Contents

Authors xiii
Preface xvii

Section I Introduction

1. Background and Introduction 3
1.1 Background Information 3
1.2 Gold Standard, Decision Threshold, Sensitivity, and Specificity 4
1.3 Kappa Statistics 5
1.4 Receiver Operating Characteristic Curve 6
1.5 Area and Partial Area under ROC Curve 10
1.6 Confidence Intervals, Regions, and Bands 11
1.7 Point of Intersection and Youden Index 12
1.8 Comparison of Two or More ROC Curves 13
1.9 Approaches to ROC Analysis 14
References 15

Section II Methods for Univariate and Multivariate Data

2. Diagnostic Rating Scales 21
2.1 Introduction 21
2.1.1 Frequent Need for Finer Scale 22
2.2 Interpreter-Free Diagnostic Systems 23
2.3 Human Interpreter as Integral Part of Diagnostic System 24
2.4 Remarks and Further Reading 25
References 26

3. Monotone Transformation Models 29
3.1 Introduction 29
3.2 General Assumptions 29
3.3 Empirical Methods 30
3.4 Nonparametric Kernel Smoothing 31

3.5 Parametric Models and Monotone Transformations to Binormal Distributions 34
3.6 Confidence Intervals 41
3.6.1 Sensitivity or Specificity at Specific Threshold 41
3.6.2 Confidence Intervals for Sensitivity (Specificity) at Given Specificity (Sensitivity) 46
3.6.3 Confidence Intervals for Area under Curve (AUC) 47
3.7 Concordance Measures in Presence of Monotone Transformations 51
3.8 Intraclass Correlation Coefficient 57
3.9 Remarks and Further Reading 59
References 59

4. Combination and Pooling of Biomarkers 65
4.1 Introduction 65
4.2 Combining Biomarkers to Improve Diagnostic Accuracy 66
4.2.1 Likelihood Ratio Approach 66
4.2.2 Uniformly Optimal Linear Combinations 67
4.2.3 Normal Linear Combinations Maximizing AUC 68
4.2.4 Distribution-Free Approach for Linear Combinations 70
4.3 ROC Curve Analysis with Pooled Samples 72
4.3.1 Normal Distributions 73
4.3.2 Gamma Distributions 75
4.3.3 Stable Distributions 77
4.3.4 Distribution-Free Approach 78
4.4 Remarks and Further Reading 79
4.4.1 Combination of Biomarkers 79
4.4.2 Pooling of Biomarkers 80
References 80

5. Bayesian ROC Methods 83
5.1 Introduction 83
5.2 Methods for Sensitivity, Specificity, and Prevalence 84
5.3 Clustered Data Structures and Hierarchical Methods 89
5.4 Assumptions and Models for ROC Analysis 92
5.5 Normality Transformation 93
5.6 Elicitation of Prior Information 95
5.7 Estimation of ROC Parameters and Characteristics 96
5.9 Remarks and Further Reading 96
References 98

Section III Advanced Approaches and Applications

6. Sequential Designs of ROC Experiments 103
6.1 Introduction 103
6.2 Group Sequential Tests Using Large Sample Theory 104
6.2.1 Approximating Test Statistics via Brownian Motion 104
6.2.2 Group Sequential Testing Procedures 106
6.2.3 Choosing Stopping Boundaries 106
6.3 Sequential Evaluation of Single ROC Curve 110
6.3.1 Binormal Model 110
6.3.2 Nonparametric Model 112
6.4 Sequential Comparison of Two ROC Curves 114
6.4.1 Binormal Model 115
6.4.2 Nonparametric Model 116
6.5 Sequential Evaluation of Binary Outcomes 118
6.5.1 Evaluating Accuracy of Diagnostic Tests 119
6.5.2 Comparison of Two Diagnostic Tests 121
6.6 Sample Size Estimation 123
6.7 Remarks and Further Reading 124
6.7.1 Inference upon Termination 124
6.7.2 Sample Size Re-Estimation 125
6.7.3 Computing Software 125
References 125

7. Multireader ROC Analysis 129
7.1 Introduction 129
7.2 Overall ROC Curve and Its AUC 130
7.3 Statistical Analysis of Cross-Correlated Multireader Data 132
7.3.1 Covariate-Free Approaches 133
7.3.1.1 Methods for Fixed Reader Inferences 133
7.3.1.2 Methods for Random Reader or Fixed Reader Inferences 136
7.3.1.3 Variance Components Estimation 141
7.3.2 Covariate Adjustment Approaches 144
7.3.2.1 Models for ROC Curves 145
7.3.2.2 Models for Summary Indices 147
7.3.2.3 Bayesian Approaches 148
7.4 Remarks and Further Reading 150
References 152
Appendix 7.A: Closed Form Formulation of DBM Approach for Comparing Two Modalities Using Empirical AUC 155
Appendix 7.B: Variance Estimators of Empirical AUCs 158

8. Free-Response ROC Analysis 163
8.1 Introduction 163
8.2 FROC Approach 165
8.2.1 Rating-Free FROC Analysis 165
8.2.2 FROC Space and Use of Guessing Process for Interpretation and Comparison 168
8.2.3 FROC Curve 171
8.2.4 Rating-Based FROC Analysis 172
8.2.5 Estimation of FROC Curve 175
8.2.6 Summary Indices 179
8.2.7 Comparison of FROC Performance Levels 184
8.3 Other Approaches of Detection–Localization Performance Assessment 187
8.3.1 Alternative FROC (AFROC) 187
8.3.2 Localization ROC (LROC) 188
8.3.3 Regions of Interest (ROI) Approach 191
8.4 Remarks and Further Reading 196
References 197

9. Machine Learning and Predictive Modeling 201
9.1 Introduction 201
9.2 Predictive Modeling 202
9.3 Cross-Validation 203
9.4 Bootstrap Resampling Methods 204
9.4.1 Bootstrap Method 1 205
9.4.2 Bootstrap Method 2 205
9.5 Overfitting and False Discovery Rate 206
9.6 Remarks and Further Reading 207
References 208

Section IV Discussions and Extensions

10. Summary and Challenges 213
10.1 Summary and Discussion 213
10.2 Future Directions in ROC Analysis 213
10.3 Future Directions in Reliability Analysis 215
10.4 Final Remarks 217

Appendix: Notation List 219

Index 223