Contents

Preface *ix*
Acknowledgments *xiii*

Introduction **1**

The General Setting 1
The General Problem 2
The Proposed Solution 3
Four Types of Evaluation 4
General Issues in Test Design 6
Some Application Vignettes 7
Organization of This Book 11

PART I
Quantitative Methods **13**

CHAPTER 1
Fundamentals of Accuracy Analysis **15**

1.1 Two Aspects of Performance 16
1.2 Analysis of Performance via the ROC Curve 19
1.3 Often Used, but Inadequate, Indices of Accuracy 24

1.4 Recommended Indices of Accuracy 29
1.5 A Computational Example 34
1.6 A Suitable Accuracy Index Based on One ROC Point 37
1.7 Indices of the Decision Criterion 39
1.8 Summary 44

CHAPTER 2
Extensions of Accuracy Analysis 46

2.1 Multiple Alternatives 46
2.2 Other Interpretations of ROC Data 54
2.3 Combining Individual Results into a Group or Average ROC 64
2.4 Summary 66

CHAPTER 3
Statistical Design of a Performance Test 68

3.1 Questions to Be Answered by the Reading Test 69
3.2 Logistic and Cost Constraints 70
3.3 General Approach to Statistical Analysis of Accuracy Data 71
3.4 The Degree of Statistical Power Required 76
3.5 Some Practical Rules for Study Design 77
3.6 Summary 79

CHAPTER 4
Statistical Treatment of Accuracy Results 80

4.1 Standard Error of the Accuracy Index 81
4.2 Confidence Interval of the Accuracy Index 84
4.3 Statistical Significance of a Difference between Modalities 86
4.4 Statistical Analysis of True-Positive Probability at a Fixed False-Positive Probability . 90
4.5 Assumptions Underlying the Recommended Statistical Procedure 91
4.6 Tests Comparing More Than Two Modalities 92
4.7 Summary 93

CHAPTER 5
Forms of Efficacy Analysis 94

5.1 Cost-Benefit Analysis: A Conceptual Starting Point 95
5.2 Impediments to Cost-Benefit Analysis 99
5.3 Tailoring the Analysis to Suit the Situation 103
5.4 Setting the Efficacy Analysis at an Appropriate Level 106
5.5 Summary 107

PART II
Experimental Methods 109

CHAPTER 6
Elements of Study Design 111

6.1 Developing Study Goals 112
6.2 Outlining the End Product 113
6.3 Planning the Collection of Data 116
6.4 Summary 117

CHAPTER 7
Procedure of a Performance Test 118

7.1 Case and Image Presentation 118
7.2 Response Format 119
7.3 Background Information Given Readers on Each Case 127
7.4 Timing of Trials and Sessions 131
7.5 Viewing Environment 132
7.6 Role of Test Administrator 132
7.7 Reader Orientation 133
7.8 Feedback to and from Readers 135
7.9 Summary 135

CHAPTER 8
Drawing the Original Case Sample 137

8.1 Functions of the Original Case Sample 137
8.2 Types of Data Collected on Each Case 138
8.3 Selecting a Representative Case Sample 140
8.4 Bias in Case Data 141
8.5 Required Size of the Original Case Sample 142
8.6 Summary 143

CHAPTER 9
Selection of Test Cases 144

9.1 Number of Cases for Adequate Representation 144
9.2 Proportional Representation of Case Types 145
9.3 Screening Out Inadequate Cases 147
9.4 Checking Representativeness of the Reading-Test Sample 150
9.5 Summary 150

CHAPTER 10
Selection of Test Readers 152

10.1 Number of Readers for Adequate Representation 152
10.2 Type of Reader 153
10.3 Summary 155

PART III
A Methodological Case Study 157

CHAPTER 11
Illustrative Evaluation of Imaging Systems 159

11.1 Case Selection 161
11.2 Response Format 164
11.3 Experimental Design 167
11.4 Test Procedures 168
11.5 Reader Ratings of Image Quality 169
11.6 Accuracy Results 170
11.7 Correlational Analyses 178
11.8 Obtaining Estimates of Pure Variance and Correlation Terms 181
11.9 Discussion: Sizing a Final Study from a Pilot Study 183

Appendixes 189

Appendix A
Sequence of Steps in Evaluation 191

Appendix B
Representativeness of a Performance-Test Sample 196

Appendix C
Illustrative Posttest Questionnaire 203

Appendix D
Computer Program for Accuracy Analysis 208

Appendix E
References to ROC Applications 233

References 243
Index 247