CONTENTS

APPROVAL ii

SELF DECLARATION AGAINST PLAGIARISM iii

ABSTRACT iv

ABSTRAK v

DEDICATION vi

ACKNOWLEDGMENTS vii

PREFACE viii

CONTENTS ix

LIST OF TABLES xii

LIST OF FIGURES xiii

LIST OF TERMS xv

LIST OF NOTATIONS xvi

1 INTRODUCTION 1

1.1 Rationale . . . 1
1.2 Theoretical Framework . . . 2
1.3 Conceptual Framework/Paradigm . . . 4
1.4 Statement of the Problem . . . 5
1.5 Objective and Hypotheses . . . 6
1.6 Assumption . . . 6
1.7 Scope and Delimitation . . . 6
1.8 Significance of the Study . . . 7

2 REVIEW OF LITERATURE AND STUDIES 9

2.1 Process Mining . . . 9
2.1.1 Perspective of Process Mining . . . 10
2.1.2 Process Discovery . . . 10
2.1.3 Conformance Checking . . . 11
2.1.4 Enhancement . . . 11
2.2 Techniques for Process Mining . . . 12

2.2.1 Alpha Miner . . . 12
2.2.2 Heuristic Miner . . . 12
2.2.3 Inductive Miner (IM) . . . 12
2.2.4 Inductive Miner-infrequent (IMi) . . . 13
2.3 Tools for Process Mining . . . 13
2.3.1 ProM . . . 13
2.3.2 PM4PY . . . 13
2.4 Process-Modeling Notations . . . 14
2.4.1 Petri Net . . . 14
2.5 Breast Cancer Treatment . . . 15
2.6 Literature Review . . . 15

3 RESEARCH METHODOLOGY 18
3.1 Planning . . . 18
3.2 Extracting data and creating event log . . . 19
3.2.1 Data selection and migration . . . 19
3.2.2 Creating event log . . . 19
3.3 Data Preprocessing . . . 21
3.3.1 Filter codes unrelated to breast cancer based on INA-CBGs . . . 21
3.3.2 Filter non-essential breast cancer treatment . . . 22
3.3.3 Merge similar treatments . . . 25
3.3.4 Remove low-frequency events . . . 26
3.3.5 Data cleaning . . . 26
3.3.6 Add start and end to treatment for each patient . . . 27
3.3.7 Merge subsequent non-primary treatment . . . 27
3.3.8 Event Log . . . 27
3.4 Treatment Groups . . . 28
3.5 Mining and Analysis . . . 28
3.5.1 Process Discovery . . . 29
3.5.2 Conformance Checking . . . 33
3.6 Evaluation . . . 35

4 PRESENTATION, ANALYSIS AND INTERPRETATION OF DATA 36
4.1 Presentation of the Data . . . 36
4.1.1 Age of Patient . . . 36
4.1.2 The Most Frequent Treatments . . . 37
4.1.3 The Most Frequent Trace Variants . . . 38
4.1.4 Average Time Transition Treatment . . . 38
4.2 Process Discovery Algorithm of General Treatment . . . 39
4.2.1 Alpha Miner . . . 39

4.2.2 Heuristic Miner . . . 39
4.2.3 Inductive Miner . . . 40
4.2.4 Inductive Miner-infrequent . . . 41
4.3 Conformance Checking of Process Mining Algorithm . . . 41
4.4 Process Discovery of Treatment Groups . . . 43
4.4.1 The Radiotherapy Treatment Group . . . 43
4.4.2 The Chemotherapy Treatment Group . . . 44
4.4.3 Both Radiotherapy and Chemotherapy Treatment Group . . . 45
4.5 Conformance Checking of Treatment Groups . . . 46
4.6 Evaluation . . . 49
4.6.1 Comparison of Treatment Groups of Inductive Miner-infrequent . . . 49
4.6.2 Comparison of Time Perspective in Each Treatment Groups . . . 51
4.6.3 The Average Waiting Time of Breast Cancer Treatment . . . 54
4.7 Discussion . . . 57
4.7.1 Time Perspective . . . 57
4.7.2 The Sequence of Treatment Perspective . . . 58

5 CONCLUSION AND RECOMMENDATIONS 59
5.1 Conclusions . . . 59
5.2 Future Research . . . 60

BIBLIOGRAPHY 61

Appendices 69

A Curriculum Vitae 71