age and number of hospital buildings. (Including such factors, if they were determined to be externally controlled, would require an amendment to the DEA model, defined in section 5.8.3.)

The major difficulty, however, would be in determining appropriate indicators. For example, would the population density of the area in which a hospital was situated be relevant, particularly as a great number of the patients being treated there usually come from outside the locality?

The opinion of the health care managers was that environment was significant, but the difficulties in measuring for its influence were also recognised.

It was also suggested that it can become too easy to use 'environmental factors' as an excuse for inefficiency, blaming sub-standard performance on location, the poor health of the local population in general or the Victorian hospital buildings, without looking for ways to improve efficiency despite these environmental influences.

As was the case with the discussion in the previous sections, environmental variables could be included if they were deemed to be appropriate to the particular situation being investigated. However, it would be important to apply them carefully so that they could not dominate the efficiency assessments – weight restriction scenarios could be employed to restrict their influence (see Ball and Roberts, 1998, for an example of this approach).

9.4.4 The Importance of Measuring Service Quality

In addition to including environmental factors, as debated in section 9.4.4 in relation to the application of DEA to hospital data, the usefulness of the DEA methodology could be extended with the inclusion of measures of service quality. This could be related to the discussion in section 2.5.3 on the increased emphasis given by the current Labour government to the necessity of improving the quality with which services are administered, not just the technical efficiency. The emphasis is on the viewpoint that efficiency and quality should go 'hand in hand' (Department of Health, 1997).

To date, all of the DEA models presented have used only the volume of activity in each treatment/patient category as the measures of output.

Such an approach has frequently been observed in the literature, primarily due to the fact that including quality factors was not regarded as a straightforward matter. There were several reasons for this; the main ones being a lack of appropriate and useful data and the need to adapt the DEA model and the data itself before such information could be included. For example, it was important that the data was of the correct dimension - many potential quality factors would need to be scaled before they could be included in the DEA model to ensure an increase in the quality variable related to an increase in efficiency. The same would be true for the environmental factors discussed previously.

Some attempts have been made to introduce quality variables into health care investigations, as observed in the DEA literature. For example, both Kleinsorge and Karney (1992) and Thanassoulis *et al.* (1995) incorporated measures of quality, although neither of these papers focused on the measurement of hospital efficiency.

In chapter six, where the investigation concentrated on efficiency at hospital level, it would have been difficult to determine how to measure quality of service, particularly if a single variable relating to the whole hospital was required. To measure 'quality of care' at such a broad level would have been extremely difficult. Many of the variables that could possibly be considered, such as 'number of complaints received' are not reported at individual hospital level in the NHS in Scotland's annual report, but at the next level up, that is, at NHS Trust level (The Scottish Office, 1996). Additionally, they also tend to relate to non-medical aspects of quality.

However, in the other potential examples of DEA application discussed in section 9.3, alternative types of DMUs have been identified, making the introduction of quality factors much more applicable. Quality measures would be particularly appropriate as a means for introducing policy perspectives into the DEA methodology. For example, where the models relate to single specialties such as Ophthalmology the quality factors could be obtained from the various performance indicators published by the Scottish Office annually. These correspond to the data published by the NHS in England in the form of league tables. The Director of Corporate Development regarded them as a good reflection on the quality of services provided at the Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust.

Therefore, it would be possible to develop models to incorporate waiting times, waiting lists, cancelled operations, re-admission rates and various other treatment targets, as appropriate, if alternative data samples were investigated. Again, the inclusion of quality variables would be dependent upon the objectives of the efficiency evaluation.

9.5 Providing Useful Information

One of the key issues identified in chapter six, in relation to the use of new approaches for performance assessment from the perspective of the health care manager, was that they should provide additional information to that already available through techniques such as performance indicators. It was important that the new methodologies provided better and more useful information, rather than just something different.

This is addressed in this section, which considers the nature of the DEA methodology, the output from the DEA models and a comparison with the methods currently employed for performance assessment, particularly performance indicators.

9.5.1 DEA and the Measurement of Technical Efficiency

The specific measurement of technical efficiency can be said to fulfil the criteria for new and additional information, as the methods currently applied for performance assessment have not directly addressed this. The efficiency index, discussed in chapter four, was seen to be a very crude statistic, adding little of value to the debate on health service efficiency.

In chapter five, the use of DEA as a measurement of other types of efficiency, such as allocative efficiency, was investigated and examples were given from the literature to support the possibilities for this. It was also stated, initially in chapter two and then repeated in chapter five, that the focus of this study would be on the measurement of technical efficiency, that is, the physical use of resource inputs in the relation to the production of physical outputs, as defined in section 2.2.1. An investigation into technical efficiency is therefore intended to identify either the over-use of specific resource inputs or the under-production of specific outputs, not simply identify relative or actual efficiency scores.

Using the results from model 8.1 (investigated in the second stage of the application process in section 8.2) and the weight restriction scenarios, the capabilities of DEA in relation to identifying elements of technical inefficiency can be demonstrated.

In the sample, forty-seven hospitals were investigated, with the average efficiency score being 87.98% and 20 of the DMUs rated as 100% relatively efficient. For each of the inefficient DMUs, the nature of the inefficiency was classified in terms of the changes that must be made to the inputs to bring about efficiency. The models were orientated to input minimisation and thus were concerned with the reduction of inputs, as opposed to increases in outputs, which would be the case with an output maximisation orientation. For example, as was highlighted in section 8.7.2.2, the DEA model showed that improvements could be made to the efficiency of DMU *4 by reducing the 'total direct costs' factor, identifying an over-use of resources.

The ability of the DEA model to identify inefficiencies in the use of resources was enhanced by the application of weight restrictions. In relation to DMU *4, the 'total direct costs' factor in the unbounded model was some 15% less for its reference set (see table 8.20) and similarly for the 'allocated costs' input. These were the areas where improvements in efficiency could be made, with the targets set according to the reference set's input levels.

In the weight restriction scenarios, the inefficiencies identified were considerably larger, as the efficiency of DMU *4 was reduced to as low as 70.7%. Possible reductions in its 'total direct costs' of up to 35% were determined. Extending the modelling process, by including additional cost factors in the inputs, would extend the ability of the DEA model to provide information on the technical inefficiencies within each hospital.

In relation to the definitions given above, as the DEA models have identified an over-use of resources for this DMU, the hospital's technical efficiency was clearly investigated, as required. The use of weight restrictions, in the development of efficiency strategies for each of the inefficient DMUs, made this investigation all the more comprehensive.

Applying the DEA methodology to the sample of acute hospitals in Scotland has produced information on their relative technical efficiency, which was previously unavailable in such a concise format, thus satisfying one of the requirements of the health care managers.

9.5.2 Relating DEA to Other Techniques

In the previous section, the ability of the DEA model to provide information on the technical efficiency of each DMU was confirmed. Its position, in relation to other performance assessment methodologies, is now addressed.

In chapter four, performance indicators were discussed and their limitations clearly presented. However, they have been used considerably within the NHS at all levels of the organisational structure, particularly those relating cost to output activity. Most people working within the NHS in a managerial capacity are familiar with performance indicators and will have used them to some degree, as discussed by Jenkins *et al.* (1987). If DEA could provide additional insights to those obtained from the performance indicators, the benefit of the technique would be more clearly observed.

For example, the indicators 'cost per patient discharge' and 'cost per patient day' are frequently used to measure inpatient activity. Relatively high values are seen to represent a degree of inefficiency. Table 9.8 below summarises the key figures for these indicators using the data from the gynaecology inpatient information used in chapter five. Reference numbers are given for the DMUs with the maximum and minimum values in each case. (Reference numbers are distinct from those used in chapters seven and eight and apply only to the gynaecology data set, referenced in appendix two).

	Mean (£)	Standard Deviation (£)	Minimum (£)	Maximum (£)
Cost Per Patient Discharge	960	183	620 (#19)	1480 (#7)
Cost Per Patient Day	306	61	190 (#19)	410 (#7)

Table 9.8: Performance Indicators on Costs for Gynaecology Inpatient Data

The data is represented graphically in figures 9.3 and 9.4 to follow, showing the values for the two indicators in turn for each of the 23 DMUs. The DMUs with the highest and lowest values have been marked on both of the figures. It can be seen that a high/low value for one of these indicators does not necessarily indicate a high/low value for the other, although this is the case for DMUs #7 and #18.

Also included in each chart are two lines to represent the points that are two standard deviations from the mean on either side. This is often used as the level to test for extreme performance, in the application of techniques such as performance indicators and regression. As can be observed, only in the figure showing cost per patient discharges does a DMU show a value outside this range.

However, had the line been drawn at one standard deviation above the mean, used by the State of Massachusetts to signify 'inefficient' practice (Sherman, 1984), several other DMUs would have been highlighted for their inefficiency.

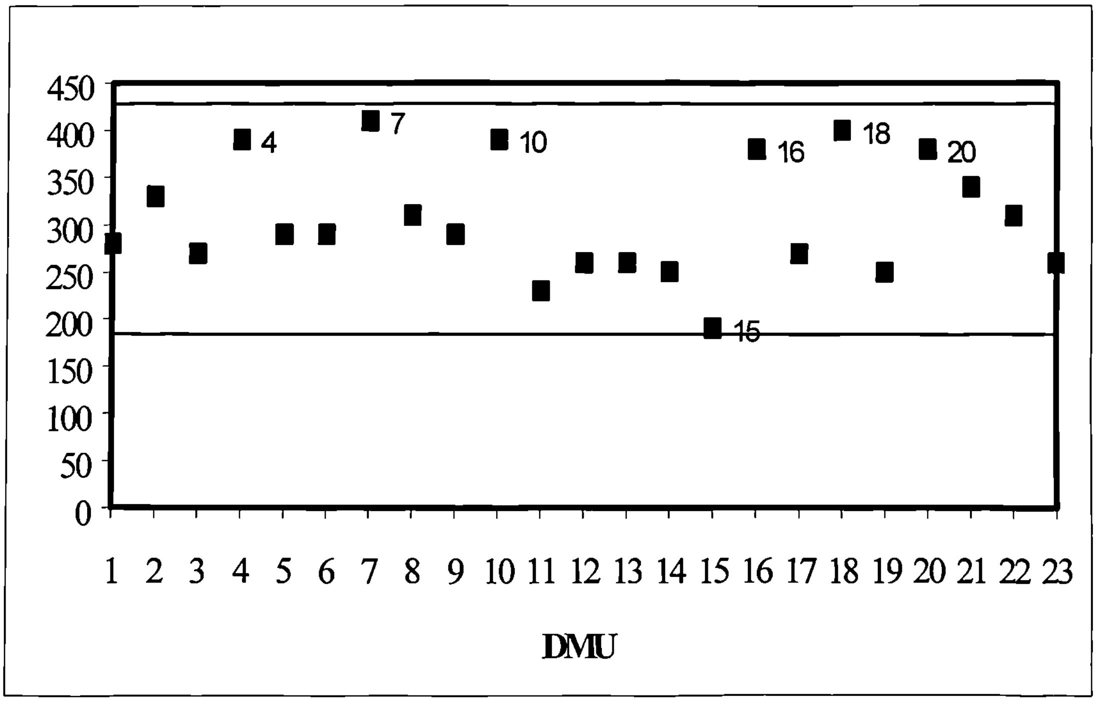


Figure 9.3: Cost per Patient Day (Gynaecology Inpatients)

Looking at figure 9.3, DMUs #4, #7, #10 and #18, amongst others, would be described as being 'worthy of further investigation' because they signify an unusually high cost per patient day for gynaecology inpatients at these hospitals. However, none of the values for these DMU lie beyond the boundary at two standard deviations above the mean. In the equivalent figure for 'cost per patient discharge' shown below, the value for DMU #7 lies outside the boundary, suggesting poor performance.

The nature of the investigation may depend upon the motivation of the investigator, although looking at patient profile, hospital type or environmental considerations are all frequently executed elements of an investigation of this type.

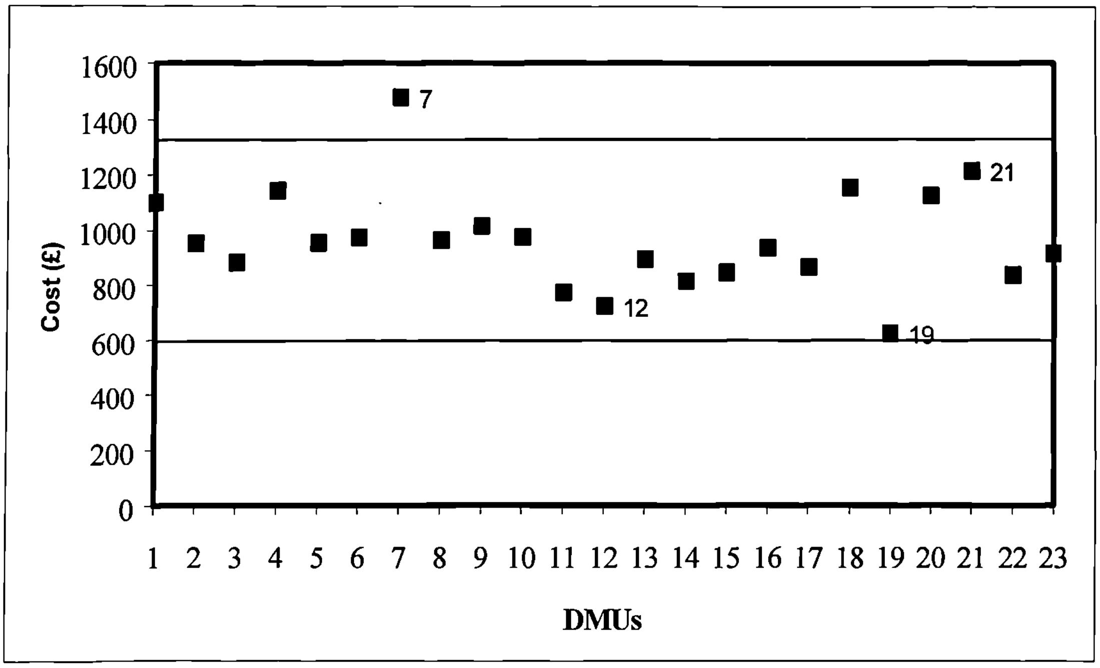


Figure 9.4: Cost per Patient Discharge (Gynaecology Inpatients)

A major limitation of this type of performance indicator evaluation (see chapter four) has been the lack of information that generally accompanies the type of chart presented in figures 9.3 and 9.4. The NHS or Department of Health produces these in report form for all trusts and hospitals. As the Financial Director of Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust observed, it would be of little value to know that the 'cost per patient day' indicator was unusually high in comparison to fifteen other hospitals, if absolutely nothing was known about those other hospitals. This would be the case if figure 9.4 were presented to the director of DMU #7, for example, which has a very high value for this indicator.

In the simplified DEA example, with just a small number of variables included, there is little difference between the results obtained from the two methods. DMU #15, which had a low value for 'cost per patient day', was rated as 100% efficient in each of the DEA models using 'inpatient days' as an output in the analysis in chapter five.

Similarly, DMU #7 was rated as the least efficient of all the DMUs, with an efficiency score of just 41.7% in the DEA model with two cost inputs and 'inpatient discharges' as the single output. In both figures 9.3 and 9.4, the indicator for DMU #7 was one of the highest.

However, as more variables were included and the some of the additional elements of the DEA methodology discussed in chapter five were incorporated, the obvious limitations of the performance indicator approach has been matched by the possibilities provided by DEA. Table 9.9 below shows the 'cost per case' values for each of the four output factors used in model 8.1 and the subsequent weighting scenarios for each of the DMUs in functional class 01. The table also shows the efficiency score obtained by each DMU in the unbounded model.

DMU	**Cost Per Case (£)**				**Efficiency Score (%)**
	TID	**DCA**	**COA**	**AEA**	
***1**	1320	262	44	31	95.2
***2**	1167	255	69	35	100
***3**	1664	459	52	67	100
***4**	1700	300	51	39	85
***5**	1325	417	62	68	95.7
***6**	1839	615	62	28	100
***7**	1589	466	49	24	100
Class Average	**1450**	**396**	**56**	**41**	**96.6**

Key: TID: 'total inpatient discharges' DCA: 'day case attendances

COA: 'consultant outpatient attendances' AEA: 'A&E attendances'

Table 9.9: Cost per Case Indicators Compared with DEA Efficiency Scores

As the table shows, three of the DMUs classified as efficient had above average values for the two primary factors (as defined in chapter eight), that is, 'total inpatient discharges' and 'day case attendances'. DMU *2 (below average for all but 'consultant outpatient discharges') and DMU *3 (above average for all but 'consultant outpatient discharges') were also efficient in each of the weight restriction scenarios, despite their very different profiles for the 'cost per case' indicators. DMU *1 was rated as 95.2% efficient in comparison, but its value for 'cost per case' for each of the output factors was significantly less than those obtained for DMU *3. Significant improvements in efficiency would be identified by the DEA model for DMU *1 in particular, which may have been overlooked if the performance indicators for cost per case had been used in isolation.

Further potential benefits for the DEA modelling process are discussed in the next section, looking at the output produced by the DEA models.

9.5.3 The Relevance of the Output from DEA Models

As was referred to previously in the discussion on performance indicators, many other examples of performance assessment techniques were limited in the type of output they produced and the subsequent usefulness of this output. In relation to performance indicators, difficult-to-interpret charts such as those presented in section 9.5.2 have not been seen to provide useful insights. Additionally, the simple allocation of an efficiency score or ranking does not really convey much useful information, that is, the observation has frequently been made that DMU m is recorded as the n^{th} most efficient in the sample.

Discussions at the DEA Symposium held in Wernigerode, Germany in 1998 (and in a subsequent conversation with Dr. A. Boussofiane of the University of Hertfordshire) focused on the issue of ranking. It was suggested that attempts to improve efficiency would not be greatly assisted by an awareness of a ranking in a list of other inefficient DMUs. It was felt that it was more important to relate each inefficient DMU to its peer group and the potential targets for improvement, than to the sample as a whole.

This confirmed the requirements specified by the Financial Director at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust, amongst others. Further to this, it was the view of the majority of health care professionals questioned that if comparative performance assessments were to be useful, they should provide detailed information on how their hospital can be compared with those to which it is most similar. Also, it was felt that the areas where the hospital was deficient, or particularly efficient, should be easily identifiable.

DMU *1, for example, was identified in the previous section as having below average values for each of the 'cost per case' indicators. Its efficiency assessment using model 8.1 (defined in section 8.2) was found to be 95.2% and it ranked 23rd out of 47 in the list of acute hospitals (see appendix 4C). This information would be regarded as meaningless by the management of DMU *4, particularly if the identities of the other hospitals in the sample were unknown and they had no opportunity to identify them. However, table 9.10 overleaf shows a detailed analysis for DMU *1 in the unbounded model, identifying the factors to be addressed in order to improve efficiency, focusing on input minimisation at this stage.

Inputs	Weight	Virtual Weight	DMU's Input	Reference Group's Input
CAP	0.001063	7.8	73	69.51931
TDC	0.0000155	92.2	59679	56833.47
TAC	0	0	32296	26730.54

KEY: CAP: 'capital charge' TDC: 'total direct costs' TAC: 'total allocated costs'

*Table 9.10: Evaluation of Input Slacks for DMU *1*

The dominant input for DMU *1 was the 'total direct costs' factor, with a virtual weight percentage of 92.2%. The figures for the reference group's inputs have demonstrated that an equivalent amount of output could have been produced with significantly less cost input in particular. Taking into account the scale of the cost figures, a reduction of over £8million in the 'total expenditure' would have corresponded to the DMUs on the best practice frontier. The target figures identified above for DMU *1 were determined from its reference set or peer group. In this case, three DMUs formed the reference set for DMU *1, these being *2, *11 and *14, with *11 the strongest contributor. Table 9.11 below shows the inputs and outputs for the reference set DMUs along side those for DMU *1.

FACTORS	DMU			
	*1	*2	*11	*14
CAP	73	21.69	41.02	61
TDC	59679	71314	28820	29597
TAC	32296	33121	15156	12026
COA	226057	257572	111314	134960
AEA	78419	62177	52902	73938
TID	52174	62654	27295	27460
DCA	13610	22091	8936	8831

Key: See Table 9.2 for a Definition of the Factor Codes

*Table 9.11: The Factor Values for DMU *1 and its Reference Set*

By identifying the reference set, DMU *1 has identified a set of other similar DMUs which were all operating relatively efficiently in comparison with the other DMUs in the data set. As was suggested by the Director of Corporate Development at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust, the identification of a small group of comparable DMUs was considered a desirable aspect of any performance assessment methodology. Other inefficient DMUs that have the same reference set could also be used to provide insights for DMU *1 as to how its efficiency could be improved. In this sense, the DEA modelling process would be considered to be a starting point for the improvement of efficiency. The next stage would be the sharing of ideas amongst the peer group of comparable DMUs.

The search for improved efficiency was enhanced in the modelling process through the imposition of weight restrictions. For the sample of DMUs used in the second stage of the analysis in chapter eight, a core group of efficient DMUs was identified, each of which remained efficient under each of the weight restriction scenarios. In chapter eight, the impact of weight restrictions was examined in detail for DMU *4 and an efficiency strategy was developed, showing that it should be possible to reduce its overall expenditure by over £30 million. The efficiency strategy was produced in graphical form, showing specific reductions in the two cost factors. The DMUs that were included in the peer group for DMU *4 were also emphasised. A similar strategy could be produced for each of the inefficient DMUs, including DMU *1, which was discussed above. For example, in scenario one, its efficiency score was reduced to 84.1%, and a reduction in overall expenditure of £17 million was proposed in order to achieve efficiency (compared with £8 million in the unbounded model).

As was suggested above, the value of the DEA model could be enhanced by further disaggregation of the cost factors, in order to highlight where the specific inefficiencies were located. Additionally, as a core group of efficient DMUs has been identified, it would be possible for the inefficient DMUs to attempt to reproduce their 'best practice'. The DMUs classified as efficient, particularly those in the core group of efficient DMUs, would still have the opportunity to improve their overall efficiency, as each of them was operating with a slightly different strategy.

Each of the efficient DMUs may well have achieved their efficient status by focusing their efforts or resources in one or two areas. Therefore, by looking at the different operating practices of the other efficient DMUs, they may well be able to determine other areas for improvement. In a sense, if each of the DMUs labelled as efficient continues to improve their overall efficiency, the position of the efficiency frontier would be moving ever outwards, as overall levels of efficiency are increased and not just relative efficiency.

The importance of the discussion in this case is that the results from a DEA modelling process were of greater use if the analysis of them focuses on the individual DMUs and the notion of reference sets, rather than on generalising about efficiency scores, averages and standard deviations. Health care managers were more interested in a technique that can provide information directly relating to the improvement of performance at their own trust or hospital, than in discussions about the overall performance of a large number of hospitals.

Clearly, a user-friendly DEA modelling tool, such as the Frontier Analysis package produced by Banxia Software, which graphically represents the areas for potential improvement and immediately identifies the reference groups for each inefficient DMU, would greatly enhance the value of the DEA output. It could be understood and investigated without the need to understand complex series of numbers. This is linked to the next section, which addresses the complexity of the DEA technique.

9.6 Avoiding Complicated Theory

The complexity of DEA to the non-practitioner was widely apparent - it has been developed using complex mathematical concepts, not easily related to the simplest models on efficiency and performance, particularly where models incorporate a large number of variables. Particular elements of the technique, such as the definition of factor weights, are especially hard to explain in a practical sense. As the discussion in chapter five also illustrated, there are a large number of alternative formulations of the DEA models (the additive and multiplicative models are particularly complex) or possible adaptations to its basic form (super-efficiency and cross-efficiency are recent developments). In the analysis presented in chapters seven and eight, the simplest form of the DEA model was utilised and weight restrictions were introduced. These were attached to the virtual inputs and outputs, in an attempt to define them in some tangible way, particularly by expressing the relative importance of each of the factors.

As was discussed in chapter six, observation at DEA conference presentations to health care evaluators, managers and researchers new to the technique and its application have cemented this belief.

The value of the technique was swallowed up by the use of technical jargon and a concentration on theoretical frameworks rather than practical development. Those to whom the technique should be of greatest interest and relevance have been overwhelmed by complex equations rather than simple diagrams and tables, and have been left thinking 'what's the point?' in relation to the use of DEA.

Other, simpler techniques, whilst severely limited in their actual value, have been accepted because their meaning is more easily understood and represented. This was certainly the case with performance indicators, in that, whilst most academic research has stressed their limitations, their use has continued to expand within the NHS in the UK.

Of all the discussions relating to the use of DEA, this was probably the most difficult to address - many of those within the Health Service to whom DEA is most relevant already have preconceived ideas about its complexity and the many difficulties associated with its application. Therefore, the way in which the technique and the results from the modelling process have been presented to non-practitioners must be addressed, particularly in the development and availability of interactive DEA modelling packages and a focus on practicality rather than theory.

In the modelling strategy developed in chapters seven and eight, there would be little value in presenting the majority of the results and analysis for discussions with health care managers. The key area, for which a clear strategy for presentation would be essential, is the efficiency strategy for each DMU.

This would include a graphical representation of the targets for decreasing the inputs (or increasing the outputs for the alternative orientation of the DEA model) and the clear identification of the reference set and its current operating practices (input levels compared with the outputs each of them produces). This would highlight to the management of each DMU the areas where improvements can be made in their technical efficiency and also give them an idea on how best to achieve them – they can use the peer group DMUs as a guide.

As was stated above, the availability of specially written and user-friendly DEA software can be used to overcome the issues surrounding the complexity of the technique.

9.7 Presenting a True Reflection of Performance

In the original analysis, presented in chapter seven, the DEA models resulted in a large number of DMUs being classified as efficient. In many cases, the efficiency calculation included a contribution from just two or three of the variables. In the unbounded model, it was suggested that each DMU was observed in its 'most favourable light' (Dyson *et al.*, 1990). However, the unbounded model also allowed a DMU to be classified as efficient by ignoring areas of weakness or on the basis of spurious ratios (as was discussed in chapter eight).

Weight restrictions were included in the models by a way of a number of scenarios, each of which represented a different operating strategy or approach to efficient practice. Just four DMUs were efficient in each of these scenarios.

Including weight restrictions prevented DMUs from concealing their inefficiencies (by allocating a zero weight to a particular factor). The inefficiencies highlighted in each of the scenarios were also much greater than in the unbounded model, for a significant number of the DMUs. Therefore, whilst it may be preferable for a DMU to be classified as efficient, the imposition of weight restrictions has highlighted the many inefficiencies that actually exist.

In chapter eight, five efficiency scenarios were identified, although there were numerous other possibilities. It would be possible, through the application of a particular set of weight restrictions, to assess each DMU according to a specific efficiency strategy, especially if the management of a particular DMU has a strategy that needs evaluating.

9.8 Relevant at a Local Level

In the previous section, looking at the introduction of weight restrictions, it was suggested that these ensured that each DMU was evaluated in its true light. However, it was also observed that including weight restrictions would allow efficiency to be assessed according to pre-determined operating strategy specific to each DMU. This would fulfil the final objective raised in chapter six, looking at DEA from the perspective of the health care manager.

The weight restriction scenarios used in chapter eight allowed efficiency to be evaluated from an externally defined perspective, specified through the values fixed for the virtual weights.

Scenario one, for example, enforced that a DMU could not ignore any of its outputs in the calculation of its efficiency score. On the other hand, scenario three focused on the primary factors ('total inpatient discharges' and 'day case attendances'). A DMU that had chosen to concentrate in all four areas of output equally might be more interested in its evaluation under scenario one than scenario three. Alternatively, one of the DMUs in the sample might have a totally different strategy in operation, for which a further set of weight restrictions would be required. Looking at its results under different scenarios might also assist each DMU in identifying a new strategy to follow, which could be connected to the DMUs identified as its peer group.

The management of each DMU would need to be involved in the development of the weight restrictions. Alternatively, the weight restrictions could be imposed at a central level, in order that an external organisation (such as the Scottish Office or NHS Executive) could assess efficiency according to its own specific standard or definition.

The inclusion of weight restrictions in the DEA modelling process has made it possible for local opinions, strategies and initiatives to be included in the DEA models. Subsequently, the DMUs have been evaluated against them. In sections 9.3 and 9.4, changes to the variable set and sample of DMUs were proposed, which may give a further opportunity to reflect local issues and strategies. Alternative weight restriction scenarios could be developed to represent a different set of local policies or issues and environmental variables may also be of benefit. However, without the inclusion of weight restrictions, it would be very difficult to address local issues or use DEA to measure specific approaches to efficient practice.

9.9 Summary

In this chapter, the illustrative examples of the DEA methodology presented in earlier chapters have been critically evaluated, using the input of health service managers and observation, in order to determine the usefulness of the methodology as a tool for efficiency assessment. The specific points for the evaluation were those raised in chapter six, with the key areas being the selection of the samples and variables to be used, the information produced by the DEA models and the reflection of local issues and concerns. It has been determined that the DEA modelling strategy has been able to fulfil many of the requirements of the health care manager.

In relation to the selection of the variables and the determination of the sample of DMUs, a number of possible alternatives were also identified. It was proposed that these may be used to enhance the applicability of the technique, to be investigated through further research.

In terms of the output produced by the modelling strategy, it was shown that additional and useful material was produced by the development of efficiency strategies, extending the information currently provided by performance indicators and other tools for performance assessment. The inclusion of weight restrictions ensured that the DMUs were evaluated in their true light and local policies, initiatives, operating strategies and environmental concerns could be incorporated into the modelling strategy. The complexity of the technique was addressed by producing graphical representations of efficiency strategies and focusing on results, as opposed to complex methodological issues.

The evaluation process used in this chapter was consistent with the evaluation of public sector performance presented in chapter two. As was noted there, according to Flynn (1986), a service cannot be efficient unless it is also effective, where efficiency is described as ‘doing the thing right’ and effectiveness as ‘doing the right things’ (Norris, 1978). In other words, DEA was only an acceptable technique for efficiency evaluation if it was ‘measuring the thing right’ and ‘measuring the right thing’, being both ‘efficient’ and ‘effective’.

It has been shown that DEA has proved to be an ‘efficient’ measure of technical efficiency, as discussed in section 9.5.1. Additionally, as the modelling strategy applied in chapters seven and eight has been seen to fulfil the requirements of the health care mangers specified in chapter six, it can also said to be an ‘effective’ approach for efficiency assessment.

In the final chapter, the key issues raised by the research are discussed in relation to the objectives specified in chapter one. The long-term future of the DEA methodology is evaluated and a number of recommendations for further study are presented.

9.10 References

1. **Ball, R. and Roberts, E.** (1998) The Relevance of DEA in the Public Sector - A Discussion Paper for Young OR 10, Young OR 10 Keynote Papers, Operational Research Society.

2. **Ball, R., Monaghan, C., Thomas, R. E. and Wagner, R.** (1997) Data Envelopment Analysis: A Practical Tool for Policy Makers?, University of Stirling Departmental Working Paper.

3. **Banker, R. D., Charnes, A., Cooper, W. W., Swarts, J. and Thomas, D. A.** (1989) An Introduction to Data Envelopment Analysis With Some of its Models and Their Uses, *Research in Governmental and Nonprofit Accounting*, Vol. 5, pp. 125 – 163.

4. **Beasley, J.** (1990) Comparing University Departments, *Omega*, Vol. 18, pp. 171-183.

5. **Belton, V.** (1992) An IDEA - Integrating Data Envelopment Analysis with Multiple Criteria Decision Analysis, In Proceedings of the Ninth International Conference on Multiple Criteria Decision Making: Theory and Applications in Business, Industry and Commerce, **Goicoechchea, A., Duckstein, L. and Zionts, S.** (eds.), pp. 71 - 79, Springer-Verlag, Berlin.

6. **Byrnes, P. and Valdmanis, V.** (1994) Analyzing Technical and Allocative Efficiency of Hospitals, In Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M (eds.) Data Envelopment Analysis: Theory, Methodology and Applications, Kluwer.

7. **Department of Health** (1997) The New NHS: Modern and Dependable.

8. **Dyson, R. G., Thanassoulis, E. and Boussofiane, A.** (1990), Data Envelopment Analysis, in "Operational Research Tutorial Papers", L.C. Hendry and R. Eglese (eds), The Operational Research Society, pp. 13 – 28.

9. **Ehreth, J. L.** (1994) The Development and Evaluation of Hospital Performance Measures for Policy Analysis, *Medical Care*, Vol. 32, No. 6, pp. 568 – 587.

10. **Flynn, N.** (1986) Performance Measures in Public Sector Services, *Policy and Politics*, Vol.14, No. 3, pp. 389 – 404.

11. **Grosskopf, S. and Valdmanis, V.** (1987) Measuring Hospital Performance: A Non-parametric Approach, *Journal of Health Economics*, Vol. 6, No. 2, pp. 89 – 107.

12. **Information and Statistics Division** (1996) Scottish Health Service Costs 1995-96, ISD, NHS in Scotland.

13. Jenkins, L., Bardsley, M., Coles, J., Wickings, I. and Leow, H. (1987) Use and Validity of NHS Performance Indicators - A National Survey, CASPE.

14. Kleinsorge, I. and Karney, D. (1992) Management of Nursing Homes Using Data Envelopment Analysis, *Socio-Economic Planning Science*, Vol. 26, pp. 57-71.

15. Norris, G (1978) The Effective University – A Management By Objectives Approach, Saxon House.

16. Ozcan, Y. A., Luke, R. and Haksever, C. (1992) Ownership and Organizational Performance: a Comparison of Technical Efficiency Across Hospital Types, *Medical Care*, Vol. 30, pp. 781-794.

17. Parkin, D. and Hollingsworth, B. (1996) Measuring Production Efficiency of Acute Hospitals in Scotland, 1991-1994: Validity Issues in Data Envelopment Analysis, Working Paper, Department of Epidemiology and Public Health, University of Newcastle.

18. Pedraja-Chaparro, F., Salinas-Jimenez, J. and Smith, P. (1999) On the Quality of the Data Envelopment Analysis Model, Journal of the Operational Research Society, Vol. 50, No. 6, pp. 636 – 644.

19. Sherman, H. D. (1984) Hospital Efficiency Measurement and Evaluation, *Medical Care*, Vol. 22, pp. 927-938.

20. Thanassoulis, E., Boussofiane, A. and Dyson, R. G. (1995) Exploring Output Quality Targets in the Provision of Perinatal Care in England Using DEA, *European Journal of Operational Research*, Volume 60, pp. 588-608.

21. The Scottish Office (1996) Scotland Getting Better, The National Health Service in Scotland Annual Report, 1995-96, The Scottish Office.

22. Wilkinson, R. (1991) The Measurement of University Performance: an Application of Data Envelopment Analysis, Ph.D. Thesis.

9.11 Additional Sources

1. Interviews at the NHS Executive Performance Directorate in Leeds (conducted in October, 1996).

2. Interviews at Blackburn, Hyndburn and Ribble Valley (BHRV) NHS Trust (conducted in July, 1996).

3. Discussions in the DEA and Health Streams at Young OR 10, Guildford, March 31st – April 2nd, 1998.

4. Discussions at the DEA Symposium, Wernigerode, Germany, 16th –18th October, 1998.

CHAPTER TEN

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

10.1 Introduction

Performance and efficiency assessment has become endemic in the National Health Service in the United Kingdom, and in many other countries around the world. Many techniques have been proposed for the measurement of these key concepts, not always with great success. The relatively new methodology of Data Envelopment Analysis (DEA), developed from linear programming ideologies, has been proposed as an alternative tool for efficiency assessment. This research has addressed these issues.

The following section draws together the main findings of the research, reviewing each of the main chapters in relation to the research question established for them in chapter one. The key objectives of the research (also outlined in chapter one) are then discussed, evaluating how well each of these has been addressed. Limitations of the research and the methodologies applied are also considered.

The final section of this chapter addresses the research question defined for this chapter, assessing the short-term and long-term future of the DEA methodology for the evaluation of health service efficiency and discussing the possibilities for further research.

10.2 Summary of Findings

In chapter one, a research question was identified for each of the main chapters of the thesis. In the review of the main points of the thesis included in this section, the research questions are answered

Chapter Two: Can performance in the National Health Service be investigated in such a way as to make use of traditional definitions of its key elements?

The key elements of public sector performance were defined to be efficiency, effectiveness and economy. However, recent changes in the principles of public sector management have extended evaluations of performance to include quality, equity and equality. In the context of this research, efficiency was defined to be of central concern in the assessment of performance, which affected the tools to be used for its measurement. Some caveats to the traditional definitions of the elements of performance were identified in the context of the NHS. This particularly related to the difficulties in distinguishing between efficiency and effectiveness in the provision of hospital services.

Chapter Three: How important is the nature of the NHS (history, structure and management philosophy) in the definition and application of performance assessment techniques?

The NHS was found to be a very complex organisation, having undergone a number of changes to its structure during its fifty-year history. It was seen that this impinged on approaches to performance assessment, in that they were developed to address the issues defined by central government to be significant at that time. In terms of financing, it was also shown that, in comparison with a number of other first-world countries, accusations of under-funding could well be justified. Despite this, it was seen that the number of patients receiving treatment has continued to increase.

Chapter Four: Are the methods currently applied to performance evaluation appropriate and do they provide useful information?

The most frequently used methods were found to be performance indicators and clinical audit. Regression analysis, benchmarking, cost-benefit analysis and QALYs have also been used to address various aspects of performance. It was seen that these methods have not directly addressed the measurement of technical efficiency, but have been widely applied to assess other aspects of performance, with a limited amount of success. Performance Indicators, for example, were found to be partial, in that they only provided a snapshot of performance and they do not readily promote 'best practice'. Benefits were observed, especially if they were used to direct the attention of management to problem areas, which could then be investigated by other means. Clinical Audit was seen to be a very varied activity, with its success often dependent upon the priorities or agendas of the people involved. The potential for other methods was clearly observed.

Chapter Five: Does DEA provide an alternative for the evaluation of efficiency in a health care environment?

It was determined that there were a number of characteristics of the DEA methodology, which suggested that it would be a useful tool for the measurement of health service efficiency. The most significant of these were related to its ability to directly measure technical efficiency and provide detailed information about the identified inefficiencies.

It was suggested that the modelling process, which was shown to be inclusive, could also record local policies and initiatives, as well as environmental influences. A number of extensions to the basic DEA methodology were also considered, such as the inclusion of weight restrictions, the measurement of allocative efficiency and the development of alternative formulations of the efficiency equations. It was observed that, whilst some of these could be shown to enhance the basic technique, the practical benefits of others had not been established, particularly in relation to health care data.

Chapter Six: What information can be learned from the literature on health care applications of the DEA methodology?

A small number of papers were identified in the Literature, each of which had used DEA-type methodologies to examine the efficiency of a small sample of hospitals, although just one of these used data from the UK. Evaluation of the approaches and the results provided insights in a number of key areas: modelling strategy, DEA model formulation, the selection of the variables, including individual perspectives and evaluating and presenting results. The information gleaned from the literature was contrasted with the perspectives of the Health Care Manager. A number of requirements for the adoption of a new methodology were specified, related to the selection of both the samples of hospitals and the variables to represent them, the quality of the information produced and the simplicity of the approach. It was felt that the methodology should present a true reflection of performance and also be relevant at a local level, able to measure and incorporate specific local initiatives.

Chapter Seven: Can a model-building process be developed for a large-scale application of the DEA methodology?

A two-stage application process was developed, which was used to measure the efficiency of a sample of seventy-four acute hospitals in Scotland. The first stage was presented in chapter seven, comprising five important elements: the definition of the sample, the formulation of the DEA model type and orientation, selection of possible inputs and outputs, the results of the preliminary analysis and the revisions to the model. In each stage, the views and perspectives of the health care manager were considered. It was found that it was necessary to define a clear objective for the efficiency analysis, as this was shown to impinge on many aspects of the application, particularly in the final selection of variables and the revised sample of DMUs (reduced to forty-seven).

Chapter Eight: What impact does the inclusion of weight restrictions have on the applicability of the DEA methodology?

The second of the two stages in the application procedure was presented in chapter eight, which was primarily focused on the inclusion of weight restrictions. The weight restrictions were included following an investigation of the results from the revised model, with particular attention given to the patterns of the factor weights and the virtual weight percentages. Weight restrictions were introduced using a number of alternative scenarios, developed to investigate a particular approach to efficient practice or operating strategy.

It was observed that the alternative scenarios could be used to develop an efficiency strategy for each DMU, giving short-term and long-term goals for reducing its inputs in relation to the targets identified by the modelling process. The main benefits of the inclusion of weight restrictions in the modelling strategy were the identification of a core group of efficient hospitals, the improvements in overall standards and the identification of efficiency strategies.

Chapter Nine: Does the DEA methodology have actual benefits in practice and how can it be enhanced?

The evaluation was from the perspective of the health care manager, using the key issues raised in chapter six. It was determined that the DEA modelling strategy was able to fulfil many of these requirements. The output from the models was shown to provide additional and useful material, through the development of efficiency strategies, which extended the information currently available. The DMUs were evaluated in their true light through the inclusion of weight restrictions, which also allowed local policies, initiatives, operating strategies and environmental concerns to be incorporated into the modelling process. Producing graphical representations of efficiency strategies and focusing on results, as opposed to complex methodological issues, reduced the complexity of the approach, although it was impossible to address this issue completely. Potential for improvement was related to the selection of the variables and the sample of DMUs. A number of alternatives were suggested, with particular attention given to using much smaller samples of hospitals and the inclusion of additional variables, measuring environmental influences or service quality.

10.3 The Research Objectives

In chapter one, four objectives for the research were identified and the success with which these have been achieved is now considered.

10.3.1 To Determine the Reasons for the Importance of Performance Assessment in Health Services in the UK

Four reasons were suggested for the elevation in the importance of performance and efficiency evaluation (discussed in chapter two), determined through recourse to the literature and information produced by a number of government agencies, such as the Department of Health, the NHS Executive and the Audit Commission. First, there are now severe restrictions on the level of public finance available and constraints have been placed on all areas of spending by government. Second, the inherent inefficiency of public sector services in general has been observed, requiring new methods to address these inefficiencies. Third, changes in the role of management within the NHS have given an increased emphasis to efficient resource use. The final reason was that a number of Government agencies, such as the Audit Commission, have adopted the pursuit of efficiency as one of their primary responsibilities.

10.3.2 To Identify the Methods Used for Performance Assessment in the NHS

In chapter three, a number of alternatives were identified, that have been used within the NHS to address various aspects of performance. Performance Indicators and Clinical Audit were seen to be the most widely used techniques, both of which have become prevalent over the last twenty years. Performance Indicators were developed in the early 1980s, whilst Clinical audit was made compulsory in the early 1990s.

A number of other techniques were investigated, including Regression analysis, benchmarking, cost-benefit analysis and QALYs, although their use has been limited. In the discussion, it was observed that, whilst the techniques did provide a degree of useful information, they were also found to have severe limitations.

10.3.3 To Investigate Alternative Methods for Performance Assessment

Data Envelopment Analysis was identified as an alternative tool that could be used to directly address the technical efficiency of health service provision, for which several particular benefits were seen. Its ability to handle incommensurate inputs and outputs, including environmental factors and quality measures, was stressed. Also, the nature of the information produced by the modelling process, such as the identification of targets for improved efficiency and a peer group of similarly operating units, was found to be very useful. The complexity of the technique was seen to provide a drawback to its widespread use, as was the lack of a comprehensive modelling strategy. A number of technical adaptations to the basic model were also observed, including weight restrictions, alternative model formulations and the measurement of other aspects of efficiency. It was suggested that these enhancements made application of the technique more difficult.

10.3.4 To Develop New Approaches for Performance Assessment

In order to enhance the applicability of the DEA methodology, a two-stage application procedure for the measurement of hospital efficiency was developed. The specifications of the process were based on the perspectives of the health care manager, from whom a number of requirements were identified.

The main requirements of the new approach were in relation to the information provided by it and its ability to represent local concerns and initiatives, with a true reflection of performance. These were primarily addressed by introducing weight restrictions into the basic DEA methodology, which allowed specific approaches to efficiency to be measured. Additionally, clear objectives for the efficiency assessment were emphasised, as this impacted on the selection of the sample of hospitals, the variables and the weight restrictions themselves. It was proposed that for each hospital, an efficiency strategy could be developed, which would identify the areas where significant improvements in efficiency could be made and produce both short-term and long-term targets. Additionally, a small set of efficient units was identified, each of which could be used to provide examples of best practice in the production of hospital services. It was intended that the application procedure was sufficiently flexible to be used in the measurement of other aspects of health service efficiency.

As has been shown, the specified objectives of the research have been addressed. However, as is to be expected, there are also a number of limitations to the approaches and analysis presented for discussion, to be discussed in the following section.

10.4 Limitations of the Research

The assessment of efficiency and performance in the Health Sector is clearly complicated by the complex organisational structure of the NHS in the UK. The techniques applied for performance assessment must be adapted to reflect these difficulties. The investigation into the DEA methodology, with particular emphasis on the measurement of hospital efficiency, was intended to address these areas.

In relation to the application of the DEA methodology to the measurement of hospital efficiency, it would clearly have been beneficial to have access to a wider range of internal data. This would have allowed for other factors to be included in the modelling process, such as an indicator of case-mix complexity for the inpatient activity. However, if the Department of Health or a number of hospitals adopted the modelling strategy, additional data would be available for inclusion in the DEA models.

Further involvement from within the Health Sector could also have enhanced the modelling strategy and analysis. A broader range of opinions and perspectives on health service performance assessment could have added an extra dimension to the evaluation of the modelling strategy or suggested other areas of perceived importance. There are many different opinions on the value of performance assessment techniques, not all of which have been incorporated. It may have been useful, for example, to address the opinions of the medical profession and incorporate these into the modelling strategies, where appropriate.

In relation to the DEA methodology and its application to the measurement of hospital efficiency, a number of further limitations can be observed. The lack of some important variables was addressed above. A further potential limitation was in the samples chosen for analysis using the DEA methodology, as was discussed in chapter nine. Methods of comparative performance measurement are reliant on the definition of a homogenous sample of operating units, whether they are hospitals, general practices or the individual doctors themselves.

The sample used, representing the acute hospitals in Scotland, was revised in an attempt to ensure that all the hospitals had similar operating strategies, related to the variables included in the model and the objectives of the research. The analysis in chapter nine highlighted potential problems with this approach and it was suggested that an alternative sample might have been more appropriate. This leads to a further limitation of the DEA methodology in the measurement of hospital efficiency, in that it could be virtually impossible to identify a sample of hospitals that would be acceptable on all sides. Additionally, the supposedly unique nature of hospitals could make it difficult to ensure that all hospitals were actually included in a sample, so that their relative efficiency could be evaluated.

The final point in this section relates to the adaptability of the modelling strategy, which was defined and tested in chapters seven and eight. Its complete flexibility has not been evaluated, through its application to a number of other samples, covering hospital departments, health care programmes or regional health authorities, for example. This, and other similar issues, is addressed in the final section of this chapter, looking at recommendations for further study.

10.5 Recommendations for Further Study

A number of areas have been addressed in this study, for which opportunities for further research are available. These relate to the future of performance assessment itself, the techniques currently used for performance assessment in the Health Service in the UK and the long-term viability of the DEA methodology.

In relation to performance and efficiency assessment, it appears unlikely that the trend for its continued application will disappear. All aspects of public and private sector services are evaluated to a certain degree. However, future research in this area could investigate whether the widespread use of performance measurement methodologies has led to actual improvements in performance, changed operating practices or altered patterns of behaviour.

In the field of performance assessment, in relation to public sector services in general and health services in particular, there is clearly scope for a great deal of further research. The techniques such as performance indicators and clinical audit have been shown to have numerous benefits, provided they are applied with care and their limitations are understood. Performance Indicators are continually evolving and their application to the measurement of outcomes is an area where there is particular scope for further research. The widespread use of clinical audit is a relatively new initiative in the NHS. Research into the actual benefits of the audit process could be beneficial, focusing on whether it can actually be shown that audit has resulted in significant improvements in performance or the standards of care provided.

The Data Envelopment Analysis methodology clearly provides scope for further research, as there are so many facets to the technique and its application. Much of the work in this field has concentrated on areas of theoretical development, for example, in the definition of alternative model formulations. It would be useful, therefore, to further address the practical applications of the methodology, focusing on the measurement of efficiency in the Health Service in the UK.

In the analysis presented in chapter eight, weight restrictions were introduced into the methodology, which resulted in the development of efficiency strategies. Further study of the weight restrictions is required, particularly in relation to the actual virtual weight percentages chosen to be the restrictions in each of the scenarios. The impact on the evaluation of efficiency of small changes to the scenarios could be investigated. Additionally, the impact and practical relevance of alternative formulations of the DEA model could also be considered.

In relation to the application procedure developed in chapters seven and eight, it would be appropriate to apply this to a number of other samples of hospitals, hospital departments or categories of patients, as was discussed in chapter nine. This could be used to determine if the modelling strategy could be applicable in a general sense. Specific research into the benefits and implications of using much smaller samples could also be carried out.

Research into the DEA methodology has tended to focus on technical developments, such as the measurement of allocative efficiency, cross-efficiencies and super efficiency. This has generally been at the expense of an application-based approach to research into efficiency assessment.

Furthermore, there are a myriad of alternative tools for performance and efficiency assessment that are more simple to understand and have already been widely applied, including performance indicators and clinical audit.

However, DEA clearly has a considerable amount to offer the health service in the UK. Even the simplest element of the DEA process, that is, the division of the sample into efficient and non-efficient DMUs, can be used as basis for further investigations into performance at a local level.

Additionally, in the analysis presented in chapters seven and eight, a methodology has been derived that can (i) identify a core group of efficient DMUs that can provide elements of best practice for the less efficient DMUs, (ii) produce efficiency strategies for the inefficient DMUs, showing where improvements can be made in outputs or savings can be made in inputs and (iii) identify different operating strategies and local policies through the imposition of weight restrictions and measure their influence on efficiency.

Therefore, if the research process for DEA itself will undergoes some radical changes, in particular becoming more application-based, there is clearly scope for DEA to become widely accepted and applied as an effective tool for efficiency assessment in the analysis of health services in the UK.

APPENDIX ONE

KEY DEFINITIONS

Appendix 1A: General Description of Functional Classification

Code	Description
01	Large general major teaching hospitals covering a full range of services (other than maternity in some cases) and with special units.
02	General hospitals with some teaching units but not necessarily wholly teaching.
03	Hospitals providing some local general services but excluding a high proportion of highly specialised units.
04	Small general hospitals with some specialised staff including a surgical unit. No maternity.
05	Small general hospitals with some specialist staff including a surgical unit but with maternity.
06	General non-teaching hospitals but not covering the full range of work within the main specialties.
07	Large teaching hospitals for children covering the full range of medicine and surgery.
08	General practitioner cottage hospitals with no maternity units and with limited surgery done either by general practitioner or visiting consultant. Centres for consulting clinics.
09	General practitioner cottage hospitals with maternity units and with limited surgery done either by general practitioner or visiting consultant. Centres for consulting clinics.
10	General practitioner cottage hospitals with maternity units and visiting consultant clinics but with no surgery of any kind.
11	Mixed specialist hospitals with maternity. No special units. Consultant type surgery undertaken.
12	Mixed specialist hospitals without maternity units. No special units. Consultant type surgery undertaken.
13	Hospitals with medical and/or surgical units but with a large chronic sick element.
14	Special orthopaedic units with active surgery. Adults and children.
15	Consultant staffed units in which surgery and accident work predominate.

Appendix 1B: Scottish Health Board Codes

Code	Name
A	Ayrshire and Arran
B	Borders
C	Argyll and Clyde
F	Fife
G	Greater Glasgow
H	Highland
L	Lanarkshire
N	Grampian
R	Orkney
S	Lothian
T	Tayside
V	Forth Valley
W	Western Isles
Y	Dumfries and Galloway
Z	Shetland

Appendix 1C: Definitions of Key Terms

PATIENT ACTIVITY

Inpatient: An inpatient is a patient who occupies an available staffed bed in hospital and remains overnight.

Inpatient Discharge: An inpatient discharge marks the end of an inpatient episode. The patient leaves hospital for a location external to the NHS or is transferred to another NHS inpatient service or dies.

Inpatient Discharges from Hospital/Specialty: Inpatient discharges from hospital/specialty are counted as the number of patients who are discharged from the hospital/specialty plus those who die.

Inpatient Case: The term case is used as a basis for comparing inpatient costs.

1. Hospital Case - A person is counted as one hospital case during the time (s)he is an inpatient in specific hospital, i.e. between date of admission and date of discharge from hospital.
2. Specialty Case - A person is counted as one specialty case during the time (s)he is an inpatient in a specific specialty within a specific hospital, i.e. between the date of admission to the specialty and the date of discharge from the specialty.

Average Duration of Stay (Hospital/Specialty): The average duration of stay (hospital/specialty) is calculated by dividing the number of inpatient days in the hospital/specialty by the number of inpatient discharges from the hospital/specialty.

Day Case: A day case is a patient who makes a planned attendance to a specialty for clinical care, sees a doctor or dentist, requires the use of a bed, or trolley in lieu of a bed. The patient is not planned to and does not remain overnight. Many of these patients require anaesthesia. The emphasis is on planned attendances and day case care should be regarded as an alternative to inpatient care.

Day Case Episode or Attendance: A day case episode is the occasion of a patient attending for day case treatment within a specialty, either at a day bed unit or in an inpatient ward or other designated area.

Day Hospital: A day hospital is a hospital which provides services on a regular day time basis for specific patient/client groups, for example, the elderly, mentally ill or mentally handicapped. Services normally provided are assessment, rehabilitation, maintenance of function and clinical treatment.

Day Patient: A day patient is a patient who attends a day hospital on a regular basis. The emphasis is on a regular attendance and day patients usually attend one or more times each week. Where there is no formally established day hospital but patients attend regularly at a ward on a day basis they should be classed as day patients.

Day Patient Attendance: A day patient attendance is the occasion of a patient attending a day hospital, or an inpatient ward for day patient care, for one day or part of a day. The attendance usually lasts at least half a day.

Consultant Outpatient Attendances: Consultant outpatient attendances cover all attendances at consultant clinics plus meetings between patients and consultants outwith clinic sessions, e.g. home visits.

PAM Outpatient Attendances: PAM outpatient attendances cover all attendances made at professional and technical departments (e.g. speech therapy, physiotherapy) by patients from outwith the hospital who are not attending as part of day patient or day case care.

New Attendances: A new attendance is the first contact between the patient and the health care professional following a referral.

New Patient: A patient is defined as new at the first contact with the health care professional following a referral.

Accident and Emergency Attendences: Accident and Emergency Attendances cover both new and return (review) attendances. An A&E attendance takes place when a patient presents, or is presented, with or without prior arrangement, at an A&E department for assessment and/or management of an injury or illness. The patient is seen by a doctor or a nurse. A&E attendances include patients who have died at the scene of an incident or whilst en route to the hospital. A&E attendances do not include patients who attend an A&E department solely to receive services provided by specialties other than A&E, PAMs or other technical staff in the A&E department.

Average (Available) Staffed Beds: Available staffed beds are beds which are resourced for inpatient or day case care. For any specialty they may be allocated beds from another specialty, or temporary beds. The figure shown is the average number of available staffed beds for the year. It is calculated by dividing the total number of available staffed beds for the year by 365.

Occupied Bed Days (Inpatient Days): Occupied beds are available staffed beds which are either being used to accommodate inpatients or reserved for patients on pass. The figure shown is the number of occupied bed days for the year.

STAFF

Whole Time Equivalent (WTE): Whole time equivalent is the number of staff expressed in relation to the standard weekly hours for a particular staff category.

FINANCE

Capital Charges: Capital charges comprise depreciation on fixed assets including buildings, installations and fittings, and medical/surgical paramedical/furniture and other equipment held on the Assets register; interest paid on monies borrowed to finance any of the projects mentioned previously; 6% return on capital.

Appendix 1D: Definitions of Types of Care and the Composition of Specialty Groups

Care Type	Specialty Group	Specialty
ACUTE	Accident and Emergency	Accident and Emergency
	Acute Other	Acute Mixed, Other Acute
	Cardio-Thoracic Surgery	Cardio-Thoracic Surgery
	Communicable Diseases	Communicable Diseases
	Dental	Orthodontics
		Paediatric Dentistry
		Restorative Dentistry
	Dermatology	Dermatology
	Ear, Nose and Throat	Ear, Nose and Throat
	General Practice	General Practice
	General Surgery	General Surgery
	Gynaecology	Gynaecology
	Intensive Therapy Unit	Intensive Therapy Unit
	Medical	Cardiology
		Clinical Pharmacology and Therapeutics
		Gastroenterology
		General Medicine
		Genito-Urinary Medicine
		Haematology
		Intermittent Haemodialysis
		Medical Oncology
		Metabolic Diseases
		Nephrology
		Nuclear Medicine
		Poisons
	Medical Paediatrics	Medical Paediatrics
	Neurology	Neurology

Care Type	Specialty Group	Specialty
ACUTE contd.	Neurosurgery	Neurosurgery
	Ophthalmology	Ophthalmology
	Oral Surgery and Medicine	Oral Surgery
		Oral Medicine
	Orthopaedics	Orthopaedics
	Plastic Surgery and Burns	Plastic Surgery and Burns
	Radiotherapy	Radiotherapy
	Rehabilitation Medicine	Rehabilitation Medicine
	Spinal Paralysis	Spinal Paralysis
	Surgical Paediatrics	Surgical Paediatrics
	Urology	Urology
Maternity	Obstetrics GP	Obstetrics GP
	Obstetrics Specialist	Obstetrics Ante-Natal
		Obstetrics Post-Natal
	Special Care Baby Unit	Special Care Baby Unit
Mental Illness	Adolescent Psychiatry	Adolescent Psychiatry
	Child Psychiatry	Child Psychiatry
	Geriatric Psychiatry	Geriatric Psychiatry
	Mental Illness	Forensic Psychiatry
		Mental Illness
		Psychotherapy
Mental Handicap	Mental Handicap	Mental Handicap
Geriatric Assessment	Geriatric Assessment	Geriatric Assessment
Geriatric Continuing Care	Geriatric Long Stay	Geriatric Long Stay
Younger Physically Disabled	Young Chronic Sick	Young Chronic Sick
Community	Community Nursing, Midwifery and Health Visiting et al.	/

APPENDIX TWO

DATA AND ANALYSIS USED IN CHAPTER FIVE

Appendix 2A: Hospitals in the Sample (Gynaecology Inpatient Services)

Hospital	Functional Class	Health Board
Aberdeen Royal Infirmary	01	N
Borders General	11	B
Crosshouse Hospital	02	A
D&G Royal Infirmary	12	Y
Eastern General, Edinburgh	02	S
Edinburgh Royal Infirmary	01	S
Falkirk Royal Infirmary	11	V
Glasgow Royal Infirmary	01	G
Hairmyres, East Kilbride	12	L
Inverclyde Royal Hopsital	02	C
Monklands Hospital	12	L
Ninewells	01	T
Perth Royal Infirmary	11	T
Queen Margaret Hospital	02	F
Raigmore, Inverness	02	H
Royal Alexandra Hospital	02	C
Southern General (SGH)	02	G
St. Johns at Howden	02	S
Stirling Royal Infirmary	11	V
Stobhill, Glasgow	02	G
Vale of Leven, Alexandria	11	C
Victoria Infirmary, Glasgow	02	G
Western/Gartnavel	01	G

Appendix 2B: Data for Analysis in Chapter Five

DMU	I_1	I_2	I_3	I_4	O_1	O_2	$I_{3(a)}$	$I_{3(b)}$	$I_{3(c)}$	$I_{3(d)}$
#1	4905	1824	3081	82	17236	4439	182	1003	1176	719
#2	1857	544	1313	25	5547	1953	242	258	543	270
#3	2330	660	1670	43	8495	2643	211	518	661	280
#4	2283	553	1730	24	5882	2008	285	321	570	552
#5	1620	518	1102	25	5568	1700	180	230	405	287
#6	1948	653	1295	28	6691	2000	296	500	338	162
#7	3384	1074	2310	31	8192	2290	680	543	337	751
#8	1858	672	1186	24	5956	1942	295	363	282	245
#9	2383	876	1507	41	8353	2367	237	497	525	249
#10	1626	797	829	18	4188	1684	177	258	264	131
#11	1685	240	1445	34	7482	2184	131	273	697	345
#12	1614	352	1262	24	6241	2253	151	333	554	223
#13	785	309	476	18	3069	885	119	207	72	79
#14	1409	344	1065	25	5641	1749	306	271	289	198
#15	472	204	268	10	2496	563	31	138	47	52
#16	1493	454	1039	24	3935	1610	163	190	493	192
#17	960	314	646	12	3553	1120	106	202	242	96
#18	1424	341	1083	19	3572	1239	186	256	400	240
#19	843	343	500	17	3316	1368	71	249	111	68
#20	1321	248	1073	19	3505	1183	156	270	496	151
#21	2208	862	1346	32	6475	1825	250	480	352	263
#22	1725	535	1190	31	5599	2073	160	305	425	301
#23	1163	394	769	21	4559	1285	172	256	220	121

Key to Factor Codes:

I_1: Total Costs (£000s)

I_2: Total Allocated Costs (£000s)

I_3: Total Direct Costs (£000s)

I_4: Average Number of Staffed Beds

O_1: Total Patient Days

O_2: Total Discharges

Additional Input Data (For section 5.5.1)

$I_{3(a)}$: Total Medical Costs (£000s)

$I_{3(b)}$: Total Nursing Costs (£000s)

$I_{3(c)}$: Total Theatre Costs (£000s)

$I_{3(d)}$: Total Other Direct Costs (£000s)

Appendix 2C: DEA Results for Model G3

DMU	Efficiency Score (%)	Rank	Reference Set	Reference Set Count	Output (O_2)	Output Target	Virtual Input for I_2	Virtual Input for I_3
#1	57.3	21	12, 19	0	4439	7753	54.9	45.1
#2	70.8	12	12, 19	0	1953	2759	46.0	54.0
#3	76.9	6	12, 19	0	2643	3435	44.9	55.1
#4	61.7	19	12, 19	0	2008	3253	39.7	60.3
#5	69.1	14	12, 19	0	1700	2459	49.2	50.8
#6	66.8	15	12, 19	0	2000	2993	50.9	49.1
#7	44.7	23	12, 19	0	2290	5128	48.9	51.1
#8	66.6	16	12, 19	0	1942	2914	53.9	46.1
#9	63.0	18	12, 19	0	2367	3755	54.5	45.5
#10	74.3	9	19	0	1684	2268	0	100
#11	100	3 (1)	11	2	2184	2184	26.7	73.3
#12	100	2 (1)	12	19	2253	2253	36.5	63.5
#13	70.2	13	12, 19	0	885	1261	57.8	42.8
#14	87.0	4	12, 19	0	1749	2011	40	60
#15	76.8	7	19	0	563	733	0	100
#16	71.9	10	12, 19	0	1610	2238	47.4	52.6
#17	76.4	8	12, 19	0	1120	1466	50	50
#18	61.2	20	12, 19	0	1239	2024	39.3	60.7
#19	100	1 (1)	19	20	1368	1368	0	100
#20	66.0	17	11, 12	0	1183	1791	33.6	66.4
#21	51.6	22	12, 19	0	1825	3539	56.9	43.1
#22	79.8	5	12, 19	0	2073	2599	48.1	51.9
#23	71.7	11	12, 19	0	1285	1792	51.3	48.7

Key to Factor Codes:

O_2: Total Discharges

I_2: Total Allocated Costs

I_3: Total Direct Costs

Appendix 2D: DEA Results for Model G8

DMU	Efficiency Score (%)	Rank	Reference Set	Reference Set Count	Scale Variance Coefficient
#1	100	3 (1)	1	11	- 0.21088
#2	76.2	16	1, 12, 19	0	- 0.49639
#3	95.0	6	1, 12, 19	0	- 0.39060
#4	61.8	22	11, 12, 19	0	0.01325
#5	71.2	19	1, 12, 19	0	- 0.59003
#6	80.0	15	1, 12, 19	0	- 0.50136
#7	55.6	23	1, 12, 19	0	- 0.28154
#8	83.1	13	1, 12, 19	0	- 0.54599
#9	89.0	9	1, 12, 19	0	- 0.42942
#10	92.3	8	1, 19	0	- 0.78375
#11	100	4 (1)	11	9	0.1976
#12	100	2 (1)	12	12	- 0.51946
#13	81.2	14	11, 15, 19	0	0.30286
#14	87.3	11	11, 12, 19	0	0.02144
#15	100	5 (1)	15	6	1.00
#16	72.0	18	11, 12, 19	0	0.01918
#17	83.5	12	11, 15, 19	0	0.26703
#18	67.3	20	11, 15, 19	0	0.20070
#19	100	1 (1)	19	18	0.21149
#20	87.8	10	11, 15	0	0.77216
#21	65.7	21	1, 19	0	- 0.48271
#22	92.5	7	1, 12, 19	0	- 0.54687
#23	74.8	17	11, 15, 19	0	- 0.21743

APPENDIX THREE

DATA AND ANALYSIS
USED IN CHAPTER SEVEN

Appendix 3A: Acute Hospitals in Scotland

Hospital Name	Functional Class	Health Board
Aberdeen City	13	N
Aberdeen Royal Infirmary	01	N
Adamson, Cupar	08	F
Arbroath Infirmary	05	T
Arran War Memorial	05	A
Balfour, Kirkwall	05	H
Ballochmyle, Mauchline	12	A
Belford, Fort William	05	H
Blairgowrie Cottage	08	T
Borders General	11	B
Caithness General, Wick	05	R
Campletown Hospital	10	C
Canniesburn, Glasgow	13	G
Chalmers, Banff	09	N
Crosshouse Hospital	02	A
D&G Royal Infirmary	12	Y
Dailburgh, South Uist	09	W
Davidson Cottage, Girvan	10	A
Dundee Royal Infirmary	01	T
Dunoon & District General	05	C
Eastern General, Edinburgh	02	S
Edinburgh City	02	S
Edinburgh Royal Infirmary	01	S
Falkirk Royal Infirmary	11	V
Forfar Infirmary	13	T
Fraserburgh	10	N
Garrick, Stranraer	04	Y
Gilbert Bain, Lerwick	05	Z
Glasgow Royal Infirmary	01	G
Hairmyres, East Kilbride	12	L
Huntly Jubilee	09	N
Insch and District War Memorial	10	N
Inverclyde Royal Hopsital	02	C
Islay, Bowmore	10	C
Kincardine O'Neil	10	N
Law, Carluke	11	L
Leanchoil, Forres	09	N
Lorn & Islands District General	05	C
Mackinnon Memorial, Skye	05	H
Mid Argyll, Lochgilphead	10	C
Monklands Hospital	12	L
Montrose Royal Infirmary	10	T
Ninewells	01	T
Perth Royal Infirmary	11	T
Peterhead Community Hospital	10	N

Hospital Name	Functional Class	Health Board
Portree, Skye	10	H
Princess Margaret Rose	14	S
Queen Margaret Hospital	02	F
Raigmore, Inverness	02	H
Randolph, Wemyss	12	F
RHSC, Edinburgh	07	S
RHSC, Yorkhill	07	G
Roodlands, Haddington	04	S
Ross Memorial, Dingwall	08	H
Rothesay Victoria Hospital	05	C
Royal Aberdeen Children's	07	N
Royal Alexandra Hospital	02	C
Seafield, Buckie	10	N
Southern General (SGH)	02	G
St. Andrew's Memorial	08	F
St. Johns at Howden	02	S
Stirling Royal Infirmary	11	V
Stobhill, Glasgow	02	G
Stonehouse	12	T
Stracathro	12	L
The Ayr Hopsital	12	A
Turner Memorial, Keith	10	N
Vale of Leven, Alexandria	11	C
Victoria Infirmary, Glasgow	02	G
Victoria Kircaldy	02	F
Western General, Edinburgh	01	S
Western Isles, Stornoway	11	W
Western/Gartnavel	01	G
Woodend General, Aberdeen	02	N

KEY: Functional Class and Health Board Codes are defined in appendix one.

Appendix 3B: Data Used for Analysis in Chapter Seven (Inputs)

DMU	CAP	AVE	TDC	TAC	TC
#1	73	1057	59679	32296	91975
#2	21.69	936	71314	33121	104435
#3	14.9	868	69489	24140	93629
#4	77.46	733	63496	26346	89842
#5	50.67	650	46399	23813	70212
#6	42.45	509	36051	23404	59455
#7	40.69	228	16693	7125	23818
#8	56.62	877	46236	24127	70363
#9	66.26	685	36218	18309	54527
#10	35.47	601	30156	15610	45766
#11	41.02	538	28820	15156	43976
#12	75	522	25459	18522	43981
#13	79.38	501	29744	10600	40344
#14	21.87	489	12301	9513	21814
#15	61	476	29597	12026	41623
#16	59.96	362	19832	11519	31351
#17	62.18	333	22815	10101	32916
#18	45.04	328	12181	6282	18463
#19	21.2	311	13309	6475	19784
#20	70	303	19598	11175	30773
#21	57.33	49	2121	1568	3689
#22	0	39	1885	913	2798
#23	24.71	138	3928	3558	7486
#24	92	110	4072	3231	7303
#25	71.78	129	3543	2262	5805
#26	24.25	90	2099	1672	3771
#27	58	73	3228	1744	4972
#28	42.48	59	2428	1456	3884
#29	83.73	57	3440	2279	5719
#30	78.05	28	2823	763	3586
#31	69.31	22	668	470	1138
#32	21.27	22	707	470	1177
#33	86.6	293	26954	12407	39361
#34	42	143	15517	7136	22653
#35	51	102	7347	4250	11597
#36	62.32	56	1085	582	1667
#37	37	55	1364	739	2103
#38	56.34	34	1568	1020	2588
#39	50.86	30	889	726	1615
#40	39.58	58	1455	969	2424
#41	71.87	55	1223	966	2189
#42	0.46	38	673	514	1187
#43	64.15	21	752	419	1171
#44	0.73	71	1249	863	2112
#45	88.95	68	1343	1068	2411
#46	22.51	62	1474	1171	2645
#47	14.62	47	923	773	1696
#48	47.8	44	1192	590	1782
#49	110.68	31	1049	1133	2182
#50	54.57	30	733	511	1244

DMU	CAP	AVE	TDC	TAC	TC
#51	0.39	30	477	381	858
#52	21.56	26	539	431	970
#53	75.61	15	302	193	495
#54	62.87	13	439	438	877
#55	71.87	11	260	148	408
#56	0	645	31229	11107	42336
#57	80.34	426	24889	8474	33363
#58	107.6	389	19158	11899	31057
#59	73.25	387	23730	9680	33410
#60	82.84	374	16066	11543	27609
#61	52.37	291	13948	8512	22460
#62	208.16	211	7629	5799	13428
#63	62.25	527	21178	14950	36128
#64	51	523	28332	14370	42702
#65	84.56	352	19240	11133	30373
#66	125.24	310	20194	11469	31663
#67	34.16	229	8642	5789	14431
#68	45.67	229	6815	4128	10943
#69	32.66	126	2945	2435	5380
#70	51.25	65	1222	940	2162
#71	44.59	186	7190	3553	10743
#72	48.79	178	2581	2778	5359
#73	36.92	51	1161	740	1901
#74	16.2	109	6972	2699	9671

Key to Factor Codes:

CAP: Capital Charge

AVE: Average Number of Staffed Beds

TAC: Total Allocated Costs (£000s)

TDC: Total Direct Costs (£000s)

TC: Total Costs (£000s)

Appendix 3B: Data Used for Analysis in Chapter Seven (Outputs)

DMU	IYS	IYM	IYG	IYO	TIY
#1	117624	133313	17236	4756	272929
#2	117034	152900	5547	2192	277673
#3	83072	157506	8495	4099	253172
#4	96903	114612	5882	3861	221258
#5	42714	107025	22440	8770	180949
#6	37085	90494	0	3983	131562
#7	45017	4088	0	1416	50521
#8	72532	139179	20287	5972	237970
#9	45546	117750	8192	25926	197414
#10	61176	68746	19918	6133	155973
#11	60320	78995	8353	2730	150398
#12	39648	47748	12502	40633	140531
#13	55619	57803	23059	5802	142283
#14	23026	82310	0	45901	151237
#15	70376	60622	6241	4840	142079
#16	35382	40858	8641	1987	86868
#17	35656	52490	0	0	88146
#18	10299	41713	14219	21292	87523
#19	16090	33002	0	29686	78778
#20	44581	26064	2496	11560	84701
#21	149	12063	11	3	12226
#22	3808	5917	0	56	9781
#23	7920	15528	807	13074	37329
#24	5409	12681	3236	3241	24567
#25	6639	10453	1041	16149	34282
#26	2912	4522	826	17445	25705
#27	5760	9276	569	1973	17578
#28	0	6352	1885	8586	16823
#29	6181	5603	843	6	12633
#30	4199	661	30	20	4910
#31	1398	0	44	2627	4069
#32	354	0	333	4451	5138
#33	33541	35851	0	4327	73719
#34	14570	16241	0	7380	38191
#35	9092	11964	0	628	21684
#36	0	0	0	18246	18246
#37	95	0	0	13031	13126
#38	171	4	13	8182	8370
#39	0	3322	0	5011	8333
#40	0	0	626	13552	14178
#41	0	0	262	12789	13051
#42	0	0	378	10247	10625
#43	771	1059	15	1887	3732
#44	0	0	477	21664	22141
#45	0	0	948	20356	21304
#46	0	3128	462	14876	18466
#47	0	2939	124	10388	13451
#48	0	0	1082	11696	12778
#49	0	0	952	6978	7930
#50	0	0	154	8218	8372

DMU	IYS	IYM	IYG	IYO	TIY
#51	0	0	170	7614	7784
#52	0	0	35	7416	7451
#53	0	0	239	3487	3726
#54	0	0	42	3249	3291
#55	0	0	470	1646	2116
#56	61949	84573	9630	19311	175463
#57	40282	56405	12140	3829	112656
#58	30095	49339	10036	15662	105132
#59	32471	60050	10430	2715	105666
#60	24872	53008	8684	15102	101666
#61	9931	33736	8793	29847	82307
#62	6726	22780	2785	20541	52832
#63	50519	58044	6475	40280	155318
#64	53216	58640	5599	25350	142805
#65	35877	58115	4559	3376	101927
#66	44685	36047	0	1095	81827
#67	18698	37584	0	6829	63111
#68	9977	23008	0	24735	57720
#69	6107	11355	0	17733	35195
#70	137	2632	0	14286	17055
#71	15315	0	0	37818	53133
#72	0	0	0	57605	57605
#73	0	0	0	12798	12798
#74	28814	0	0	0	28814

Key to Factor Codes:

IYS: Inpatient Days, Surgical

IYM: Inpatient Days, Medical

IYG: Inpatient Days, Obstetrics and Gynaecology

IYO: Inpatient Days, Other

TIY: Total Inpatient Days

Appendix 3B: Data Used for Analysis in Chapter Seven (Outputs)

DMU	IDS	IDM	IDG	IDO	TID
#1	24151	22241	4439	1343	52174
#2	29691	30455	1953	555	62654
#3	13826	21451	2643	327	38247
#4	18006	16745	2008	1255	38014
#5	8900	17197	7245	1728	35070
#6	8584	12375	0	1061	22020
#7	8488	406	0	782	9676
#8	11346	12355	5856	1725	31282
#9	8889	12856	2290	2116	26151
#10	11582	10209	5395	1010	28196
#11	12677	11164	2367	1087	27295
#12	8967	8117	5189	3258	25531
#13	10735	9757	6047	1374	27913
#14	3209	3274	0	338	6821
#15	12747	10540	2253	1920	27460
#16	6905	7011	3472	422	17810
#17	7154	9703	0	0	16857
#18	2488	3941	4801	283	11513
#19	5764	2638	0	910	9312
#20	7829	3817	563	1035	13244
#21	137	755	15	3	910
#22	946	904	0	29	1879
#23	1535	1697	305	69	3606
#24	1097	1344	793	61	3295
#25	997	30	297	658	1982
#26	394	67	299	848	1608
#27	1158	960	229	5	2352
#28	0	981	379	21	1381
#29	1208	813	298	7	2326
#30	790	95	19	4	908
#31	269	0	19	270	558
#32	55	0	134	543	732
#33	9376	9520	0	581	19477
#34	5626	6207	0	577	12410
#35	4199	3784	0	191	8174
#36	0	0	0	348	348
#37	68	0	0	258	326
#38	119	4	13	403	539
#39	0	220	0	224	444
#40	0	0	225	873	1098
#41	0	0	107	589	696
#42	0	0	109	274	383
#43	441	33	11	234	719
#44	0	0	182	261	443
#45	0	0	331	418	749
#46	0	152	162	1149	1463
#47	0	79	49	431	559
#48	0	0	289	112	401
#49	0	0	363	906	1269
#50	0	0	79	905	984

DMU	IDS	IDM	IDG	IDO	TID
#51	0	0	105	394	499
#52	0	0	12	304	316
#53	0	0	78	144	222
#54	0	0	18	335	353
#55	0	0	144	122	266
#56	10504	9843	3883	1646	25876
#57	7916	7931	3983	706	20536
#58	5550	6335	2919	992	15796
#59	5900	7182	3797	560	17439
#60	4374	6155	3055	823	14407
#61	2697	3444	2632	360	9133
#62	1285	1970	729	258	4242
#63	8922	7022	1825	1135	18904
#64	10888	10281	2073	2843	26085
#65	7795	9082	1285	800	18962
#66	11004	7402	0	477	18883
#67	3390	3829	0	4	7223
#68	1692	2314	0	173	4179
#69	250	982	0	472	1704
#70	42	152	0	79	273
#71	3286	0	0	899	4185
#72	0	0	0	267	267
#73	0	0	0	450	450
#74	2869	0	0	0	2869

Key to Factor Codes:

IDS: Inpatient Discharges, Surgical

IDM: Inpatient Discharges, Medical

IDG: Inpatient Discharges, Obstetrics and Gynaecology

IDO: Inpatient Discharges, Other

TID: Total Inpatient Discharges

Appendix 3B: Data Used for Analysis in Chapter Seven (Outputs)

DMU	COA	AEA	DCA	DPA	POA
#1	226057	78419	13610	9586	186993
#2	257572	62177	20091	18749	0
#3	265709	91626	12431	14717	0
#4	236585	74136	11317	10294	185916
#5	187745	4887	10764	16439	88388
#6	106569	18092	19664	2611	142780
#7	63329	46719	5758	0	64500
#8	141634	44627	13249	3353	176863
#9	125593	51429	10133	12687	120484
#10	105664	25335	7349	2052	94047
#11	111314	52902	8936	6728	126107
#12	93177	43437	14027	6840	134153
#13	128834	66457	15228	0	97713
#14	3627	0	909	4147	31233
#15	134960	73938	8831	0	113130
#16	81839	37070	4931	0	84334
#17	123241	36670	9751	1601	102597
#18	17886	43	3350	4980	34068
#19	15733	0	2477	4967	28040
#20	91295	35244	8718	5154	19912
#21	11222	4	932	2526	20472
#22	5948	9320	950	0	8446
#23	10701	6370	491	981	19937
#24	9464	6041	1263	0	10700
#25	8148	2986	567	0	14880
#26	6669	7540	207	2374	22569
#27	9215	8623	766	0	11098
#28	15469	12979	151	0	24092
#29	10702	7774	1092	0	20777
#30	2381	1093	336	0	2288
#31	2992	1249	284	0	3237
#32	2545	4849	126	0	11818
#33	79026	36481	4945	1490	41298
#34	41384	31733	3441	0	56886
#35	42554	21625	1376	2103	31653
#36	1742	2827	0	1166	10947
#37	4134	4354	1345	0	10132
#38	10903	6800	777	0	20019
#39	5047	3721	0	0	8246
#40	6201	10281	194	265	4470
#41	2080	3447	138	3243	4114
#42	2027	1695	0	0	7108
#43	1364	466	50	0	1381
#44	1677	2568	0	0	7315
#45	3938	14513	0	0	3976
#46	12726	5504	25	2043	5580
#47	4013	3280	0	0	12665
#48	5791	2449	0	0	6863
#49	4998	10169	0	0	9304
#50	6468	4459	0	2023	3863

DMU	COA	AEA	DCA	DPA	POA
#51	913	2946	41	395	2360
#52	860	861	0	0	3061
#53	0	911	0	356	46
#54	2837	1441	0	0	3078
#55	0	0	0	0	0
#56	94874	41287	9392	2439	99692
#57	86453	34902	8486	0	106777
#58	85584	26643	6586	3836	96152
#59	90394	42599	7966	0	134826
#60	35946	11634	6319	3043	32287
#61	37289	21402	3985	4933	32951
#62	19730	5789	1148	1895	12728
#63	81536	39739	6810	6582	72862
#64	121895	59293	11208	9831	110806
#65	58652	27796	5833	3295	37902
#66	82290	42230	11460	0	119026
#67	9627	5675	3439	2358	17419
#68	31945	4277	4872	0	27053
#69	10073	0	76	0	10850
#70	3335	0	482	3494	9690
#71	29493	0	3901	0	9864
#72	0	0	0	0	0
#73	6220	4344	0	0	8655
#74	21912	0	824	0	38424

Key to Factor Codes:

COA: Total Consultant Outpatient Attendances

AEA: Total A&E Attendances

DCA: Total Day Case Attendances

DPA: Total Day Patient Attendances

POA: Total PAM Outpatient Attendances

Appendix 3C: Results from Model 7.1

DMU	Efficiency Score	Rank	Reference Set	Reference Set Count
#1	35	96.0	2 10 13 15 22	0
#2	17 (1)	100.0	2	4
#3	25 (1)	100.0	3	1
#4	23 (1)	100.0	4	2
#5	15 (1)	100.0	5	5
#6	54	81.8	2 22 34 35 46	0
#7	10 (1)	100.0	7	6
#8	56	80.7	5 10 13 15 51 56	0
#9	49	86.2	5 13 17 22 35 46	0
#10	12 (1)	100.0	10	5
#11	32	99.3	10 15 22 35	0
#12	19 (1)	100.0	12	4
#13	5 (1)	100.0	13	14
#14	69	56.6	22 35 51	0
#15	9 (1)	100.0	15	6
#16	33	97.6	5 13 18 22 35	0
#17	14 (1)	100.0	17	5
#18	7 (1)	100.0	18	9
#19	44	90.0	7 22 35 51	0
#20	39	93.2	4 15 17 35 46	0
#21	53	82.9	35 46	0
#22	3 (1)	100.0	22	16
#23	43	91.2	18 35	0
#24	52	83.8	18 22 28 35	0
#25	65	71.7	12 13 35 43 51	0
#26	51	84.3	12 35 43 49 50 51	0
#27	58	78.6	13 22 35 49	0
#28	8 (1)	100.0	28	9
#29	59	78.3	13 35 49	0
#30	47	88.0	7 13 15	0
#31	41	91.9	35 43 50	0
#32	20 (1)	100.0	32	3
#33	50	86.2	2 17 22 34 35	0
#34	16 (1)	100.0	34	4
#35	1 (1)	100.0	35	29
#36	72	46.5	45 50	0
#37	68	60.4	22 28 45 50	0
#38	26 (1)	100.0	38	1
#39	55	80.9	28 35 46 50	0
#40	27 (1)	100.0	40	1
#41	71	46.6	18 49 50 51 55	0
#42	31 (1)	100.0	42	1
#43	18 (1)	100.0	43	4
#44	62	75.6	51 56	0
#45	11 (1)	100.0	45	5
#46	6 (1)	100.0	46	13
#47	66	70.6	28 45 46 50 51	0
#48	28 (1)	100.0	48	1
#49	13 (1)	100.0	49	5
#50	2 (1)	100.0	50	17

DMU	Efficiency Score	Rank	Reference Set	Reference Set Count
#51	4 (1)	100.0	51	15
#52	70	55.3	46 50 51	0
#53	64	74.0	32 45 50 55	0
#54	38	93.3	35 50	0
#55	21 (1)	100.0	55	3
#56	22 (1)	100.0	56	3
#57	29 (1)	100.0	57	1
#58	42	91.5	13 18 28 35 46	0
#59	48	87.4	5 13 34 35	0
#60	37	94.9	18 50 51 65	0
#61	61	76.8	13 18 28 32 51	0
#62	67	60.7	18 28 35 50	0
#63	40	93.2	10 12 22 35 51	0
#64	30 (1)	100.0	64	1
#65	24 (1)	100.0	65	2
#66	34	96.6	7 22 35	0
#67	46	88.7	22 35	0
#68	60	77.8	17 35 46	0
#69	57	78.7	35 46 50	0
#70	73	38.4	35 46	0
#71	36	96.0	7 35 50 51	0
#72	74	11.2	50 51	0
#73	63	74.9	13 22 28 46 50	0
#74	45	89.8	7 13	0

Appendix 3D: Results from Model 7.11

DMU	Efficiency Score	Rank	Reference Set	Reference Set Count
#1	93.7	27	15, 17, 22	0
#2	100	17 (1)	2	3
#3	100	11 (1)	3	5
#4	84.9	40	2, 3, 17, 22	0
#5	95.7	25	2, 3, 17	0
#6	100	19 (1)	6	1
#7	88.5	34	13, 15, 22, 28, 46	0
#8	74.6	50	15, 17, 22	0
#9	78.5	47	15, 17, 22, 35	0
#10	93.9	26	15, 22, 35	0
#11	97.2	24	15, 22, 35	0
#12	100	8 (1)	12	9
#13	100	4 (1)	13	16
#14	52.7	66	22, 35, 51	0
#15	100	7 (1)	15	11
#16	89.9	32	15, 22, 35	0
#17	100	5 (1)	17	15
#18	90.3	31	13, 22, 35	
#19	69.2	58	13, 22, 35	0
#20	92.2	29	17, 22, 46, 48	0
#21	83.5	41	38, 46, 50, 68	0
#22	100	1 (1)	22	28
#23	83	42	35, 51	0
#24	74.5	51	12, 50	0
#25	50.4	67	12, 22, 35, 50	0
#26	71.2	55	22, 35, 46, 50, 51	0
#27	67	60	13, 22, 35, 50	0
#28	100	14 (1)	28	4
#29	67.5	59	22, 37, 50, 66	0
#30	45.2	71	13	0
#31	87.6	36	50, 66, 68	0
#32	100	22 (1)	32	1
#33	71.7	54	13, 22, 35	0
#34	79.7	46	13, 22, 35	0
#35	100	2 (1)	35	26
#36	42.8	72	22, 45, 50	0
#37	100	11 (1)	37	4
#38	100	15 (1)	38	3
#39	70.1	56	45, 46, 50	0
#40	100	21 (1)	40	1
#41	48.5	69	22, 49, 50	0
#42	78.3	49	3, 17, 22	0
#43	81.3	44	13, 50	0
#44	41.2	73	22, 46, 51	0
#45	100	13 (1)	45	4
#46	100	6 (1)	46	12
#47	69.6	57	28, 45, 46, 51	0
#48	78.3	48	17, 50	0
#49	100	18 (1)	49	2
#50	100	3 (1)	50	24

DMU	Efficiency Score	Rank	Reference Set	Reference Set Count
#51	100	9 (1)	51	6
#52	47.9	70	35, 50	0
#53	57.7	63	13, 50	0
#54	73.2	53	50	0
#55	85.9	38	13, 50	0
#56	100	20 (1)	56	1
#57	92	30	13	0
#58	85.6	39	12, 13, 35, 46	0
#59	80.1	45	13, 15, 17, 35, 46	0
#60	86.6	37	12, 35, 50	0
#61	63.9	62	12, 13, 22, 35, 50	0
#62	49	68	12, 35, 50	0
#63	87.6	35	15, 17, 22, 35	0
#64	99.7	23	15, 17, 22, 35	0
#65	92.6	28	12, 22, 35	0
#66	100	16 (1)	66	3
#67	81.8	43	12, 22, 35, 50	0
#68	100	10 (1)	68	6
#69	52.8	65	35, 46, 50	0
#70	54.8	64	37, 38, 68	0
#71	88.6	33	13, 17, 37, 68	0
#72	9	74	35, 50	0
#73	74	52	17, 28, 46, 50	0
#74	66.7	61	3, 17	0

Appendix 3E: DMUs Excluded from the Analysis in Chapter Eight

DMU	Functional Class	Health Board	Reason for Exclusion
#14	2	N	No 'A&E Attendances'
#18	2	S	Very Small Number of 'A&E Attendances'
#19	2	S	No 'A&E Attendances'
#21	4	S	Very Small Number of 'A&E Attendances'
#22	4	Y	Missing Data for 'Capital Charge'
#36	8	T	No 'Day Case Attendances'
#39	8	H	No 'Day Case Attendances'
#42	9	N	No 'Day Case Attendances'
#43	9	W	Very Small Number of 'Day Case Attendances'
#44	10	N	No 'Day Case Attendances'
#45	10	N	No 'Day Case Attendances'
#47	10	C	No 'Day Case Attendances'
#48	10	T	No 'Day Case Attendances'
#49	10	N	No 'Day Case Attendances'
#50	10	A	No 'Day Case Attendances'
#51	10	N	Very Small Number of 'Day Case Attendances'
#52	10	C	No 'Day Case Attendances'
#53	10	N	No 'Day Case Attendances'
#54	10	H	No 'Day Case Attendances'
#55	10	N	No 'Day Case Attendances', 'A&E Attendances' or 'Consultant Outpatient Attendances'
#56	11	L	Missing Data for 'Capital Charge'
#69	12	A	No 'A&E Attendances' and Small Number of 'Day Case Attendances'
#70	12	F	No 'A&E Attendances'
#71	13	G	No 'A&E Attendances'
#72	13	N	No 'Day Case Attendances', 'A&E Attendances' or 'Consultant Outpatient Attendances'
#73	13	T	No 'Day Case Attendances'
#74	14	S	No 'A&E Attendances'

KEY: Functional Class and Health Board Codes are defined in appendix one.

APPENDIX FOUR

DATA AND ANALYSIS

USED IN CHAPTER EIGHT

Appendix 4A: Hospitals in Refined Sample (Model 8.1)

Hospital Name	Functional Class	Health Board
Aberdeen Royal Infirmary	01	N
Adamson, Cupar	08	F
Arbroath Infirmary	05	T
Arran War Memorial	05	A
Balfour, Kirkwall	05	H
Belford, Fort William	05	H
Borders General	11	B
Caithness General, Wick	05	R
Campletown Hospital	10	C
Chalmers, Banff	09	N
Crosshouse Hospital	02	A
D&G Royal Infirmary	12	Y
Dundee Royal Infirmary	01	T
Dunoon & District General	05	C
Edinburgh Royal Infirmary	01	S
Falkirk Royal Infirmary	11	V
Gilbert Bain, Lerwick	05	Z
Glasgow Royal Infirmary	01	G
Hairmyres, East Kilbride	12	L
Huntly Jubilee	09	N
Inverclyde Royal Hopsital	02	C
Lorn & Islands District General	05	C
Mackinnon Memorial, Skye	05	H
Monklands Hospital	12	L
Ninewells	01	T
Perth Royal Infirmary	11	T
Queen Margaret Hospital	02	F
Raigmore, Inverness	02	H
RHSC, Edinburgh	07	S
RHSC, Yorkhill	07	G
Rothesay Victoria Hospital	05	C
Royal Aberdeen Children's	07	N
Royal Alexandra Hospital	02	C
Southern General (SGH)	02	G
St. Andrew's Memorial	08	F
St. Johns at Howden	02	S
Stirling Royal Infirmary	11	V
Stobhill, Glasgow	02	G
Stonehouse	12	T
Stracathro	12	L
The Ayr Hopsital	12	A
Vale of Leven, Alexandria	11	C
Victoria Infirmary, Glasgow	02	G
Victoria Kircaldy	02	F
Western General, Edinburgh	01	S
Western Isles, Stornoway	11	W
Western/Gartnavel	01	G

Appendix 4B: Data Used for Analysis in Chapter Eight

DMU (Code in Chapter Seven)	CAP	TDC	TAC	TID	COA	AEA	DCA
*1 (#1)	73	59679	32296	52174	226057	78419	13610
*2 (#2)	21.69	71314	33121	62654	257572	62177	20091
*3 (#3)	14.9	69489	24140	38247	265709	91626	12431
*4 (#4)	77.46	63496	26346	38014	236585	74136	11317
*5 (#5)	50.67	46399	23813	35070	187745	4887	10764
*6 (#6)	42.45	36051	23404	22020	106569	18092	19664
*7 (#7)	40.69	16693	7125	9676	63329	46719	5758
*8 (#8)	56.62	46236	24127	31282	141634	44627	13249
*9 (#9)	66.26	36218	18309	26151	125593	51429	10133
*10 (#10)	35.47	30156	15610	28196	105664	25335	7349
*11 (#11)	41.02	28820	15156	27295	111314	52902	8936
*12 (#12)	75	25459	18522	25531	93177	43437	14027
*13 (#13)	79.38	29744	10600	27913	128834	66457	15228
*14 (#15)	61	29597	12026	27460	134960	73938	8831
*15 (#16)	59.96	19832	11519	17810	81839	37070	4931
*16 (#17)	62.18	22815	10101	16857	123241	36670	9751
*17 (#20)	70	19598	11175	13244	91295	35244	8718
*18 (#23)	24.71	3928	3558	3606	10701	6370	491
*19 (#24)	92	4072	3231	3295	9464	6041	1263
*20 (#25)	71.78	3543	2262	1982	8148	2986	567
*21 (#26)	24.25	2099	1672	1608	6669	7540	207
*22 (#27)	58	3228	1744	2352	9215	8623	766
*23 (#28)	42.48	2428	1456	1381	15469	12979	151
*24 (#29)	83.73	3440	2279	2326	10702	7774	1092
*25 (#30)	78.05	2823	763	908	2381	1093	336
*26 (#31)	69.31	668	470	558	2992	1249	284
*27 (#32)	21.27	707	470	732	2545	4849	126
*28 (#33)	86.6	26954	12407	19477	79026	36481	4945
*29 (#34)	42	15517	7136	12410	41384	31733	3441
*30 (#35)	51	7347	4250	8174	42554	21625	1376
*31 (#37)	37	1364	739	326	4134	4354	1345
*32 (#38)	56.34	1568	1020	539	10903	6800	777
*33 (#40)	39.58	1455	969	1098	6201	10281	194
*34 (#41)	71.87	1223	966	696	2080	3447	138
*35 (#46)	22.51	1474	1171	1463	12726	5504	25
*36 (#57)	80.34	24889	8474	20536	86453	34902	8486
*37 (#58)	107.6	19158	11899	15796	85584	26643	6586
*38 (#59)	73.25	23730	9680	17439	90394	42599	7966
*39 (#60)	82.84	16066	11543	14407	35946	11634	6319
*40 (#61)	52.37	13948	8512	9133	37289	21402	3985
*41 (#62)	208.16	7629	5799	4242	19730	5789	1148
*42 (#63)	62.25	21178	14950	18904	81536	39739	6810
*43 (#64)	51	28332	14370	26085	121895	59293	11208
*44 (#65)	84.56	19240	11133	18962	58652	27796	5833
*45 (#66)	125.24	20194	11469	18883	82290	42230	11460
*46 (#67)	34.16	8642	5789	7223	9627	5675	3439
*47 (#68)	45.67	6815	4128	4179	31945	4277	4872

Key To Factor Codes
CAP: Capital Charge
TAC: Total Allocated Costs (£000s)
TDC: Total Direct Costs (£000s)
TID: Total Inpatient Discharges
COA: Consultant Outpatient Attendances
AEA: A&E Attendances
DCA: Day Case Attendances

Appendix 4C: Results from Model 8.1

DMU	Efficiency Score	Ranking	Reference Set Count	Reference Group
*1	95.2	23	0	2 11 14
*2	100	7(1)	6	2
*3	100	10(1)	3	3
*4	85	32	0	2 3 14 16
*5	95.7	21	0	2 3 16
*6	100	16(1)	2	6
*7	100	17(1)	1	7
*8	77.5	38	0	2 6 13 16
*9	79	37	0	2 14 16 43
*10	100	20(1)	1	10
*11	100	6(1)	7	11
*12	100	2(1)	13	12
*13	100	3(1)	12	13
*14	100	4(1)	10	14
*15	91.1	26	0	11 14 30
*16	100	5(1)	7	16
*17	93.8	24	0	13 16 35 47
*18	83.9	33	0	12 30
*19	77.2	39	0	12 30
*20	52	45	0	12 30
*21	86.2	30	0	14 30 33
*22	71.5	41	0	12 13 27 30
*23	100	18(1)	1	23
*24	71	42	0	27 31 35 45
*25	45.2	47	0	13
*26	90.1	28	0	35 45 47
*27	100	9(1)	4	27
*28	73.8	40	0	11 13 30
*29	83.3	34	0	11 14 30
*30	100	1(1)	18	30
*31	100	14(1)	2	31
*32	100	19(1)	1	32
*33	100	15(1)	2	33
*34	53.8	44	0	12 27 30
*35	100	8(1)	6	35
*36	92	25	0	13
*37	85.6	31	0	12 13 30 35
*38	80.1	36	0	13 14 16 30 35
*39	86.9	29	0	12 30
*40	66	43	0	12 13 14 30
*41	51.4	46	0	12 30
*42	90.3	27	0	11 12 14 30 43
*43	100	11(1)	3	43
*44	95.2	22	0	11 12 13 30
*45	100	12(1)	3	45
*46	82.2	35	0	12 13 30
*47	100	13(1)	3	47

Appendix 4D: Results from Weight Restriction Scenarios

	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5	
DMU	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
*1	84.1	12	82.9	15	79.9	13	84.7	16	84.9	17
*2	100	3(1)	100	2(1)	100	2(1)	100	5(1)	100	5(1)
*3	100	4(1)	100	4(1)	100	4(1)	100	6(1)	100	7(1)
*4	72.2	23	70.8	27	70.7	20	70.8	29	71.2	30
*5	21.2	46	53	38	51.4	37	53.9	38	86.3	16
*6	62.3	30	95.7	9	81	12	89.4	11	96.8	11
*7	85.7	10	85.5	12	85.4	9	86.1	14	87.4	15
*8	72.6	22	75.4	23	72.6	19	76.6	23	77.3	26
*9	74.9	21	75.7	22	73.7	17	76.4	24	77.1	27
*10	76.7	17	86.2	11	81.3	11	86	15	88.2	14
*11	96.5	6	97.1	7	96.6	6	97.6	8	98.1	10
*12	91.7	8	97.6	6	92.6	7	100	4(1)	100	3(1)
*13	100	1(1)	100	1(1)	100	1(1)	100	1(1)	100	1(1)
*14	97.4	5	97.9	5	97.9	5	97.9	7	98.6	8
*15	76.2	18	72.7	24	69	22	76.3	25	76.4	28
*16	88.9	9	93.4	10	91.5	8	95.4	9	98.2	9
*17	85	11	84	13	76.6	15	88.5	13	89.7	13
*18	50.6	38	48	39	37.5	40	49.7	39	49.9	39
*19	57.6	33	64.2	33	49.9	38	69.4	30	69.4	33
*20	37.6	42	41	41	34.3	42	45.4	41	46.1	41
*21	46.1	39	40.7	42	35.5	41	42.2	43	42.8	43
*22	67.5	28	64.8	31	59.3	29	67.5	33	67.9	34
*23	31.3	44	26	46	25.5	46	26.1	46	26.2	46
*24	68.2	26	64.5	32	56.2	32	71	28	71	31
*25	23.7	45	33.1	45	28.7	45	35.4	45	36.5	45
*26	71.6	25	78.4	18	65.5	25	89.5	10	90.1	12
*27	67	29	60.8	35	56	33	61.2	36	61.9	37
*28	57.3	34	56.6	37	52.6	36	57.5	37	57.7	38
*29	67.6	27	67.6	29	67.3	23	69.2	31	70.5	32
*30	80	15	70.9	26	61.9	27	73.7	27	73.7	29
*31	60.8	32	61.4	34	53.2	35	63.9	35	63.9	36
*32	82.4	13	70.5	28	60.4	28	78.5	21	78.5	23
*33	53.7	37	46.4	40	44.2	39	46.9	40	47.2	40
*34	37.8	41	35.1	44	30.3	44	36.9	44	37.3	44
*35	9.7	47	7.7	47	7.4	47	8	47	8	47
*36	75.5	20	81.2	16	79	14	82.2	19	83.3	19
*37	75.5	19	77.1	20	66.2	24	83.1	17	84.6	18
*38	77.7	16	76.4	21	73.4	18	78.1	22	78.2	24
*39	54.3	36	71.7	25	58.4	30	75.2	26	78.7	22
*40	62.3	31	60.6	36	54.8	34	64.5	34	64.5	35
*41	34	43	39.4	43	32.1	43	43.5	42	45.8	42
*42	80.5	14	78.6	17	74.1	16	82.3	18	82.4	20
*43	100	2(1)	100	3(1)	100	3(1)	100	3(1)	100	6(1)
*44	72	24	77.5	19	64.1	26	78.7	20	78.9	21
*45	96	7	96	8	82.8	10	100	2(1)	100	4(1)
*46	43.2	40	65	30	57.1	31	68.7	32	77.9	25
*47	57.1	35	83.8	14	70.2	21	88.9	12	100	2(1)