Prediction Model Performance Metrics for the Validation of Multi-State Clinical Prediction Models

MA Barrowman GP Martin N Peek M Lambie M Sperrin

Contents

1 Introduction **2**

2 Motivating Data Set **2**

3 Current Approaches **2**

3.1 Baseline Models . . . 2
3.2 Notation . . . 2
3.3 Patient Weighting . . . 2
3.4 Accuracy - Brier Score . . . 2
3.5 Discrimination - c-statistic . . . 2
3.6 Calibration - Intercept and Slope . . . 2

4 Extension to Multi-State Models **2**

4.1 Trivial Extensions . . . 2
4.2 Accuracy - Multiple Outcome Brier Score . . . 2
4.3 Discrimination - Polytomous Discriminatory Index . . . 2
4.4 Calibration - Multinomial Intercept, Matched and Unmatched Slopes . . . 2

5 Application to Real-World Data **2**

5.1 Accuracy . . . 2
5.2 Discrimination . . . 2
5.3 Calibration . . . 2

6 Discussion **2**

References **2**

Last updated: 21 Apr

1 Introduction

2 Motivating Data Set

3 Current Approaches

3.1 Baseline Models

3.2 Notation

3.3 Patient Weighting

3.4 Accuracy - Brier Score

3.5 Discrimination - c-statistic

3.6 Calibration - Intercept and Slope

4 Extension to Multi-State Models

4.1 Trivial Extensions

4.2 Accuracy - Multiple Outcome Brier Score

4.3 Discrimination - Polytomous Discriminatory Index

4.3.1 Computational Limitations

4.4 Calibration - Multinomial Intercept, Matched and Unmatched Slopes

5 Application to Real-World Data

5.1 Accuracy

5.2 Discrimination

5.3 Calibration

6 Discussion

References