# Contents

Preface ................................................................. vii  
Acknowledgements ................................................ xi  

1 Introduction ....................................................... 1  
1.1 Prognosis and Prediction in Medicine ......................... 1  
1.1.1 Prediction Models and Decision-Making ................. 1  
1.2 Statistical Modelling for Prediction .......................... 2  
1.2.1 Model Uncertainty .......................................... 3  
1.2.2 Sample Size .................................................. 4  
1.3 Structure of the Book ........................................... 5  
1.3.1 Part I: Prediction Models in Medicine .................... 5  
1.3.2 Part II: Developing Valid Prediction Models .......... 6  
1.3.3 Part III: Generalizability of Prediction Models ........ 6  
1.3.4 Part IV: Applications ...................................... 7  
1.3.5 Questions and Exercises ................................... 7  

Part I Prediction Models in Medicine  

2 Applications of Prediction Models ................................. 9  
2.1 Applications: Medical Practice and Research ............... 11  
2.2 Prediction Models for Public Health .......................... 12  
2.2.1 Targeting of Preventive Interventions ................. 12  
2.2.2 Example: Incidence of Breast Cancer ................. 12  
2.3 Prediction Models for Clinical Practice ....................... 13  
2.3.1 Decision Support on Test Ordering ...................... 13  
2.3.2 Example: Predicting Renal Artery Stenosis ............ 14  
2.3.3 Starting Treatment: the Treatment Threshold .......... 15  
2.3.4 Example: Probability of Deep Venous Thrombosis .... 16  
2.3.5 Intensity of Treatment .................................... 16  
2.3.6 Example: Defining a Poor Prognosis Subgroup in Cancer 18
2.3.7 Cost-Effectiveness of Treatment
2.3.8 Delaying Treatment
2.3.9 Example: Spontaneous Pregnancy Chances
2.3.10 Surgical Decision-Making
2.3.11 Example: Replacement of Risky Heart Valves

2.4 Prediction Models for Medical Research
2.4.1 Inclusion and Stratification in an RCT
2.4.2 Example: Selection for TBI Trials
2.4.3 Covariate Adjustment in an RCT
2.4.4 Gain in Power by Covariate Adjustment
2.4.5 Example: Analysis of the GUSTO-III Trial
2.4.6 Prediction Models and Observational Studies
2.4.7 Propensity Scores
2.4.8 Example: Statin Treatment Effects
2.4.9 Provider Profiling
2.4.10 Example: Ranking Cardiac Outcome

2.5 Concluding Remarks

3 Study Design for Prediction Models
3.1 Study Design
3.2 Cohort Studies for Prognosis
3.2.1 Retrospective Designs
3.2.2 Example: Predicting Early Mortality in Oesophageal Cancer
3.2.3 Prospective Designs
3.2.4 Example: Predicting Long-Term Mortality in Oesophageal Cancer
3.2.5 Registry Data
3.2.6 Example: Surgical Mortality in Oesophageal Cancer
3.2.7 Nested Case–Control Studies
3.2.8 Example: Perioperative Mortality in Major Vascular Surgery

3.3 Studies for Diagnosis
3.3.1 Cross-Sectional Study Design and Multivariable Modelling
3.3.2 Example: Diagnosing Renal Artery Stenosis
3.3.3 Case–Control Studies
3.3.4 Example: Diagnosing Acute Appendicitis

3.4 Predictors and Outcome
3.4.1 Strength of Predictors
3.4.2 Categories of Predictors
3.4.3 Costs of Predictors
3.4.4 Determinants of Prognosis
3.4.5 Prognosis in Oncology
3.5 Reliability of Predictors ........................................ 42
  3.5.1 Observer Variability ........................................ 42
  3.5.2 Example: Histology in Barrett’s Oesophagus ............ 42
  3.5.3 Biological Variability ...................................... 43
  3.5.4 Regression Dilution Bias .................................. 43
  3.5.5 Example: Simulation Study on Reliability of a
    Binary Predictor ............................................. 43
  3.5.6 Choice of Predictors ...................................... 44

3.6 Outcome .......................................................... 44
  3.6.1 Types of Outcome ........................................... 44
  3.6.2 Survival Endpoints ........................................ 45
  3.6.3 Example: Relative Survival in Cancer Registries ........ 45
  3.6.4 Composite End Points ...................................... 46
  3.6.5 Example: Mortality and Composite End Points
    in Cardiology ................................................. 46
  3.6.6 Choice of Prognostic Outcome ............................... 46
  3.6.7 Diagnostic End Points ..................................... 47
  3.6.8 Example: PET Scans in Oesophageal Cancer .............. 47

3.7 Phases of Biomarker Development ............................... 47

3.8 Statistical Power .................................................. 48
  3.8.1 Statistical Power to Identify Predictor Effects ........... 49
  3.8.2 Examples of Statistical Power Calculations .............. 49
  3.8.3 Statistical Power for Reliable Predictions ................ 50

3.9 Concluding Remarks ............................................ 51

4 Statistical Models for Prediction .................................. 53
  4.1 Continuous Outcomes .......................................... 53
    4.1.1 Examples of Linear Regression ............................ 54
    4.1.2 Economic Outcomes ...................................... 54
    4.1.3 Example: Prediction of Costs ............................ 54
    4.1.4 Transforming the Outcome ................................ 54
    4.1.5 Performance: Explained Variation ....................... 55
    4.1.6 More Flexible Approaches ................................. 55
  4.2 Binary Outcomes ................................................ 57
    4.2.1 $R^2$ in Logistic Regression Analysis .................... 58
    4.2.2 Calculation of $R^2$ on the Log Likelihood Scale ....... 58
    4.2.3 Models Related to Logistic Regression ................... 60
    4.2.4 Bayes Rule ............................................... 61
    4.2.5 Example: Calculations with Likelihood Ratios .......... 62
    4.2.6 Prediction with Naïve Bayes ............................. 63
    4.2.7 Examples of Naïve Bayes ................................ 65
    4.2.8 Calibration and Naïve Bayes ............................... 65
    4.2.9 Logistic Regression and Bayes ............................ 65
    4.2.10 More Flexible Approaches to Binary Outcomes ........ 65
4.2.11 Classification and Regression Trees .......................... 67
4.2.12 Example: Mortality in Acute MI Patients .................. 67
4.2.13 Advantages and Disadvantages of Tree Models .......... 67
4.2.14 Trees as Special Cases of Logistic Regression Modelling .................................................. 69
4.2.15 Other Methods for Binary Outcomes ...................... 70
4.2.16 Summary on Binary Outcomes .............................. 71
4.3 Categorical Outcomes ............................................. 71
4.3.1 Polytomous Logistic Regression .............................. 72
4.3.2 Example: Histology of Residual Masses ..................... 72
4.3.3 Alternative Models ............................................ 73
4.3.4 Comparison of Modelling Approaches ....................... 74
4.4 Ordinal Outcomes .................................................. 74
4.4.1 Proportional Odds Logistic Regression ....................... 75
4.4.2 Alternative: Continuation Ratio Model ....................... 77
4.5 Survival Outcomes ................................................ 77
4.5.1 Cox Proportional Hazards Regression ....................... 77
4.5.2 Predicting with Cox ............................................ 78
4.5.3 Proportionality Assumption ................................... 78
4.5.4 Kaplan–Meier Analysis ....................................... 79
4.5.5 Example: NFI After Treatment of Leprosy .................. 79
4.5.6 Parametric Survival .......................................... 80
4.5.7 Example: Replacement of Risky Heart Valves ............. 80
4.5.8 Summary on Survival Outcomes ............................. 81

4.6 Concluding Remarks .............................................. 81

5 Overfitting and Optimism in Prediction Models .................. 83
5.1 Overfitting and Optimism ........................................ 83
5.1.1 Example: Surgical Mortality in Oesophagectomy .......... 84
5.1.2 Variability within One Centre ............................... 84
5.1.3 Variability between Centres: Noise vs. True Heterogeneity .................................................. 85
5.1.4 Predicting Mortality by Centre: Shrinkage ......... 87
5.2 Overfitting in Regression Models ............................... 87
5.2.1 Model Uncertainty: Testimation ......................... 87
5.2.2 Other Biases ................................................. 89
5.2.3 Overfitting by Parameter Uncertainty ..................... 90
5.2.4 Optimism in Model Performance ......................... 90
5.2.5 Optimism-Corrected Performance ...................... 92
5.3 Bootstrap Resampling ............................................ 92
5.3.1 Applications of the Bootstrap ......................... 93
5.3.2 Bootstrapping for Regression Coefficients ................ 93
5.3.3 Bootstrapping for Optimism Correction ................ 94
5.3.4 Calculation of Optimism-Corrected Performance .... 95
5.3.5 Example: Stepwise Selection in 429 Patients 96
5.4 Cost of Data Analysis ................................. 97
5.4.1 Example: Cost of Data Analysis in a Tree Model 98
5.4.2 Practical Implications .......................... 98
5.5 Concluding Remarks ................................. 99

6 Choosing Between Alternative Statistical Models ................. 101
6.1 Prediction with Statistical Models ....................... 101
6.1.1 Testing of Model Assumptions and Prediction ........ 102
6.1.2 Choosing a Type of Model .......................... 102
6.2 Modelling Age–Outcome Relationships .................... 103
6.2.1 Age and Mortality After Acute MI .................... 103
6.2.2 Age and Operative Mortality ......................... 103
6.2.3 Age–Outcome Relationships in Other Diseases ........ 106
6.3 Head-to-Head Comparisons .............................. 107
6.3.1 StatLog Results ................................. 107
6.3.2 GUSTO-I Modelling Comparisons .................. 108
6.3.3 GUSTO-I Results ................................. 109
6.4 Concluding Remarks .................................. 110

Part II Developing Valid Prediction Models

7 Dealing with Missing Values ............................... 113
7.1 Missing Values in Predictors ............................ 115
7.1.1 Inefficiency of Complete Case Analysis ............ 116
7.1.2 Interpretation of Analyses with Missing Data ........ 117
7.1.3 Missing Data Mechanisms .......................... 117
7.1.4 Summary Points .................................. 118
7.2 Regression Coefficients Under MCAR, MAR, and MNAR .. 118
7.2.1 R Code ........................................ 120
7.3 Missing Values in Regression Analysis ................... 121
7.3.1 Imputation Principle ................................ 121
7.3.2 Simple and More Advanced Single Imputation Methods .................................. 122
7.3.3 Multiple Imputation ................................ 123
7.4 Defining the Imputation Model ............................ 124
7.4.1 Transformations of Variables ......................... 125
7.4.2 Imputation Models for SI .......................... 125
7.4.3 Summary Points .................................. 126
7.5 Simulations of Imputation Under MCAR, MAR, and MNAR . 126
7.5.1 Multiple Predictors ................................ 127
7.6 Imputation of Missing Outcomes ........................ 128
7.7 Guidance to Missing Values in Prediction Research .......... 129
7.7.1 Patterns of Missingness ................................. 129
7.7.2 Simple Approaches ..................................... 130
7.7.3 Maximum Fraction of Missing Values Before Omitting a Predictor ..................................... 131
7.7.4 Single or Multiple Imputation for Predictor Effects? ..................................................... 131
7.7.5 Single or Multiple Imputation for Predictions? ............................................................. 132
7.7.6 Reporting of Missing Values in Prediction Research ....................................................... 133
7.8 Concluding Remarks ........................................ 134
7.8.1 Summary Statements ...................................... 135
7.8.2 Currently Available Software and Challenges ......................................................... 136

8 Case Study on Dealing with Missing Values .................. 139
8.1 Introduction .................................................. 139
8.1.1 Aim ..................................................... 139
8.1.2 Patient Selection ............................................ 140
8.1.3 Selection of Potential Predictors .......................................... 140
8.1.4 Coding and Time Dependency of Predictors ......................................................... 141
8.2 Missing Values in the IMPACT Study ....................... 142
8.2.1 Missing Values in Outcome .................................. 142
8.2.2 Quantification of Missingness of Predictors ......................................................... 143
8.2.3 Patterns of Missingness ...................................... 144
8.3 Imputation of Missing Predictor Values ...................... 147
8.3.1 Correlations Between Predictors .................................. 147
8.3.2 Imputation Model .............................................. 147
8.3.3 Distributions of Imputed Values .................................... 149
8.4 Estimating Adjusted Effects ..................................... 149
8.4.1 Adjusted Analysis for Complete Predictors: Age and Motor Score .................................. 151
8.4.2 Adjusted Analysis for Incomplete Predictors: Pupils ............................................. 154
8.5 Multivariable Analyses ........................................ 155
8.6 Concluding Remarks ........................................... 155

9 Coding of Categorical and Continuous Predictors ............ 159
9.1 Categorical Predictors ......................................... 159
9.1.1 Examples of Categorical Coding .................................. 160
9.2 Continuous Predictors ........................................... 161
9.2.1 Examples of Continuous Predictors .................................. 161
9.2.2 Categorization of Continuous Predictors .................................. 162
9.3 Non-Linear Functions for Continuous Predictors ............ 163
9.3.1 Polynomials .................................................. 164
9.3.2 Fractional Polynomials .......................................... 164
9.3.3 Splines ...................................................... 165
9.3.4 Example: Functional Forms with RCS or FP .............................................. 166
9.3.5 Extrapolation and Robustness ..................................... 166
9.4 Outliers and Truncation .................................................. 167
  9.4.1 Example: Glucose Values and Outcome of TBI .......... 168
9.5 Interpretation of Effects of Continuous Predictors .......... 170
  9.5.1 Example: Predictor Effects in TBI ....................... 171
9.6 Concluding Remarks .................................................. 172
  9.6.1 Software ..................................................... 172

10 Restrictions on Candidate Predictors ......................... 175
  10.1 Selection Before Studying the Predictor–Outcome
        Relationship ..................................................... 175
    10.1.1 Selection Based on Subject Knowledge ................. 175
    10.1.2 Example: Too Many Candidate Predictors ............. 176
    10.1.3 Meta-Analysis for Candidate Predictors ............... 176
    10.1.4 Example: Predictors in Testicular Cancer .......... 176
    10.1.5 Selection Based on Distributions .................... 177
  10.2 Combining Similar Variables ................................. 177
    10.2.1 Example: Coding of Comorbidity ..................... 178
    10.2.2 Assessing the Equal Weights Assumption .......... 178
    10.2.3 Logical Weighting ..................................... 179
    10.2.4 Statistical Combination ............................... 180
  10.3 Averaging Effects ............................................. 180
    10.3.1 Example: Chlamydia Trachomatis Infection Risks ... 180
    10.3.2 Example: Acute Surgery Risk Relevant for
        Elective Patients? ....................................... 180
  10.4 Case study: Family History for Prediction of a
        Genetic Mutation ........................................... 181
    10.4.1 Clinical Background and Patient Data ............... 181
    10.4.2 Similarity of Effects ................................ 182
    10.4.3 CRC and Adenoma in a Proband ..................... 184
    10.4.4 Age of CRC in Family History ..................... 185
    10.4.5 Full Prediction Model for Mutations ............... 186
  10.5 Concluding Remarks ........................................... 187

11 Selection of Main Effects ......................................... 191
  11.1 Predictor Selection ............................................ 191
    11.1.1 Reduction Before Modelling ......................... 191
    11.1.2 Reduction While Modelling ........................... 192
    11.1.3 Collinearity ............................................ 192
    11.1.4 Parsimony ............................................. 193
    11.1.5 Should Non-Significant Variables Be Removed? ... 193
    11.1.6 Summary Points ....................................... 194
  11.2 Stepwise Selection ............................................ 194
    11.2.1 Stepwise Selection Variants ....................... 194
    11.2.2 Stopping Rules in Stepwise Selection ............ 195
11.3 Advantages of Stepwise Methods ........................................ 196
11.4 Disadvantages of Stepwise Methods ................................. 197
  11.4.1 Instability of selection ........................................ 197
  11.4.2 Biased Estimation of Coefficients ............................. 199
  11.4.3 Bias of Stepwise Selection and Events Per Variable ........ 199
  11.4.4 Misspecification of Variability ............................... 201
  11.4.5 Exaggeration of P-Values .................................. 204
  11.4.6 Predictions of Worse Quality Than from a Full Model .... 204
11.5 Influence of Noise Variables ....................................... 205
11.6 Univariate Analyses and Model Specification ..................... 206
  11.6.1 Pros and Cons of Univariate Pre-Selection .................. 207
  11.6.2 Testing of Predictors within Domains ........................ 207
11.7 Modern Selection Methods ........................................ 207
  11.7.1 Bootstrapping for Selection .................................. 208
  11.7.2 Bagging and Boosting ...................................... 208
  11.7.3 Bayesian Model Averaging (BMA) ............................ 208
  11.7.4 Practical Advantages of BMA ................................. 209
  11.7.5 Shrinkage of Regression Coefficients to Zero ............. 210
11.8 Concluding Remarks ................................................ 210

12 Assumptions in Regression Models: Additivity and Linearity .......... 213
  12.1 Additivity and Interaction Terms ................................. 213
  12.1.1 Potential Interaction Terms to Consider ..................... 214
  12.1.2 Interactions with Treatment ................................ 214
  12.1.3 Other Potential Interactions ................................. 215
  12.1.4 Example: Time and Survival After Valve Replacement ........ 216
  12.2 Selection, Estimation and Performance with Interaction Terms ... 216
  12.2.1 Example: Age Interactions in GUSTO-I ..................... 217
  12.2.2 Estimation of Interaction Terms ............................. 217
  12.2.3 Better Prediction with Interaction Terms? ................... 219
  12.2.4 Summary Points ........................................... 220
  12.3 Non-linearity in Multivariable Analysis ............................ 220
  12.3.1 Multivariable Restricted Cubic Splines (RCS) ............... 220
  12.3.2 Multivariable Fractional Polynomials (FP) .................. 221
  12.3.3 Multivariable Splines in GAM ................................ 222
  12.4 Example: Non-Linearity in Testicular Cancer Case Study .......... 222
  12.4.1 Details of Multivariable FP and GAM Analyses ............. 224
  12.4.2 GAM in Univariate and Multivariable Analysis ............. 224
  12.4.3 Predictive Performance .................................... 226
  12.4.4 R code for Non-Linear Modelling ........................... 227
  12.5 Concluding Remarks ............................................. 227
  12.5.1 Recommendations ............................................ 228
13 Modern Estimation Methods

13.1 Predictions from Regression and Other Models
13.2 Shrinkage
13.2.1 Uniform Shrinkage
13.2.2 Uniform Shrinkage in GUSTO-I
13.3 Penalized Estimation
13.3.1 Penalized Maximum Likelihood Estimation
13.3.2 Penalized ML in Sample4
13.3.3 Shrinkage, Penalization, and Model Selection
13.4 Lasso
13.4.1 Estimation of Lasso Model
13.4.2 Lasso in GUSTO-I
13.4.3 Predictions after Shrinkage
13.4.4 Model Performance after Shrinkage
13.5 Concluding Remarks

14 Estimation with External Information

14.1 Combining Literature and Individual Patient Data
14.1.1 Adaptation Method 1
14.1.2 Adaptation Method 2
14.1.3 Estimation
14.1.4 Simulation Results
14.1.5 Performance of Adapted Model
14.1.6 Improving Calibration
14.2 Example: Mortality of Aneurysm Surgery
14.2.1 Meta-Analysis
14.2.2 Individual Patient Data Analysis
14.2.3 Adaptation Results
14.3 Alternative Approaches
14.3.1 Overall Calibration
14.3.2 Bayesian Methods: Using Data Priors to Regression Modelling
14.3.3 Example: Predicting Neonatal Death
14.3.4 Example: Mortality of Aneurysm Surgery
14.4 Concluding Remarks

15 Evaluation of Performance

15.1 Overall Performance Measures
15.1.1 Explained Variation: $R^2$
15.1.2 Brier Score
15.1.3 Example: Performance of Testicular Cancer Prediction Model
15.1.4 Overall Performance Measures in Survival
Part III Generalizability of Prediction Models

19 Patterns of External Validity

19.1 Determinants of External Validity
19.1.1 Case-Mix
19.1.2 Differences in Case-Mix
19.1.3 Differences in Regression Coefficients
19.2 Impact on Calibration, Discrimination, and Clinical Usefulness
19.2.1 Simulation Set-Up
19.2.2 Performance Measures
19.3 Distribution of Predictors
19.3.1 More- or Less-Severe Case-Mix According to X
19.3.2 Example: Interpretation of Testicular Cancer Validation
19.3.3 More or Less Heterogeneous Case-Mix According to X
19.3.4 More- or Less-Severe Case-Mix According to Z
19.3.5 More or Less Heterogeneous Case-Mix According to Z
19.4 Distribution of Observed Outcomes $Y$
19.5 Coefficients $\beta$
19.5.1 Coefficient of Linear Predictor $< 1$
19.5.2 Coefficients Different
19.5.3 R Code
19.5.4 Influence of Different Coefficients
19.5.5 Other Scenarios of Invalidity
19.5.6 Summary of Patterns of Invalidity
19.6 Reference Values for Performance
19.6.1 Calculation of Reference Values
19.6.2 R Code
19.6.3 Performance with Refitting
19.6.4 Examples: Testicular Cancer and TBI
## 21 Updating for Multiple Settings

21.1 Differences Between Settings

- 21.1.1 Testing for Calibration-in-the Large
- 21.1.2 Illustration of Heterogeneity in GUSTO-I
- 21.1.3 Updating for Better Calibration-in-the Large
- 21.1.4 Empirical Bayes Estimates
- 21.1.5 Illustration of Updating in GUSTO-I
- 21.1.6 Testing and Updating of Predictor Effects
- 21.1.7 Heterogeneity of Predictor Effects in GUSTO-I
- 21.1.8 R Code for Random Effect Analyses

21.2 Provider Profiling

- 21.2.1 Indicators for Differences Between Centres
- 21.2.2 Ranking of Centres
- 21.2.3 Example: Provider Profiling in Stroke
- 21.2.4 Testing of Differences Between Centres
- 21.2.5 Estimation of Differences Between Centres
- 21.2.6 Uncertainty in Differences
- 21.2.7 Ranking of Centres
- 21.2.8 Essential R Code for Provider Profiling
- 21.2.9 Guidelines for Provider Profiling

21.3 Concluding Remarks

- 21.3.1 Bibliographic Notes

## 22 Prediction of a Binary Outcome: 30-Day Mortality After Acute Myocardial Infarction

22.1 GUSTO-I Study

- 22.1.1 Acute Myocardial Infarction
- 22.1.2 Treatment Results from GUSTO-I
- 22.1.3 Prognostic Modelling in GUSTO-I

22.2 General Considerations of Model Development

- 22.2.1 Research Question and Intended Application
- 22.2.2 Outcome and Predictors
- 22.2.3 Study Design and Analysis

22.3 Seven Modelling Steps in GUSTO-I

- 22.3.1 Data Inspection
- 22.3.2 Coding of Predictors
- 22.3.3 Model Specification
- 22.3.4 Model Estimation
- 22.3.5 Model Performance
- 22.3.6 Model Validation
- 22.3.7 Presentation

---

### Part IV Applications

- **22 Prediction of a Binary Outcome: 30-Day Mortality After Acute Myocardial Infarction**
  - 22.1 GUSTO-I Study
  - 22.2 General Considerations of Model Development
  - 22.3 Seven Modelling Steps in GUSTO-I
### 22.4 Validity
- **22.4.1 Internal Validity: Overfitting**
- **22.4.2 External Validity: Generalizability**
- **22.4.3 Summary Points**

### 22.5 Translation into Clinical Practice
- **22.5.1 Score Chart for Choosing Thrombolytic Therapy**
- **22.5.2 Predictions for Choosing Thrombolytic Therapy**
- **22.5.3 Covariate Adjustment in GUSTO-I**

### 22.6 Concluding Remarks

### 23 Case Study on Survival Analysis: Prediction of Secondary Cardiovascular Events

#### 23.1 Prognosis in the SMART Study

#### 23.2 General Considerations in SMART
- **23.2.1 Research Question and Intended Application**
- **23.2.2 Outcome and Predictors**
- **23.2.3 Study Design and Analysis**

#### 23.3 Data Inspection Steps in the SMART Cohort

#### 23.4 Coding of Predictors
- **23.4.1 Extreme Values**
- **23.4.2 Transforming Continuous Predictors**
- **23.4.3 Combining Predictors with Similar Effects**

#### 23.5 Model Specification
- **23.5.1 Selection**

#### 23.6 Model Estimation, Performance, Validation, and Presentation
- **23.6.1 Model Estimation**
- **23.6.2 Model Performance**
- **23.6.3 Model Validation: Stability**
- **23.6.4 Model Validation: Optimism**
- **23.6.5 Model Presentation**

#### 23.7 Concluding Remarks

### 24 Lessons from Case Studies

#### 24.1 Sample Size
- **24.1.1 Example: Sample Size and Number of Predictors**
- **24.1.2 Number of Predictors**
- **24.1.3 Potential Solutions**

#### 24.2 Validation
- **24.2.1 Examples of Internal and External Validation**

#### 24.3 Subject Matter Knowledge

#### 24.4 Data Sets
- **24.4.1 GUSTO-I Prediction Models**
- **24.4.2 Modern Learning Methods in GUSTO-I**
24.4.3 Modelling Strategies in Small Data Sets from GUSTO-I ........................................... 453
24.4.4 SMART Case Study .................................................. 453
24.4.5 Testicular Cancer Case Study ................................. 455
24.4.6 Abdominal Aortic Aneurysm Case Study .............. 455
24.4.7 Traumatic Brain Injury Data Set ............................... 459
24.5 Concluding Remarks................................................ 459

References ........................................................................ 463
Index .............................................................................. 487
Clinical Prediction Models
A Practical Approach to Development, Validation, and Updating
Steyerberg, E.
2009, XXVIII, 500 p., Hardcover
ISBN: 978-0-387-77243-1