Decision-Support for Rheumatoid Arthritis Using Bayesian Networks: Diagnosis, Management, and Personalised Care

Ali Fahmi

School of Electronic Engineering and Computer Science
Queen Mary University of London

This dissertation is submitted for the degree of
Doctor of Philosophy

Submitted: May 2021
Amended: November 2021

I would like to dedicate this thesis to my loving parents and sister ...

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Ali Fahmi
Submitted: May 2021
Amended: November 2021

Acknowledgements

I would like to thank my primary PhD supervisor, Dr William Marsh, for his supervision and support. William has taught me how to think critically and how to be precise in my scientific speaking and writing. I also acknowledge William's patience and the way he admires people's abilities and collaborates with them. I want to thank other members of the supervision panel, Professor Norman Fenton and Professor Martin Neil. Their comments helped to improve the quality of my PhD research.

I would like to thank Professor Tjeerd van Staa, Dr Evangelia (Lina) Kyrimi, Dr Christopher Joyner, and Clare who proofread my thesis and helped me to improve it. Six colleagues of mine Lina, Dr Amy MacBraye, Professor Paul Curzon, Dr Hamit Soyel, Dr Mariana Raniere Neves, and Dr Haoyuan Zhang were always collaborative and friendly that I am so grateful for. I also thank the internal and external examiners of my PhD viva, Dr Huy Phan and Professor Paul Krause, for providing helpful comments on my thesis.

I really appreciate the support, love, and encouragement that my parents and sister, and my brother-in-law gave to me. I also thank my friends Manish, Amin, Ali, Pegah, Amin, Akram, Azam, Bita, Sunduz, Kubra, Sukru, Aytac, Mariana, Fahad, Sherry, Ipek, Mohammad, Saeedeh, and others who accompanied me and encouraged to keep it up both before and throughout my PhD.

I want to thank all previous supervisors and lecturers of mine who taught me and helped me to come to the UK and do my PhD: Professor Cengiz Kahraman, Professor Burc Ulengin, Dr Hasan Rezazadeh, Dr Tolga Kaya, Professor Behzad Salmani, Dr Mohsen Naderpour, and Professor Jie Lu.

Abstract

Bayesian networks (BNs) have been widely proposed for medical decision support. One advantage of a BN is reasoning under uncertainty, which is pervasive in medicine. Another advantage is that a BN can be built from both data and knowledge and so can be applied in circumstances where a complete dataset is not available. In this thesis, we examine how BNs can be used for the decision support challenges of chronic diseases. As a case study, we study Rheumatoid Arthritis (RA), which is a chronic inflammatory disease causing swollen and painful joints. The work has been done as part of a collaborative project including clinicians from Barts and the London NHS Trust involved in the treatment of RA. The work covers three stages of decision support, with progressively less available data.

The first decision support stage is diagnosis. Various criteria have been proposed by clinicians for early diagnosis but these criteria are deterministic and so do not capture diagnostic uncertainty, which is a concern for patients with mild symptoms in the early stages of the disease. We address this problem by building a BN model for diagnosing RA. The diagnostic BN model is built using both a dataset of 360 patients provided by the clinicians and their knowledge as experts in this domain. The choice of factors to include in the diagnostic model is informed by knowledge, including a model of the care pathway which shows what information is available for diagnosis. Knowledge is used to classify the factors as risk factors, relevant comorbidities, evidence of pathogenesis mechanism, signs, symptoms, and serology results, so that the structure of BN model matches the clinical understanding of RA.

Since most of the factors are present in the dataset, we are able to train the parameters of the diagnostic BN from the data. This diagnostic BN model obtains promising results in differentiating RA cases from other inflammatory arthritis cases. Aware that eliciting knowledge is time-consuming and could limit the uptake of these techniques, we consider two alternative approaches. First, we compare its diagnostic performance with an alternative BN model entirely learnt from data; we argue that having a clinically meaningful structure allows us to explain clinical scenarios in a way that cannot be done with the model learnt purely from data. We also examine whether useful knowledge can be retrieved from existing

medical ontologies, such as SNOMED CT and UMLS. Preliminary results show that it could be feasible to use such sources to partially automate knowledge collection.

After patients have been diagnosed with RA, they are monitored regularly by a clinical team until the activity of their disease becomes low. The typical care arrangement has two challenges: first, regular meetings with clinicians occur infrequently at fixed intervals (e.g., every six months), during which time the activity of the disease can increase (or 'flare') and decrease several times. Secondly, the best medications or combinations of medications must be found for each patient, but changes can only be made when the patient visits the clinic. We therefore develop this stage of decision support in two parts: the first and simplest part looks at how the frequency of clinic appointments could be varied; the second part builds on this to support decisions to adjust medication dosage. We describe this as the 'self-management' decision support model.

Disease activity is commonly measured with Disease Activity Score 28 (DAS28). Since the joint count parts of this can be assessed by the patient, the possibility of collecting regular (e.g., weekly) DAS28 data has been proposed. It is not yet in wide use, perhaps because of the overheads to the clinical team of reviewing data regularly. The dataset available to us for this work came from a feasibility study conducted by the clinical collaborators of one system for collecting data from patients, although the frequency is only quarterly. The aim of the 'self-management' decision support system is therefore to sit between patient-entered data and the clinical team, saving the work of clinically assessing all the data. Specifically, in the first part we wish to predict disease activity so that an appointment should be made sooner, distinguishing this from patients whose disease is well-managed so that the interval between appointments can be increased. To achieve this, we build a dynamic BN (DBN) model to monitor disease activity and to indicate to patients and their clinicians whether a clinical review is needed. We use the data and a set of dummy patient scenarios designed by the experts to evaluate the performance of the DBN.

The second part of the 'self-management' decision support stage extends the DBN to give advice on adjustments to the medication dosage. This is of particular clinical interest since one class of medications used (biological disease-modifying antirheumatic drugs) are very expensive and, although effective at reducing disease activity, can have severe adverse reactions. For both these reasons, decision support that allowed a patient to 'taper' the dosage of medications without frequent clinic visits would be very useful. This extension does not meet all the decision support needs, which ideally would also cover decision-making about the choice of medications. However, we have found that as yet there is neither sufficient data nor knowledge for this.

The third and final stage of decision support is targeted at patients who live with RA. RA can have profound impacts on the quality of life (QoL) of those who live with it, affecting work, financial status, friendships, and relationships. Information from patient organisations such as the leaflets prepared by the National Rheumatoid Arthritis Society (NRAS) contains advice on managing QoL, but the advice is generic, leaving it up to each patient to select the advice most relevant to their specific circumstances. Our aim is therefore to build a BN-based decision support system to personalise the recommendations for enhancing the QoL of RA patients. We have built a BN to infer three components of QoL (independence, participation, and empowerment) and shown how this can be used to target advice. Since there is no data, the BN is developed from expert knowledge and literature. To evaluate the resulting system, including the BN, we use a set of patient interviews conducted and coded by our collaborators. The recommendations of the system were compared with those of experts in a set of test scenarios created from the interviews; the comparison shows promising results.

Table of contents

List of figures **xiv**

List of tables **xx**

Nomenclature **xxiv**

1 An Introduction to the Research Project and Objectives **1**

1.1 Challenges in Modelling of Diagnosis, Self-Management, and Personalised Care 4

1.2 Research Objectives 5

1.3 Author's Contributions 5

1.4 Structure of Thesis 6

1.5 Publications and Awards 7

2 Introduction and Application of Bayesian Networks **9**

2.1 Principles of Bayesian Networks 10

2.2 Inference with Bayesian Networks 11

2.3 Time-Based Bayesian Networks 13

2.3.1 Dynamic Bayesian Networks 13

2.3.2 Other Time-Based Bayesian Networks 16

2.4 Inference with Dynamic Bayesian Networks 18

2.5 Building Bayesian Networks 20

2.5.1 Structure Learning 20

2.5.2 Parameter Learning 24

2.6 Medical Application of Bayesian Networks 25

2.6.1 Diagnosis 26

2.6.2 Personalised Care and Treatment 27

2.6.3 Reasoning and Explanation 27

2.7 Clinical Decision Support 28

2.7.1 Rheumatoid Arthritis Applications 29
2.7.2 Other Applications 30
2.8 Summary 31

3 Medical Ontologies 32
3.1 What is an Ontology? 32
3.2 Developing Ontologies 34
3.3 Semantic Web and Ontology Languages 35
3.3.1 Resource Description Framework 36
3.3.2 Web Ontology Language 36
3.4 Medical Ontologies 38
3.4.1 Systematized Nomenclature of Medicine-Clinical Terms 38
3.4.2 Unified Medical Language System 42
3.5 Bayesian Network Models and Ontologies 44
3.5.1 Contrasting Care Pathways, Knowledge Graphs, and Bayesian Networks 44
3.5.2 Building Bayesian Networks from Ontologies 45
3.6 Summary 47

4 Case Study: Diagnosis and Treatment of Rheumatoid Arthritis 48
4.1 PAMBAYESIAN Project 49
4.2 Rheumatoid Arthritis 49
4.2.1 Diagnosis 50
4.2.2 Management and Treatment 51
4.2.3 Quality of Life 53
4.3 Models of Care Pathways for Rheumatoid Arthritis 54
4.3.1 Diagnosis 56
4.3.2 Pathway for Rheumatoid Arthritis Initial Management 58
4.3.3 Pathway for Rheumatoid Arthritis Ongoing Management 60
4.3.4 Pathway for Personalised Care for Living with Rheumatoid Arthritis 62
4.4 Description and Analysis of Available Data from Studies on Rheumatoid Arthritis 64
4.4.1 Introduction to Studies 64
4.4.2 Description and Analysis of Personal Information, Risk Factors and Comorbidities 67
4.4.3 Description and Analysis of Signs, Symptoms, Serology Results, and Disease Activity 69

4.4.4 Interpolation of Signs, Symptoms, Serology Results, and Disease Activity Records of PEAC Dataset . . . 79
4.4.5 Description and Analysis of Treatment . . . 82
4.5 Selected Decision Support Points . . . 87
4.6 Summary . . . 89

5 Building Bayesian Network Models for Diagnosis of Rheumatoid Arthritis 90
5.1 Introduction . . . 90
5.2 Description of Variables for Diagnosis of Rheumatoid Arthritis . . . 92
5.2.1 Personal Information, Risk Factors, and Comorbidities . . . 92
5.2.2 Disease Manifestation . . . 95
5.2.3 Pathogenesis . . . 97
5.2.4 Intervention . . . 98
5.2.5 Diagnosis . . . 98
5.3 Bayesian Network Model for Diagnosis of Rheumatoid Arthritis . . . 99
5.3.1 Data . . . 99
5.3.2 Structure . . . 99
5.3.3 Parameterisation . . . 101
5.4 Results and Evaluation . . . 102
5.4.1 Cross-Validation . . . 103
5.4.2 Evaluation with Scenarios of Dummy Patients . . . 108
5.4.3 Review of Inaccurate Cases . . . 109
5.4.4 Explanation of One Case of Rheumatoid Arthritis . . . 109
5.4.5 Reasoning with Absent Variables . . . 110
5.5 Summary . . . 112

6 Building Medical Bayesian Networks from Care Pathways and Knowledge Graphs 113
6.1 Introduction . . . 114
6.2 Analysing Care Pathways and Creating Knowledge Graphs from Medical Ontologies . . . 115
6.3 Case Study: Diagnosis of Rheumatoid Arthritis . . . 116
6.3.1 Systematized Nomenclature of Medicine-Clinical Terms . . . 116
6.3.2 Unified Medical Language System . . . 120
6.4 Discussion . . . 125
6.5 Summary . . . 125

7 Building Dynamic Bayesian Network Model for Self-Management of Rheumatoid Arthritis: Appointment Scheduling 127
7.1 Introduction . . . 128
7.2 DBN Variables for Self Management . . . 129
7.2.1 Evidence Variables . . . 129
7.2.2 Latent Variables . . . 130
7.2.3 Advice Variable . . . 133
7.3 DBN Model for Self Management of Rheumatoid Arthritis: Appointment Scheduling . . . 133
7.3.1 Structure . . . 134
7.3.2 Data . . . 135
7.3.3 Parameterisation . . . 136
7.4 Using the DBN for Decision-Support . . . 137
7.4.1 Inputs . . . 138
7.4.2 Dynamics of Disease State and Overall Flare . . . 139
7.4.3 Prediction of Advice on Appointment Scheduling . . . 140
7.5 Evaluation . . . 140
7.5.1 Evaluation with PEAC Data . . . 141
7.5.2 Evaluation with Scenarios of Dummy Patients . . . 146
7.6 Summary . . . 147

8 Building Dynamic Bayesian Network Models for Self-Management of Rheumatoid Arthritis: Medication Review 149
8.1 Introduction . . . 150
8.2 Medications for Rheumatoid Arthritis . . . 151
8.2.1 Conventional Disease-Modifying Anti-Rheumatic Drugs . . . 151
8.2.2 Targeted Synthetic and Biological Disease-Modifying Anti-Rheumatic Drugs . . . 152
8.2.3 Steroids . . . 154
8.3 Strategies, Regimens, and Decisions for Treatment of Rheumatoid Arthritis 154
8.3.1 Treatment Strategies . . . 155
8.3.2 Treatment Regimens . . . 156
8.3.3 Treatment Decision-Making . . . 156
8.4 Description of Variables for Treatment of Rheumatoid Arthritis . . . 158
8.4.1 Evidence Variables . . . 158
8.4.2 Latent Variables . . . 160
8.4.3 Advice Variable . . . 161

8.5 DBN Model for Self-Management of Rheumatoid Arthritis: Medication Review . . . 162
8.5.1 Structure . . . 162
8.5.2 Data . . . 164
8.5.3 Parameterisation . . . 165
8.6 Using the DBN for Decision-Support . . . 166
8.6.1 Inputs . . . 166
8.6.2 Dynamics of Adverse Medication Events and Tolerance for DMARDs 166
8.6.3 Dynamics of Disease State and Overall Flare . . . 168
8.6.4 Prediction of Advice on Medication Review . . . 169
8.7 Evaluation . . . 169
8.7.1 Evaluation with PEAC Data . . . 170
8.7.2 Evaluation with Scenarios of Dummy Patients . . . 174
8.8 Summary . . . 175

9 Building a Bayesian Network Model for Personalised Care for Living with Rheumatoid Arthritis 176
9.1 Introduction . . . 177
9.2 Description of Variables for Personalised Care for Living with Rheumatoid Arthritis . . . 178
9.2.1 Personal and Environmental Variables . . . 178
9.2.2 Disease Manifestation and Disease Activity Variables . . . 179
9.2.3 Characteristics of Quality of Life and Lifestyle Choices . . . 180
9.2.4 Advice Variable . . . 182
9.3 Data Collection Questions and Information Retrieval . . . 184
9.3.1 Personal, Disease Duration, and Environmental Variables . . . 184
9.3.2 Disease Manifestations . . . 184
9.3.3 Characteristics of Quality of Life . . . 185
9.3.4 Lifestyle Choices . . . 185
9.4 BN Model for Personalised Care for Living with Rheumatoid Arthritis . . . 186
9.4.1 Structure . . . 186
9.4.2 Parameterisation . . . 188
9.5 Rule-based System for Personalised Care for Living with Rheumatoid Arthritis 190
9.5.1 Advice Generation . . . 190
9.5.2 Advice Modification . . . 192
9.6 Evaluation . . . 193
9.6.1 Scenario Analysis . . . 193

9.6.2 Sensitivity Analysis 195
9.7 Summary 195

10 Summary and Future Directions 198
10.1 Contributions 198
10.2 Future Directions 202

References 206

Appendix A Analysis, Interpolation, and Expansion of Data of Pathobiology of Early Arthritis Cohort 223

Appendix B Bayesian Networks for Diagnosis of Rheumatoid Arthritis 234
B.1 Elicitation of Parameters from Experts 234
B.1.1 Initial Rules 234
B.1.2 Secondary Rules 237
B.2 Observations of Dummy Patient Scenarios 240

Appendix C Bayesian Network and Rule-Based System for Personalised Care for Living with Rheumatoid Arthritis 242
C.1 Data Collection Questions and Information Retrieval 242
C.2 Evaluation of Bayesian Network Model for Personalised Care for Living with Rheumatoid Arthritis Using Scenarios 250
C.3 Using Tornado Graphs for Sensitivity Analysis of Bayesian Network Model for Personalised Care of Rheumatoid Arthritis 254

List of figures

1.1 Thesis structure. 7

2.1 A simple DAG 10
2.2 A simple BN for a viral infection. 11
2.3 Inference with a simple BN model for a viral respiratory infection. 12
2.4 A simple DBN model with n time slices. 14
2.5 A simple DBN for a viral infection. 15
2.6 Simplified notation of the simple DBN for a viral infection with a initial and a terminal conditions. 15
2.7 A simple NPEDT 16
2.8 A simple CTBN 17
2.9 Inference with a simple DBN for a viral respiratory infection maintained as a sliding window of two time slices and having observed evidence variables. 19
2.10 Medical idioms proposed by [101] for a simple BN model for a viral respiratory infection. 21
2.11 Combination of medical idioms proposed by [101] for a simple BN model for a viral respiratory infection. 22

3.1 Terms of a vocabulary of a simple ontology for a viral respiratory infection. 33
3.2 Terms of a vocabulary and taxonomy of a simple ontology for viral respiratory infection with an individual instance called Ali. 34
3.3 Simple ontology of viral respiratory infections written in OWL. 37
3.4 Simple ontology of viral respiratory infections visualised by OntoGraph in Protege. 38
3.5 A chain of concepts and their relationship in SNOMED CT 39
3.6 SNOMED CT domain hierarchies and their subtypes. 40

4.1 A hand with affected joints by RA. 50
4.2 NICE pathway for treatment of RA 52

4.3 Cover page of NRAS booklet on Living Better with RA. 53
4.4 Information on disease activity and QoL in the NRAS booklet on Living Better with RA. 54
4.5 Pathway for diagnosis of RA. 57
4.6 Pathway for initial management of RA. 59
4.7 Pathway for RA ongoing management. 62
4.8 Pathway for personalised care of RA. 63
4.9 Dashboard of visits in clinicians interface of BioT app. 65
4.10 BioT app pages for counting tender joints. 66
4.11 BioT app pages on medication tracking and treatment feedback from patients. 66
4.12 Pair plot of scatter plots of Age, BMI, Smoking, and Alcohol in the baseline of PEAC data with kernel density estimation of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 70
4.13 Scatter plots joint with histogram plots and kernel density estimation of Smoking, CCP, and RhF in the baseline of PEAC data, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 72
4.14 Pairplot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the baseline of PEAC data with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 74
4.15 Scatter plot of DAS28ESR and DAS28CRP joint with histogram plot of each and their kernel density estimation, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 75
4.16 Scatter plot of DAS28ESR and DAS28CRP with marginal histogram of each of them. 76
4.17 Pair plot of scatter plots of monitoring and blood results and kernel density estimation of each monitoring or blood result. 77
4.18 Scatter plots of monitoring and blood results with marginal histogram plots of each variable. 78
4.19 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the baseline and first follow-up visit (3 months) and interpolation of 11 weeks between the baseline and first follow-up. 80
4.20 Flare variables computed using disease state values of the baseline, first follow-up visit (3 months), and interpolation of 11 weeks between the baseline and first follow-up. 82

4.21 Discretised expanded csDMARD values between the baseline and first follow-up visit with 11 weeks in between. 84
4.22 Decisions on medication made by rheumatologists in baseline and each follow-up visit. 85
4.23 Medication decisions in BioT study made in Time 3 to Time 11. 86

5.1 BN model for RA diagnosis built from experts' knowledge. 100
5.2 Learnt BN model for RA diagnosis built using SEM algorithm. 102
5.3 ROC curves of knowledge-based BN (black) and BN model learnt with SEM (grey). 104
5.4 Performance evaluation of the knowledge-based BN model with a range of thresholds from 0.0 to 0.9. 105
5.5 Performance evaluation of the learnt BN model with a range of thresholds from 0.0 to 0.9. 105
5.6 Tornado graph of RA state of Diagnosis variable of BN model for diagnosis of RA with inputs of evidence variables and two absent variables of 'Early Menopause' and 'Pregnancy or postpartum'. 111

6.1 A simplified model of care pathway for diagnosis of RA. 116
6.2 Candidate BN model for RA diagnosis built based on knowledge retrieved from SNOMED CT. 120
6.3 Candidate BN model for RA diagnosis built based on UMLS knowledge. . 124

7.1 DBN model for self-management of RA for appointment scheduling. 134
7.2 Unrolled DBN model for self-management of RA. 138
7.3 Fragment of 'Disease State 4' and 'Overall Flare 4' and its parent variables. 139
7.4 The total difference of the probabilities of the Low state of the 'Disease State' variables in the initial DBN model for self-management of RA with sliding windows of 2, 3, 4, 5, and 6 length. 141
7.5 ROC curve of DBN model for self-management of RA for appointment advice. 142
7.6 Performance analysis of the DBN model for self-management of RA for appointment advice with a range of thresholds from 0.0 to 0.9. 143
7.7 Probability of appointment advice corresponding to the interpolated DAS28 data. 145

8.1 DBN model for self-management of RA for medication review. 163
8.2 Unrolled DBN model for self-management of RA for medication review. . . 167

8.3 Fragment of 'Tolerance for DMARDs' variable and its parent variables in Week 5. 168
8.4 Fragment of 'Overall Disease Control' variable and its parent variables. . . 169
8.5 The total difference of the probabilities of the Low state of the 'Disease State' variables in the extended DBN model for self-management of RA with sliding windows of 2, 3, 4, 5, and 6 length. 170
8.6 ROC curve of DBN model for self-management of RA for medication review. 173
8.7 Performance analysis of the DBN model for self-management of RA for medication review with a range of thresholds from 0.0 to 0.9. 173

9.1 QoL components change and advice 183
9.2 BN model for personalised care for living with RA. 187
9.3 Tornado graph of the 'Independence & empowerment & participation' state of the 'Advice Priority' variable of BN model for personalised care for living with RA. 196

A.1 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 224
A.2 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 225
A.3 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 226
A.4 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. 227
A.5 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the first follow-up visit (3 months) and second follow-up visit (6 months) and interpolation of 11 weeks between the first and second follow-ups. 228

A.6 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the second follow-up visit (6 months) and third follow-up visit (9 months) and interpolation of 11 weeks between the second and third follow-ups. 229
A.7 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the third follow-up visit (9 months) and fourth follow-up visit (12 months) and interpolation of 11 weeks between the third and fourth follow-ups. 230
A.8 Flare variables computed using disease state values of the first follow-up visit (3 months), the second follow-up visit (6 months), and interpolation of 11 weeks between the first and second follow-ups. 231
A.9 Flare variables computed using disease state values of the second follow-up visit (6 months), the third follow-up visit (9 months), and interpolation of 11 weeks between the second and third follow-ups. 231
A.10 Flare variables computed using disease state values of the third follow-up visit (9 months), the fourth follow-up visit (12 months), and interpolation of 11 weeks between the third and fourth follow-ups. 231
A.11 Discretised expanded csDMARD values between the first and second follow-up visits with 11 weeks in between. 232
A.12 Discretised expanded csDMARD values between the second and third follow-up visits with 11 weeks in between. 232
A.13 Discretised expanded csDMARD values between the third and fourth follow-up visits with 11 weeks in between. 233

B.1 Initial rules to label demographic factors and parameterise 'Demographic Risks' variable provided by Dr. Amy MacBrayne. 234
B.2 Initial rules to label personal factors and parameterise 'Personal Risks' variable provided by Dr. Amy MacBrayne. 235
B.3 Initial rules to label medical background factors and parameterise 'Medical Background Risks' variable provided by Dr. Amy MacBrayne. 235
B.4 Initial rules to label lifestyle factors and parameterise 'Lifestyle Risks' variable provided by Dr. Amy MacBrayne. 236
B.5 Initial rules to label hormonal pathogenesis factors and parameterise 'FSH Effect on RA' variable provided by Dr. Amy MacBrayne. 236
B.6 Initial rules to label two antibodies and parameterise Serostatus variable provided by Dr. Amy MacBrayne. 237
B.7 Secondary rules to label personal factors and parameterise 'Personal Risks' variable provided by Dr. Amy MacBrayne. 237

B.8 Secondary rules to label demographic factors and parameterise ‘Demographic Risks’ variable provided by Dr. Amy MacBrayne. 238
B.9 Secondary rules to label lifestyle factors and parameterise ‘Lifestyle Risks’ variable provided by Dr. Amy MacBrayne. 238
B.10 Secondary rules to label hormonal pathogenesis factors and parameterise ‘FSH Effect on RA’ variable provided by Dr. Amy MacBrayne. 239

C.1 Tornado graph of the Independence state of the ‘Advice Priority’ variable. . 254
C.2 Tornado graph of the Empowerment state of the ‘Advice Priority’ variable. . 255
C.3 Tornado graph of the Participation state of the ‘Advice Priority’ variable. . . 256
C.4 Tornado graph of the ‘Independence & empowerment’ state of the ‘Advice Priority’ variable. 257
C.5 Tornado graph of the ‘Independence & participation’ state of the ‘Advice Priority’ variable. 258
C.6 Tornado graph of the ‘Empowerment & participation’ state of the ‘Advice Priority’ variable. 259

List of tables

3.1 Examples of relationships in SNOMED CT 39
3.2 SNOMED CT hierarchies. 41
3.3 Contrasts of Properties of care pathways, knowledge graphs, and BNs. . . . 45

4.1 Summary of personal information, risk factors, and comorbidities of PEAC data. 69
4.2 Description of signs, symptoms, serology results, and disease activity records in PEAC data at first, second, third, and fourth follow-ups. 73
4.3 Description of signs, symptom, serology results, and disease activity records of BioT data in 11 times. 73
4.4 Summary of treatment data in PEAC data at baseline, first, second, third, and fourth follow-ups. 83
4.5 Summary of treatment data in BioT data in eleven times of data collection. . 85
4.6 Selected decision support points in pathways for diagnosis, initial management, ongoing management, and personalised care of RA. 88

5.1 Summary of personal information, risk factors, and comorbidity variables for diagnosis of RA 94
5.2 Summary of disease manifestation variables associated with RA 96
5.3 Summary of pathogenesis variables associated with RA development 97
5.4 Summary of intervention variable associated with RA development 98
5.5 Summary of diagnosis variable 98
5.6 Comparison of diagnosis models performance 104
5.7 Confusion matrix of knowledge-based BN model. 106
5.8 Confusion matrix of learnt BN model. 106
5.9 Alternative confusion matrix of knowledge-based BN model. 106
5.10 Alternative confusion matrix of learnt BN model. 106
5.11 Prediction accuracy of other IA separately 107

5.12 Accuracy of RA prediction in 5 bins 107
5.13 Scenario-based evaluation of knowledge-based BN and learnt BN. 108
5.14 Reasoning with knowledge-based BN model on absent variables: 'Early Menopause' and 'Pregnancy or Postpartum'. 110

6.1 Refined concepts of rheumatoid arthritis and their relationships retrieved from SNOMED CT. 119
6.2 Atoms of rheumatoid arthritis and their properties retrieved from UMLS . . 122

7.1 Summary of evidence variables for self-management of RA 130
7.2 Summary of latent variables for self-management of RA. 132
7.3 Summary of the advice variable for self-management of RA: appointment scheduling. 133
7.4 Confusion matrix of DBN model for self-management of RA for appointment advice with threshold of 40%. 143
7.5 Confusion matrix of DBN model for self-management of RA for appointment advice with threshold of 30%. 144
7.6 Confusion matrix of DBN model for self-management of RA for appointment advice using dummy patient scenarios. 146
7.7 Alternative confusion matrix of DBN model for self-management of RA for appointment scheduling advice using dummy patient scenarios. 147
7.8 Accurately predicted monitoring advice separated based on the follow-up visit numbers. 147

8.1 Summary of evidence variables of medication and AME for treatment of RA. 158
8.2 Summary of latent variables of medication and AME for self-management of RA. 160
8.3 Summary of the advice variable for self-management of RA: medication review. 161
8.4 Confusion matrix of DBN model for self-management of RA for medication review advice - first follow-up visit. 171
8.5 Confusion matrix of DBN model for self-management of RA for medication review advice - second follow-up visit. 171
8.6 Confusion matrix of DBN model for self-management of RA for medication review advice - third follow-up visit. 172
8.7 Confusion matrix of DBN model for self-management of RA for medication review advice - fourth follow-up visit. 172

8.8 Aggregated confusion matrix of DBN model for self-management of RA for medication review advice. 172
8.9 Confusion matrix of DBN model for self-management of RA for medication review advice - dummy patient scenarios. 174

9.1 Summary of personal and environmental factors. 179
9.2 Summary of disease manifestation and disease activity variables. 180
9.3 Summary of characteristics of QoL and lifestyle choices. 182
9.4 Summary of advice variable. 183
9.5 Priority of regions based on combination of states of independence, empowerment, and participation 189
9.6 Prioritised regions with the NRAS advice. 190
9.7 Further advice for evidence variables with undesired state being observed. . 191
9.8 Further advice for latent variables with undesired state being estimated. . . 192
9.9 Modification of redundant advice for evidence variables with desired state being observed. 192
9.10 Modification of redundant advice for latent variables with desired state being estimated. 193
9.11 BN outputs and expected states of Flare, 'Current Disease Activity', 'Overall Disease Activity', and 'Disease Acceptance'. 194
9.12 BN outputs and expected states of independence, participation, empowerment, and QoL. 194
9.13 BN advice and expected states of advice priority. 195

B.1 Scenario of dummy patient 1 - observations of personal factors, demographic factors, lifestyle factors, medical background, comorbidities, signs, symptoms, and serology results, and expected outcome of diagnosis. 240
B.2 Scenario of dummy patient 1: Observations and expected outcome 241

C.1 Questions to collect personal information and 'Weather Sensitivity' data from patients. 243
C.2 Questions to collect disease manifestation data from patients 244
C.3 Questions to collect quality of life characteristics data from patients. 246
C.4 Questions to collect lifestyle choice data from patients. 247
C.5 AtTRA 10's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration 250
C.6 AtTRA 17's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration 251

C.7 AtTRA 24's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration 252
C.8 AtTRA 38's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration 253

Nomenclature

Acronyms / Abbreviations

ABA Abatacept

ABV Alcohol By Volume

ACR American College of Rheumatology

ADA Adalimumab

ALP Alkaline Phosphatase

ALT Alanine Aminotransferase

AME Adverse Medication Events

API Application Programming Interface

AST Aspartate Aminotransferase

AUI Atom Unique Identifier

AUROC Area Under Receiver Operating Characteristic

AZA Azathioprine

BDeu Bayes Dirichlet Equivalent Uniform

bDMARDs Biological Disease-Modifying Anti-Rheumatic Drugs

BMI Body Mass Index

BN Bayesian Network

CA Crystal Arthropathy

CaMML Causal Minimum Message Length

CCP Cyclic Citrullinated Peptide

CDS Clinical Decision Support

CIM Conditional Intensity Matrix

CPG Clinical Practice Guideline

CPT Conditional Probability Table

CRP C-Reactive Protein

csDMARDs Conventional Disease-Modifying Anti-Rheumatic Drugs

CTBN Continuous Time Bayesian Networks

CTD Connective Tissue Disease

CTZ Certolizumab

CUI Concept Unique Identifiers

DAG Directed Acyclic Graph

DAS Disease Activity Score

DBN Dynamic Bayesian Network

EECS School of Electronic Engineering and Computer Science

EM Expectation-Maximisation

ESR Erythrocyte Sedimentation Rate

ETN Etanercept

EULAR European League Against Rheumatism

FSN Fully Specified Name

GH Global Health

GMB Golimumab

GS Grow-Shrink

HAQ Health Assessment Questionnaire

HCQ Hydroxychloroquine

HIV Human Immunodeficiency Viruses

HTML Hypertext Markup Language

IA Inflammatory Arthritis

IA Intra-Articular

IAMB Incremental Association Markov Blanket

IHTSDO International Health Terminology Standards Development Organisation

IM Intramuscular

ITBN Irregular-Time Bayesian Networks

LEF Leflunomide

MAR Missing At Random

MCMC Markov Chain Monte Carlo

MCP Metacarpophalangeal

MMPC Max-Min Parents & Children

MonoA Monoarthritis

MTP Metatarsophalangeal

MTX Methotrexate

NICE National Institute for Health and Care Excellence

NLM National Library of Medicine

NPT Node Probability Table

NRAS National Rheumatoid Arthritis Society

OA Osteoarthritis

OMERACT Outcome Measures in Rheumatology Clinical Trials

ORF Original Release Format

OWL Web Ontology Language

PAG Patient Advisory Group

PAMBAYESIAN PAtient Managed decision support using Bayesian networks

PEAC Pathobiology of Early Arthritis Cohort

PIP Proximal Interphalangeal

PPI Patient and Public Involvement

PsA Psoriatic Arthritis

QMUL Queen Mary University of London

QoL Quality of Life

RA Rheumatoid Arthritis

RDF Resource Description Framework

RhF Rheumatoid Factor

ROC Receiver Operating Characteristic

RRF Rich Release Format

RSAB Root Source Abbreviation

RS Recommender System

RTX Rituximab

SCTID SNOMED CT Identifier

SEM Structural Expectation-Maximisation

SJC Swollen Joints Count

SMD School of Medicine and Dentistry

SNOMED CT Systematized Nomenclature of Medicine-Clinical Terms

SSZ Sulfasalazine

STM State and Transition Model

TAD Thyroid Autoimmune Disease

TB Tuberculosis

TCZ Tocilizumab

TJC Tender Joints Count

TNF Tumour Necrosis Factor-alpha

UA Undifferentiated Arthritis

UMLS Unified Medical Language System

UML Unified Modeling Language

URI Universal Resource Identifier

UTS UMLS Terminology Service

VAS Visual Analog Scale

W3C World Wide Web Consortium

WBC White Blood Cell

XML Extensible Markup Language

Chapter 1

An Introduction to the Research Project and Objectives

Machine learning methods have shown dramatic success in recent years, with numerous applications in medicine [159]. Particular success has been achieved with deep neural networks applied to images [94], particularly for diagnosis, or to large collections of patient data from electronic health record systems. Although deep neural networks have emerged as the most successful method for learning complex patterns in a large dataset, many other machine learning methods are available, including linear and generalised regression models, decision trees, ensemble methods such as random forests, support vector machine, and Bayesian networks (BNs) [158], allowing the user to trade-off the complexity of the model against the size of the dataset available.

This thesis is primarily concerned with a possible approach to enhance the management of chronic diseases. Since this would involve new patterns of medical care, there is, as yet, limited data. This precludes the use of those methods, such as deep neural networks, whose success depends on large datasets. Instead, the available data should be combined with knowledge at least until new and larger datasets become available. Given the importance of many decisions taken in medicine, machine learning methods that give probabilities to different outcomes are desirable as this allows an artificial intelligence system to operate simply and support decision-making by clinicians, in a way that is not so easy when using a machine learning method that gives a single most probable outcome.

BNs are a machine learning method that matches these criteria. BNs can be wholly or partly learnt from data or wholly created from knowledge, so are usable however much data is available. They are particularly suited to problems of reasoning under uncertainty, though limited in their application to very high dimensional data (such as images). BNs are popular decision support systems in medical applications [1, 110, 99] perhaps because

they can handle medical uncertainty with explainable models that can help clinicians to understand the basis of a prediction and also for their flexibility in being built from various sources including data and knowledge elicited from experts, literature, ontologies, or other sources [150, 54, 31, 190].

BNs built from data and/or knowledge create decision support systems that can improve medical services including diagnosis, treatment, and personalised care [110, 216]. With the expansion of communication using smart phones and the advent of care apps, personalised healthcare is getting more popular among care providers and patients. Personalised care implies the prediction of disease activity and progression using data remotely provided by patients [216], especially those with chronic diseases. This enables clinicians to provide more efficient care and also allows patients to be actively involved in the management of their own disease.

To achieve the self-management of chronic diseases, researchers from the School of Electronic Engineering and Computer Science (EECS) and the School of Medicine and Dentistry (SMD) from Queen Mary University of London (QMUL) proposed a joint research project on patient-managed medical decision support to monitor and control chronic conditions. The PAtient Managed decision support using Bayesian networks (PAMBAYESIAN) project intends to create clinical decision support systems for remote monitoring of chronic diseases with several case studies including one in Rheumatoid Arthritis (RA) [172, 152]. This thesis focuses on the RA case study of the PAMBAYESIAN project. It aims to assist clinicians in the diagnosis of RA, monitoring of disease activity and progression of RA, and help patients to manage their own disease. These capabilities could allow clinicians to taper or reduce the dose of medications, reducing side effects and saving the cost, especially of the expensive category of medications called biological disease-modifying anti-rheumatic drugs (bDMARDs).

RA is the most frequent inflammatory arthritis (IA), which affects 1% of the United Kingdom's population [184]. It is a chronic and autoimmune disease causing swollen, inflamed, and painful joints. Inadequate evidence of the onset of inflammatory arthritis diseases can result in diagnostic delays [77]. Rheumatology researchers have proposed various diagnostic criteria for RA, such as the American College of Rheumatology (ACR) 1987 criteria [9] and its revision jointly with the European League Against Rheumatism (EULAR) known as the 2010 criteria [6]. The latter are point-based and consider four domains, namely joint involvement, serology results, acute-phase reactants and duration of symptoms [93]. Similar guidelines from the National Institute for Health and Care Excellence (NICE) of the UK consider serology results, joint involvement, and functional ability [138].

Although the 2010 criteria have improved the early diagnosis of RA compared to the 1987 criteria, there is still a need to improve the accuracy of RA diagnosis [77].

Monitoring and treatment of RA aim to reduce patients' disease activity and ideally put them into 'remission' [138]. To achieve this target, clinicians prescribe various medications including conventional disease-modifying anti-rheumatic drugs (csDMARDs), bDMARDs, and steroids [138, 140]. The NICE guideline recommends reviewing serology results and disease activity monthly until the target of remission or low disease activity is achieved [138]. Medications may also be changed if the disease activity worsens persistently, i.e. if flares occur [17].

Apart from clinical care, patients need personalised care for living with RA since the disease can affect the psychological, social, and physical aspects of their lives. The National Rheumatoid Arthritis Society (NRAS) prepares various contents to provide advice for RA patients to enhance their quality of life (QoL); however, the advice is generic and not targeted to individual patients. The estimation of QoL can help to recommend personalised advice to the patients to adjust their lifestyle to their disease activity.

In this thesis, we build BN models for the diagnosis, self-management, and personalised care of RA using clinical expertise and data from our colleagues at the Centre for Experimental Medicine & Rheumatology of SMD. For expertise, we elicit clinical knowledge of our colleagues by having regular meetings. We were able to use two datasets collected in two separate studies called Pathobiology of Early Arthritis Cohort (PEAC) [157] and BioT DAS App project. The PEAC study recruited patients with inflammatory arthritis diseases including RA and gathered patients' clinical and treatment data quarterly. The BioT project, however, was a mobile app study collecting the records of RA patients to support remote monitoring and management of medications, particularly the bDMARDs [168, 112]. Neither dataset exactly matches our need to build a decision support using dynamic BNs (DBNs). The issue with the PEAC dataset is that it is quarterly recorded, which is not as frequently as we wanted (e.g., weekly), and the issue with the BioT dataset is that it contains a limited number of cases. Therefore, we decided to transform the PEAC dataset to make it suitable for building DBN models. We also built a BN model for personalised care, but the required data were not available or may not exist. Instead, we used experts' judgement and information extracted from the interviews with patients and published in the literature to create the decision support system.

1.1 Challenges in Modelling of Diagnosis, Self-Management, and Personalised Care

We need to deal with certain challenges in the modelling of medical diagnosis, self-management, and personalised care, which are listed below:

- **Remote monitoring as a decision support:** The main challenge in the modelling of self-management is that remote monitoring systems are not useful without providing decision support. For decision support, there is limited self-management data available. In our research, we have access to the PEAC and BioT datasets. The quarterly collected PEAC dataset and (almost) fortnightly collected BioT data provided us with the required data for building BNs. However, the former dataset is not as frequent as we need for self-management modelling and the latter has not enough cases needed for self-management modelling.

- **Knowledge acquisition**: In our research, we found it challenging to obtain domain knowledge or find established literature without having access to domain experts. There are various structured methods proposed to help modellers to elicit knowledge from experts to build causal BN models. One useful approach is to create medical idioms as suggested by [125, 101] and build medical BNs. Another way of acquiring medical knowledge is to outline clinical care pathways or caremaps to visually represent the process of care delivery as [55] explained. However, it may not be easy to find the idioms or to have access to clear care pathways. Another source of medical knowledge is the ontologies which store clinical terms, their properties, and semantic relations between terms. They allow knowledge acquisition to specify a set of customised terms and their relations, which is known as ontology [74]. Ontologies can help modellers to create meaningful structures for BN models automatically.

- **Personalisation of care and unavailability of data**: It is challenging to create a decision support that can properly identify the person and can recommend personalised advice to care. This required us to view the care of chronic diseases from the view point of a person with those diseases, which makes it different from the remote monitoring of chronic diseases by clinicians. The unavailability of any retrospective data of personalised care also adds to this challenge. This is because we cannot refer to the analysis of data, but we can only focus on patients' experiences, experts' knowledge, or literature.

1.2 Research Objectives

The overall objective of this thesis is that probabilistic models can be used to provide decision support for a chronic disease such as RA, covering diagnosis, self-management, and personalised care. The research objectives of this thesis address these challenges:

- Objective 1: Show how data and knowledge elicited from experts, literature, and clinical guidelines were used to build BN models for efficient prediction of medical diagnosis.
- Objective 2: Investigate how the knowledge of medical care processes is captured from care pathways and retrieved from pre-existing medical ontologies and translated into candidate BN models for medical diagnosis.
- Objective 3: Show how data, elicited knowledge from experts, clinical guidelines, and literature are used to build DBN models for self-management of chronic diseases and how these DBN models are used to support clinicians in monitoring and medication review.
- Objective 4: Investigate how to use the knowledge gathered from experts, interviews with patients, and literature help to build a BN model to estimate the QoL, and how to recommend personalised advice to improve the experience of living with RA.

1.3 Author's Contributions

The research objectives were achieved in collaboration with other researchers and clinicians working on the PAMBAYESIAN project. My specific novel contributions are as follows:

- For Objective 1, I show how to develop a BN model for the diagnosis of RA as described in Section 5.3. I show how a combination of data and knowledge help to build a BN model that is more efficient than another BN model built from data only (Section 5.4).
- For Objective 2, I show how to create BN models using knowledge represented in a care pathway and further knowledge retrieved from two medical ontologies of SNOMED CT and UMLS (Sections 6.3.1 and 6.3.2)
- For Objective 3, I show how to specify the variables of two DBN models for self-management of RA chronic disease from data, knowledge, guidelines, and literature.

I show how the initial DBN model is built containing monitoring observations and how it is able to estimate disease activity and flare (Section 7.3). By extending the initial DBN model, we create a DBN model with medication and adverse medication events are involved in the model. I explain how the adverse medication events, disease state, and flare help us to estimate the tolerance for DMARDs and disease control (Section 8.5). I then show how to do reasoning and elaborate how the estimation of disease state and flare with the initial DBN model can support clinicians in appointment scheduling (Section 7.4). I show the development of a DBN-based decision support system for self-management of RA that can help clinicians to review medications (Section 8.6). I do reasoning with the initial and extended DBN models to evaluate their ability to provide advice on appointment scheduling and medication review and support clinicians in monitoring patients remotely (Sections 7.5 and 8.7).

- For Objective 4, I show how to investigate the use of knowledge acquired from experts, coded interviews with RA patients, and established literature to build a BN model to estimate QoL (Section 9.4). The outcomes of the BN model are used in a rule-based system to first generate a set of advice and then modify the advice to obtain personalised advice that is linked to the recommendations of the NRAS to improve patients' experience of living with RA (Section 9.5).

1.4 Structure of Thesis

This thesis is formed of ten chapters as shown in Figure 1.1 and explained as follows:

Chapter 2 describes Bayes' theorem and the principles of BN models and reasoning with BNs. It introduces time-based BN models, mainly DBNs. This chapter includes a literature review on building BNs from data and knowledge. The chapter also reviews the literature on medical applications of BNs and clinical decision support.

Chapter 3 explains ontologies, the web ontology language, and ontology tools. It introduces the medical ontologies of SNOMED CT and UMLS. The chapter consists of a comparison between ontologies and BNs and a review of the literature on BNs built from ontologies.

Chapter 4 introduces the case study of RA. It presents four models of care pathways for RA applications. It contains a description and analysis of data collected in the PEAC study and Bio-T DAS app study. After data analysis, the selected decision support points of care pathways are explained in the chapter.

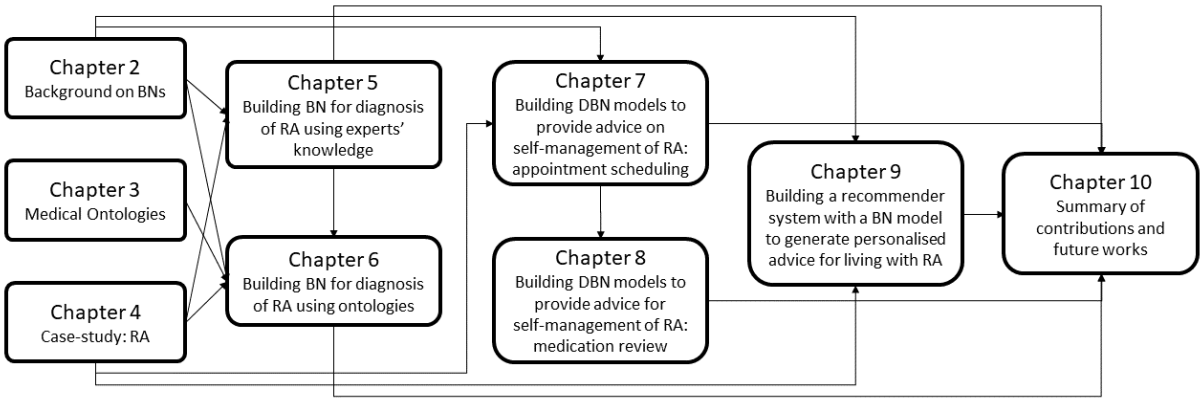


Fig. 1.1 Thesis structure.

The contributions of this thesis start in Chapter 5. Addressing the first research objective, this chapter introduces the diagnosis of RA and clinical decision support models that help in the diagnosis of RA. It describes the relevant variables for the diagnosis of RA.

Chapter 6 addresses the second research objective, with an analysis of care pathways and knowledge graphs created from medical ontologies. The chapter shows how to build BN models from care pathways and knowledge graphs created by retrieving knowledge from SNOMED CT and UMLS ontologies. It evaluates the BN models by comparing them with the BN models of Chapter 5.

Chapter 7 proposes a DBN model for self-management of RA. It includes a description of the relevant variables, the structure of the BN model, and the parameter learning. The model provides decision support for appointment scheduling. Together with Chapter 8, this chapter addresses the Objective 3.

Chapter 8 builds on Chapter 7. After describing medications for the treatment of RA, it proposes an extended DBN model which provides decision support for medication review.

Chapter 9 explains a BN model to estimate QoL and recommends personalised advice for living with RA based on a set of recommendations given by the NRAS. This chapter addresses the final objective.

Chapter 10 summaries the contributions made in Chapter 5, 6, 7, 8, and 9, and discusses the future directions of research.

1.5 Publications and Awards

The research presented in this thesis has resulted in the following journal publications:

- Kyrimi, E., Mclachlan, S., Dube, K., Neves, M. R., **Fahmi, A.**, and Fenton, N., (2021). A Comprehensive Scoping Review of Bayesian Networks in Healthcare: Past, Present and Future, *Artificial Intelligence in Medicine*, DOI: https://doi.org/10.1016/j.artmed.2021.102108.

- Kyrimi, E., Dube, K., Fenton, N., **Fahmi, A.**, Neves, M. R., Marsh, W., and Mclachlan, S., (2021). Bayesian Networks in Healthcare: What is preventing their adoption?, *Artificial Intelligence in Medicine*, DOI: https://doi.org/10.1016/j.artmed.2021.102079.

Our research has led to the following conference papers:

- **Fahmi, A.**, MacBrayne, A., Kyrimi, E., Mclachlan, S., Humby, F., Marsh, W., and Pitzalis, C., Causal Bayesian networks for medical diagnosis: a case-study in rheumatoid arthritis, 8^{th} IEEE International Conference on Healthcare Informatics, 30 Nov - 3 Dec 2020, e-conference. **This paper won the Best Student Paper award and received $500 prize**.

- **Fahmi, A.**, Soyel, H., Marsh, W., Curzon, P., MacBrayne, A., and Humby, F., From personalised predictions to targeted advice: improving self-management in rheumatoid arthritis, EFMI STC2020: Integrated Citizen centered digital health and social care - Citizens as data producers and service co-creators, 26-27 Nov 2020, e-conference.

We also presented our research in an online conference:

- **Fahmi, A.**, MacBrayne, A., Marsh, W., and Humby, F., A dynamic Bayesian network model for self-monitoring of rheumatoid arthritis, at STEMM MIT CSAIL AI in Healthcare Summit 2020, 1-2 Oct 2020. Awarded the Best Student Poster Presentation by the STEMM Global Scientific Community at STEMM MIT CSAIL AI in Healthcare Summit 2020.

Chapter 2

Introduction and Application of Bayesian Networks

Bayesian networks (BNs) are directed acyclic graphs consisting of a set of random variables, dependencies between variables, and underlying conditional probabilities. In this chapter, we describe BNs, how to build BNs, and a review of the application of BNs in medicine and clinical decision support (CDS) systems, which help us to contribute to Objectives 1, 3, and 4 covered later in Chapters 5, 7, 8, and 9. First, we introduce the components of BNs including variables, states of variables, and dependencies between variables in Section 2.1. We also explain the parameters of BN variables and Bayes' theorem and the inference with BNs in Section 2.2. Section 2.3 gives an introduction to time-based BNs - also known as temporal extensions of BNs, namely, dynamic BNs (DBNs), network of probabilistic events in discrete time (NPEDT), continuous time BNs, and irregular-time BNs (ITBNs). We briefly explain the inference with DBNs using inference algorithms and elaborate how a segment of DBN called 'window' are used for inference with DBNs.

We review various methods to build BNs in Section 2.5. This section separates the structure and parameters of BNs. Section 2.6 presents a literature review of the medical applications of static and time-based BNs. It separates diagnosis, personalised care and treatment, and reasoning and explanation of medical BNs. In Section 2.7, I have produced a novel and comprehensive review of the state-of-the-art of CDS for Rheumatoid Arthritis (RA) application (which is the case study of this thesis explained in Chapter 4) and for other diseases.

2.1 Principles of Bayesian Networks

BNs are probabilistic graphical models that represent a set of random variables with dependencies and conditional probabilities between them as directed acyclic graphs (DAGs) [158]. Figure 2.1 shows a simple DAG with two ovals of H and E respectively representing a hypothesis and an evidence variables. The variable H is considered as independent and the directed arc from H to E represents a dependency of E on H meaning E has a direct dependence on H and that E is dependent on H [60]. Variable H is called the 'parent' variable of the 'child' variable E. The child variable E has an underlying joint probability distribution, constituting its conditional dependency on the parent variable H and specifies its parameters, whereas the probabilities or parameters of H are independent.

Fig. 2.1 A simple DAG.

Mathematical background of BNs is built on the probability theory and the conditional probability. Dependent variables of BNs are formulated by a conditional probability rule using Bayes' theorem, which computes conditional probabilities [60]. Consider the simple DAG of Figure 2.1 to be a BN model. The conditional probability of E given H using Bayes' theorem is as follows:

$$P(E|H) = \frac{P(H|E)P(E)}{P(H)} \tag{2.1}$$

Figure 2.2 displays a simple BN model for a viral respiratory infection that we created just to illustrate BNs. We use AgenaRisk software to create the model [3]. It has five variables: Smoking, Tuberculosis, 'Lung Functioning', Fever, and Cough. We consider Smoking and Tuberculosis to have no parent variables, acting as risk factors, which can affect 'Lung Functioning' as a dependent variable. 'Lung Functioning' can be measured by two variables of Fever and Cough. We define two states of Yes and No for Smoking and Tuberculosis variables. We consider two states of Good and Poor for 'Lung Functioning' variable and two states of Yes and No for Fever and Cough variables. Prior probability values of two states of Smoking and Tuberculosis are shown in two tables next to each variable. The conditional probability tables (CPTs) or node probability tables (NPTs) of dependent variables 'Lung Functioning', Fever, and Cough are indicated next to their variables. All probability values are made up to demonstrate a simple BN model.

The full joint probability distribution of the simple BN model is shown in Equation 2.2 considering Smoking and Tuberculosis variables as the parents of 'Lung Functioning', and 'Lung Functioning' as the parent variable of Fever and Cough variables. We represent each variable with the initial letter of its name to simplify the equation.

$$P(S,T,L,F,C) = P(S).P(T).P(L \mid S,T).P(F \mid L).P(C \mid L) \tag{2.2}$$

2.2 Inference with Bayesian Networks

Inference - also called belief updating or probability propagation - with BNs refers to using BNs to update beliefs in a BN model by giving observations to an evidence variable and calculating the posterior probability of the hypothesis variables [96]. Therefore, BNs provide a naturally suitable basis for inference [110] and this allows reasoning with them.

Figure 2.3 shows the probabilities of the simple BN model for a viral respiratory infection, the observed variables, and the updated probabilities of unobserved variables. Prior probabilities of parent variables (Smoking and Tuberculosis) and conditional probabilities of children variables ('Lung Functioning', Fever, and Cough) are shown in Figure 2.3a. The probability of good functioning of the lungs is almost 68%. Given yes to smoking as an observation, the posterior probability of good functioning of the lungs falls to 34%, as displayed in Figure 2.3b. We can see the updated probabilities of Fever and Cough variables since the new information on smoking flows from Smoking variable to Fever and Cough via 'Lung Functioning'

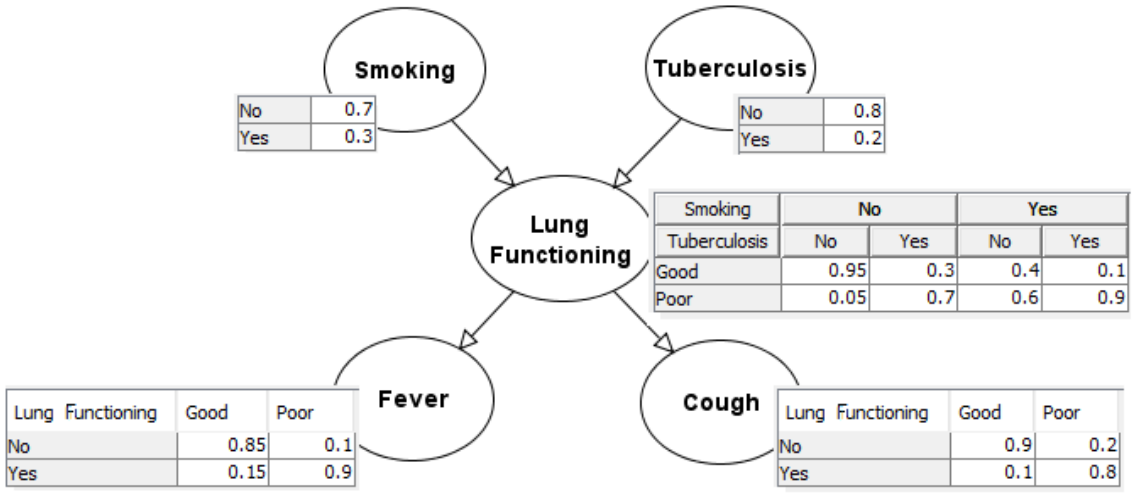
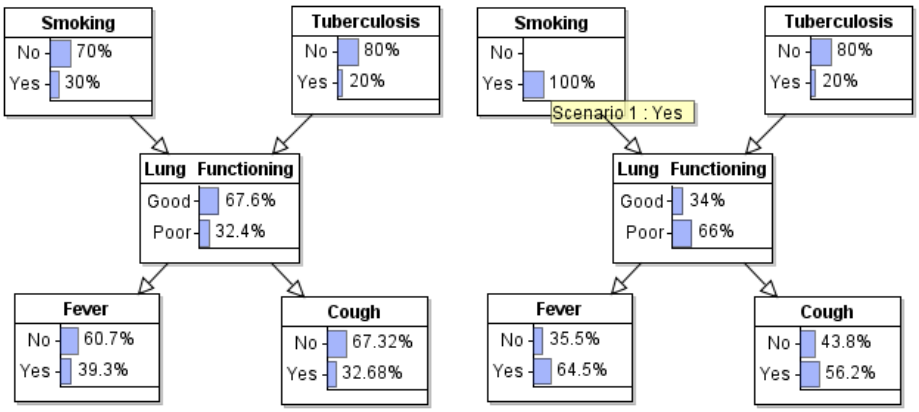
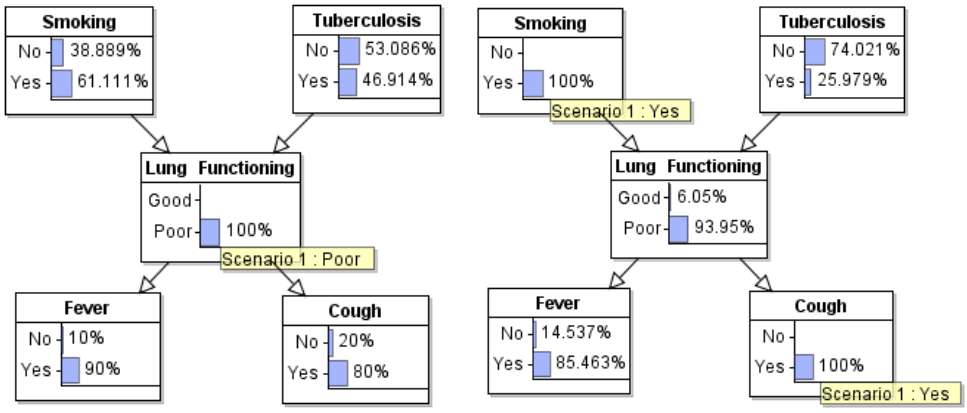


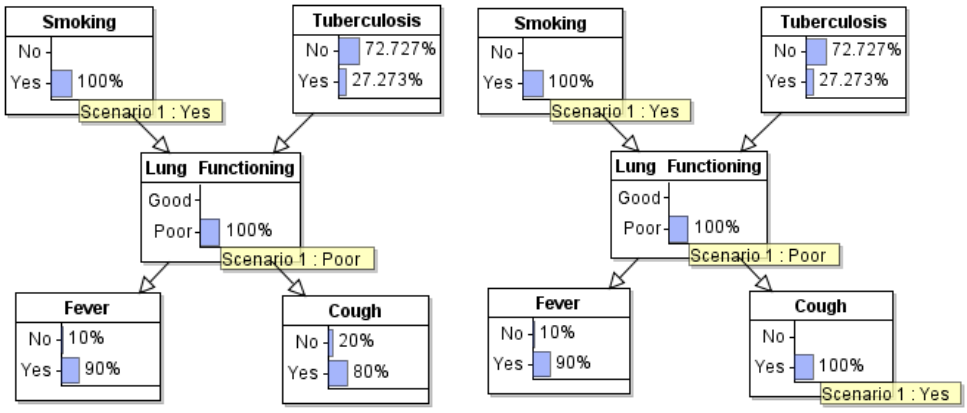
Fig. 2.2 A simple BN for a viral infection.



(a) No observation. (b) Smoking observed Yes.



(c) Lung Functioning observed Poor. (d) Smoking and Cough observed Yes.



(e) Smoking and Lung Functioning observed Yes and Poor, respectively. (f) Smoking, Lung Functioning, and Cough observed Yes, Poor, and Yes, respectively.

Fig. 2.3 Inference with a simple BN model for a viral respiratory infection.

variable. Smoking and Fever or Cough are known as dependency-connected or d-connected. On the other hand, Tuberculosis stays unchanged given observation to Smoking.

When an observation is entered in a BN, updating of the beliefs may occur in all directions, not just following the arcs of the DAG. As demonstrated in Figure 2.3c, an observation on 'Lung Functioning' updates the probabilities of Smoking and Tuberculosis variables in an indirect way. It means a person with poor Lung Functioning is more likely to be a smoker and/or suffer from tuberculosis.

A comparison between Figures 2.3b and 2.3d reveals that the update of belief goes from Cough variable upwards to Tuberculosis and also to Fever variable, given the observation of Yes to Cough variable. Although Fever and Cough have no direct dependency, the flow between Cough and Fever is through 'Lung Functioning'. This mutual parent variable did not exist between Smoking and Tuberculosis, therefore an observation on Smoking did not change the posterior probability of Tuberculosis as shown in Figure 2.3b.

Giving observations to the Smoking and 'Lung Functioning' variables updates the probability of Tuberculosis, as shown in Figure 2.3e, which is different from Smoking and 'Lung Functioning' being observed separately as displayed in Figures 2.3b and 2.3c, respectively. Despite the change of posterior probabilities of Tuberculosis and Fever when Cough was observed in Figure 2.3d, Fever and Tuberculosis stayed unchanged in Figure 2.3f compared to Figure 2.3e while new observation is given to Cough. This is because the observation on 'Lung Functioning' variable blocks the update of beliefs to go from Cough to Fever and Tuberculosis.

2.3 Time-Based Bayesian Networks

Time-based BNs allow BNs to include variables whose value varies with time. Various time-based BNs have been developed to represent the time dimension. These methods can be different in terms of structure, granularity, and computation and inference requirements. Some well-known time-based BN models are explained below:

2.3.1 Dynamic Bayesian Networks

DBNs were developed to involve time variables in BNs [42, 142]. They are the most widely used time-based BNs [150]. DBNs are discrete models which form a stochastic process in each time slice and repeat it over time. Simple DBNs cannot represent processes that evolve at different time granularities, but they show the whole system in the finest possible granularity.

A dynamic process via a set of stochastic variables connects the parts of model in different time slices. DBNs consist of (1) variables in each time slice, (2) intra-slice arcs representing

dependencies between variables of the same time slice, and (3) inter-slice or temporal arcs which connect the variables of two time slices. Assume time is represented with t and discretised into $t = 1, 2, ..., n$. Figure 2.4 displays a simple DBN model with n time slices and a hypothesis (H) and an evidence (E) in each time slice. The arc from H variables and their corresponding E variables are intra-slice and those between H variables are inter-slice.

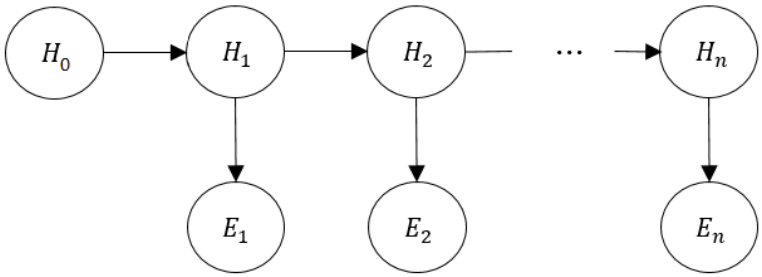


Fig. 2.4 A simple DBN model with n time slices.

Both intra-slice and inter-slice arcs quantify as conditional probabilities associated with intra-slice and inter-slice parents. Imagine H and E variables in each time are collectively called X. The distribution of X is parameterised as follows:

$$P(X_{1:n}) = \prod_{t=1}^{n-1} P(X_{t+1}|X_{0:t}) \tag{2.3}$$

DBNs are supposed to be time invariant, i.e., the model structure does not change between two time slices. Therefore, the word 'dynamic' in DBNs refers to the system being dynamic temporally, rather than the change in the network structure. Another assumption of DBNs is that they use the Markovian property (expressed in Figure 2.4), i.e., the conditional probability of each variable depends on its parent variable(s) in the same time or those of the previous time slice, not any variables from earlier time slices. Markovian property simplifies Equation 2.3 into:

$$P(X) = \prod_{t=1}^{n-1} P(X_{t+1}|X_t) \tag{2.4}$$

We illustrate DBN models by a simple DBN for a viral respiratory infection, as displayed in Figure 2.5. The model includes three variables of 'Lung Functioning' and two symptoms of Fever and Cough referring to high body temperature and continuous coughs. 'Lung Functioning' has two states of Good and Poor, and both symptoms are binary having two states of Yes and No. This figure demonstrates two time slices represented by 1 and 2, but it

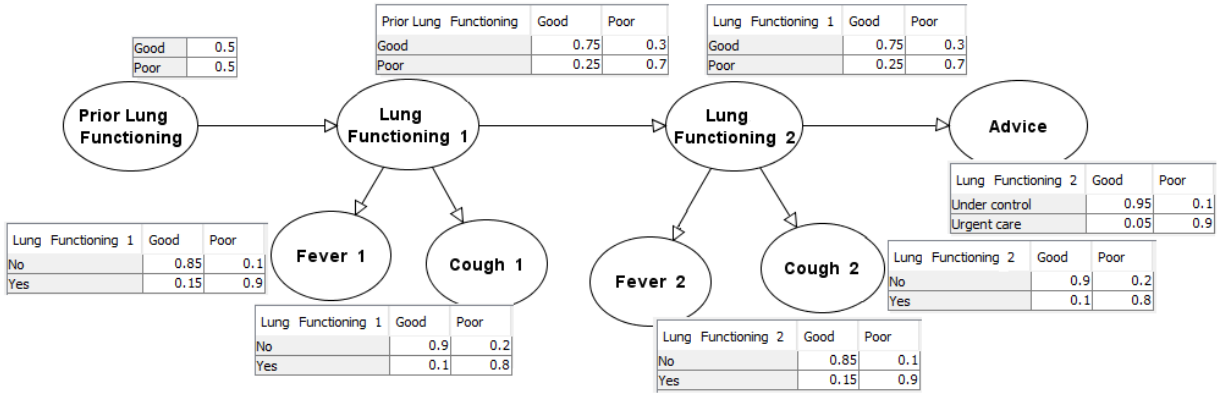


Fig. 2.5 A simple DBN for a viral infection.

can continue after time slice 2 with the same structure and NPTs. Variables of each time slice are differentiated by an additional time slice number. In Figure 2.5, NPT of each variable is shown next to the variable. NPTs of each variable in each time slice are repeated, e.g., fever variables 'Fever 1' and 'Fever 2' have identical NPTs. A variable called 'Prior Lung Functioning' is added to the model representing the initial lung functioning. It has two states of Good and Poor just like other lung functioning variable in time slices 1, 2, or others.

DBN models continue chaining until a terminal condition is met. Terminal conditions can be one or multiple variables with parent variables within the last time slice. Figure 2.6 shows a terminal condition called Advice representing a recommendation for decision-makers (e.g., clinicians) regarding the status of disease. It has two states of 'Urgent care' and 'Under control'.

This structure can be simplified by placing prior variables, variables involved in the dynamic process, and the terminal variable in three areas of initial conditions, temporal plate,

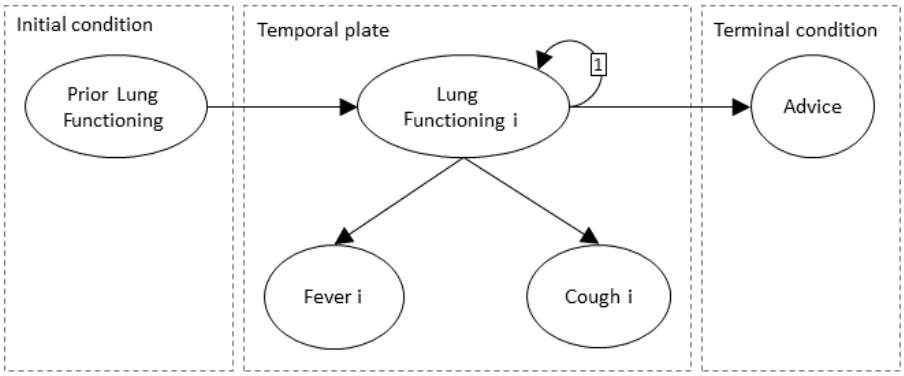


Fig. 2.6 Simplified notation of the simple DBN for a viral infection with a initial and a terminal conditions.

and terminal conditions, respectively. Figure 2.6 shows a simplified notation of the example of a simple DBN model. In this figure, the variables 'Lung Functioning i', 'Fever i', and 'Cough i' represent the lung functioning and two symptoms of fever and cough in the time slice *i*. The variable 'Lung Functioning i' has a first-order dependency which represents the temporal connections between lung functioning variables in two consecutive time slices.

We implement all BN and DBN models in AgenaRisk software. For implementation of the simplified notation of DBNs, we consider DBN models as BNs by chaining the BN models of each time and linking them using temporal dependencies.

2.3.2 Other Time-Based Bayesian Networks

Networks of Probabilistic Events in Discrete Time

DBNs are fixed time models which may not match with the real world cases. An alternative for DBNs can be a time-based BN model without fixed time. One of these models is called Networks of Probabilistic Events in Discrete Time or NPEDT proposed by [65]. This model has temporal random variables which represent the presence or absence at that time instant. The connections between variables represent the causal and temporal relationships between them. Temporal noisy gates help for knowledge acquisition and modeling of uncertain temporal knowledge, as well as the reduction of parameters from exponential to linear [107]. As shown in Figure 2.7, $E_1, E_2, \ldots, E_n$ are the parents of the variable H_n and they represent the temporality of E_i in a range between 1 to n.

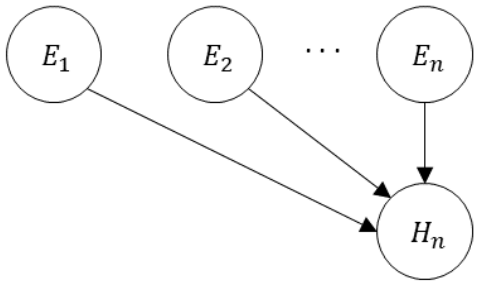


Fig. 2.7 A simple NPEDT model with n time slices.

NPEDT approach is more appropriate for domains that involve temporal fault diagnosis and prediction since it leads to less complex networks than those obtained from the formalism of DBNs, by assuming that each event occurs only once. Although not widely applied to medical reasoning, NPEDT models have been used in some cases such as diagnosis and prognosis of nasopharyngeal cancer in [64].

Continuous Time Bayesian Networks

CTBNs are another alternative to the fixed time DBNs. The states of CTBN variables change continuously, rather than in discrete time. CTBNs have two main components [149]: an initial distribution specified as a Bayesian network and a continuous transition model which has a conditional intensity matrix (CIM) for each variable to change the states. Time is explicitly represented in CTBNs and they are able to represent processes evolving at different granularities.

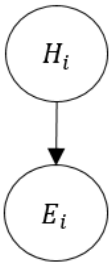


Fig. 2.8 A simple CTBN model.

A CIM for variables H_1, H_2, ..., H_n and E_1, E_2, ..., E_n represents the change of variable states throughout time. The CIM for E_i of the simple CTBN shown in Figure 2.8 is represented with $Q_{E_i|parents(E_i)}$ and calculated with Equation 2.5 as follows:

$$Q_{E_i|parents(E_i)} = \begin{bmatrix} -q_1(H_i,E_i) & q_{12}(H_i,E_i) & \cdots & q_{12}(H_i,E_i) \\ q_{21}(H_i,E_i) & -q_2(H_i,E_i) & \cdots & q_{2k}(H_i,E_i) \\ \vdots & \vdots & \ddots & \vdots \\ q_{k1}(H_i,E_i) & q_{k2}(A_i,E_i) & \cdots & -q_k(H_i,E_i) \end{bmatrix} \tag{2.5}$$

where the intensity q_i gives the probability of leaving state h_i by H_i or e_i by E_i and the intensity q_{ij} gives the probability of transitioning from h_i of H_i to h_j of H_j or e_i of E_i to e_j of E_j. A practical example of CIM calculation for a simple CTBN is presented in [107].

Inference in CTBNs has been tried to be done by exact algorithms like full amalgamation, but due to the exponential number of states to the number of variables, approximate algorithms such as clique tree inference, marginalisation, or expectation propagation can perform better [149].

CTBNs are appropriate models for domains where data have no natural time slices, such as in the diagnosis of cardiogenic heart failure. They are applied to some medical cases such as monitoring of chronic obstructive pulmonary disease [108], analysing colon cancer [181], and diagnosis and prognosis of cardiogenic heart failure [69].

Irregular-time Bayesian networks

ITBNs generalise DBNs in a way that each time slice may span over a time interval [170, 149]. ITBNs have the ability to compute probabilities given an evidence from the far past in one step. ITBNs structure can be learnt the same way as any BNs, when fixed time offsets between nodes are known or when the structure is fully observed.

Inference in ITBNs mostly involves either estimating unobserved nodes in a time slice or finding the representative time point of a time slice. Learning parameters of the ITBNs do not require to specify a temporal granularity or constant observation rate. ITBNs are more appropriate especially when the time point of interest is far into the future. They are also able to represent processes that evolve during time. ITBNs are applied in various medical applications, such as monitoring the glucose level of patients [170] and ovarian follicles detection [169].

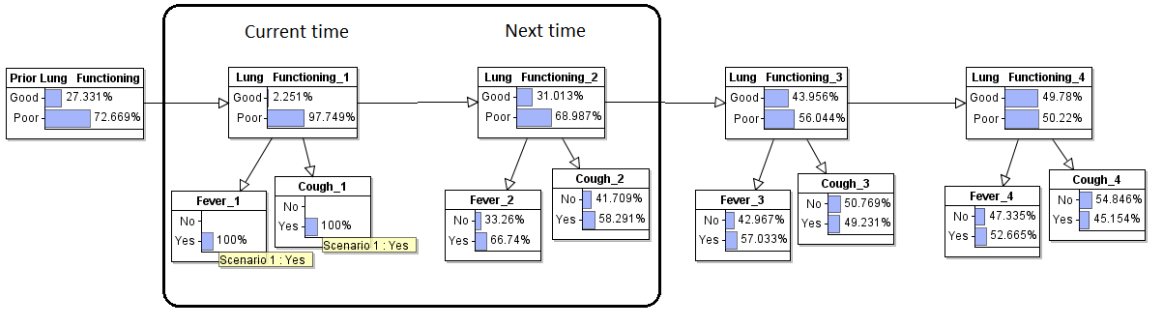
2.4 Inference with Dynamic Bayesian Networks

We focus on DBN models since they are popular time-based BN models. In this section, we introduce how to reason with DBNs. Inference with DBNs is essentially similar to the inference with BNs (explained in Section 2.2)–giving observations to evidence variables and updating beliefs. However, the belief updating in DBNs obtains the posterior probability of all non-evidence variables (H_i variables in Figure 2.4) in the time slice of the evidence variable and later time slices [96]. Updating the beliefs in a DBN model can help experts to do reasoning.

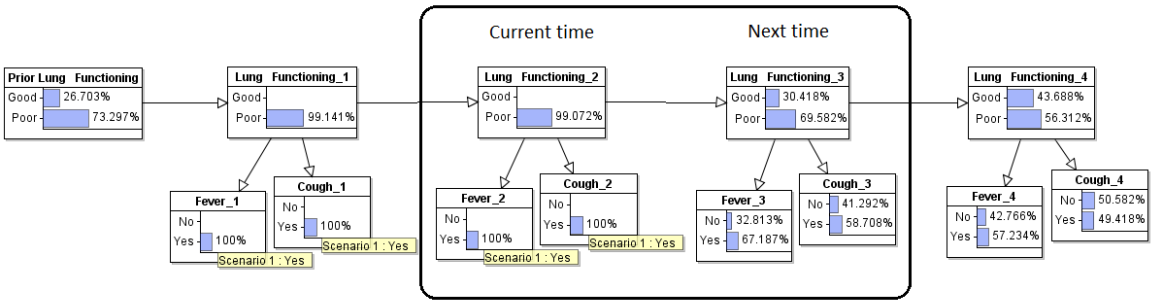
DBNs chain time slices and can grow longer quickly. It is recommended to limit the length of DBN time slices using a fixed size sliding window of time slices. The sliding window maintains a specific number of time slices and drops off an old time slice when the inference process moves forward with time [96].

In [96], Korb and Nicholson explain four steps of the algorithm of DBN updating process as 1) sliding, 2) prediction, 3) rollup, and 4) estimation. The sliding window moves from the initial state (Figure 2.9a) to one time slice ahead (Figure 2.9b). The model predicts the posterior probability of the hypothesis variable of the previous time, given the observation to the evidence variable of the same time. This posterior probability is used as the prior probability for the hypothesis variable of the current time that is known as rollup. Finally, the observation is given to the new current time and the model updates belief of the hypothesis variable of the next time.

There are two main categories of inference algorithms for DBNs: exact and approximation. Exact algorithms using clustering can be used in DBNs with a short chain of two



(a) Sliding window in the beginning.



(b) Sliding window moving one time forward.

Fig. 2.9 Inference with a simple DBN for a viral respiratory infection maintained as a sliding window of two time slices and having observed evidence variables.

time slices. However, since exact inference is computationally complex [37], approximate algorithms using stochastic simulation are more suitable for inference with DBNs with longer than two time slices or densely connected [96]. Stochastic simulation generates a large number of cases from the network and allows to calculate the posterior probability of the target variable. Various methods are proposed to use stochastic simulation, including arc reversal method [194], filtering method (e.g., Murphy's method [123]), and ignoring weak dependencies in the stochastic process(e.g., Boyen and Koller's method [19]).

As mentioned in Section 2.3.1, we consider DBNs as multiple BN models chained to each other and we use AgenaRisk software to implement it. We need to define the number of time steps of the temporal plate which is shown in Figure 2.6. This time step represents the sliding window required for inference with DBNs. The size of the sliding window can be defined by adding time slices one-by-one and measuring the changes of posterior probabilities of the

target variable (e.g., 'Lung Functioning' in Figure 2.6). We then determine the size of the sliding window, once the probabilities stabilise (described in Sections 7.5 and 8.7).

We implemented DBN models with sliding windows in AgenaRisk [3] by assuming the models as BNs that are chained within a window. When time moves forward, the window also slides forward and one time slice drops out of the window. The posterior probabilities of the variable(s) with temporal connections in the dropped time slice are kept as the prior probabilities for the network within the window. The variable with prior probabilities is shown on the left side (initial conditions) of Figure 2.6.

2.5 Building Bayesian Networks

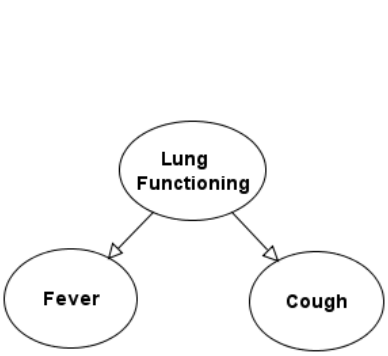
BNs are characterised by a graph and parameters (described in Section 2.1), which can be specified from knowledge and data. Knowledge and data help to perform two main tasks for building BN models: determining the BN structure and specifying its parameters. Required knowledge and data can be obtained by eliciting experts' knowledge, investigating literature, using available data, knowledge retrieval from ontologies, or any combination of these. We explain how to specify the structure and parameters of BNs from expertise, data, ontologies, and hybrid approaches.

2.5.1 Structure Learning

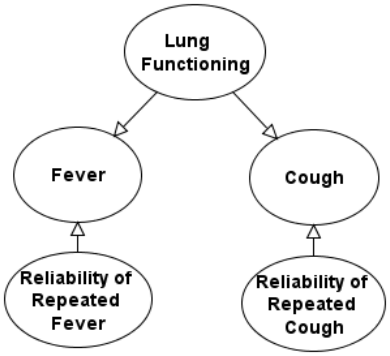
Using Experts' Knowledge

Expert elicitation methods have been classically conducted to extract knowledge directly from domain experts. Elicited knowledge is used for building the structure of BN models. As examples, BN modelling using expert knowledge was applied to build a prognostic support system through DBN [214], questioning from the main expert and external experts was used to construct naive BN models [56], and an expert-based BN modelling was developed for structure learning [141].

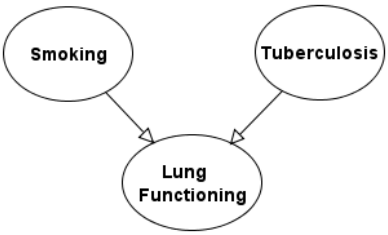
Building the structure of small BN models may be easy, but making large models is complicated and needs appropriate approaches to be applied. One method to build large-scale BN models is to create BN building blocks known as idioms that consist of a small number of variables acting as a library of patterns of definitional/synthesis, cause-consequence, measurement, induction, or reconciliation [125]. A medical extension of these idioms suggests specific idioms for medical BN development that include manifestation, manifestation reliability, risk factor, pathogenesis, comorbidity common cause, comorbidity common symptomology, complication, treatment, treatment reliability, counterfactual treatment, and



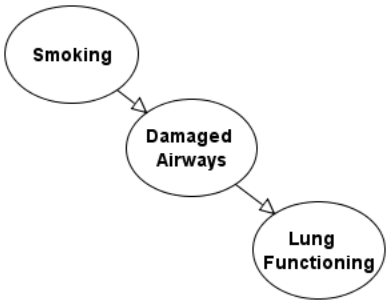
(a) Manifestation idiom.



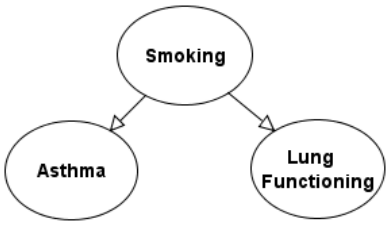
(b) Manifestation reliability idiom.



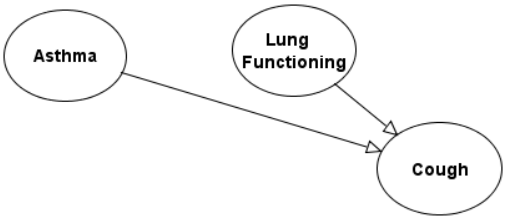
(c) Risk factor idiom.



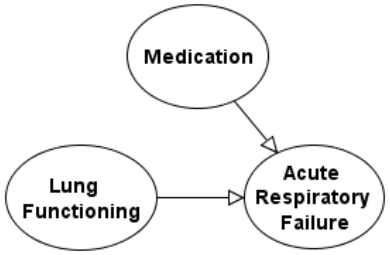
(d) Pathogenesis idiom.



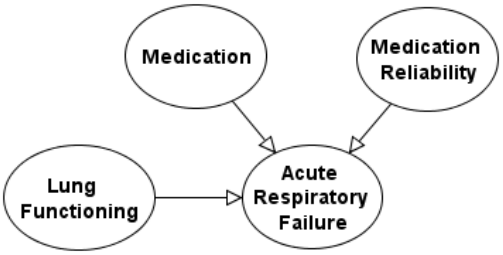
(e) Common cause idiom.



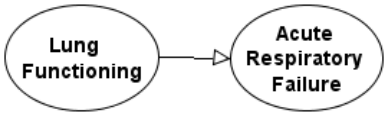
(f) Common symptomology idiom.



(g) Treatment idiom.



(h) Treatment reliability idiom.



(i) Complication idiom.

Fig. 2.10 Medical idioms proposed by [101] for a simple BN model for a viral respiratory infection.

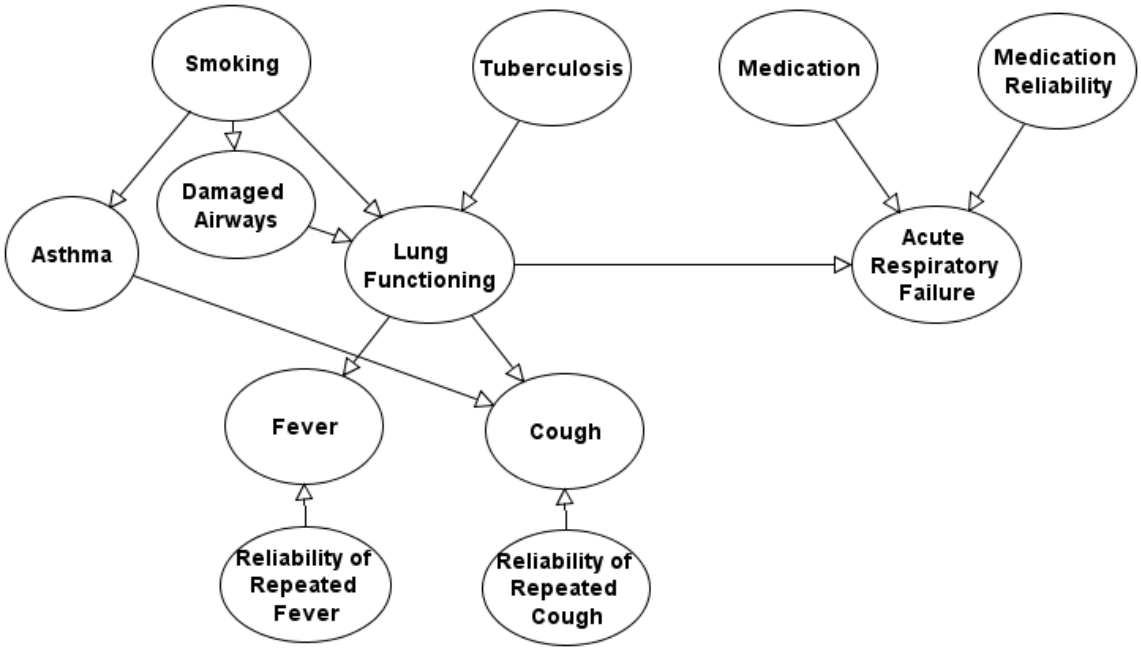


Fig. 2.11 Combination of medical idioms proposed by [101] for a simple BN model for a viral respiratory infection.

the combination of medical idioms [101]. Examples of medical idioms for building a BN model for a viral respiratory infection are presented in Figure 2.10 and a combination of those medical idioms is displayed in Figure 2.11.

Apart from the idioms, some suggest to use pre-defined levels for building the structure of BNs based on the two-level structure [165]. Later, [190] built BN models for the diagnosis of neurological diseases from expert's knowledge using a three-level structure containing background information, diseases, and symptoms, signs, and test results in each level.

To incorporate the element of time into BN modelling, [24] suggested using temporal abstractions for the analysis and interpretation of longitudinal data of monitoring patients with chronic diseases proposed by [15]. Temporal abstractions are defined as a knowledge-based process for abstracting raw temporal data into higher level interval-based concepts [150]. In [24], first, a pairwise temporal relationship is specified between two events which is called fluent, then fluent relations are transformed into BNs. These BN models can maintain the reasoning ability of BNs and address two limitations of alternative temporal BN models, DBNs, namely, the need for representing time at different levels of granularity; however, it is unclear how long the hierarchy of the suggested BN models can continue. Temporal abstractions were later extended to DBNs [150]. They explained two main categories of temporal abstractions: basic and complex. Basic temporal abstractions entail states,

trends, and persistence, and used these abstractions to specify the variables and temporal dependencies of DBN models.

Some researchers employ qualitative approaches or knowledge representation models to elicit knowledge from domain experts. To build DBN models, [143] suggests using the state and transition models (STMs), which are qualitative descriptive methods for modelling ecological processes. In [191], an online surveying tool is used to acquire the knowledge of a large number of experts to build BNs, instead of eliciting knowledge from individual experts. Using a STM for intensive care, [161] proposed a hierarchical DBN model and logistic regression to estimate the temporal pattern and predict the outcome in intensive care .

Using Data

The structure of BN models can be learnt purely from data. Algorithms for structure learning from data are classified into three groups of score-based, constraint-based, and hybrid. Score-based algorithms list a set of candidate models with assigned parent variables for each variable and then optimise the assignment of parents or score the candidate models [182]. Score-based algorithms commonly score candidate models using posterior conditional probabilities by Bayes Dirichlet equivalent uniform (BDeu) or using the Bayesian information criterion (BIC) [183, 185]. There are several other functions to score the candidate models such as log-likelihood.

Greedy search algorithm is a widely applied score-based algorithm [188]. As a greedy algorithm, Hill Climbing searches locally to create an initial DAG and performs a number of steps to reach the greatest improvement in its scoring function [66]. CaMML or causal minimum message length performs two-phase search with simulated annealing method and then Markov Chain Monte Carlo (MCMC) method and uses the minimum message length (MML) metric to score the candidate models [61]. By incorporating experts' views as prior knowledge, [61] applied CaMML score-based structure learning algorithm to develop BN models. In a similar study, [174] proposed a hybrid approach using elicited knowledge as priors and CaMML algorithm to develop DBN models.

Although the problem of missing values affects different categories of structure learning algorithms differently, it is still an issue for ending up with a BN. As an extension of expectation-maximisation or EM (explained in Subsection 2.5.2), structural EM (SEM) can handle missing values by performing a two-step iteration of measuring expected sufficient statistics (E-step) and searching for a network with maximum likelihood (M-step) as if the data were complete [183, 62]. [2] proposed an algorithm to translate the input dataset into complete data and performs structure learning by augmenting the network with variables that

take into account all possible completions of the data. Further structure learning algorithms with the ability to deal with incomplete data are explained in [183].

Constraint-based algorithms are based on causal graphical models and inductive causation algorithm [217]. These algorithms use conditional independence constraints and assume conditional independence to imply graphical separation, i.e., no arc between two variables. Common constraints or tests are the mutual information test and the exact Student's *t* test for correlation [185]. PC is the well-known constraint-based algorithm that creates a complete graph with undirected dependencies and then tests the dependencies and deletes the null ones [204]. The final model is a partially directed graph which can represent the underlying DAG. Other constraint-based algorithms are Grow-Shrink (GS), Incremental Association Markov Blanket (IAMB), and Max-Min Parents & Children (MMPC).

Hybrid structure learning algorithms mix score-based and constraint-based approaches, which respectively help hybrid algorithms to identify optimal models and restrict the space of candidate models [185].

A package in R programming language called bnlearn makes many structure learning algorithms available, including hill climbing, PC, GS [187], as well as SEM which is able to deal with missing values [186]. SEM in bnlearn has several arguments, mainly `impute`,`maximize`,and `fit`,which respectively employ an imputation method for the expectation step, a score-based algorithm for the maximisation step, and a parameter learning method in the maximisation step [186]. SEM is broadly used for structural learning of BN model, e.g., [212, 173, 55].

2.5.2 Parameter Learning

Using Experts' Knowledge

Expert elicitation methods have been classically conducted to acquire knowledge directly from domain experts and build BN models using expert knowledge such as [164, 227], and some others recommend to use expertise combined with structure learning [61]. Although these can be done to make simple BN models, complex BN models with large data and knowledge entry require a framework to simplify the building of models.

Using Data

The parameters of a BN model can be learnt from data. One way of parameter learning from data is to find the parameters that maximise the likelihood of the model given the available data. Data are used for parameter learning of BN models in [190], but the structure of the models are built from expert's knowledge. In [174] and [61], BN parameters are trained from

the available data. These modelling approaches are also called hybrid since they employ multiple sources for structure and parameter learning.

Some datasets are not complete and require dealing with missing values. The EM algorithm enables parameter estimation in probabilistic models such as BNs, if there is a limited percentage of missing values and if they are missing at random (MAR) [175, 44, 98]. It has two main steps called E-step and M-step referring to expectation and maximisation, respectively. Assume θ^* is the parameter that EM searches to maximise its posterior probability as shown below:

$$\theta^* = \underset{\theta}{argmax} P(\theta|U) \tag{2.6}$$

where θ is marginalised over the data U is given. In the E-step, the algorithm chooses a local lower bound to the posterior distribution, then in the M-step, it optimizes the bound. The iteration between the two steps continues and the objective function ($P(\theta|U)$) monotonically increases until it converges to a local maximum of the objective function [43, 49]. EM algorithm is widely applied for parameter learning of BN models with data containing missing values, e.g., [148, 150, 226, 212, 36, 35, 227].

The AgenaRisk software allows us to incorporate knowledge in parameter learning using the EM algorithm. The percentage of knowledge incorporation can be customised for discrete variables using the following equation:

$$\alpha = r \times \alpha_{knowledge} + (1 - r) \times \alpha_{data} \tag{2.7}$$

where $\alpha_{knowledge}$ represents the parameters elicited from experts, α_{data} is the parameters learnt from data, and r is the relative weight of expert knowledge and available data, which we calculate it as follows:

$$r = \frac{\#\ of\ missing\ values}{\#\ of\ overall\ values} \tag{2.8}$$

2.6 Medical Application of Bayesian Networks

Once the structure of BNs is built and their parameters are learnt, they can be applied to particular domains. BNs are widely developed in various medical applications of diagnosis, prediction of specific medical outcomes, and analysis of treatment [59]. There are many static BNs applied to diagnosis and treatment and various time-based BNs for monitoring of diseases.

2.6.1 Diagnosis

Many BN models have been developed for medical diagnosis using data, knowledge, medical literature, ontologies or other sources of medical evidences, or a combination of these sources (as explained in Section 2.5).

Some BN models are purely learnt using the knowledge elicited from domain experts, such as [121] that presents a BN model for diagnosing preeclampsia, whereas some other BN models are built purely from data. For example, [120] combines structured data with information extracted from medical notes (i.e., text) to construct a BN model for dementia diagnosis. In [203], four data-driven classifier models, namely, rules, logistic regression, tree, and Bayes net, are trained to recognise primary hyperparathyroidism.

[190] developed multiple BN models for the diagnosis of Alzheimer's disease, dementia, and mild cognitive impairment using both knowledge and data, and compared their performances with those of other classifiers learnt entirely from data. The performance of knowledge-based BN models with AUROCs (or area under the receiver operating characteristic (ROC) curve) for Alzheimer's disease, dementia, and mild cognitive impairment are 0.86, 0.96, and 0.97, respectively, which is comparable to that of the best data-driven classifier with AUROCs of 0.90, 0.98, and 0.97, respectively. In contrast, [229] analysed and identified the most important variables and learnt a BN model for the diagnosis of Alzheimer's. To differentiate dengue from other acute febrile illness, [176] builds a BN model from knowledge elicited from domain expert, literature, and available data. Using a hybrid approach, [56] built a BN for diagnosing musculoskeletal disorders of the shoulder from expert knowledge, medical literature, and a set of retrospective data. The expert elicitation used two stages of initial modelling and then a review by a panel of orthopaedic specialists.

Some studies suggest retrieving and learning medical knowledge from medical dictionaries or ontologies. [21] proposed a two-level BN model for diagnosis built by automatically extracting medical knowledge from an ontology; the diagnostic part is a supplement to a decision network to evaluate available tests using utility values specified by a physician. In [230], multiple models including BNs, C4.5, and support vector machine, are automatically developed to diagnose mild cognitive impairment using a hybrid approach including knowledge extraction from the SNOMED ontology.

We reviewed many studies on BN models for diagnostic applications; however, we did not come across any BN model for the diagnosis of RA that we intend to focus on. This shows the need to build a BN model for diagnosing RA using different sources, including experts' knowledge, available data, or ontologies.

2.6.2 Personalised Care and Treatment

Personalisation of care focuses on monitoring disease progression based on the interpretation of patient data using probabilistic models such as BNs [216]. One of the first attempts to create a BN model for personalised care was a temporal BN to forecast disease progression of preeclampsia [216] using physiological knowledge and patients data. This study proposes a generic structure of state and temporal BNs for diagnosis and remote monitoring of medical conditions.

As an early attempt to use machine learning in personalised medicine, [80] built a fuzzy inference system to predict the effect of regular aerobic exercise on blood pressure using evidence published in the medical literature. In contrast, [220] used multiple classifiers, including relational probability trees and relational functional gradient boosting algorithms to predict myocardial infarction as a personalised tool.

To investigate the interrelationships between sleep duration and physical activity in patients with self-reported stroke, [189] reviewed the literature, analysed a large dataset and built multiple BN models using data-driven algorithms and found an augmented naive Bayes model to be the best model. In [108], two time-based BN models of DBN and CTBN are developed to self-manage chronic obstructive pulmonary disease. They analyse a regular dataset and another irregular one, and apply structure learning algorithms which resulted in outperformance of DBN model in regular data and CTBN in irregular data.

The visual representation and inference ability of BNs make them appealing for medical treatment applications. In [192], a BN model is developed for treatment selection and personalised survival estimation of lung cancer. The performance of this model is compared with that of a BN model learnt from data using the CaMML algorithm. The AUROC of the knowledge-based and the learnt BN models are 0.75 and 0.81, respectively. The predictive performances of the BNs are compared with those of the complex classifiers, including naive Bayes, logistic regression, and C4.5 decision tree.

Although we retrieved and reviewed multiple BN models for personalised care and treatment of different diseases, we did not find any DBN models applied to the management of RA chronic disease. This reveals that the state-of-the-art on BNs for personalised care and treatment applications lacks BNs (including DBNs) for personalised care and treatment of RA.

2.6.3 Reasoning and Explanation

Probabilistic reasoning using probabilistic models can quantify and provide more precise answers for clinical problems, unlike deterministic approaches with rule-based systems [193].

BNs provide the formalism to do reasoning on conditional uncertainty when partial evidence is available [158]. The reasoning with BNs enables modellers to explain the model since creating an explainable model is one of the objectives of building BNs [102, 100]. Explanation of reasoning with BN models can provide the results obtained by the model, or the results not obtained by the model, or hypothetical reasoning [100], which could lead to evaluate the performance of the model.

In [100], an algorithm is proposed to explain the reasoning with BNs. The algorithm considers an observation being given to the evidence variables or intermediate variables, then identifies the changes in the model. Finally, it produces three levels of explanation: the impact of evidence variables on the target variable given observations to evidence variables (level 1), information flow given observations to the intermediate variables (level 2), and the impact of the significant evidence variable on the intermediate variables (level 3).

As mentioned in Subsections 2.6.1 and 2.6.2, we did not come across any BN models for diagnosis, personalised care, or treatment of RA, and therefore we did not find any study of BN models that reasons with the model or explains it. Thus, not only there are needs to build BN models for diagnosis, personalised care, and treatment of RA, but also we can do reasoning with these models and explain how they perform diagnosis, personalised care, or treatment.

2.7 Clinical Decision Support

Clinical decision support or CDS tools perform high-level cognitive functions helping clinical stakeholders [190], mainly clinicians and patients, to collect data and make decisions on diagnosis, treatment, self-management, and other clinical processes. CDSs can contribute to reasoning, learning, and decision-making using inference engines that can be based on BNs, a set of rules, or semantic networks [95, 190].

BNs are a popular technique for decision support in medicine [1, 110] perhaps because they can be built using a combination of data and knowledge elicited from domain experts or literature [31, 150, 190, 54]. BNs are also favourable CDSs because of several algorithms developed to make probabilistic inference [95, 100]. The use of knowledge is an advantage when there is a lack of sufficient good-quality data, but essential when the application requires a causally coherent model.

We review some CDSs applied to various clinical applications, treatment and self-management of chronic diseases and RA in the next subsections.

2.7.1 Rheumatoid Arthritis Applications

Various machine learning methods have been applied for the diagnostic and treatment of RA. We find different classifiers applied to diagnose RA, however, to the best of our knowledge, there is no BN model for RA diagnosis. Other classifiers are used for RA diagnosis, mainly by exploiting data. In [179], a combination of data and expert knowledge is used to develop an RA diagnostic tool. The method uses fuzzy logic and neural networks. A team of orthopaedic surgeons and rheumatologists were consulted to determine the relevant variables including symptoms, medical tests and disease severity measurement. Another study [198] used multiple machine learning algorithms including decision trees (using a variety of algorithms) and support vector machine, using a variant of the AdaBoost, showing that their algorithm outperforms others in distinguishing RA from other rheumatic diseases (at highest 81% versus 85% accuracy). Separately, [67] also applied AdaBoost using social and medical data to separate RA from fibromyalgia. In [195], decision trees are trained using personal and medical data of patients using a feature selection algorithm. The classifier is used to differentiate RA patients from non-RA ones.

Some studies investigate the relationships between clinical, demographic, laboratory, and genetic factors on each other in patients with RA. In [109], RA patients are classified based on their pain, then multiple supervised classifiers (e.g., random forest, principal component analysis, and naive Bayes) are used to discover connections between disease manifestation and personal factors which can help to predict persistent pain in RA patients. Another study focused on identifying epistatic risk factors with PTPN22 gene using machine learning methods of random forest and logistic regression [20]. Using random survival forest, [105] predicted the mortality of RA patients considering their demographic and clinical factors.

Machine learning is also applied to the treatment of RA. A Gaussian process regression model is used to differentiate RA patients who respond to anti-tumor necrosis factor (anti-TNF) drugs from non-responder ones, using demographic, clinical, and genetic markers of patients [75]. Another study employed several machine learning methods, namely, logistic regression, least absolute shrinkage and selection operator (LASSO), random forest, and extreme gradient boosting (XGBoost) to identify insufficient response to a disease-modifying antirheumatic drug called methotrexate [72]. In [85], authors resent a literature review on machine learning methods applied to arthritis diseases and stated that the major unmet need for arthritic patients particularly people with RA is a personalised treatment approach.

This literature review expresses the lack of BN-based CDSs for various RA applications, including diagnosis or management of RA. This gap can be filled by building BN models and supporting clinicians in diagnosing and management, and also assisting RA patients to self-manage their own disease.

2.7.2 Other Applications

CDSs are developed for various clinical applications including the treatment of arthritis, cancer, mental disorders, and chronic diseases. In [104], an artificial neural network model is developed to distinguish early-TNF users from non-early-TNF users in the treatment of ankylosing spondylitis. [22] developed a CDS to communicate with patients with polyarticular juvenile idiopathic arthritis to monitor them using a set of activities, called algorithms, for each phase of arthritis, namely, new diagnosis, remission, medication tapering, and flare. In [156], a CDS is proposed for pregnant women with systemic lupus erythematosus using logistic regression for feature selection and a neural network for prediction of pregnancy outcomes: live births, spontaneous abortions, and second-trimester abortions. [73] suggested a set of criteria to manage the order of a blood test known as Erythrocyte Sedimentation Rate (ESR).

For treatment selection and personalised survival estimation of lung cancer, [192] proposed a CDS with a BN model using expert knowledge, and its performance is compared with that of a learnt BN model. In [193], the BN-based CDS for treatment selection and survival estimation of lung cancer [192] is extended and a software architecture of the CDS prototype is proposed. The prototype has two sides called server and client. The server side has an ontology, a database, and a BN component, which process predictions and provide decision support through a user interface in the client side.

Some of the CDSs are called recommender systems (RSs) which filter information to provide the anticipated recommendations [28, 180]. This enables RSs to personalise information using the rating or preferences of the person estimated by one or multiple prediction methods [28].

For mental health applications, [30] suggested using a set of rules to create an ontology of depression and then a BN model provides a diagnosis. [83] developed a CDS for the diagnosis and treatment of depression by gathering information and communicating with the users. In [23], a BN model with a meaningful structure is proposed that captures the latent variables related to psychotic depression, indicating the dynamics of disease progression.

For chronic diseases, [70] proposed an RS to create a profile for patients and capture spatial and temporal information to recommend appropriate exercise advice. A similar system provides recommendations for the diet of patients with chronic diseases using a rule-based mechanism [33].

Some RSs employ ontologies in Web Ontology Language (OWL) technologies to capture users' preferences and filter information accordingly. They are widely used for diabetes applications. An OWL ontology-based CDS is proposed by [7] to recommend treatment for diabetes using clinical practice guidelines (CPGs) and considering various factors such

as signs, symptoms, and serology results. Similarly, [84] proposed a CDS tool for the diagnosis and treatment of diabetes. [113] developed two ontologies for patients and drugs to recommend personalised treatment for diabetes. Apart from diseases, ontology-based RSs are also developed to store user profiles and provide personalised food recommendations [4, 206].

2.8 Summary

In this chapter, we introduced the principles of BNs and Bayes' theorem, explained the inference with BNs, and illustrated them with an example of a simple BN model. We introduced time-based BN models. We focused on DBNs and explained how to do inference with DBNs using sliding windows. We also touched upon other time-based BNs, namely, NPEDT, CTBNs, and ITBNs.

We presented the structure learning and parameter learning methods to build BN model. These help us to build BN models and/or reason with them to contribute to Objectives 1 and 4. They also provide the required materials to build DBN models and reason with them to contribute to Objective 3.

We reviewed the state-of-the-art of BN models for medical applications, considering diagnosis, personalised care and treatment, and reasoning and explanation of BNs. Finally, we also presented a literature review on CDSs for RA application as well as other medical conditions. Our novel literature review shows the lack of BN models for diagnosis, management, and personalised care of RA and the need for our contributions in this thesis. The contributions to Objectives 1 and 4 are presented in Chapters 5 and 9, respectively. Chapters 7 and 8 together present our contribution to Objective 3.

Chapter 3

Medical Ontologies

In Chapter 2, we have explained how the structure and parameters of BNs are learnt from experts' knowledge and data that can help us to contribute to Objectives 1, 3, and 4. In the current chapter, we investigate whether the work of knowledge elicitation can be reduced by retrieving knowledge from ontologies; this is covered later in Chapter 6, corresponding to the Objective 2. This chapter provides relevant background on ontologies. Ontologies consist of a large set of concepts and their relationships with other concepts. While ontologies represent the knowledge of a domain, their main aim is to share knowledge and make it reusable in a human-understandable and machine-readable language. Medical ontologies such as Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) and Unified Medical Language System (UMLS) are a collection of medical terms, their codes, synonyms, definitions, and various properties.

In this chapter, we present a background on ontologies in Section 3.1. Section 3.2 presents a step-by-step method to develop ontologies. In Section 3.3, we describe the semantic web and two languages for programming ontologies: Resource Description Framework (RDF) and OWL. Section 3.4 covers two pre-existing medical ontologies of SNOMED CT and UMLS, including their components and capabilities. In Section 3.5, we contrast care pathways, knowledge graphs (i.e., a fraction of knowledge retrieved from pre-existing ontologies), and Bayesian networks (BNs). In this section, I provide a novel review of the state-of-the-art of building BNs from ontologies.

3.1 What is an Ontology?

The word "ontology" is taken from philosophy, meaning the study of being; however, in computer science, ontology refers to the specification of a conceptualisation meaning a formal and declarative representation of an abstract and simplified part of the world [68, 74].

Ontologies try to provide a common language to share knowledge in a domain including vocabulary, classification, relations (hierarchies and constraints), and domain axioms [68].

Vocabulary or terminology is the collection of terms in an ontology [68]. These terms indicate the concepts or classes of a domain of knowledge [144]. Vocabulary can be in the form of a controlled vocabulary, a glossary, or a thesaurus [68]. A controlled vocabulary refers to a finite list of terms and their unambiguous interpretations such as a catalog. A glossary is a list of terms and their meanings in natural language, thus their ambiguity makes them unsuitable for machines to read and interpret. A thesaurus is a vocabulary with some additional semantics as relations between terms.

We explain the vocabulary for a simple ontology for viral infection (similar to the idea of the simple BN model that we made in Section 2.1). Figure 3.1 shows a set of terms of the vocabulary of the simple ontology for a viral respiratory infection. Each term is shown in a rectangular shape representing the concept. There is no order or hierarchy between the concepts.

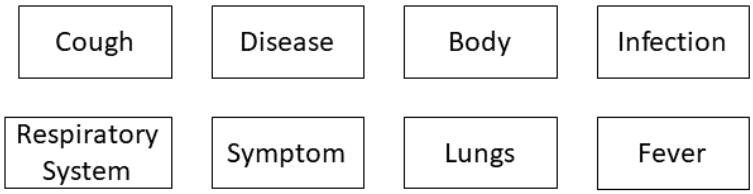


Fig. 3.1 Terms of a vocabulary of a simple ontology for a viral respiratory infection.

A taxonomy provides the classification and hierarchy of terms in an ontology by defining groups and subgroups to separate the terms [68]. Taxonomy is represented by the *is a* or subtype relations as well as other semantic relations [10] such as *causes* or *has*. These properties–also called slots–elaborate the features and attributes of the concepts. Ontologies have a set of constraint relations between concepts or restrictions on slots–also known as facets [144]. The concepts of the simple ontology for viral respiratory infection are ordered as displayed in Figure 3.2. Taxonomic hierarchy and relations between the concepts are added using *is a* for hierarchy and semantic relations of *manifests as*, *has*, *affects*, and *causes*.

A vocabulary and a taxonomy together form a conceptual framework of an ontology to discuss, analyse, and retrieve information [68]. In addition to a conceptual framework, ontologies may have a set of individual instances of classes constituting a knowledge base [144], which the knowledge of a domain is gathered, organised, and stored [68].

Figure 3.2 depicts the conceptual framework of the simple ontology for a viral infection with an individual instance called Ali. A dashed curved line links the instance of Ali to the concept of 'Infected lungs' representing that Ali has infected lungs.

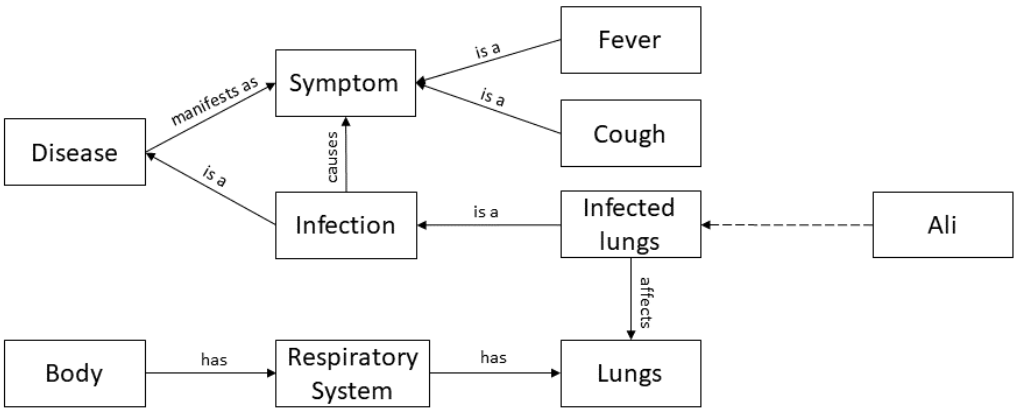


Fig. 3.2 Terms of a vocabulary and taxonomy of a simple ontology for viral respiratory infection with an individual instance called Ali.

3.2 Developing Ontologies

There are various methods to design and develop ontologies. In [10], eight general principles of ontology design are proposed that mainly include realism, perspectivism, fallibilism, and adequatism, respectively referring to describing the reality, with accuracy in description, ability to revise, and entities preserved without being reduced.

The following steps present the approach for designing ontologies originally proposed by [10] and modified based on the notion used in [144].

- **Step 1: Demarcating subject**
  This step focuses on determining the subject matter of the ontology by defining the domain and scope.

- **Step 2: Gathering information**
  Available information on the domain ontology should be gathered by identifying the general terms used in existing ontologies and textbooks and specifying the concepts of the ontology.

- **Step 3: Ordering terms**
  Gathered terms are ordered in a hierarchy and the restrictions are specified in order to form the taxonomy of the ontology. In this step, the slots are defined and their values

are specified, and if available, the values of the individual instances are given to the slots.

- **Step 4: Regimenting result**
  Resulting concepts and taxonomy should provide logical, philosophical, and scientific coherence, compatibility with neighbouring ontologies, and human understandability, especially through the formulation of human-readable definitions of the concepts. Adding or removing certain concepts or hierarchy revision may be needed to create coherence. Regimentation is an iterative process and may be repeated multiple times.

- **Step 5: Formalising regimented representation**
  Regimented representation is formalised in a machine-readable language in a way that the result can be implemented in a computable framework.

Figure 3.1 illustrates Steps 1 and 2 of developing ontologies in the example of the simple ontology for a viral respiratory infection. It shows the collection of terms related to the infection. Figure 3.2 visualises Steps 3 and 4 of developing ontologies, representing a hierarchy of terms, restrictions between them, and also an instance called Ali, which are all connected in a meaningful manner.

3.3 Semantic Web and Ontology Languages

Semantic web is an evolution of the World Wide Web adding further semantic richness to the web technology such as Extensible Markup Language (XML) and Hypertext Markup Language (HTML) [68]. Semantic Web is also known as Web 3.0 to differentiate its technologies from the technologies of social networking, blogging, and similar, which are known as Social Web or Web 2.0.

The idea of semantic web is taken from ontologies that the semantics of things within the web pages are given by their relations, so that they become machine-readable and machine-interpretable [68]. Just like ontologies, the contents of the Web would be able to not only represent information, but also contribute to automation, integration, and reuse purposes [71, 68].

Web technologies like XML and HTML are made by the World Wide Web consortium (W3C). XML enable the specification and markup of machine-readable documents, and unlike HTML, XML allows to annotate documents of arbitrary structure. The current technologies are RDF and OWL [218, 68], which are both based on XML. The following subsections further explain RDF and OWL languages.

3.3.1 Resource Description Framework

Resource Description Framework or RDF is a web technology to describe facts about Web resources. The difference between XML and RDF is that the former does not provide any interpretation of data and does not contribute to the semantic aspect [68]. However, RDF and its extension, RDF Schema, can provide a standard model for the interpretation of data.

The RDF describes a graph of nodes and arcs by identifying the universal resource identifiers (URIs). RDF statements are triples consisting of subject, predicate, and object. RDF models can provide only a domain-neutral mechanism which is suitable for the description of individual resources; however, RDF Schema can encode ontologies [68]. RDF Schema adds the notion of classes to RDF, which allows them to form hierarchies of classes and properties and allows defining simple constraints on properties.

The root of hierarchies in RDF Schema is the concept of *Resource*. RDF Schema has a frame-based modelling, e.g. *Class*, *subClassOf*.

3.3.2 Web Ontology Language

Web Ontology Language or OWL is lastly suggested by the W3C and it is based on description logics which is a subset of first-order logic [46]. It is built on top of RDF and RDF Schema. OWL describes the terms of an ontology and the relations between the terms using not only things like *Property*, *subPropertyOf*, but also uses *ObjectProperty*, *DatatypeProperty* [68].

The relationships between classes of an ontology are represented by OWL properties, which are mainly object properties, data properties, and annotation properties [81]. Object properties define the relationships between two classes, but data properties focus on the relationships between classes and instance. Annotations add more information about the class, instance, and object or data type properties. One of the main features of OWL is its richness in describing relations between classes, properties, and instances [68].

Simple ontology of viral respiratory infection is written in OWL as shown in Figure 3.3. This ontology contains three classes of *ObservableEntity*, *Disease*, and *Body*. The class *ObservableEntity* has one subclass called *Symptom*, which has two subclasses of *Fever* and *Cough*. The class *Disease* has a subclass of *Infection* that itself has a subclass of *InfectedLungs*.

The class *Body* has a subclass of *RespiratorySystem* which has a subclass called *Lungs*. Two object properties of *affects* and *causes* are defined. The former links *InfectedLungs* to *Lungs* and the latter relates *InfectedLungs* to *Fever* and *Cough*. An individual instance called *Ali* is defined that is linked to *InfectedLungs* meaning his lungs are infected.

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.semanticweb.org/...#"
xml:base="http://www.semanticweb.org/..."
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:xml="http://www.w3.org/XML/1998/namespace"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <owl:Ontology rdf:about="ViralInfection"/>
    <owl:Class rdf:about="ViralInfection#ObservableEntity"/>
    <owl:Class rdf:about="ViralInfection#Disease"/>
    <owl:Class rdf:about="ViralInfection#Body"/>
    <owl:Class rdf:about="ViralInfection#Symptom">
        <rdfs:subClassOf rdf:resource="ViralInfection#ObservableEntity"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#Fever">
        <rdfs:subClassOf rdf:resource="ViralInfection#Symptom"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#Cough">
        <rdfs:subClassOf rdf:resource="ViralInfection#Symptom"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#Infection">
        <rdfs:subClassOf rdf:resource="ViralInfection#Disease"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#InfectedLungs">
        <rdfs:subClassOf rdf:resource="ViralInfection#Infection"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#RespiratorySystem">
        <rdfs:subClassOf rdf:resource="ViralInfection#Body"/>
    </owl:Class>
    <owl:Class rdf:about="ViralInfection#Lungs">
        <rdfs:subClassOf
        rdf:resource="ViralInfection#RespiratorySystem"/>
    </owl:Class>
    <owl:ObjectProperty rdf:about="ViralInfection#affects">
        <rdfs:domain rdf:resource="ViralInfection#InfectedLungs"/>
        <rdfs:range rdf:resource="ViralInfection#Lungs"/>
    </owl:ObjectProperty>
    <owl:ObjectProperty rdf:about="ViralInfection#causes">
        <rdfs:domain rdf:resource="ViralInfection#InfectedLungs"/>
        <rdfs:range rdf:resource="ViralInfection#Cough"/>
        <rdfs:range rdf:resource="ViralInfection#Fever"/>
    </owl:ObjectProperty>
    <owl:NamedIndividual rdf:about="ViralInfection#Ali">
        <rdf:type rdf:resource="ViralInfection#InfectedLungs"/>
    </owl:NamedIndividual>
</rdf:RDF>
```

Fig. 3.3 Simple ontology of viral respiratory infections written in OWL.

One of the well-known tools to build ontologies is the Protege which provides a graphical user interface to define the concepts and properties of an ontology. It also enables users to visualise their ontology using plug-ins such as OntoGraf.

Figure 3.4 depicts the simple ontology of viral respiratory infection that is visualised by OntoGraph in Protege. In this figure, rectangular shapes represent classes, subclasses, and an instance. Blue arcs are hierarchical links, dashed arcs are object properties, and purple arc links instance to the corresponding class.

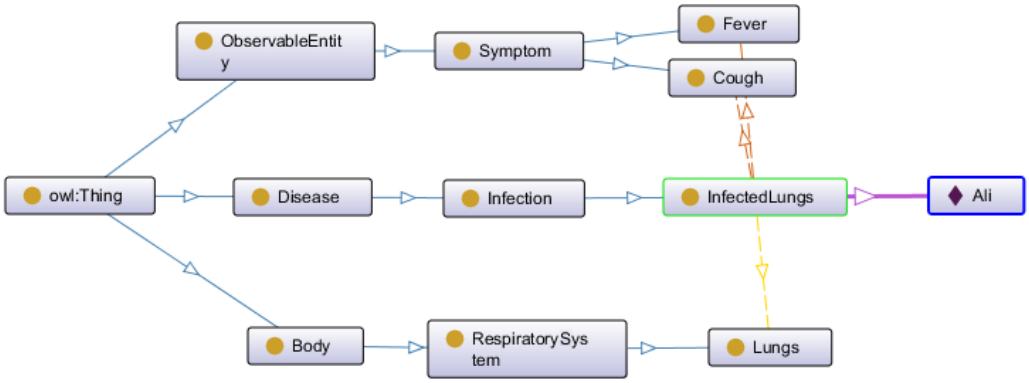


Fig. 3.4 Simple ontology of viral respiratory infections visualised by OntoGraph in Protege.

3.4 Medical Ontologies

There are various medical ontologies developed in different countries or built by international groups. In this section, we describe SNOMED CT and UMLS.

3.4.1 Systematized Nomenclature of Medicine-Clinical Terms

Systematized Nomenclature of Medicine - Clinical Terms or SNOMED CT is a standardised and multilingual collection of clinical terminology. SNOMED CT is owned, managed, licensed, maintained, and distributed by the SNOMED International, which is a non-profit organisation headquartered in London, UK. The SNOMED International is the trading name of the International Health Terminology Standards Development Organisation (IHTSDO) that was established in 2007 by nine member countries of Australia, Canada, Denmark, Lithuania, Sweden, the Netherlands, New Zealand, UK, and US. SNOMED CT has four components: 1) concepts, 2) descriptions and terms, 3) relationships, and 4) common features of components [82]. These components are described below:

Concepts

Concepts are clinical ideas which are known by a unique SNOMED CT identifier. A concept is associated with (1) a unique fully specified name (FSN) specifying the meaning represented by the concept, and (2) a set of descriptions, each of which represents the same concept, other human-readable terms and supports alternative representations such as synonyms and translations into different languages.

Description and Terms

Terms are character strings that consist of words, phrases, and other human-readable representations that convey the meanings of concepts. A description is a term connected to a particular concept. Descriptions are two types: either a FSN or a synonym. Each description may be a preferred term which is the synonym of a term more commonly used in another dialect including British and American for English language. Preferred terms are indicated by references to the description from the Language Reference Set for that language or dialect. Language Reference Set or refset is a standard format for maintaining and distributing a set of references to SNOMED CT components.

Each concept has at least one FSN which is an unambiguous description of a concept. Concepts with more than one FSN are having one synonym which is the preferred term in that language or dialect.

Relationships

Relationships are the logical associations between two concepts. A relationship contains a relationship identifier and a corresponding relationship type. Relationship type is a synonym for attribute name. Two examples are as follows:

Table 3.1 Examples of relationships in SNOMED CT

sourceId	relationshipId	destinationId
3723001|arthritis|	116680003 |is a|	399269003 |Arthropathy|
80166006 | Streptococcus pyogenes (organism)|	246075003 |causative agent|	43878008 |Streptococcal sore throat (disorder)|

Above-mentioned example of arthritis and its relationship can be visualised as shown in Figure 3.5.

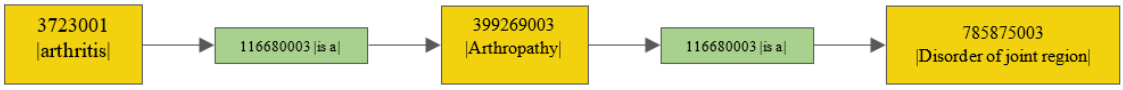


Fig. 3.5 A chain of concepts and their relationship in SNOMED CT

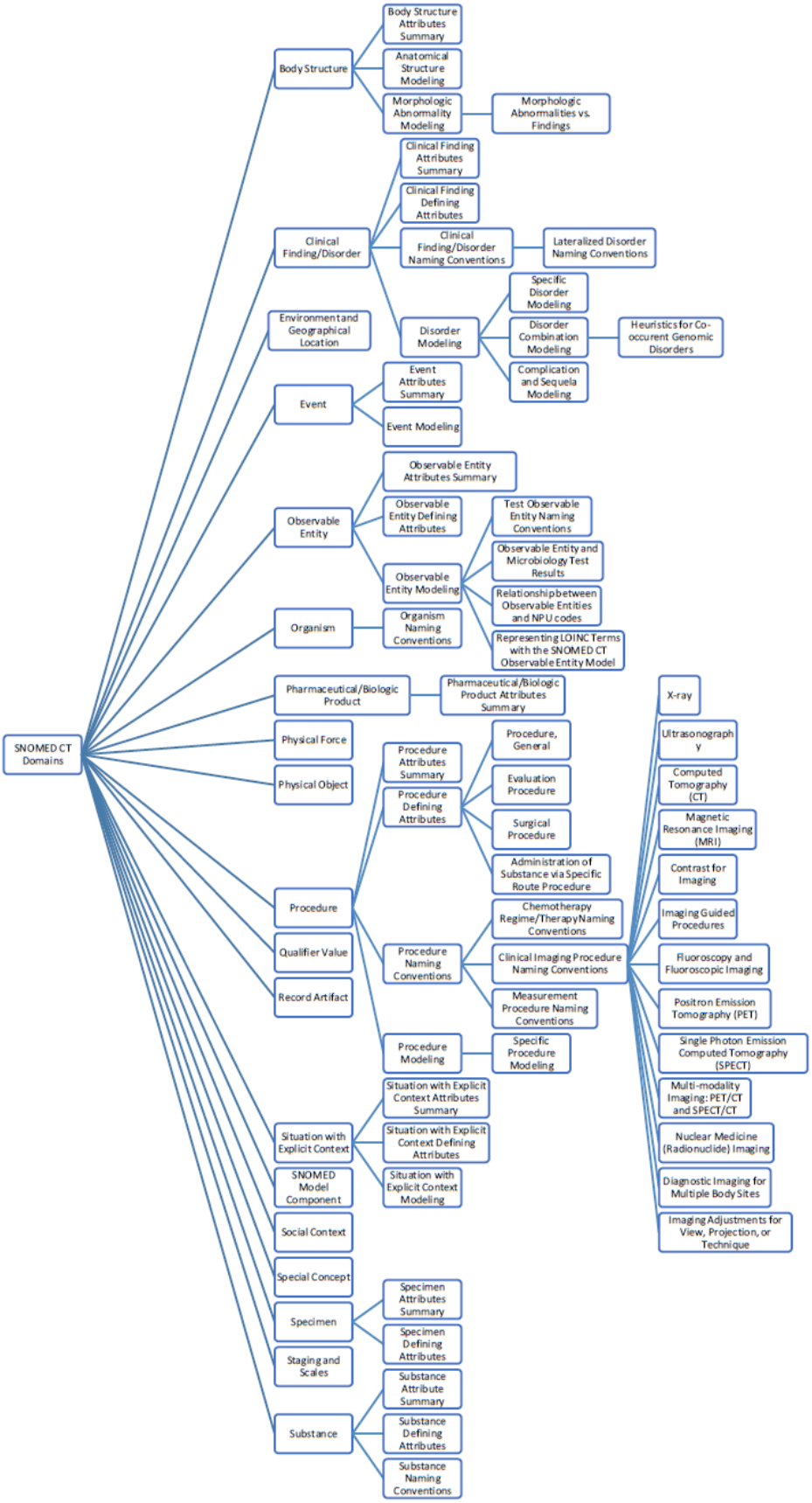


Fig. 3.6 SNOMED CT domain hierarchies and their subtypes.

Arthropathy has several synonyms, namely, Joint disease, Disorder of joint, Arthritis, and Joint disorder, that are all sharing the same Id () and connected by a 116680003 |is a| relationship.

There are totally 122 different relationships in SNOMED CT, e.g., *is a*, *causative agent*, *associated morphology*, and *clinical finding*. Subtypes and concept model attribute concepts are listed in the SNOMED CT concept enumeration for relationship typeId.

Common features of components

The common features of SNOMED CT components include identification and history management. SNOMED CT concepts can be identified and referenced using numeric identifiers called SNOMED CT Identifier (SCTID). History management is considered because the concepts of SNOMED CT evolve in each release. These changes vary from new concepts, new descriptions, new relationships, and new reference set members, as well as updates and retirement of any of these components.

Some selected SNOMED CT attributes have a hierarchical relationship to each other called attribute hierarchies. SNOMED CT concepts are organised into 19 distinct hierarchies, each of which cover different types of concepts [134]. Figure 3.6 displays the domain

Table 3.2 SNOMED CT hierarchies.

SNOMED CT hierarchy	Semantic tag
Body structure	body structure
Clinical finding	finding
Environment or geographical location	environment / location
Event	event
Observable entity	observable entity
Organism	organism
Pharmaceutical / biologic product	product
Physical force	physical force
Physical object	physical object
Procedure	procedure
Qualifier value	qualifier value
Record artifact	record artifact
Situation with explicit context	situation
SNOMED CT Model Component	metadata
Social context	social concept
Special concept	special concept
Specimen	specimen
Stages and scales	staging scale
Substance	substance

hierarchies and their subtypes of SNOMED CT. As shown in this figure, one general attribute is the parent of one or multiple specific subtypes of that attribute [202]. Concepts defined using the more general attribute can inherit concepts modeled with more specialised subtypes of that attribute.

SNOMED CT hierarchies contain a semantic tag in parentheses, which represents the type of semantic of the concept. Semantic tags disambiguate different concepts that may be referred to by the same commonly used terms. Table 3.2 shows these hierarchies and semantic tags. The SNOMED CT hierarchies stated in this table match with the set of SNOMED CT hierarchies expressed in Figure 3.6 (second layer of hierarchies from left hand side).

3.4.2 Unified Medical Language System

UMLS is a set of files and software that brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems [18]. It has three knowledge sources namely Metathesaurus, Semantic Network, and SPECIALIST Lexicon and Lexical Tools.

UMLS provides access to users through online browser of UMLS Terminology Service (UTS) which has 4 groups of applications including UTS Methatesaurus Browser, UTS Semantic Browser, UTS SNOMED CT Browser, and other applications of the US National Library of Medicine (NLM) [210]. Other services of UMLS are: the UMLS Knowledge Source, a UMLS REST Application Programming Interface (API), and a UMLS SOAP API. Above-mentioned browsers and APIs are explained as follows:

UTS Methatesaurus Browser

UTS Metathesaurus Browser is an online service with three search options: Term, CUI or concept unique identifier, and Code. It allows users to choose which UMLS release to browse in. Search type can be selected, i.e. user can search the exact term or approximate match, and other options. Users can also choose the source of browsing. The outcome is provided in 3 tabs: basic view, report view, and raw view.

Metathesaurus users may select from two relational formats: the Rich Release Format (RRF), first introduced in 2004, and the Original Release Format (ORF). Metathesaurus is linked to the other UMLS Knowledge Sources – the Semantic Network and the SPECIALIST Lexicon. All concepts in the Metathesaurus are assigned to at least one Semantic Type from the Semantic Network.

The Metathesaurus contains concepts, concept names, and other attributes from more than 100 terminologies, classifications, and thesauri, some in multiple editions. One of its

primary purposes of Metathesaurus is to connect different names for the same concept from many different vocabularies. A concept is a meaning and a meaning can have many different names. A key goal of Metathesaurus construction is to understand the intended meaning of each name in each source vocabulary and to link all names from all source vocabularies that mean the same thing.

Each concept or meaning in the Metathesaurus has a unique and permanent CUI. An atom is the smallest unit of naming in a source of UMLS which includes a specific string with specific code values and identifiers from a specific source. AUI or Atom Unique Identifier is an identifier for each atom. An atom has Root Source Abbreviation (RSAB) which is an NLM-assigned source abbreviation stripped of its version information.

The Metathesaurus includes many relationships between different concepts. Most of these relationships come from individual source vocabularies and some are added by NLM during Metathesaurus construction. Some have been contributed by Metathesaurus users to support certain types of applications.

UTS Semantic Browser

UTS Semantic Network Browser has an interface for the UMLS Semantic Network which is a set of broad subject categories, or semantic types, that provide a consistent categorization of all concepts represented in the UMLS, and a set of useful and important relationships, or semantic relations, that exist between semantic types.

UTS SNOMED CT Browser

UTS SNOMED CT Browser gives access to SNOMED CT US Edition. SNOMED CT US is one of the sources of UTS Metathesaurus Browser.

UMLS Knowledge Source

UMLS Knowledge Source is an offline browser. The most recent release (2020AB) includes the UMLS Metathesaurus, Semantic Network, Specialist Lexicon and Lexical Tools, database load scripts, and MetamorphoSys for installation and customisation of UMLS subsets and browsing data. It provides a flexible customising platform such as choosing semantics, languages, and sources.

UMLS APIs

UMLS REST API is freely available in GitHub. It can also be reached using Postman service which has some samples online. Another UMLS API is called Soap API which uses Maven and its codes are mainly available in java. UMLS SOAP API Home Page is here.

3.5 Bayesian Network Models and Ontologies

BNs and ontologies are two complementary tools to create either an uncertain ontology or BNs based on ontologies. BN models and ontologies are inherently different types of graphs, although they can be transformed into each other. BNs are probabilistic graphs, representing the associations between variables of BNs; however, ontologies are deterministic, demonstrating terms and semantic relations between them.

Care pathways are a means of expressing care processes and representing medical knowledge, which contain very limited but purposeful knowledge compared to the medical ontologies like SNOMED CT and UMLS that we described in Section 3.4. These ontologies can be browsed and a fraction of their knowledge can be retrieved. We call this fraction of ontologies as knowledge graphs to distinguish a retrieved fraction of ontologies from the entire ontology. The following subsections present a contrast between care pathways, knowledge graphs as an ontology from a specific domain, and BNs, and the application of ontologies in building BN models.

3.5.1 Contrasting Care Pathways, Knowledge Graphs, and Bayesian Networks

Care pathways are flow diagrams of care activities arranged on pathways; the notation is consistent with elements of Unified Modeling Language (UML) such as the Activity Diagram [119]. In a care pathway, care activities are arranged sequentially. A potentially divergent path is represented by a decision point where clinicians (or patients) choose how care progresses [119]. Care pathways are useful to elicit expert knowledge as they give a clear visual representation of the sequence of care activities and decisions, typically extracted from clinical practice guidelines (CPGs) or treatment protocols [117] for the diagnosis or management of a particular condition [119].

Knowledge graphs are the same as ontologies that acquire and integrate the information of ontologies [52, 13]. They are a special type of information network that represents the knowledge of a domain in RDF-style [197], and allows them to mine, organise, and effectively manage knowledge from large-scale data [32]. Like ontologies, knowledge graphs are hierarchical networks consisting of concepts or classes, their properties, attributes, and instances. The connections between concepts of a knowledge graph are semantic relations, e.g., *is a* is the most common relation that links subclasses to their superclass concept. As explained in Subsection 3.2, developing an ontology requires defining the concepts, arranging them in a taxonomic hierarchy, defining slots, i.e., mainly relationships between concepts,

defining allowed values for slots, and finally, if instances are given, allocating values for them.

As described in Section 2.1, a BN is a directed acyclic graph consisting of a set of variables, dependencies between them, and parameters [60]. The qualitative part consists of a set of nodes that represent random or uncertain variables. The connections between these nodes are directed arcs that represent causal or influential relationships between the connected variables. The quantitative part consists of a set of conditional probabilities that model the dependencies between connected variables.

The structure of care pathways, knowledge graphs, and BNs have similarities and differences. All of them show the primary events for the given problem. All are directed graphs, however, BNs are acyclic while care pathways and knowledge graphs may contain cycles. Care pathways represent the sequential flow of care and knowledge graphs express the hierarchy of concepts and their semantic relations, while BNs capture the causal or influential relationships between variables.

Information of the care pathway nodes impacts the progression of care, which influences the semantic structure of the knowledge graph, while instantiating information in BN nodes updates the probability of each conditionally dependent variable, regardless of its position in the care pathway or the knowledge graph. Table 3.3 presents the contrasts of the properties of care pathways, knowledge graphs, and BNs. This table shows the similarities and differences of these models and the potential of transforming care pathways and knowledge graphs into BN models.

Table 3.3 Contrasts of Properties of care pathways, knowledge graphs, and BNs.

Property	Care pathways	Knowledge graphs	BNs
Graph type	Flow diagram	Knowledge graph	Directed model
Process flow	Cyclic	Cyclic	Acyclic
Nodes	Care activity/decision	Concepts/classes	Random/uncertain variables
Arcs	Sequence	Attributes	Causality/association/influence
Information flow	Care progression	Semantic relations	Conditional dependency

3.5.2 Building Bayesian Networks from Ontologies

Ontologies are a means of representing structured knowledge and help to share a common understanding of a domain of knowledge, knowledge reuse, and reasoning [45]. Ontologies are deterministic and lack support to process uncertainty [103], whereas BNs are equipped with conditional probabilities underneath their variables. In other words, ontologies and BNs can operate as two complementary tools to create either an uncertain ontology or BNs based

on ontologies. In this thesis, we focus on the latter: creating BNs from ontologies. [78] pioneered a BN modelling method for medical applications using ontology of diseases to identify the causes and effects, and finally quantify the variables using elicited probabilities. Similarly, [45] suggested a method to automatically build BNs from ontologies by identifying relevant variables, their values, defining dependencies between variables, as well as assigning conditional probabilities. To train a BN model, [207] suggested applying ontologies and detection of interdependencies using low-level methods for ontology mapping.

There are many studies on BN modelling based on OWL and using rule-based algorithms to create BN models. [46] extended BNs to OWL through additional probabilistic markups and a set of translation rules to switch probabilistically annotated OWL into BNs. Based on the BN modeling using OWL, [153] developed a BN modeling method using the source and target ontologies which obtain two translated BNs and probabilities for CPTs learnt by classification techniques. Using the semantic of OWL, [40] built a language to translate probabilistic ontologies and developed a multi-entity BN by integrating Bayesian probability theory and the first-order logic to express a varying number of entities, different entity types and repeated patterns.

Different versions of OWL algorithms have been proposed, such as OntoBayes, PR-OWL, BayesOWL, and OMEN. OntoBayes is an ontology-driven BN modeling using the probabilities and dependencies of OWL [225]. PR-OWL or probability OWL is proposed to deal with the uncertainty of semantic web as an extension OWL to provide probabilistic inference and Bayesian learning [41]. BayesOWL provides a framework to convert OWL ontologies into BNs using semantic web ontologies [47]. It exploits new OWL classes variables, prior probabilities, and conditional probabilities, and a set of rules to map OWL ontology taxonomy into BN model, which led to present an efficient NPT quantification algorithm. As a probabilistic ontology mapping tool, [103] suggested the ontology mapping enhancer (OMEN) that processes the source ontologies, a set of initial mappings, and initial probabilities of matching concepts.

OWL is also used to organise medical evidence from different sources. Using the semantics and vocabulary of UMLS, [53] gathered a knowledgebase containing objects of the problem domain and the semantic links, which maps the knowledgebase on BN models. Ontologies are not only for building BN models from, but they can also be used to organise clinical evidence. For example, [228] proposes an ontology based on OWL to store evidence, organise evidence, mainly by classifying them into supporting evidence, conflicting evidence, and excluded. It finally translates the evidence into BN components: variables, dependencies, and parameters.

This literature review demonstrates the limited number of studies on creating BN models from ontologies for medical applications. It also shows the need for novel methods to transform ontologies and build BN models applied to medicine, particularly diagnosis. This gap can be filled by using care pathways to acquire the initial knowledge, and then creating knowledge graphs from the retrieved medical knowledge stored in pre-existing medical ontologies (e.g., SNOMED CT and UMLS) from the concepts of the care pathway. The retrieved knowledge can help us to build the structure of BN models for medical applications.

3.6 Summary

In this chapter, we introduced ontologies and their components, and we illustrated them with an example of a simple ontology. We explained a step-by-step method to develop ontologies. We described the Semantic Web and ontology languages of RDF and OWL, and illustrate how the simple ontology can be written in the OWL language.

Next, we introduced two medical ontologies, namely, SNOMED CT and UMLS, and we explained their features. These medical ontologies are the sources of medical knowledge that we need to build BN models and contribute to the Objective 2 of this thesis. We described BNs and ontologies comparatively by contrasting care pathways, knowledge graphs that are built from ontologies, and BN models. We then review the state-of-the-art of building BNs from ontologies. The contrast between care pathways, knowledge graphs, and BNs shows the possible transformation from knowledge graphs to BNs, and the novel literature review on BNs and ontologies reveals the gap in the literature to build meaningful structures for candidate BN models from medical ontologies. Our contribution to build BN models from ontologies is presented later in Chapter 6.

Chapter 4

Case Study: Diagnosis and Treatment of Rheumatoid Arthritis

Rheumatoid arthritis (RA) is an inflammatory autoimmune disease, causing swollen and painful joints, and characterised by fluctuating inflammatory activity [9]. RA is a chronic disease requiring lifelong treatment with disease-modifying antirheumatic drugs (DMARDs), since there is no cure [90]. The fluctuating activity of RA disease requires patients to be monitored continuously and provided more care at home with portable or wearable devices and fewer visits to medical centres, as stated in [152]. RA is a case study of the PAMBAYESIAN project (introduced in Chapter 1) and intends to develop Bayesian network (BN) models for diagnosis and self-management of RA. The main focus of this case study is on building BNs for self-management and treatment applications corresponding to the Objective 3 of this thesis. Another focus of this case study is on the personalised care of RA which accords with Objective 4 of this thesis. Diagnosis is a side focus of this case study which corresponds to Objectives 1 and 2 of this thesis.

We introduce the PAMBAYESIAN project in Section 4.1, followed by an explanation of RA condition in Section 4.2. Four models of care pathways for diagnosis, initial management, ongoing management, and personalised care of RA are presented in Section 4.3. In Section 4.4, we describe the available datasets for RA case study collected in the Pathobiology of Early Arthritis Cohort (PEAC) study and BioT project. We introduce a set of selected decision support points in Section 4.5 that are the components of the care pathways selected after analysing the available data. Finally, we summarise the chapter in Section 4.6.

4.1 PAMBAYESIAN Project

The PAtient Managed decision-support using Bayesian networks or PAMBAYESIAN project is a joint research at Queen Mary University of London (QMUL). The researchers from the School of Electronic Engineering and Computer Science and the School of Medicine and Dentistry (SMD) of QMUL try to create multiple decision support tools, mainly for the self-management of chronic diseases, namely RA, diabetes in pregnancy, and atrial fibrillation [172, 152]. Each of these chronic diseases is a case study of the PAMBAYESIAN project. As already established, this thesis covers the main part of our research on the RA case study. We target three areas of support for RA comprising of diagnosis, self-management, and personalised care. The self-management application intends to provide a remote monitoring capability that enables clinicians to taper or reduce the dose of medications, especially the biological disease-modifying antirheumatic drugs or bDMARDs. These drugs are not only more toxic than the other conventional DMARDs (csDMARDs), but also they are expensive.

For each area of decision support for RA, we create one or multiple BN models to support clinicians and persons. We employ the available data and acquire medical knowledge from experts to build BN models. The available data are described in Section 4.4. The knowledge is acquired from multiple experts: three rheumatologists from the Centre for Experimental Medicine & Rheumatology of SMD, namely, Dr Frances Humby, Dr Amy MacBrayne, and Dr Gloria Ribera Lliso, and two computer scientists Professor Paul Curzon and Dr Hamit Soyel. Dr Frances Humby is an honorary consultant rheumatologist, Dr Amy MacBrayne is a clinical research fellow, and Dr Gloria Ribera Lliso is a postdoctoral researcher. Their medical knowledge helped us to build the diagnostic and self-management models, as presented in Chapters 5, 6, 7, and 8. Dr Frances Humby is the senior rheumatologist and Dr Amy MacBrayne provided the majority of expertise on the diagnosis and management of RA; therefore they are called the "senior expert" and the "main expert" in the above-mentioned chapters. For the personalised care application (presented in Chapter 9), Professor Paul Curzon and Dr Hamit Soyel contributed, in addition to Dr Frances Humby and Dr Amy MacBrayne. Since Dr Hamit Soyel provides the majority of expertise on the human interaction with computers and personalisation, we call him the "main expert" in that chapter.

4.2 Rheumatoid Arthritis

Many people suffer from chronic diseases and joint pain caused by inflammatory arthritis (IA). RA is the most frequent IA, affecting 1% of the UK population [184]. RA is an autoimmune disease that causes inflammation in the body joints [12] mainly in hands, feet,

and wrists. The major symptoms of RA are joint pain and swelling, which can emerge as a simple pain or morning stiffness. Other common symptoms of RA can be listed as follows: anaemia, flu-like symptoms (i.e., feeling generally ill or feeling hot and sweating), mental symptoms (e.g., fatigue, depression, and irritability) [12]. Less common symptoms of RA are weight loss, inflammation in the eyes, rheumatoid nodules, and inflammation of other parts of the body. Figure 4.1 shows a hand of a person with RA whose joints are affected [146].

Fig. 4.1 A hand with affected joints by RA.

4.2.1 Diagnosis

To diagnose RA and prevent deformation of joints, rheumatologists have proposed diagnostic criteria, such as American College of Rheumatology (ACR) 1987 [9] and, in 2010, its revision jointly with the European League Against Rheumatism (EULAR) [6]. The 2010 criteria are point-based and consider four domains: joint involvement, serology results, acute-phase reactants and duration of symptoms [93]. The National Institute for Health and Care Excellence (NICE) recommends a set of criteria for referral from primary care to specialist including small joints of hands and feet and more than one joint being affected [138]. For further investigation, the guideline recommends to carry out blood test for rheumatoid factor (RhF) and antibody cyclic citrullinated peptide (CCP), and X-ray scanning of hand and feet joints [138]. Although the NICE guideline and 2010 criteria for diagnosis of RA have

improved the early identification of RA compared with the 1987 criteria, there is still the need to improve the accuracy of RA diagnosis [77].

4.2.2 Management and Treatment

After diagnosis of RA, the disease and its progression need to be continuously monitored to achieve a low and stable disease activity. Disease activity is commonly measured with Disease Activity Score 28 (DAS28) that is an extension of another measurement called DAS [114]. DAS28 is a composition measurement of swollen joints count (out of 28), tender joints count (out of 28), blood markers of ESR or CRP, and usually global health (GH). The equation of DAS28 with ESR and GH is as follows:

$$DAS28ESR = 0.56 \times \sqrt{TJC28} + 0.28 \times \sqrt{SJC28} + 0.7 \times ln(ESR) + 0.014 \times GH \quad (4.1)$$

DAS28 with CRP and GH is similar to DAS28 ESR with a slight difference as shown below:

$$\begin{aligned} DAS28CRP = & 0.56 \times \sqrt{TJC28} + 0.28 \times \sqrt{SJC28} \\ & + 0.36 \times ln(CRP + 1) + 0.014 \times GH + 0.96 \end{aligned} \quad (4.2)$$

In some certain periods, symptoms get worse which is called flare-up or flare. There is no consensus on a standard definition of flares as revealed in multiple attempts of standardisation and validation by the Outcome Measures in Rheumatology Clinical Trials (OMERACT) in 2009 [17], in 2011 [16], in 2013 [213], in 2016 [25], and most recently in 2020 [124]. Although many researchers have tried to provide an inclusive definition for flare, there is no consensus among either rheumatologists or patients. An international research project called the RA Flare Definition Working Group, organised by the OMERACT, targets to standardise the flare definition and understand the characteristics and the impacts of flare. OMERACT defined flare as "any worsening of disease activity that would, if persistent, lead to initiation or change of therapy" [17]. Recently, OMERACT has identified the features of flare – including fatigue, stiffness, symptom persistence, systemic features, and participation – in addition to the existing core set of RA features [16]. Unlike OMERACT, some other researchers have tried to define flare based on objective signs of disease or disease activity, particularly van der Maas, et al. [213] defined flare as 1.2 unit increase in DAS28 or more than 0.6 increase in DAS28 if it is already greater than 3.2.

Based on their DAS28, clinicians classify patients into four groups of remission, low, moderate, and high respectively for DAS28 values less than or equal to 2.6, greater than 2.6

and less than or equal to 3.2, greater than 3.2 and less than or equal to 5.1, and greater than 5.1. These groups help clinicians to track disease activity and make a decision on medications in a way that prevents flares from happening. The lack of a standardised definition of flares makes it difficult to predict flares; however, it is known that regular treatment helps to decrease the number of flares and their length, which leads to protect the joints against damage in long-term [127].

Medications for RA mainly encompass csDMARDs, targeted synthetic and biological disease-modifying antirheumatic drugs (ts/bDMARDs), and steroids. The NICE provides a pathway for treatment of RA [140] that can assist clinicians to prescribe proper medications to patients. The NICE guideline on drug treatment for RA suggests a pathway for treatment strategies [140]. Adults with RA, first receive treatment based on the treat-to-target strategy, as shown in Figure 4.2. This strategy aims to treat active RA to achieve a target of remission (e.g., DAS28 less than 2.6) or low disease activity (e.g., DAS28 less than 3.2), if remission cannot be achieved.

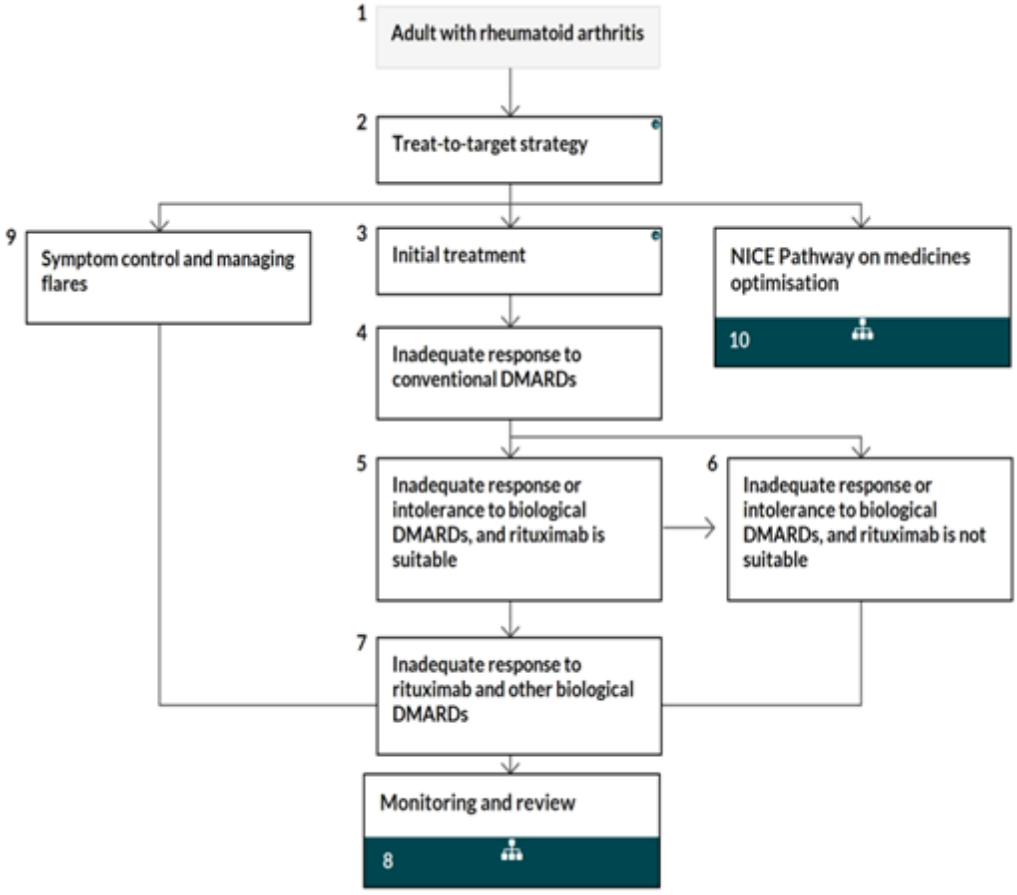


Fig. 4.2 NICE pathway for treatment of RA

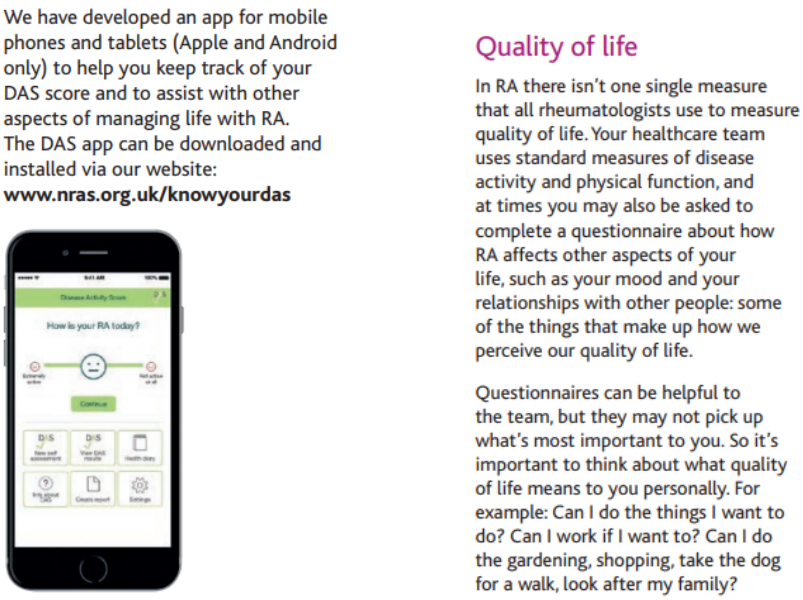
Apart from this pathway, the NICE provides multiple guidelines describing csDMARDs, bDMARDs, and steroids [135, 137]. These guidelines help clinicians to make a decision on medication prescription mainly by expressing their contraindications in case of comorbidities. For example, [135] states the contraindication of adalimumab bDMARD in people with active tuberculosis. Although guidelines help clinicians to avoid prescribing wrong medications considering their contraindications, they lack precision information on which medications are effective in which patients. This leads clinicians to do trial and error to find the right medication or combination of medications for each patient that can effectively reduce their disease activity stably. Trial and error of medications exposes patients to toxic drugs which can cause various adverse reactions, also wastes resources to search and find proper medications for each patient.

4.2.3 Quality of Life

In addition to treatment, rheumatologists provide information about RA to patients and highlight the need for lifelong care. The information about RA is prepared by the National Rheumatoid Arthritis Society (NRAS) and usually printed in the form of leaflets or booklets. Patients also become familiar with a large amount of information in the website of NRAS.

Fig. 4.3 Cover page of NRAS booklet on Living Better with RA.

For example, Figure 4.3 displays the cover page of the NRAS booklet on Living Better with RA which contains a 72-page document on advice and information to better live with RA [145]. This booklet provides information on how to keep a track record of disease activity using an available mobile app shown in Figure 4.4a. It describes QoL and how it is measured as shown in Figure 4.4b.



(a) Disease activity. (b) QoL.

Fig. 4.4 Information on disease activity and QoL in the NRAS booklet on Living Better with RA.

This document and similar other documents and online contents contain advice for patients to improve their lives and maintain their quality of life (QoL).However, the information is not personalised and patients get loads of text contents which may not be as useful as expected by the clinicians and NRAS.

4.3 Models of Care Pathways for Rheumatoid Arthritis

A clinical care pathway or caremap represents the clinical steps to manage a condition. By integrating management plans, care pathways provide management goals, sequences, and timing of steps [154], which are usually extracted from clinical guidelines and experts' consensus. Care pathways are graphical representation of the sequence of patient care activities required in managing a specific medical condition [39, 118, 177, 178]. Care

pathways provide a visual prompt that can enable effective knowledge transfer between the medical and decision experts [118, 117].

We form a group of domain experts to elicit their knowledge and create the models of care pathways for diagnosis and management of RA. To elicit expert knowledge on RA diagnosis and management, we arranged four meetings with our senior rheumatology expert, Dr Frances Humby. In the first meeting, she explained the general points about RA diagnosis, follow-up, and treatment. Second meeting was devoted to a detailed description of care process. We drew the primary generic care pathways of RA diagnosis, initial management, and ongoing management based on the elicited knowledge. We tried to remove too many details and considered the guidelines and pathways suggested by the NICE such as [138, 140]. We met with our senior rheumatology expert for two more times to improve the illustration of these pathways by asking open-ended questions. These provided us with a more detailed information of RA care process and enabled us to clarify some obscure clinical activities or decisions.

Along with meetings with expert for knowledge elicitation, we were invited twice to attend patient and public involvement (PPI) meetings called Patient Advisory Group (PAG), where a group of patients and one or more clinicians and researchers gather to share their experiences and recent findings.

We noticed the need for further investigation of patients' reported factors such as fatigue and weather on their QoL, which are less regarded in the elicited knowledge from our rheumatology experts. Therefore, we used the expertise of our two computer scientist experts who were involved in a qualitative study to interview with RA patients and learn about the experiences in living with RA. They also go through the established literature on RA and QoL to find the main factors of RA management from patients' point of view. We gain valuable insight into patients' efforts and challenges to manage their RA condition by joining PAGs and eliciting knowledge from our experts.

Our colleague Dr Scott McLachlan drew three models of care pathways for diagnosis, initial management, and ongoing management of RA. These pathways are based on the medical knowledge elicited from rheumatology experts and NICE guidelines. To facilitate chasing the flow of information, activity and decision boxes are orderly shown by A and D and numerated in each pathway. The numerations of the diagnosis pathway continue in the initial management pathway and those of initial management continue in the ongoing management pathway.

We suggest the following rules to sketch care pathways with simple structure to express initial medical knowledge required for experts and non-experts.

1. Pathways encompass of (1) start and exit points, (2) activity nodes, (3) decision nodes, and (4) arrows represented by circles, rectangular shapes, diamond shapes, and directed lines, respectively.

2. Looped nodes express the repetition of the care process from the first node.

3. Each looped node require a unique exit point.

4. Solid arrows display the direction of medical care stream.

We redraw three pathways for diagnosis, initial management, and ongoing management of RA by applying the above rules. For this, we acquire further medical knowledge from the experts in rheumatology and computer science, patient experiences, and established literature. Apart from the three care processes, we create a pathway for personalised care for living with RA based on the knowledge published in the literature and explained by patients.

4.3.1 Diagnosis

A model of the pathway for diagnosis of RA is displayed in Figure 4.5 [1]. It depicts activity nodes and decision points, corresponding to data records and the flow of information. The following parts describe the activities and decision points of the pathway for RA diagnosis:

Activity 01: Diagnosis process begins by receiving a referral from a general practitioner (GP) along with basic patient personal and demographic data such as ethnicity, sex, religion, and age, which are recorded.

Activity 02: Rheumatologists collect patient's history including symptomology, joint details, other medical conditions, and any medication at the moment. This provides them with risk factors, e.g., family history (FHx) and occupation, and disease manifestation factors, e.g., joint pain, swollen and affected joints, fever, morning stiffness for more than 30 minutes, systematic flu-like features, and fatigue. Information on risk and manifestation factors are recorded.

Activity 03: Clinicians perform examinations to find RA signs and symptoms including three or more tender and swollen joint areas, distribution of joints affected, symmetrical joint involvement in hands and/or feet, and positive squeeze at metacarpophalangeal (MCP) or metatarsophalangeal (MTP) joints. MCP joints situate between the metacarpal bones and the proximal phalanges of the digits and MTPs are the joints between the metatarsal bones of the foot and the proximal bones of the toes. Examination results are directed to signs and symptoms data.

[1] Our colleague, Dr Scott McLachlan, created this model of pathway originally and we revised its content and edited its structure visually.

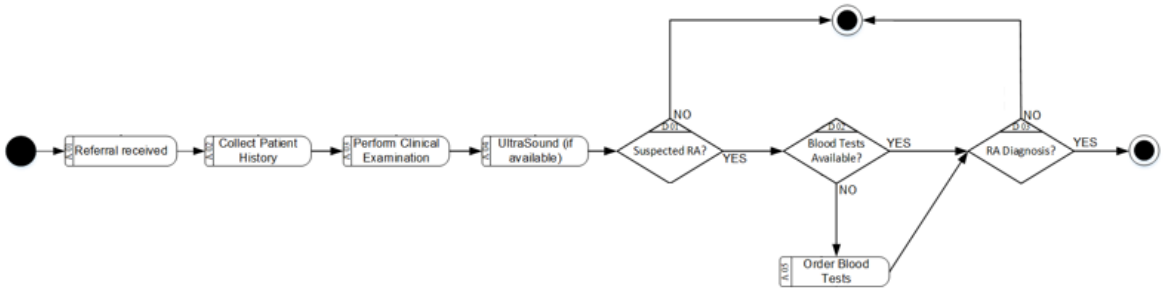


Fig. 4.5 Pathway for diagnosis of RA.

Activity 04: If ultrasound scan results are available, rheumatologists can check joint inflammation and detect affected joints. Ultrasound results are recorded.

Ultrasound is not necessary for all patients, but it is helpful in about 50 percent of patients where antibody CCP is negative but joints are inflamed. For example, if a patient comes with 2 sore joints of CRP a bit high but no RhF, then an ultrasound scan is used to see if actually RA exists or not. Ultrasound scan is not available in every hospital, and not every rheumatologists are trained to use ultrasound.

Decision 01: In the first decision point, rheumatologists suspect whether RA exists or not. If yes, they go one step forward, and if not, they exclude the person from the diagnosis process.

Decision 02: The blood tests of a patient with suspected RA are needed to investigate if the inflammation markers (ESR and CRP) and antibodies (RhF and CCP) are positive or not. If blood tests are not available, rheumatologists direct the patient to Activity 05. This blood analysis includes the results of kidney test and liver function tests (LFT) to make sure if the patient can use DMARDs which may affect those with kidney and liver issues.

Activity 05: Patients receive blood tests. Based on the type of the test, it may last a couple of hours to some days to prepare the results. The majority of blood tests including full blood count (FBC), liver test, kidney test, and CRP and ESR measurements take a couple of hours to prepare. However, the RhF and CCP tests probably last about 48 to 72 hours. Lab test results are finally recorded.

Decision 03: Rheumatologists try to diagnose RA using two criteria: (1) if the joints are tense and swollen and (2) if the blood test has been positive for more than 6 weeks. If RA is diagnosed, the patient highly likely suffers from RA and will be directed to get initial treatment. If RA is not diagnosed, the doctor excludes the patient from the RA diagnosis process. The decision outcome is registered in the diagnosis data. To diagnose RA, blood tests for ESR and CRP should be updated because they change a lot during one month period

and they should be repeated. However, RhF and CCP antibody tests are not that sensitive and they can be valid from a year before diagnosis.

Generally, rheumatologists prefer CRP rather than ESR. In inflammatory conditions, both ESR and CRP tend to rise, but ESR stays up and takes a long time to decrease. On the other hand, CRP just increases in similar conditions and decreases by taking antibiotics. These two, CRP and ESR, support each other and rheumatologists can check both. For example, there are some patients whom their ESR just tends to be more reflective than their CRP. This should also be taken into consideration that a CRP test is more expensive than an ESR one. Overall, rheumatologists prefer CRP to ESR.

4.3.2 Pathway for Rheumatoid Arthritis Initial Management

A model of care pathway for initial management of RA is shown in Figure 4.6 [2]. Following parts describe the activities and decision points of this pathway:

Activity 06: Management or treatment of RA begins, if RA is diagnosed in the Decision 03 called 'RA Diagnosis?' represented in the diagnosis pathway (Figure 4.5).

Activity 07: Rheumatologists document the initial therapy and required information. Patients get an explanation about lifestyle interventions such as weight control, smoking, Mediterranean diet, patient education and self-management programmes, occupational therapy, doing exercise, sleep promotion, appropriate foot care, and thermotherapy.

Activity 08: Rheumatologists measure the severity of disease using different measurements such as DAS28 and Health Assessment Questionnaire (HAQ). There are other disease assessment methods such as CDAI, SDAI, ultrasound, and X-ray measures. DAS28 is commonly used because it is easy to use in clinic and validated in routine clinical care. Further description of DAS28 is provided in Section 4.2.

Decision 04: Rheumatologists make a decision whether patients need steroids based on the severity of their symptoms. Patients with active RA and painful joints are eligible to receive steroids. Compared with DMARDs, steroids are effective quickly (in some hours or days), but they are very toxic. Steroids are given in the form of injection by professionals, but other forms of tablet, dissolvable, liquid, and syrup are also available.

Although rheumatologists tend to give steroids to all patients, diabetic patients have problems to use them. For example, steroids can increase blood sugar, but it is still given to diabetic patients. It can also induce diabetes. Injection of steroids is preferred compared to tablets because they are more effective. Inflammation rises over nights and their peak is usually at about 2am, which wakes up patients with severe pain and stiffness, and it takes a

[2] Our colleague, Dr Scott McLachlan, created this model of pathway originally and we revised its content and edited its structure visually.

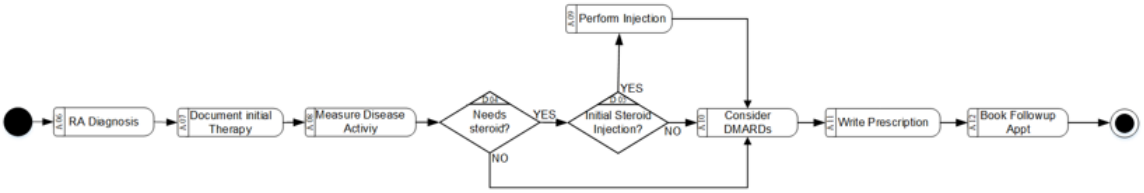


Fig. 4.6 Pathway for initial management of RA.

long time to clear. Since everything gets worse at night, rheumatologists ideally prescribe steroid tablets to be used at night. However, these tablets keep patients up and they get less sleep. Alternatively, patients can take steroids in the morning, which they would need to take a higher dose. If rheumatologists give them an intro-muscular injection, it would be slow modifying release, which is over 24 hours with lower dose compared to tablets.

Rheumatologists would like to know if patients received steroid previously or it would be their initial steroid. If initial, patients will be directed to Activity 10. If not, Decision 05 will be processed. The details of decision on the need for steroids are recorded.

Decision 05: If patients receive their initial steroids injection, they are directed to Activity 09. Otherwise, they go to Activity 10.

Activity 09: Initial steroid injection should be performed by a practitioner.

Activity 10: Rheumatologists select initial csDMARDs such as methotrexate (MTX) and sulfasalazine (SSZ) considering the severity of signs and symptoms, patient history, and comorbidities. Patients with lung disease, kidney disease should not use csDMARDs. There are other categories of DMARDs, i.e., tsDMARDs and bDMARDs, which are not considered in initial management. DMARDs can cause side-effects on liver, lungs, kidneys, and other comorbidities such as heart disease, skin disease, and overlap connective issue disease. Further description of DMARDs is provided in Section 4.2 of this chapter and in Section 8.2.

Activity 11: Rheumatologists write a prescription of DMARDs that they considered in Activity 09. Prescription details data are recorded.

Activity 12: Rheumatologist books the follow-up appointment to visit the patient which is generally 3 months later. For patients given bDAMRDs, it can be every 6 months.

4.3.3 Pathway for Rheumatoid Arthritis Ongoing Management

A model of pathway for ongoing management of RA is demonstrated in Figure 4.7 [3]. This pathway has three entries of 'Clinic Visit', 'Review Patient's Test Results', and 'Patient Feels Flare' that are respectively represented by Activity 13, Activity 26, and Activity 27. We, first, describe the entry from Activity 13 and then elaborate the other two entries orderly. The activities and decision points of pathway for ongoing management of RA are described as follows:

Activity 13: Ongoing management of RA starts by clinic visit of the patient which was processed in the Activity 12 called 'Book Followup Appt' represented in the pathway for initial management of RA (Figure 4.6).

Activity 14: Rheumatologists calculate DAS28 and check blood test results. Details of disease activity and blood results are stored.

Decision 05: Rheumatologists categorise patients considering DAS28 into three categories: High ($DAS28 > 5.1$), Moderate ($DAS28 > 3.2$ and $DAS28 \leq 5.1$) and Low ($DAS28 > 2.6$ and $DAS28 \leq 3.2$), and Remission ($DAS28 \leq 2.6$). Rheumatologists are liable to categorise patients based on their DAS28 and using their discretion as well as patient's observations of their disease activity. Categories of disease activity create three branches in the process:

1. Remission ($DAS28 \leq 2.6$):

   Activity 15: RA remission is diagnosed.

   Decision 07: Rheumatologists decide to change the medication or not, which denotes to escalation of drugs, adjust patients on bDMARDs to tapering their medication. If medication is not changed, Activity 16 is processed. If rheumatologists decide to change medications, Activity 17 is processed.

   Activity 16: Rheumatologists schedule next clinic appointment to review blood test results.

   Activity 17: Rheumatologists write prescription of changed medication. Details of prescription are recorded.

   Activity 18: Rheumatologists make decision on frequency of monitoring considering blood test results and also new medications.

2. Moderate and Low Disease Activity ($2.6 < DAS28 \leq 5.1$):

[3] Our colleague, Dr Scott McLachlan, created this model of pathway originally. We revised its content, added an additional branch (gray nodes and arrows), and edited its structure visually.

Activity 19: Disease is categorised as moderate and low.

Activity 20: Rheumatologists review blood test results of inflammation markers and functioning of liver and other organs. Blood tests are usually conducted in a local clinic close to patients' residence place and rheumatologists receive the results using the available systems of electronic health records.

Decision 08: If rheumatologists find an issue in blood test results, Activity 21 is processed. Otherwise, Activity 16 is processed.

Activity 19: Patients get a phone call.

Decision 08: Clinicians investigate if patients can receive steroid injection. If so, Activity 22 will be performed. Otherwise, medication change may be needed in Decision 07.

Activity 22: Patients attend a clinic and clinicians perform steroid injection.

3. High Disease Activity ($DAS28 > 5.1$):

   Activity 23: A patient with DAS28 greater than 5.1 is categorised as high disease.

   Activity 24: Clinicians review blood test results of highly active RA patients regularly.

   Decision 10: Rheumatologists checks the history of medication to know if the patient is on bDMARDs. If so, patient is led to Decision 08 for possible medication change. If patient is not on bDMARDs, rheumatologist directs to Decision 11.

   Decision 11: Rheumatologists review if patients have a history of failure of two prior csDMARDs and patient is still on a highly active RA. If so, rheumatologists direct patient to Activity 25. Otherwise, the patient will be directed to Decision 07 for possible change of medications, i.e. csDMARDs or steroids.

   Activity 25: Rheumatologist considers biologics or bDMARDs. About 30 to 40 percent of patients respond to bDMARDs. And if they don't respond after 3 months, then rheumatologists consider other bDMARDs. In other words, they do trial and error. Before bDMARDs prescription, rheumatologists screen patients' blood test, hepatitis, Tuberculosis (TB), and Human Immunodeficiency Viruses (HIV). There are 18 different biologic drugs comprising first-line bDMARDs, and second line bDMARDs. The main criterion to choose bDMARDs is patients' comorbidities. bDMARDs are often combined with csDMARDs and steroids.

   bDMARDs producer companies deliver drugs and train patients through a helpline. Since patients cannot reach hospital often, they are advised to not accept drug delivery if they feel much better. However, they tend to accept drug delivery and stock them.

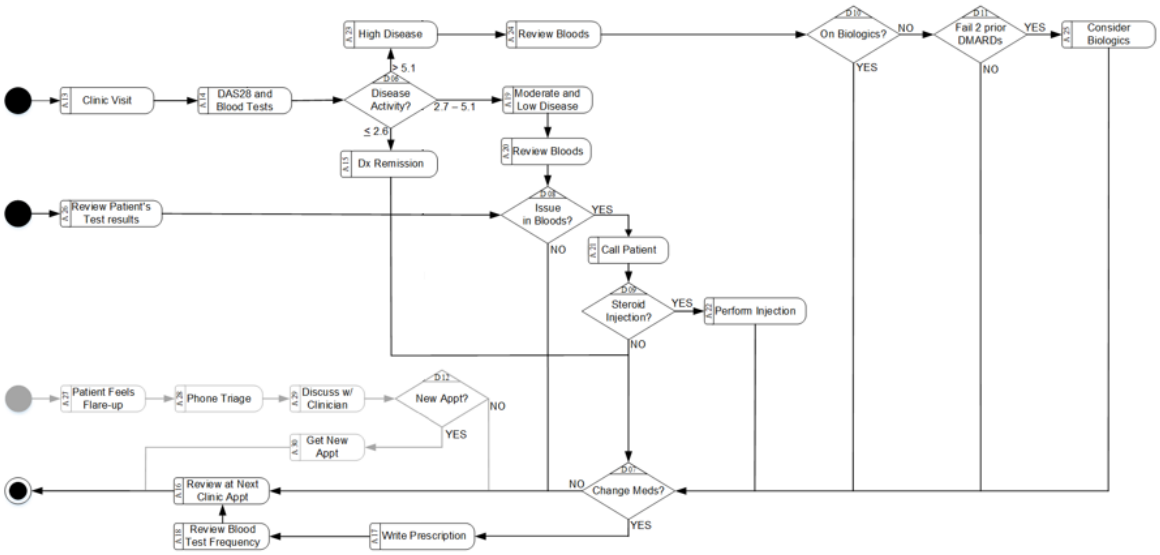


Fig. 4.7 Pathway for RA ongoing management.

The second entry into the process of ongoing management of RA is through Activity 26, as described below:

Activity 26: Patient's blood test results are reviewed in a local clinic close to the patient's residency, either by a GP nurse or the hospital phlebotomist.

Another entry into the process of ongoing management of RA is from Activity 27. This entry and its subsequent activity and decision nodes are coloured gray because they are implemented with limitations through mobile health apps or other novel technologies.

Activity 27: Patients experience a worsening of their symptoms or flare and their symptoms are sent to a digital health centre.

Activity 28: Health service does phone triage by asking questions to the patient to know the level of urgency.

Activity 29: Patients discuss with a clinician to get proper advice on treatment and lifestyle.

Decision 12: Clinician decides to set a new appointment or not, if the patient's flare is severe and prevents the patients' normal life.

Activity 30: Patient receives an urgent appointment.

4.3.4 Pathway for Personalised Care for Living with Rheumatoid Arthritis

In the current clinical practice, patients receive some information in the form of leaflets prepared by NRAS in the diagnosis visit or later in monitoring visits. As mentioned in

Subsection 4.2.3, information on living with RA and managing QoL is not personalised. Personalisation of care refers to having choices and control to select care planning and delivery [133]. Personalised care for living with RA is not properly done, except for the provision of information and sources of information.A proper information on personalisation for living with RA can lead patients to improve their QoL in three aspects of physical, social, and psychological by balancing their disease activity and lifestyle factors.

The process of providing personalised care for living with RA can be depicted as a pathway shown in Figure 4.8. The numeration of activities and decision indicates their order, but it is not a continuation of diagnosis, initial and ongoing management pathways as they were. The reason is because the pathway for personalised care for living with RA is a new care process that we suggest adding to the current clinical practice, yet it is not implemented. The activity and decision nodes of this pathway are described as follows.

Activity 01: After diagnosing with RA, patients get familiar with the system and learn how to use it.

Activity 02: Patients create an account and enter their personal information including age, sex, and body mass index (BMI). Patient's location can also help to determine environmental factors which may affect their experience.

Activity 03: Patients provide their observations - including their disease manifestation, QoL characteristics, and lifestyle choices - by answering a set of questions.

Activity 04, 05, and 06: Based on the information and observations, the system estimates three components of QoL corresponding to the social, psychological, and physical aspects of QoL in each time.

Activity 07: The assessment of physical, psychological, and social aspects of QoL helps to estimate the independence, empowerment, and participation components of QoL, respectively. The estimation of QoL components are mapped to a set of advice. This advice targets any aspect of QoL that is declined and provides personalised advice to improve them.

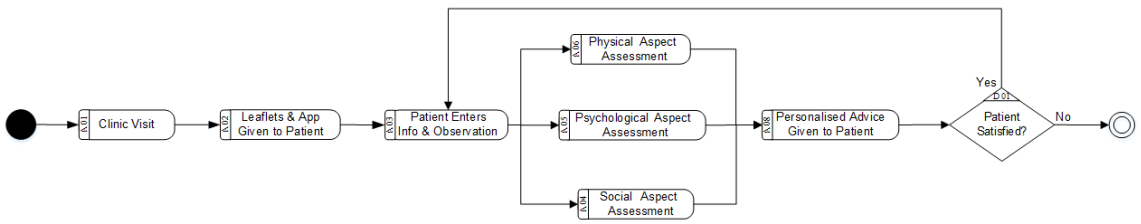


Fig. 4.8 Pathway for personalised care of RA.

Decision 01: In the end of the cycle, if the patient is satisfied with the personalised advice, they use it again by entering their observation (return to Activity 03). The system invokes their personal information and estimates their QoL components. If the patient is not satisfied, they withdraw from the system.

4.4 Description and Analysis of Available Data from Studies on Rheumatoid Arthritis

We are provided with two sets of data: PEAC data and BioT App data. These two datasets and their contents are described in the following subsections.

4.4.1 Introduction to Studies

Pathobiology of Early Arthritis Cohort (PEAC) is an ongoing study collecting data of people diagnosed with inflammatory arthritis diseases, particularly RA [157]. The study has been running since 2009. The PEAC data include patients' personal, demographic, lifestyle, medical background, monitoring information–namely signs, symptoms, and serology results–, diagnosis, medication, and adverse events. They are collected in quarterly clinic visits for five times: baseline, first follow-up (3-month), second follow-up (6-month), third follow-up (9-month), and fourth follow-up (12-month). There is an extra follow-up in the first month, but it only includes medication and disease activity data, not monitoring records. Therefore, we do not consider it.

Rheumatologists make diagnosis based on the risk factors, signs, symptoms, and serology results. In the PEAC study, patients are either diagnosed with RA or other IA conditions including psoriatic arthritis (PsA), monoarthritis (MonoA), undifferentiated arthritis (UA), and other sorts of inflammatory arthritis (IA). Diagnosis data is recorded in two variables: Diagnosis and 'Revised Diagnosis'. The former is the initial diagnosis made in the first clinic visit which in some cases was more than 10 years ago, and they were diagnosed using an old diagnosis criteria called 1987 criteria [9]. The new diagnosis criteria called 2010 criteria was released which could improve the early diagnosis of RA [6]. From 2012 onward, patients were classified by 2010 criteria in PEAC study. However, 'Revised Diagnosis' refers to any diagnosis different than the initial diagnosis made after gaining more evidence about the condition. This mainly can happen to those diagnosed with UA.

Baseline data are gathered in the first visit by a rheumatologist. It contains the records of 384 patients, which drops to 373 cases since 11 cases have no recorded diagnosis. Of them,

226 are RA, 79 cases are UA, 49 cases are PsA, 12 cases are MonoA, and 7 cases are other arthritis.

In the original baseline dataset, there are 152 columns, which 139 of them are unique. We keep interesting columns in consultation with our main rheumatology expert comprising of personal information, risk factors, comorbidities, signs, symptoms, serology results, diagnosis, disease activity measurements, and the PEAC identification numbers (PEAC IDs) of patients. A total of 30 columns remain with 494 missing values (almost 4%). We drop any case (row) with more than 4 missing values; thus 360 cases remain with almost 4% missing values. We use the final subset of the baseline data for building BN models for the diagnosis of RA described in Chapter 5.

Follow-ups data have identical number of columns. In these data, we keep cases that are initially diagnosed with RA since we only consider monitoring and treatment of RA cases. As mentioned before, 226 RA cases exist in the baseline, which diminished to 215. We keep the records of 215 RA cases and drop the records of other arthritis cases. Other follow-up visits, namely 3-, 6-, 9-, and 12-month, have 181, 176, 157, and 146 cases, respectively. The reduction of cases is mainly because some patients give up attending the study or they are directed to other studies. The final subsets of follow-up data will be manipulated and ultimately used for building dynamic BN (DBN) models for self-management of RA (Chapters 7 and 8).

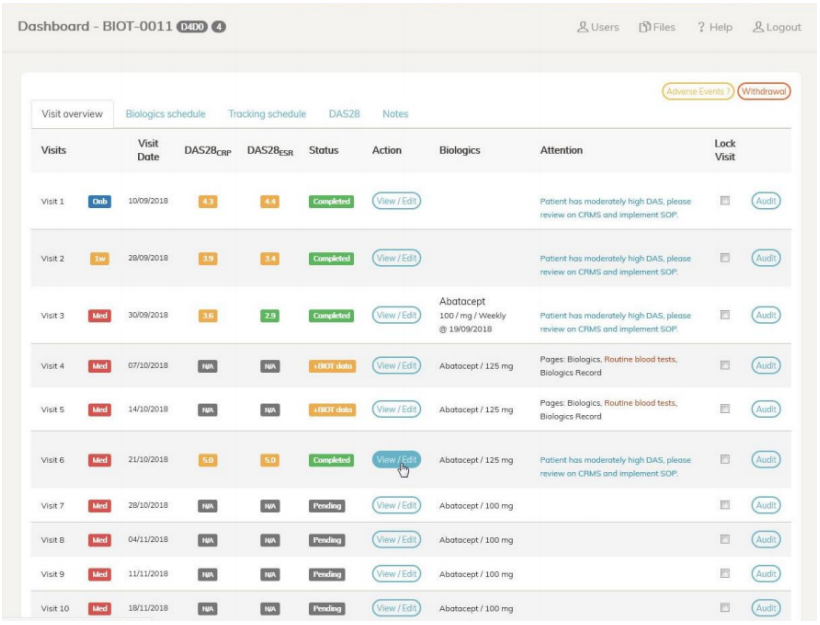
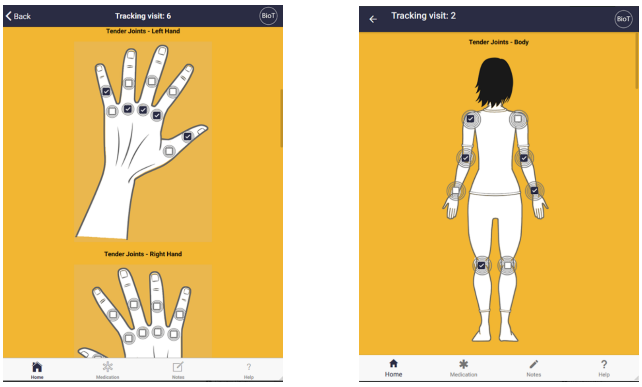


Fig. 4.9 Dashboard of visits in clinicians interface of BioT app.

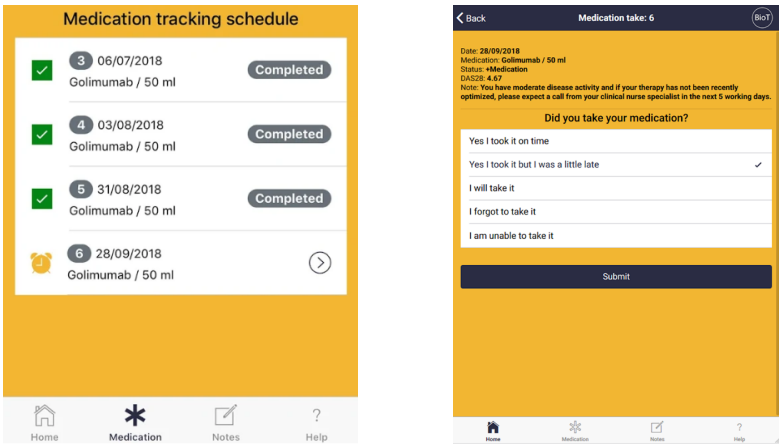
BioT study is on a mobile app co-designed by clinicians from the Centre for Experimental Medicine & Rheumatology of SMD and patients using an iterative collaborative design process [112]. It collects recorded joint counts and tracks medications as a monitoring tool to support clinicians in tapering biological drugs [112]. These drugs are expensive and some cases are misused or wasted. The BioT app intends to enhance remote monitoring and biological drug management.



(a) Small tender joints. (b) Large tender joints.

Fig. 4.10 BioT app pages for counting tender joints.

The app has an interface for clinicians and another interface for patients. Clinicians can see the profile of patients with a trend diagram of their DAS28. The app demonstrates a dashboard (like Figure 4.9) with details of patient's visits, bDMARDs schedule, tracking schedule, DAS28, and other notes.



(a) Medication tracking. (b) Medication adherence.

Fig. 4.11 BioT app pages on medication tracking and treatment feedback from patients.

The interface for patients interacts with patients by receiving their observations and providing them with simple advice such as reminding them to take their medicine. As shown in Figures 4.10a and 4.10b, the app respectively visualises small and large joints to the patients to determine their tender joints. Similar pages allow patients to specify their swollen joints.

BioT app displays a list of medication tracking schedules and reminds patients to take their medications, as shown in Figure 4.11a. The submission of taken medication can be done only on the due date and afterwards up to a maximum 7 days. The app also questions patients if they adhered their medication or not, as demonstrated in Figure 4.11b.

4.4.2 Description and Analysis of Personal Information, Risk Factors and Comorbidities

We have classified personal and risk factors into four groups: personal, lifestyle, demographic, and medical background. Personal information include Age and Sex. Age has a number of outlier values, e.g., one age record is '10' but we know that all patients of the PEAC study are at least 18 years old. Sex records include Male and Female categories and there are four missing values.

Lifestyle factors encompass BMI, Alcohol, and Smoking. There are outlier values in the records of them which we deleted them based on our rheumatologist's judgement. For example, any BMI value less than 8 is dropped. For alcohol consumption, patients are supposed to claim how much alcohol they consume weekly and then rheumatologists calculate the units of alcohol consumption. According to [129], units are computed by the percentage of alcohol by volume (ABV) and its volume as follows:

$$\text{Alcohol units} = \text{strength(ABV)} \times \text{volume(L)} \tag{4.3}$$

In some cases, the number of pints or the number of glasses per week are recorded, which we convert them to the units of alcohol consumption by multiplying them into 2.3 units [50].

Smoking data are gathered in three variables: current smoker, previous smoker, and packsyear. The first two are binary (Yes and No), whereas the latter contains numbers of 'packsyear' for patients who were or are smokers. Pack-year or packsyear is calculated by multiplying the number of smoked packs of cigarettes by the number of years of smoking as follows:

$$\begin{aligned}\text{Number of pack-years} =&(\text{number of cigarettes smoked per day}/20)\\ &\times \text{number of years smoked}\end{aligned} \tag{4.4}$$

Demographic factors include education, ethnicity, and occupation. Education is recorded as the number of years of studying, e.g., 12 or 15 representing A-level and Bachelor's degree, respectively. Clinicians record the ethnicity of patients identified by themselves. Based on NHS's ethnic category code [130], we classified ethnicity into 5 groups: White, Asian or Asian British, Black or Black British, Mixed and Others. The Mixed and Others categories are originally separate, but with the confirmation of our main expert, we merge them since each of them is infrequent.

Patients state their occupation in the clinic visit and there is a variety of occupation records. We divide occupation into Industrial and Non-industrial, due to the effect of the industrial particle on getting diagnosed with RA [88, 106, 155]. Retirement from industrial or non-industrial occupations are assigned to their relevant categories of Industrial or Non-industrial. In addition, we separated the unemployed people and assigned them to Unemployed category.

Family history (FHx) of patients are recorded based on patients' claim. Some records include a question mark which refers to the uncertainty of the patient whether their family had RA or not. We remove these cases they are not reliable. For consider RA FHx as a binary variable (Yes and No), and any record of 'RA', 'rheumatoid arthritis', or 'felty syndrome' as a relevant family history of RA.

We consider the FHx of other inflammatory arthritis (IA) as a binary variable (Yes and No). We consider the following cases as relevant IA FHx: 'psoriatic arthritis', 'monoarthritis', 'inflammatory arthritis', 'juvenile arthritis', and 'undifferentiated arthritis' or 'unspecified arthritis'. Although skin psoriasis is not an IA disease, we add any FHx record of 'skin psoriasis' as an IA FHx due to its close association with psoriatic arthritis disease. Similar to RA FHx, any record with a question mark refers to the uncertainty of the patient in recalling their family history, which we remove them.

Comorbidities of crystal arthropathy (CA), connective tissue disease (CTD), osteoarthritis (OA), other inflammatory arthritis (IA), skin psoriasis, and thyroid autoimmune disease (TAD) are recorded in the PEAC data. We group any record of 'Gout' as a comorbidity of CA. We categorise the records of 'Secondary Sjogrens Syndrome' and 'Psoriasis' as CTD and skin psoriasis, respectively. The records that contain 'Psoriatic', 'Undifferentiated', 'Juvenile', 'Reactive', 'Spondyloarthritis', or 'Enteropathic' terms are considered as 'Other IA' cases. 'Hypothyroidism', 'Hyperthyroidism', and 'Graves disease' are grouped as TAD.

Table 4.1 presents a summary of risk factors and comorbidities recorded of the PEAC data specifying the mean and range of continuous variables and a category of the categorical variables and its percentage. Missing values and their percentage is also provided. All figures belong to the subset of baseline PEAC dataset containing 360 with 4% missing values, as explained in Subsection 4.4.1.

Table 4.1 Summary of personal information, risk factors, and comorbidities of PEAC data.

Variable	Statistics	# Missing
Age	51(18-97)	4(1%)
Sex	Female(63%)	0(0%)
Education	School(45%)	101(28%)
Ethnicity	White(44%)	36(10%)
Occupation	Industrial(8%)	53(15%)
BMI	27.80(16.41-53.82)	56(16%)
Alcohol	3.62(0-65)	45(13%)
Smoking	6.98(0-69)	46(13%)
RA FHx	Yes(19%)	0(0%)
Other IA FHx	Yes(4%)	0(0%)
Comorbidity OA	Yes(3%)	0(0%)
Comorbidity TAD	Yes(6%)	0(0%)
Comorbidity Skin Psoriasis	Yes(8%)	0(0%)
Comorbidity Other IA	Yes(0%)	0(0%)
Comorbidity CTD	Yes(0%)	0(0%)
Comorbidity CA	Yes(2%)	0(0%)

Figure 4.12 demonstrates a pair plot of scatter plots of risk factors with continuous records, namely, Age, Smoking, BMI, and Alcohol, with colour encoding of diagnosis records (RA, UA, MonoA, PsA, and Others). The figures on the diagonal depict the kernel density estimation of histograms of each risk factor with encoded colours of diagnosis.

4.4.3 Description and Analysis of Signs, Symptoms, Serology Results, and Disease Activity

Rheumatologists examine patients' joints to find tender and swollen joints. In the PEAC data, there are two records of tender joint count (TJC) and swollen joints count (SJC). They examine 28 joints: 20 small joints (MCP and proximal interphalangeal (PIP) joints of each finger) and 8 large joints (wrists, elbows, shoulders, and knees). TJC is examined by pressing the joints. If patients feel pain while pressing, their joints are considered as tender. For SJC, rheumatologists observe the same 28 joints and count the swollen ones.

Symmetrical patter is a sign of RA and differentiates it from other IA diseases. In the initial visit for diagnosis, rheumatologists check the pattern of swollen joints in small joints to determine whether the swollen joints are symmetrical or not. They record a binary (Yes and No) variable called symmetrical swollen joints.

GH, Fatigue, Pain, Health Assessment Questionnaire (HAQ), and Stiffness are the common symptoms of RA collected in the PEAC study. The data of GH, Fatigue, and Pain

are originally the Visual Analog Scale (VAS) for patient's global assessment, tiredness, and pain, respectively, i.e., clinicians ask patients to specify a number between 0 and 100 to express their global health, feeling of fatigue or tiredness, and feeling of pain, respectively. 0 refers to the minimum (desirable) and 100 refers to the maximum (undesirable) feeling of unhealthiness, tiredness, and pain. Alternatively, patients are given an almost 10cm (9.8cm exactly) ruler to point out their level of pain, fatigue, and global health. Then, rheumatologists convert the pointed number into a percentage.

Patients are also asked whether they feel morning stiffness for more than 30 minutes. A binary answer (Yes or No) is recorded. HAQ measures the disability of patients calculated

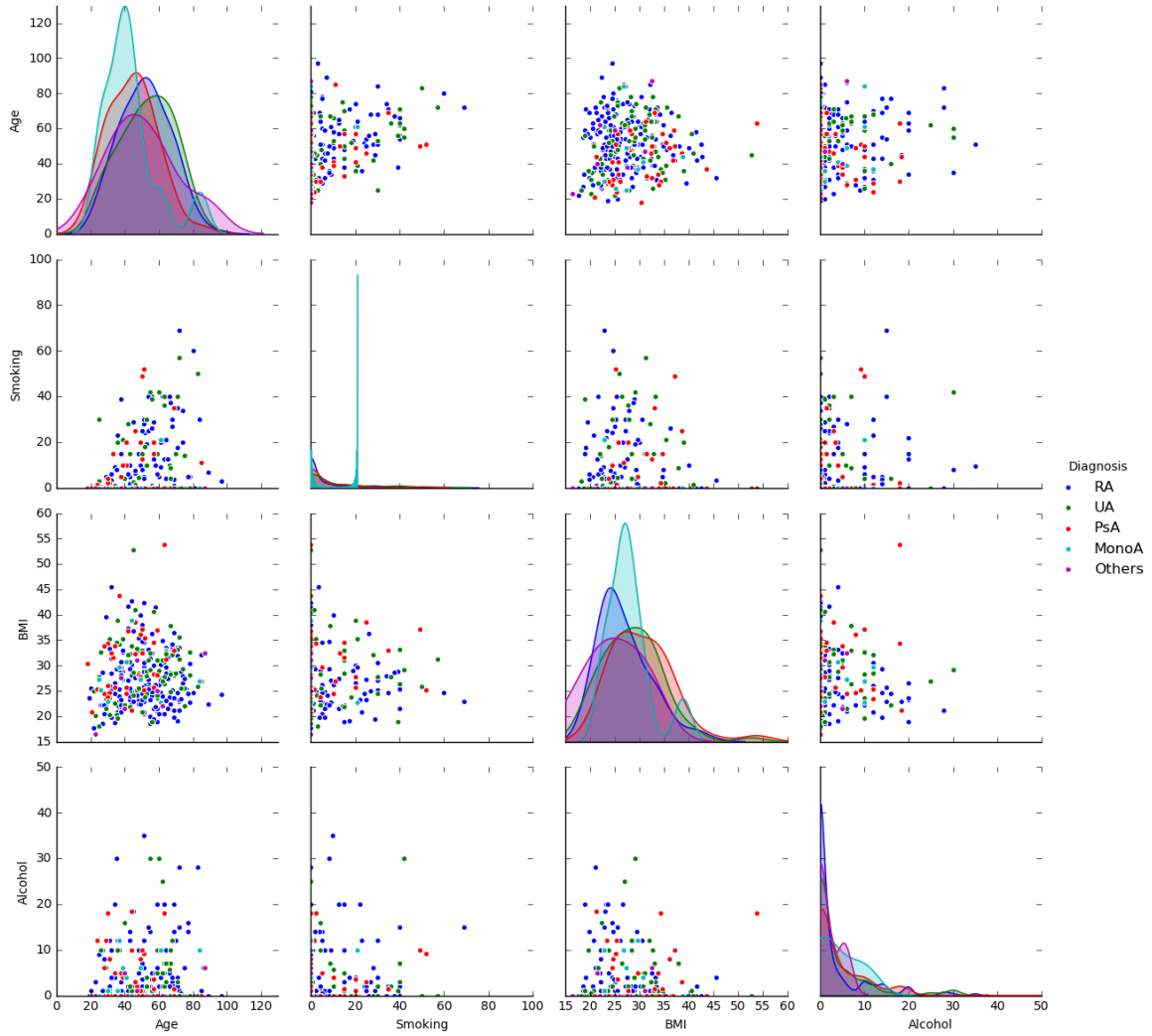


Fig. 4.12 Pair plot of scatter plots of Age, BMI, Smoking, and Alcohol in the baseline of PEAC data with kernel density estimation of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

by asking a set of questions suggested in [76]. HAQ values are between 0 and 3, which are commonly categorised into three groups of Mild, Moderate, and Severe.

RhF titre is an antibody found in the blood test result and it is measured by a unit called the international unit per millilitre (IU/ml). We deal with some inaccurate records such as '<11' or 'Neg' that mean less than 11 and less than 20 units of RhF, respectively. With the confirmation of our rheumatology expert, we respectively replace the inaccurate records with random numbers smaller than 11 and 20. We observe a relatively greater frequency in 180 and 600 that is because these values are the cut-off levels of most of the labs. This is a minor issue since rheumatologists tend to categorise the records of RhF titre into Negative (≤ 20), Low positive ($20 >$ and ≤ 60), and High positive ($60 <$).

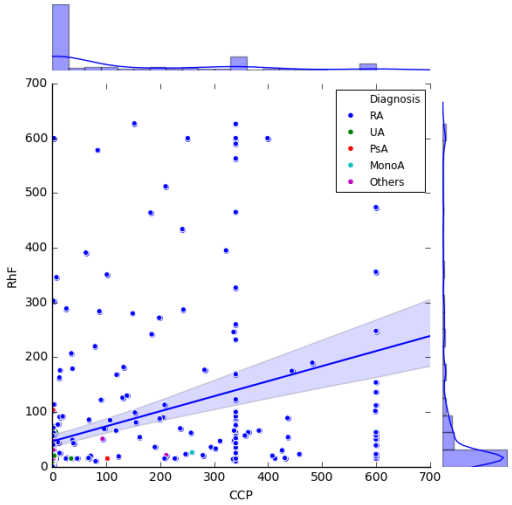
CCP titre is an antibody reported in the blood test results measured in U/ml unit. The minimum possible value of CCP is 0 and similar to RhF, the common maximum value is the lab cut-off level which is 600. There are no outlier values, but there is an accumulation of records in 340 and 600 values, which is because the Bart's Health NHS Trust and most of the UK labs apply these two values as their cut-off levels. Like RhF, we can ignore the issue with the accumulation of records in the cut-off points since rheumatologists usually categorise the records of CCP titre into Negative (≤ 10), Low positive ($10 >$ and ≤ 30), and High positive ($30 <$).

Figure 4.13a demonstrate the scatter plot of CCP and RhF with joint histogram and kernel density estimation of each and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others. We learn from our main expert that smoking has association with the release of RhF and CCP antibodies [160, 151, 14], as a pathogenesis mechanism of RA development which we further describe in Subsection 5.2.3. Therefore, we draw scatter plots of RhF and CCP with Smoking in Figures 4.13b and 4.13c. Histogram and kernel density estimations are plotted and the colours of scatter plots encode diagnosis categories.

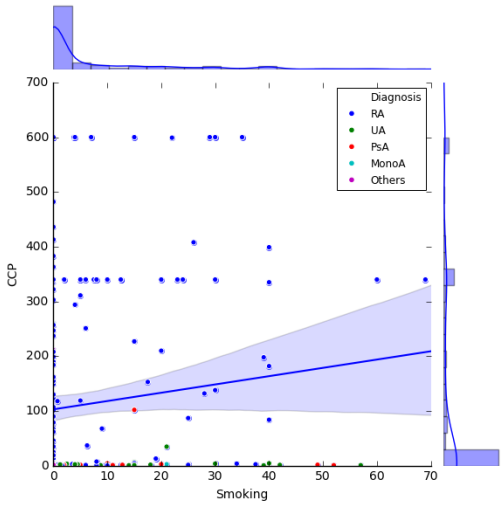
There are two inflammation markers collected in the PEAC study. One of them is Erythrocyte sedimentation rate or ESR that is the distance of dropping of red blood cells and its measurement unit is mm/hour. C-reactive protein or CRP is the other inflammation marker that is a response to any sort of inflammation in the body and its measurement unit is milligram per litre (mg/lit). In the PEAC data, there are some inaccurate records of '<5' and 'Neg' which both mean less than 5mg/lit. We replace those inaccurate records with random numbers between 0 and 4 as it is confirmed by our rheumatology expert.

In the PEAC study, disease activity is estimated using DAS28 (introduced in Subsection 4.2.2). Both types of DAS28 with ESR and CRP are recorded depending on their availability. The records of GH are also considered in the calculations of DAS28.

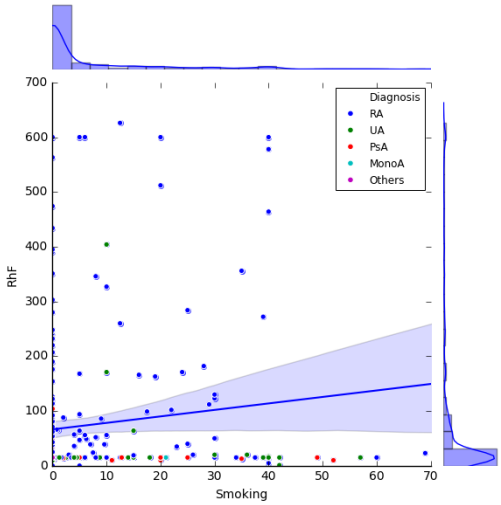
Table 4.2 describes the mean, range, number of missing values, and missing value percentages of signs (TJC and SJC), symptoms (GH, Fatigue, Pain, HAQ, and Stiffness), serology results (CCP, RhF, ESR, and CRP), and disease activity (DAS28ESR and DAS28CRP) collected at the baseline, and the first, second, third, and fourth follow-up visits. Baseline data comprises of 360 cases (226 RA), and the first, second, third, and fourth follow-up visits include 181, 176, 157, and 146 RA cases only.



(a) CCP-RhF.



(b) Smoking-CCP.



(c) Smoking-RhF.

Fig. 4.13 Scatter plots joint with histogram plots and kernel density estimation of Smoking, CCP, and RhF in the baseline of PEAC data, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

Table 4.2 Description of signs, symptoms, serology results, and disease activity records in PEAC data at first, second, third, and fourth follow-ups.

Variable	Baseline		First follow-up		Second follow-up		Third follow-up		fourth follow-up	
	Statistics	Missing	Statistics	Missing	Statistics	Missing	Statistics	Missing	Statistics	Missing
TJC	9.75(0-28)	0(0%)	5.97(0-9)	0(0%)	5.60(0-28)	0(0%)	6.01(0-28)	0(0%)	6.30(0-27)	0(0%)
SJC	5.90(0-26)	0(0%)	2.74(0-28)	0(0%)	2.57(0-16)	0(0%)	2.78(0-20)	0(0%)	2.21(0-22)	0(0%)
GH	60.51(0-100)	0(0%)	38.73(0-100)	1(0%)	36.38(0-100)	2(1%)	40.24(0-100)	0(0%)	39.99(0-100)	0(0%)
Fatigue	38.92(0-100)	0(0%)	37.53(0-100)	1(0%)	37.70(0-100)	2(1%)	38.05(0-99)	0(0%)	39.30(0-97)	1(0%)
Pain	50.02(0-100)	0(0%)	31.54(0-100)	1(0%)	31.81(0-97)	2(1%)	34.28(0-100)	0(0%)	35.36(0-100)	1(0%)
HAQ	1.29(0-3)	0(0%)	0.86(0-2.88)	1(0%)	0.87(0-2.63)	2(1%)	0.89(0-2.75)	0(0%)	0.92(0-3)	1(0%)
Stiffness	Yes(61%)	0(0%)	Yes(7%)	0(0%)	Yes(6%)	0(0%)	Yes(7%)	8(5%)	Yes(8%)	4(3%)
CCP	122(0-600)	28(8%)	-	-	-	-	-	-	-	-
RhF	76.27(0-627)	13(4%)	-	-	-	-	-	-	-	-
ESR	32.52(2-124)	3(1%)	19.09(0-88)	12(7%)	18.39(2-95)	8(4%)	20.23(2-114)	8(5%)	20.30(2-106)	4(3%)
CRP	16.97(0-170)	3(1%)	8.56(0-79)	17(9%)	8.32(0-75)	13(7%)	9.28(0-66)	12(8%)	8.65(0-93)	11(8%)
DAS28ESR	5.19(0.78-8.92)	4(1%)	3.66(0.49-7.94)	13(7%)	3.56(0.29-7.69)	8(4%)	3.72(1.98-7.91)	8(5%)	3.70(0.49-8.12)	4(3%)
DAS28CRP	4.90(1.75-8.12)	5(1%)	3.65(0.96-7.38)	17(9%)	3.48(1.20-7.48)	13(7%)	3.66(0.96-7.07)	12(8%)	3.63(0.96-7.08)	11(8%)

Table 4.3 Description of signs, symptom, serology results, and disease activity records of BioT data in 11 times.

Variable	1		2		3		4		5		6		7		8		9		10		11	
	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n	Statistics	n
ESR	7(0-30)	32	8(0-29)	24	10(1-29)	21	9(0-26)	20	9(0-41)	22	8(0-40)	19	9(2-40)	20	11(0-40)	18	12(2-40)	19	8(2-29)	20	11(2-36)	20
CRP	4(1-38)	32	4(1-38)	30	3(1-24)	27	2(1-13)	23	3(1-24)	24	2(1-10)	20	2(1-10)	23	2(1-10)	21	2(1-10)	21	2(1-10)	20	2(1-8)	21
TJC	3(0-18)	32	4(0-28)	30	4(0-17)	26	4(0-10)	24	4(0-13)	25	4(0-12)	20	4(0-13)	23	4(0-10)	18	5(0-17)	20	4(0-10)	20	3(0-8)	21
SJC	1(0-11)	32	4(0-28)	30	4(0-28)	26	2(0-10)	24	3(0-10)	25	2(0-10)	20	3(0-13)	23	2(0-8)	18	3(0-10)	20	3(0-9)	20	3(0-9)	21
GH	24(0-91)	32	28(0-81)	30	23(0-64)	26	25(0-69)	24	27(0-83)	25	23(0-60)	20	27(0-85)	23	21(0-60)	18	26(0-77)	20	22(0-72)	20	21(0-73)	21
DAS28ESR	3(0-6)	24	3(0-7)	22	3(0-5)	20	3(0-5)	18	3(0-5)	21	3(1-5)	18	3(1-5)	20	3(2-5)	14	3(1-5)	17	3(1-5)	20	3(1-5)	20
DAS28CRP	3(1-6)	32	3(1-6)	28	3(1-6)	24	3(1-5)	23	3(1-5)	23	3(1-5)	19	3(1-5)	23	3(2-4)	17	3(2-5)	20	3(1-4)	19	3(1-4)	20

A pairplot of scatter plots of the components of DAS28, namely, TJC, SJC, GH, ESR, and CRP, in the baseline is displayed in Figure 4.14. The kernel density estimation of histogram of each variable is plotted on the diagonal of the pairplot. Colour encoding of diagnosis records separates RA, UA, MonoA, PsA, and Other arthritis cases. Similar pairplots of TJC, SJC, GH, ESR, and CRP in the first, second, third, and fourth follow-up visits with kernel density estimations of histograms and colour encoding of DAS28ESR are presented in Figures A.1, A.2, A.3, and A.4.

Disease activity measurements of DAS28ESR and DAS28CRP in the PEAC dataset are closely correlated, as Figure 4.15 displays a scatter plot of DAS28ESR and DAS28CRP with

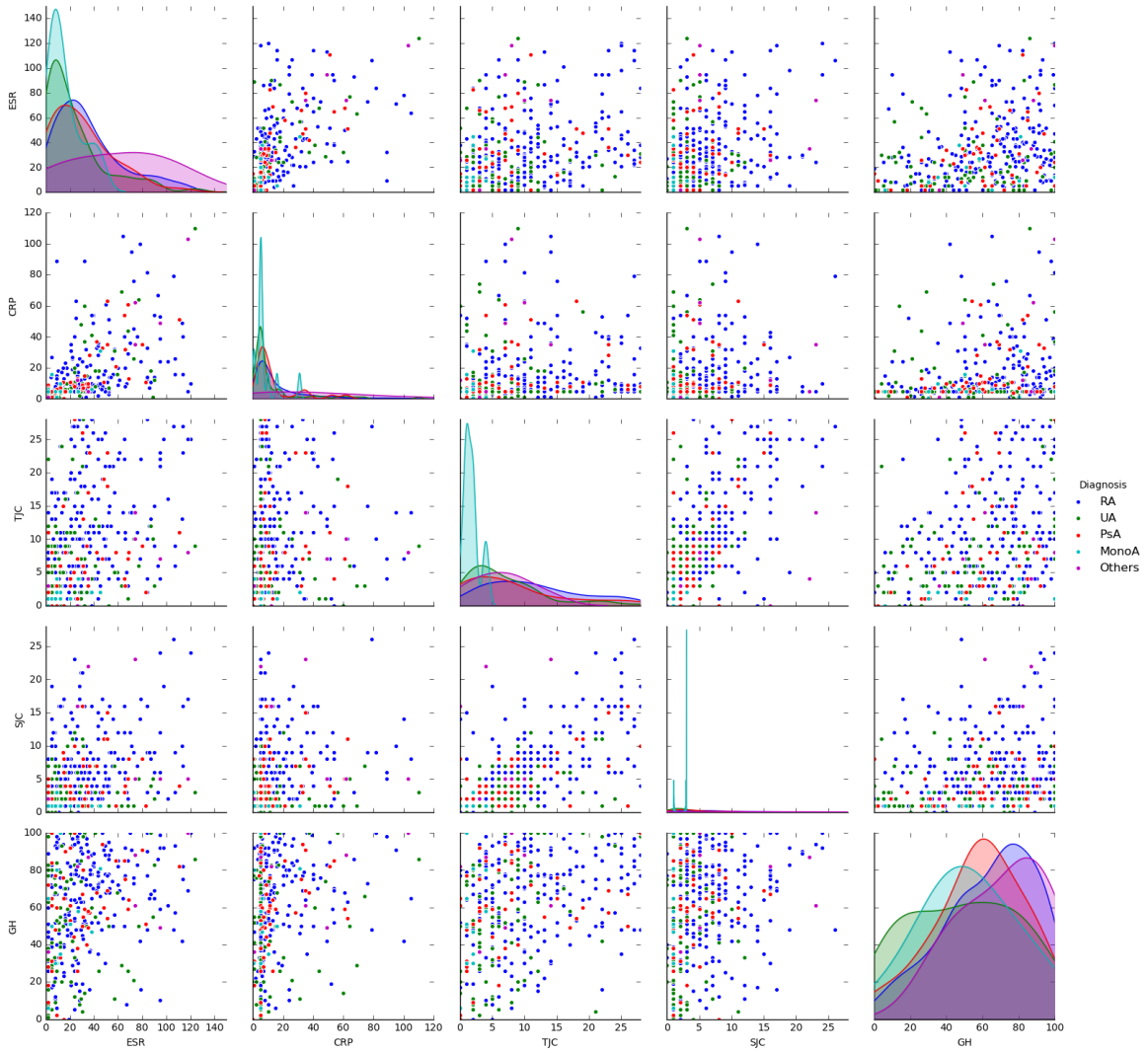


Fig. 4.14 Pairplot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the baseline of PEAC data with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

joint plots of their histograms and kernel density estimation of each. The colours of scatter plot encode the diagnosis of RA, UA, MonoA, PsA, and Others.

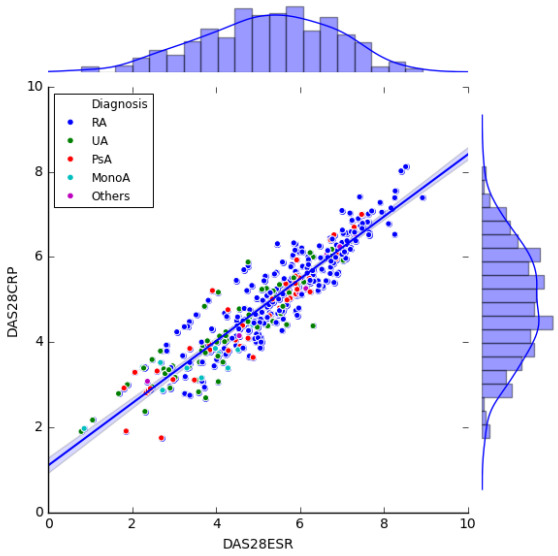


Fig. 4.15 Scatter plot of DAS28ESR and DAS28CRP joint with histogram plot of each and their kernel density estimation, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

BioT Data

Recruited patients of the BioT app study input their TJC, SJC, and VAS or GH in a weekly manner and their ESR or CRP is checked in a local clinic every 2-4 weeks [111]. They also need to declare whether they took their medication or not.

The dataset consists of the records of 32 patients, whom input their records for a maximum of 11 times during about 4 months. Some patients left the study and have partial records. It includes two signs (TJC and SJ), one symptom (GH), two blood test results (CCP, CRP, and ESR), two disease activity measurements (DAS28ESR and DAS28CRP), medications, status of medication, patient response, clinician action outcome, and nurse action outcome. In the beginning of the study, clinicians collect personal information of patients, namely, height, weight, BMI, pulse, temperature, and blood pressure are measured.

Records of TJC, SJC, GH, CCP, CRP, ESR, DAS28ESR, DAS28CRP, and personal information of patients are continuous data, but medication and medication status are categorical data. Patient response, clinician action outcome, and nurse action outcome are text.

There are four types of dates, namely, performed date of joints counting, performed date of GH scaling, date of blood tests, date of action by clinicians. The dates of joints counting

and GH scaling are identical in all cases. Blood test dates should be maximum 15 days apart to the dates of joints counting and GH scaling, so that disease activity measurements would be valid.

Table 4.3 describes the signs, symptom, serology results, and disease activity scores.

As displayed in Figure 4.16, we prepared a scatter plot of DAS28ESR and DAS28CRP to show their similarity. The plot has a colour encoding of DAS28CRP records of High, Moderate, Low, and Remission. We chose DAS28CRP since CRP and subsequently DAS28CRP have more available records than ESR and DAS28ESR. This may be because CRP is more preferred than ESR (see Section 4.3.1).

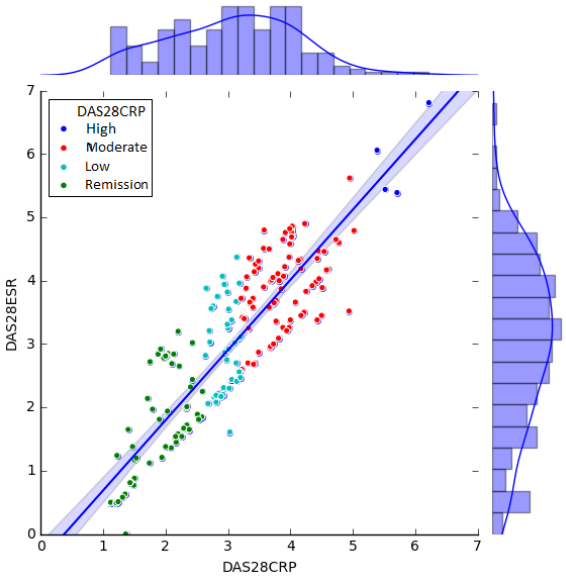


Fig. 4.16 Scatter plot of DAS28ESR and DAS28CRP with marginal histogram of each of them.

A pair plot of monitoring records and blood results and kernel density estimation of each monitoring or blood result is shown in Figure 4.17. In the subplots of this figure, cases are separated based on their DAS28CRP into four groups of remission, low, moderate, and high.

Figures 4.18a, 4.18b, 4.18c, 4.18d, and 4.18e depict the scatter plots of DAS28CRP with ESR, CRP, TJC, SJC, and GH, respectively. Each plot has two histograms of its corresponding variables as marginal plots. Scatter plots differentiate the cases with respect to DAS28CRP in four groups of remission, low, moderate, and high. Among these plots, GH and TJC show the highest correlation with DAS28CRP. CRP and SJC look more correlated with DAS28CRP compared to ESR and TJC, respectively.

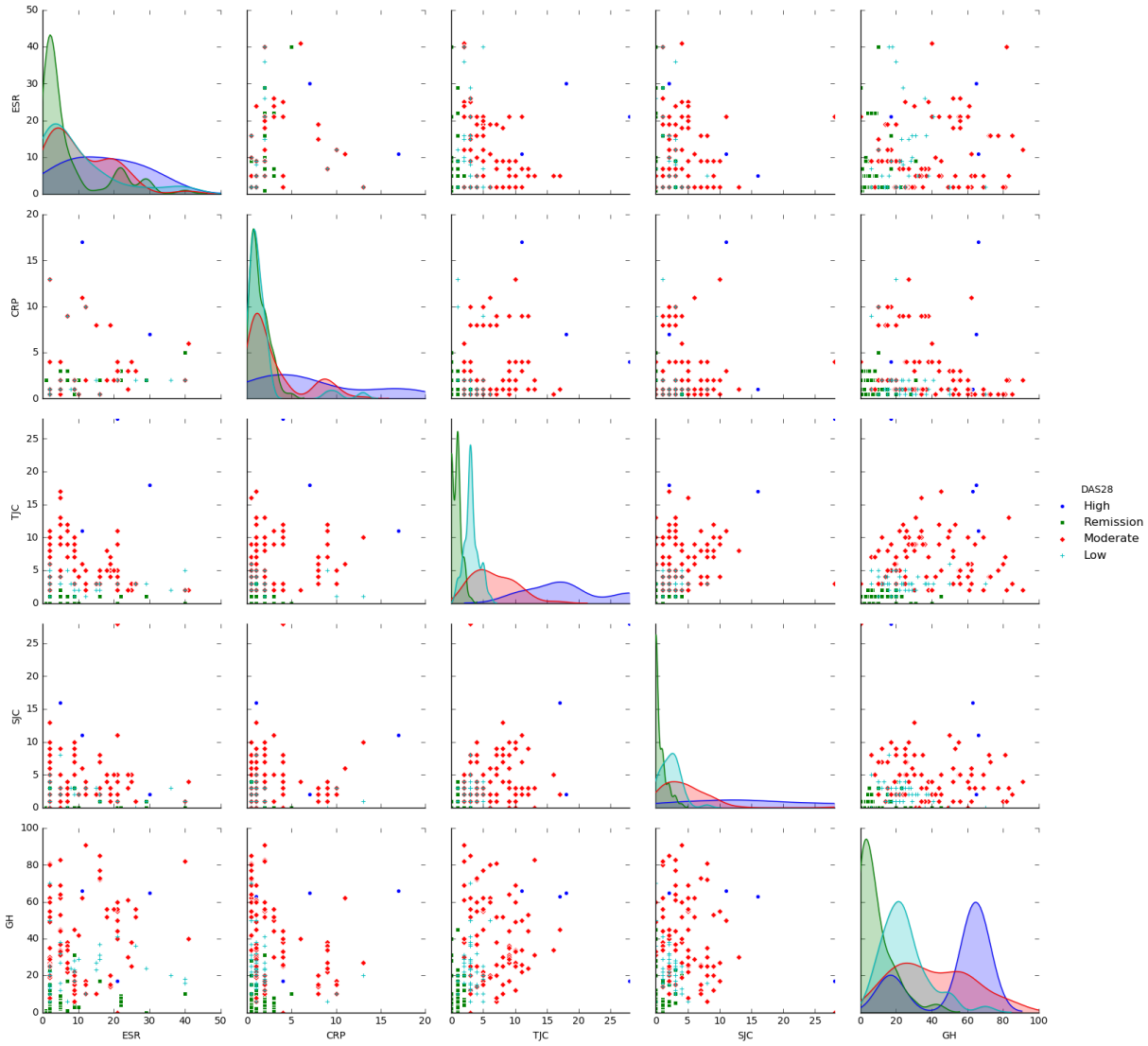
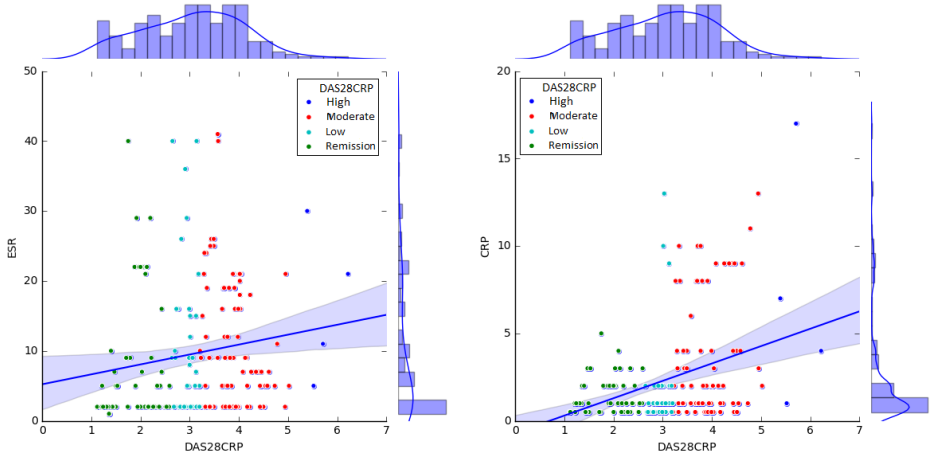
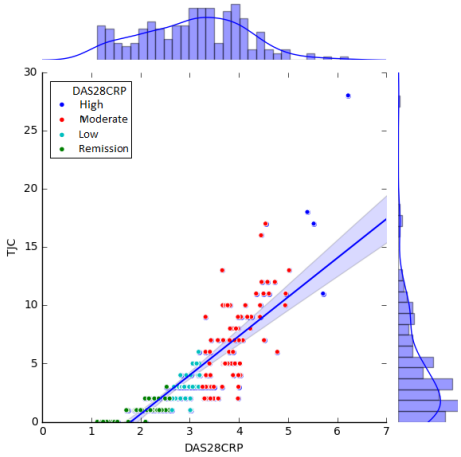


Fig. 4.17 Pair plot of scatter plots of monitoring and blood results and kernel density estimation of each monitoring or blood result.

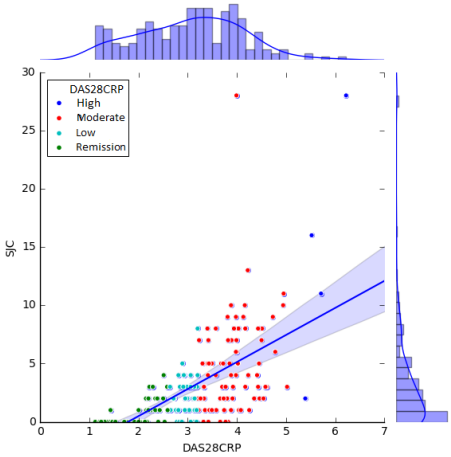


(a) DAS28CRP and ESR

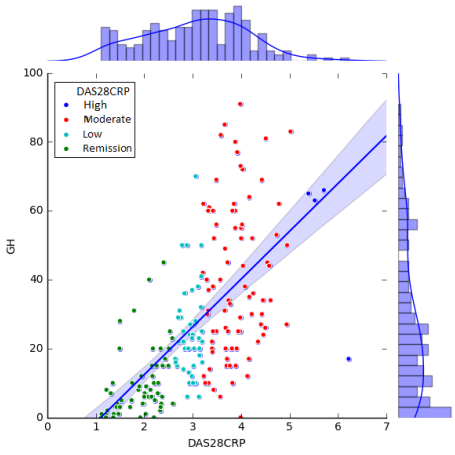
(b) DAS28CRP and CRP



(c) DAS28CRP and TJC



(d) DAS28CRP and SJC



(e) DAS28CRP and GH

Fig. 4.18 Scatter plots of monitoring and blood results with marginal histogram plots of each variable.

4.4.4 Interpolation of Signs, Symptoms, Serology Results, and Disease Activity Records of PEAC Dataset

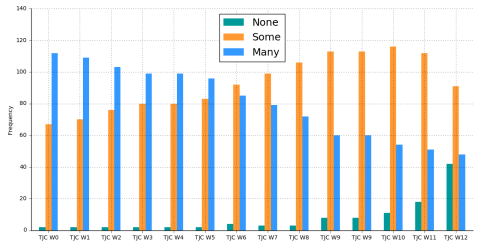
The PEAC dataset is collected quarterly, but we need more frequent data (e.g. weekly) to do reasoning with DBN models for self-management and treatment of RA, as we describe in Chapters 7 and 8. To create a set of weekly data from the PEAC dataset, we interpolate the records of signs, symptoms, and serology results. This allows us to transform the dataset into a granular one with weekly frequency. The interpolation is only applied to the cases with the diagnosis of RA. As mentioned in Subsection 4.4.1, the number of patients diminishes in each follow-up visit after diagnosis since patients leave the cohort. We interpolate the values between two consecutive follow-up visits using the linear interpolation of the 1-dimension interpolate class of SciPy software [34]. Our assumption is that the three months difference between two visits include exactly 13 weeks. Therefore, we interpolate 11 values between the two visits of PEAC data. We interpolate signs (TJC and SJC), symptoms (Pain, fatigue, GH, and HAQ), and serology results (CRP and ESR).

After interpolation, we categorise the actual records and interpolated values using medically meaningful thresholds provided by our main expert. We discretise TJC records into the same three categories but slightly differently: None ($= 0$), Some ($1 \leq$ and ≤ 7), and Many ($8 \leq$ and ≤ 28). Interpolated TJC values are decimal numbers, but the count of joints should be integer. We discretise interpolated TJC values into the same categories of None (≤ 0.5), Some ($0.5 <$ and ≤ 7.5), and Many ($7.5 <$ and ≤ 28). We discretise SJC records using three categories: None (=0), Some ($1 \leq$ and ≤ 5), and Many ($6 \leq$ and ≤ 28). Interpolated SJC values are decimals and we discretise them into the same categories: None (≤ 0.5), Some ($0.5 <$ and ≤ 5.5), and Many ($5.5 <$ and ≤ 28).

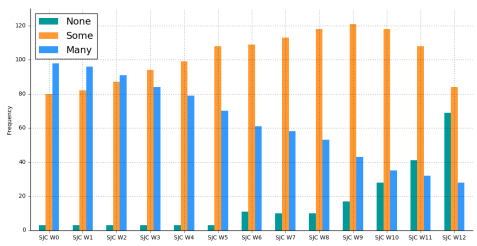
Pain, Fatigue, and GH records and interpolated values are between 0 and 100. We categorise them into three groups of Low (< 33.33), Medium ($33.33 \leq$ and < 66.66), and High ($66.66 \leq$). HAQ records and interpolation values are between 0 and 3, and we discretise them into three common categories: Mild ($0 \leq$ and ≤ 1), Moderate ($1 <$ and ≤ 2), and Severe ($2 <$ and ≤ 3).

CRP values are discretised into three categories of Normal (> 0 and ≤ 5), Moderate (> 5 and ≤ 30), and High (> 30). Similarly, ESR values are discretised into three categories of Normal (> 0 and ≤ 20), Moderate (> 20 and ≤ 50), and High (> 50).

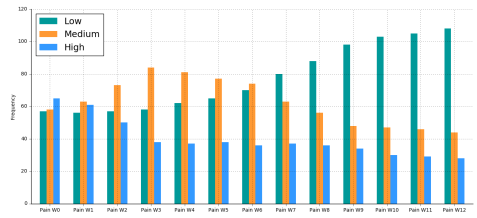
We calculate the DAS28 values for each week using interpolated TJC, SJC, ESR, and GH values. We call it 'Disease State' and discretise its values into three categories of Low (≤ 2.6), Moderate (> 2.6 and ≤ 5.1), and High (> 5.1).



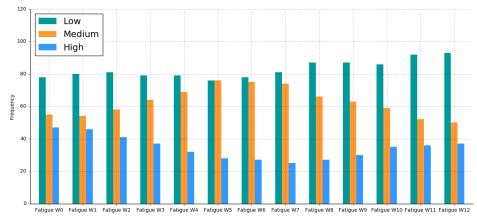
(a) TJC.



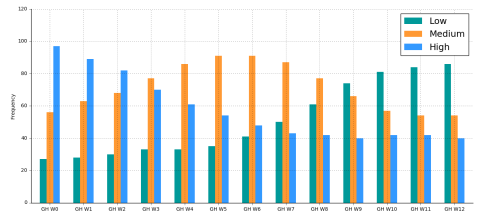
(b) SJC.



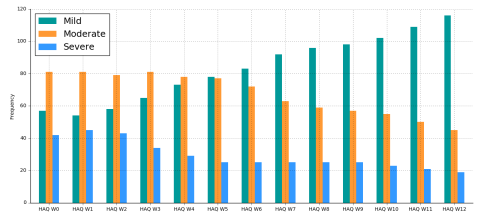
(c) Pain.



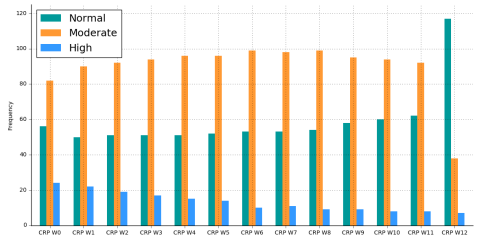
(d) Fatigue.



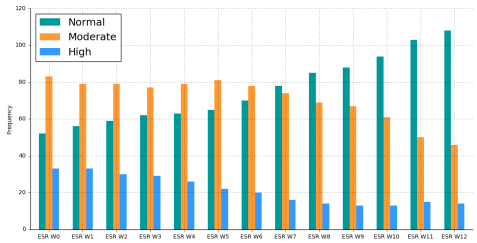
(e) GH.



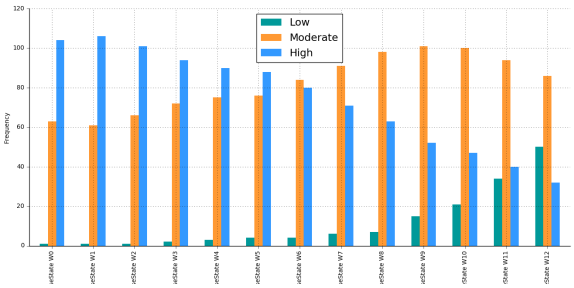
(f) HAQ.



(g) CRP.



(h) ESR.



(i) Disease State.

Fig. 4.19 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the baseline and first follow-up visit (3 months) and interpolation of 11 weeks between the baseline and first follow-up.

Figure 4.19 displays the bar plots of records of TJC, SJC, Pain, Fatigue, GH, HAQ, CRP, ESR, and 'Disease State' collected in the baseline and the first follow-up visit (3 months) and interpolated values of 11 weeks between the baseline and the first follow-up visit.

There is a sharp increase of the frequency of Normal CRP cases in Week 12, as shown in Figure 4.19g. It is because the threshold of normal CRP is 5 and many CRP records equal 5 in Week 12 or the first follow-up visit. Since 5 is the threshold, it is commonly used as a representative of 5 or any value less than 5 by the lab operators; therefore, a sharp increase appears in CRP records.

Bar plots of TJC, SJC, Pain, Fatigue, GH, HAQ, CRP, ESR, and 'Disease State' between the first and second follow-ups, between the second and third follow-ups, and between the third and fourth follow-ups are presented in Figures A.5, A.6, and A.7.

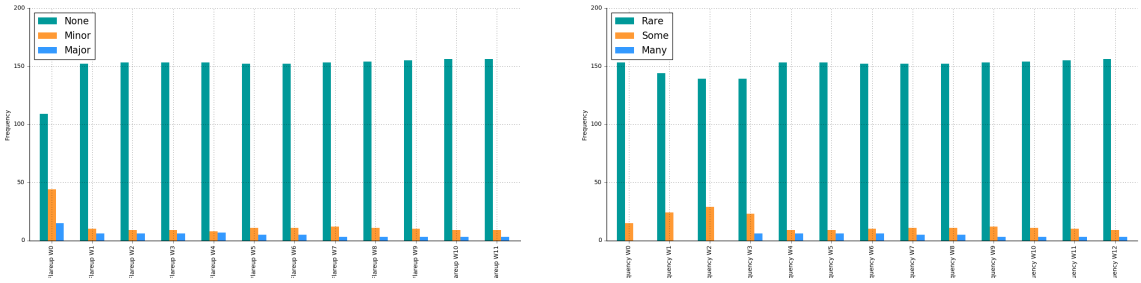
Using weekly 'Disease State' values, we define two new variables, which are missing in the PEAC dataset, called Flare and 'Flare Frequency'. Flare refers to any worsening of symptoms (explained in Subsection 4.2.2); therefore it is a relative concept measured by comparing the disease state in two consecutive times, i.e., weeks. Based on van der Maas et al.'s [213] suggestion, we define three states of Flare variable, namely, None, Minor, and Major, by comparing the disease state in a week with that of the previous week.

Major flare happens in three occasions: 1) disease state is less than 3.2 and the rise of disease state compared to the last week is at least 1.2, 2) disease state is equal to or greater than 3.2 and its rises is 0.6, and 3) disease state is equal or greater than 5.1 and increase 0.1 unit. A Minor flare is when: 1) disease state is less than 3.2 and grows between 0.1 and 1.2 unit, 2) disease state is equal to or greater than 3.2 and less than 5.1 and increases between 0.1 and 0.6, and 3) disease state is at least 5.1 and rises between 0 and 0.1. None state of flare is in case 1) disease state is less than 3.2 and changes at most 0.1 unit (can be a negative change, i.e. improvement of disease state), 2) disease state to be equal to or greater than 3.2 and less than 5.1 and changes at most 0.1 unit, and 3) disease state is equal or greater than 5.1 and does not increase at all, i.e. at most 0 unit of change.

An issue with our definition of flare is that in the baseline that we have no information about the previous week's disease state to compare with the baseline's. Only in the baseline (week 0), we consider a major flare to be any disease state greater than 7.3, a minor flare to be any disease state greater than 6.2 and less than or equal to 7.3, and a none state to be any disease state less than or equal to 6.2.

'Flare Frequency' in a week is the frequency of major or minor flares in the last four week. To compute the values of flare frequency, we define a weighted sum of minor flares (with the weight of 1) and the major flares (with a weight of 2). Then, we discretise the values of flare frequency into Rare (≤ 1), Some ($2 \leq$ and ≤ 5), and Many ($6 \leq$).

Figure 4.20 shows the bar plot of Flare and 'Flare Frequency' in the baseline, the first follow-up visit, and interpolated values between the two times. Similarly, the bar plots of Flare and 'Flare Frequency' between the first and second follow-ups, between the second and third follow-ups, and between the third and fourth follow-ups are presented in Figures A.8, A.9, and A.10.



(a) Flare. (b) Flare Frequency.

Fig. 4.20 Flare variables computed using disease state values of the baseline, first follow-up visit (3 months), and interpolation of 11 weeks between the baseline and first follow-up.

4.4.5 Description and Analysis of Treatment

DMARDs are a group of medications classified into three groups of conventional (csDMARDs), targeted synthetic (tsDMARDs), and biological (bDMARDs) that are commonly used to treat RA. Clinicians usually start the treatment of RA by prescribing csDMARDs and later they may prescribe ts/bDMARDs to eligible patients. In some cases, clinicians give a combination of two or three medications. There are three main csDMARDs: Methotrexate (MTX), Sulfasalazine (SSZ), and Hydroxychloroquine (HCQ). Other csDMARDs like Leflunomide (LEF) and Azathioprine (AZA) are rarely prescribed. Detailed description of csDMARDs is presented in Subsection 8.2.1.

If the disease stays active after prescribing two csDMARDs, clinicians may investigate to prescribe ts/bDMARDs. Of them, bDMARDs are either first-line or second-line, which the former ones are prescribed first and if failed, the latter ones would be prescribed. Etanercept (ETN) and Certolizumab (CTZ) are frequently prescribed first-line drugs and Rituximab (RTX) is widely given second-line drugs. ts/bDMARDs are further explained in Subsection 8.2.2.

Some patients are previously prescribed steroids by their GP to relieve the symptoms and inflammation indicators. Steroids are given in the form of tablets (also called oral; e.g., prednisolone or pred), intramuscular (IM) injections (including depomedrone or depo), and intra-articular (IA) injection. More explanation of steroids is given in Subsection 8.2.3.

csDMARDs of MTX, SSZ, HCQ, LEF, and AZA are prescribed in the baseline and follow-up visits of the PEAC study. The majority of patients receive MTX, SSZ, or HCQ, and they may receive them combined; therefore, we exclude LEF and AZA cases. Among the remaining csDMARDs, MTX is taken weekly, but SSZ and HCQ are taken daily.

ts/bDMARDs are prescribed as well, but they are given only after the third follow-up visit. In the first and second follow-up visits, some patients are referred to other specialists to be checked whether they could take ts/bDMARDs or not since these drugs are very toxic and can cause severe adverse reactions. In clinical practice, clinicians are supposed to inform patients about the effects and side-effects of ts/bDMARDs, and patients are entitled to accept or refuse. Clinicians may consider prescribing ts/bDMARDs but postpone it.

In the PEAC study, some patients are prescribed steroids to get their disease activity lower in a short period of time. These patients are those who experience flares or high disease activity. Table 4.4 summarises the frequency and percentage of persons who received csDMARDs, ts/bDMARDs, and Steroids in the baseline and follow-up visits.

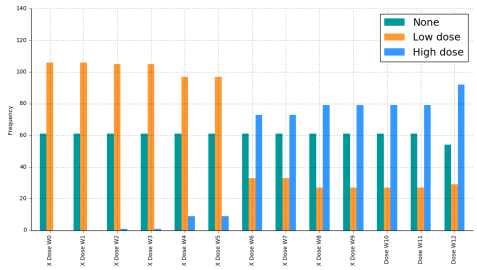
Table 4.4 Summary of treatment data in PEAC data at baseline, first, second, third, and fourth follow-ups.

Category	Medication	Baseline	1st follow-up	2nd follow-up	3rd follow-up	4th follow-up
csDMARDs	MTX	118(64%)	121(67%)	125(71%)	99(63%)	96(66%)
	SSZ	93(51%)	83(46%)	70(40%)	40(157%)	65(45%)
	HCQ	44(24%)	63(35%)	60(34%)	53(34%)	43(29%)
	LEF	0(0%)	0(0%)	2(1%)	2(1%)	3(2%)
	AZA	0(0%)	0(0%)	1(1%)	1(1%)	0(0%)
ts/bDMARDs	ETN	0(0%)	0(0%)	0(0%)	2(1%)	4(3%)
	CTZ	0(0%)	0(0%)	1(1%)	1(1%)	4(3%)
	RTX	0(0%)	0(0%)	0(0%)	2(1%)	1(1%)
	(Referred)	0(0%)	3(2%)	24(14%)	14(9%)	13(9%)
	(Refused)	0(0%)	0(0%)	1(1%)	1(1%)	1(1%)
Steroids	Pred	77(72%)	90(50%)	46(26%)	29(18%)	27(18%)
	IM depo	47(26%)	11(6%)	9(5%)	9(6%)	7(5%)
	IA	1(1%)	4(2%)	4(2%)	1(1%)	2(1%)

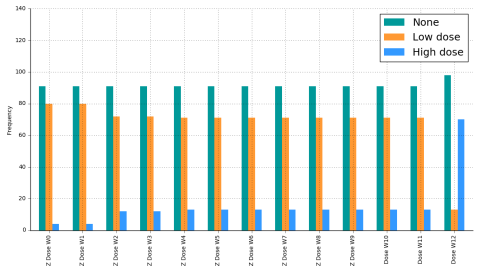
We expanded the PEAC data of csDMARDs of MTX, SSZ, and HCQ using their initial dose and escalation. We exclude LEF and AZA since they are recorded infrequently. We expand the medication data assuming them. This expansion transforms the quarterly collected medication data into weekly and matches the expanded csDMARDs data with the interpolated monitoring data explained in Subsection 4.4.4. Like interpolation, we assume a 3-month period to have 13 weeks. After expansion, we categorise the values of MTX into three states of None ($= 0mg$), 'Low dose' ($5mg \leq$ and $\leq 12.5mg$), and 'High dose' ($15mg \leq$

and $\leq 25mg$). Similarly, we categorise SSZ into three states of None ($= 0mg$), 'Low dose' ($500mg$ and $1000mg$), and 'High dose' ($1500mg$ and $2000mg$). HCQ has three states of None ($= 0mg$), 'Low dose' ($200mg$), and ($400mg$).

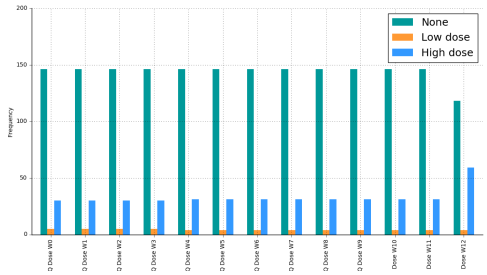
Figure 4.21 shows the frequency of discretised categories of expanded MTX, SSZ, and HCQ records into weekly between the baseline and the first follow-up visit. Similar plots of the discretised expanded MTX, SSZ, and HCQ between the first and second follow-ups, between the second and third follow-ups, and between the third and fourth follow-ups are displayed in Figures A.11, A.12, and A.13, respectively.



(a) MTX.



(b) SSZ.



(c) HCQ.

Fig. 4.21 Discretised expanded csDMARD values between the baseline and first follow-up visit with 11 weeks in between.

In the BioT study, all patients are on ts/bDMARDs, mainly ETN, TCZ, and CTZ, but other drugs including ABA, ADA, and GMB are also prescribed. As shown in Table 4.5, the frequency of medications is almost stable as time passes forward, except in four cases in Times 8 and 10. Those cases can potentially imply medication tapering; they may not be dose reduction since patients in this study are on a stable state with mild or no severe adverse reaction to medications. Note that medication records begin from Time 3, not 1, perhaps because they were on a training period without taking medication in the first two times.

Clinicians make a decision on medication in each visit. Their decision is mainly based on disease activity and adverse events of previous medications. Their goal is to put patients in remission or at least low disease activity with the minimum adverse reaction to medication.

Table 4.5 Summary of treatment data in BioT data in eleven times of data collection.

ts/bDMARDs	Time								
	3	4	5	6	7	8	9	10	11
ETN	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(21%)	6(21%)	6(23%)
TCZ	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(21%)	6(21%)	6(23%)
CTZ	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(20%)	6(21%)	6(21%)	5(19%)
ABA	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(12%)
ADA	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(12%)
GMB	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	2(7%)	2(7%)	0(0%)
Benpali	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(10%)	3(12%)

Thus, they may escalate the dosage of medication or combine multiple medications to achieve their goal. If the disease activity of a patient gradually diminishes towards remission, clinicians may decide to continue the same dosage. In case a patient achieves remission stably, clinicians may taper the dosage of their medication. However, if a patient experiences any severe adverse events such as organ damage, symptomatic side-effects, or infections, clinicians will change the medication immediately. Albeit, they tend to reduce the medication dosage in case of mild or moderate adverse events. Various adverse medication events (AMEs) of different medications are described in detail in Section 8.2.

The dosage of medicines in the PEAC dataset reveals clinicians' decisions on dose escalation, same regimen continuation, dose tapering, or reduction. For example, Figures 4.21a, 4.21a, and 4.21a respectively show the decision of dose escalation of MTX, SSZ, and

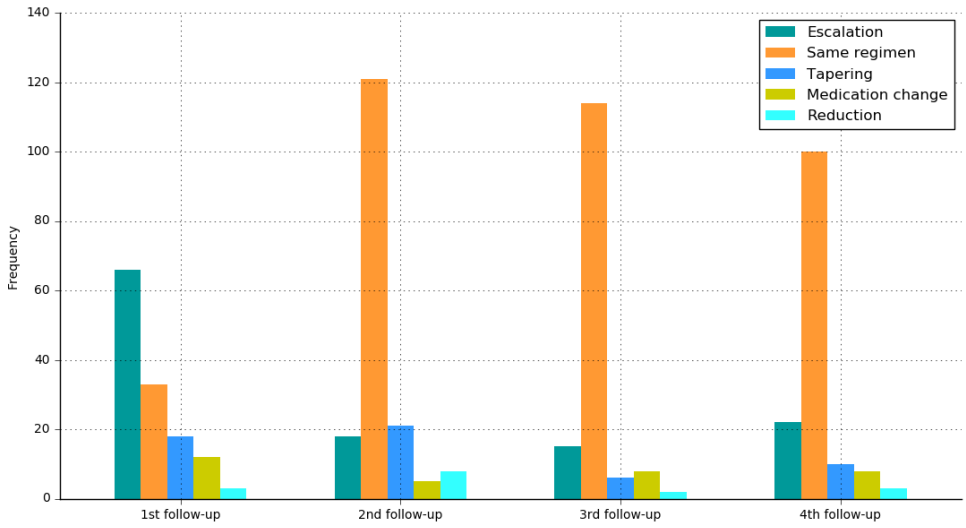


Fig. 4.22 Decisions on medication made by rheumatologists in baseline and each follow-up visit.

HCQ in Week 12 or the first follow-up. In the PEAC data, clinicians write notes in each follow-up visit about AMEs or any decision by patients (e.g., family planning or travelling abroad). These notes include AMEs that can reveal any decision on medication change. We go through the AMEs written in each visit and compared the medication dosage of the visit day and a week before to extract the decision of clinicians. We consider five categories of decisions as mentioned above: Escalation, 'Same regimen', Tapering, 'Medication change', and Reduction. Although tapering and reduction have equal meaning, dose tapering refers to decreasing dose for patients in remission stably, whereas dose reduction is the decreasing of medication dosage for those who experience a mild or moderate adverse event. Figure 4.22 demonstrates the medication decisions in the first, second, third, and fourth follow-up visits. Clearly, the Escalation decision is the most frequent decision in the first visit and the 'Same Regimen' becomes the main decision in the next visits. The Tapering and Reduction decisions slightly rise in the second visit, then diminish in the third visit and again rise in the fourth visit, whereas the 'Medication Change' decision falls in the second visit and then grow slightly by the fourth visit.

Despite PEAC data, BioT data consists of patients who are all on ts/bDMARDs, rather than csDMARDs. ts/bDMARDs are single dose drugs (except GMB), and we cannot extract clinicians' decisions on medication by comparing their dosage. However, we compare the medications taken in one time with those in the previous time. We also go through the clinicians and nurse actions to find any implication of their decision on medications. The

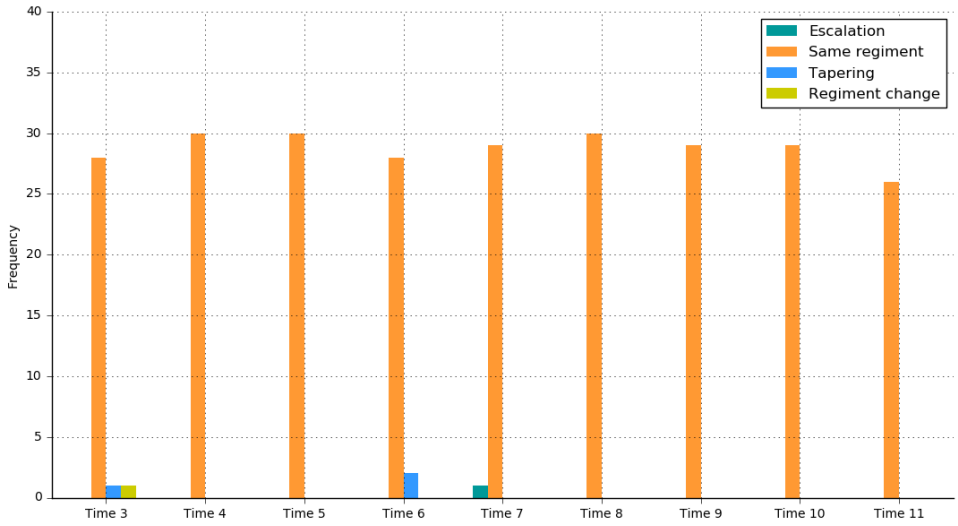


Fig. 4.23 Medication decisions in BioT study made in Time 3 to Time 11.

majority of decisions are the continuation of the same regimen, but there are rare tapering, escalation and regimen change cases as well.

As mentioned before, medication records in the BioT data begin from Time 3. Although we mainly extract medication decisions from comparing medication records, we decided to translate any record of a medication in Time 3 as a decision on continuation of the same regimen. This is firstly because we know that these patients are diagnosed since a long time before and they received treatment previously. Secondly, we came across a tapering decision and a regimen change decision in Time 3. It reveals the availability of medication information before Time 3 that leads clinicians to make those two decisions.

Figure 4.23 displays the frequency of medication decisions made in Time 3 to Time 11. Unlike multiple diverse medication decisions made in the PEAC data, there are rare decisions recorded or made in BioT data.

4.5 Selected Decision Support Points

The pathways for diagnosis, initial management, ongoing management, and personalised care for living with RA express the potential points of decision-making that a decision-maker may need to be supported. These points can be represented as decision points (diamond shapes) or activity nodes (rectangular shapes) that provide medical knowledge on the target of decision-making in any particular care process. We may acquire further knowledge from domain experts or volunteer patients, established literature, available clinical guidelines, or pre-existing ontologies as complementary knowledge. The complementary knowledge may be needed to annotate additional information or a revision of the model of care pathway, as we revised the original pathway and added the gray branch in the pathway for ongoing management of RA (Figure 4.7). We match the learnt knowledge with the available data to select some of the decision support points and build meaningful models and finally create a decision support tool.

To specify the selected decision support points for RA applications, we go through the care pathways (presented in Section 4.3) to find the decision points and match them with the data analysed (presented in Section 4.4). A BN model can be built for each of these decision points to support decision-making. In the model of pathway for diagnosis of RA (Figure 4.5), we find the 'RA Diagnosis?' node that is the target of the diagnosis process and matches with the diagnosis data of PEAC dataset (described in Subsection 4.4.2). Having the initial knowledge and required data, we build BN models for the diagnosis of RA as explained in Chapters 5 and 6.

We come across multiple points that may need decision support in the model of care pathway for initial management of RA (Figure 4.6), namely, 'Needs steroids?', 'Consider DMARDs', and 'Book Follow-up Appointment' nodes. We can find multiple other points in the model of care pathway for ongoing management of RA (Figure 4.7), mainly 'Consider Biologics', 'Change Medications?', 'New Appointment?', and 'Review Blood Test Frequency'. We try to match these points with the original PEAC data and their interpolation (described in Subsections 4.4.3 and 4.4.4) and expanded medication data of the PEAC study and medication data of the BioT study (described in Subsection 4.4.5).

We select the 'Book Follow-up Appointment' node from the initial management pathway and 'New Appointment?' node of the ongoing management pathway to build a basic DBN model for self-management of RA, as presented in Chapter 7. The proposed model for self-management excludes any treatment variables and allows us to do initial reasoning in the observational level. Later, we select 'Change Medication?' node of the ongoing management pathway which matches with the trajectories of medication decisions in the PEAC and BioT datasets (shown in Figures 4.22 and 4.23). We extend DBN model for self-management by adding variables of medication and adverse events and further enrich the model to capture the latent variables on medication tolerance and disease control leading to a final advice on medication review which enables reasoning in the interventional level, described in Chapter 8.

By going through the pathway for personalised care (Figure 4.8), we identify a node called 'Personalised Advice Given to Patient' that is the target of the pathway and potentially needs to receive decision support. We select it as an initial source of knowledge and later

Table 4.6 Selected decision support points in pathways for diagnosis, initial management, ongoing management, and personalised care of RA.

Pathway	Decision Support Point	Description of Required Decision Support
Diagnosis	RA Diagnosis?	Differentiating RA cases from other inflammatory arthritis cases considering risk factors, signs, symptoms, serology results, and comorbidities
Initial Management	Book Follow-up Appointment	Arrangement of appointments frequency based on disease activity and prescribed medications
Ongoing Management	New Appointment?	A new appointment is needed or patient can self-manage.
	Change Meds?	Any changes of medication type or dosage considering ongoing disease activity, signs, symptoms, serology results, and medications
	Review Blood Test Frequency	Prediction of next needed blood test and appointment considering disease activity and prescriptions
Personalised Care	Personalised Advice Given to patient	A set of personalised advice to target physical, psychological, and social aspects of QoL for improving independence, empowerment, and participation components of patient's QoL

focus on the established literature and elicit experts' judgement as a replacement of data, since there is no available data on personalised care for RA. We propose a BN model built from expert's knowledge and published information in established literature as explained in Chapter 9.

Table 4.6 lists the selected decision support points that are extracted from the models of care pathway for diagnosis, initial management, and ongoing management that are matched with available or manipulated data. It also includes the selected decision support point for personalised care to live with RA.

4.6 Summary

In this chapter, we introduced the PAMBAYESIAN project and RA disease, which is a case study of the project. We explained the diagnosis, management and treatment of RA, and QoL of RA patients. We described models of care pathways for diagnosis, initial management, and ongoing management of RA. We also outlined a model of care pathway for living with RA that helps to manage QoL of RA patients. We then described and analyse the available data collected in two studies called PEAC and BioT. We first introduced these two studies and then described their data consisting of 1) personal information, risk factors, comorbidities, and 2) signs, symptoms, serology results, disease activity. These data help us to build BN models for RA diagnosis (Objective 1) explained in Chapter 5, and DBN models for self-management of RA (Objective 3) presented in Chapters 7 and 8. To evaluate the DBN models for self-management of RA, we need frequent data that is not available; therefore, we interpolated the PEAC data to increase its frequency. We explained the interpolation of signs, symptoms, serology results, and disease activity. Next, we describe and analyse the treatment data recorded in both PEAC and BioT studies, and we investigate clinicians' decisions on medications. We will use the interpolated data and treatment data later in Chapters 7 and 8 to evaluate the DBN models for self-management of RA (Objective 3).

Finally, we investigated the care pathways and used the analysed data to select decision support points among activity nodes and decision points of the models of care pathway. These selected decision support points direct us to build BN models in Chapters 5, 6, 7, 8, and 9.

Chapter 5

Building Bayesian Network Models for Diagnosis of Rheumatoid Arthritis

In Chapter 4, we introduced six selected decision support points that Bayesian networks (BNs) can be applied to (Section 4.5). In the current chapter, we address the first selected decision support point, 'RA Diagnosis?'. We show how we build a BN model for diagnosing RA using data and knowledge elicited from experts, literature, and clinical guidelines, corresponding to the Objective 1. We show how we specify the variables related to the diagnosis of RA, and how we determine their states and parameters. The rest of this chapter is organised as follows: Section 5.1 introduces the diagnosis of RA and a BN model for RA diagnosis. Section 5.2 describes the variables for diagnosis of RA, namely, (1) personal information, risk factors, and comorbidities, (2) disease manifestations, (3) pathogenesis, (4) intervention, and (5) diagnosis. The structure of BN model is presented in Section 5.3.2, followed by the parameterisation of BN Section 5.3.3. Section 5.4 explains the results and compares the performance of the BN model with a model learnt entirely from data. We evaluate the BN model with scenarios of dummy patient data designed by our rheumatology experts. We investigate the performance of BN model by reviewing the inaccurate cases and we then explain one RA case and reasoning with absent variables. Finally, Section 5.5 summaries the chapter.

5.1 Introduction

Machine learning methods have been widely used to make prediction of future states and provide decision support. Various methods, including naive Bayes, logistic regression, and rule-based systems, have been developed for medical applications, as we reviewed in

Section 2.6.1. Although these methods provide robust outcomes, they lack to provide a meaningful structure and explain the prediction of the model. Therefore, their usefulness as decision support is questioned as clinicians tend to use methods that are transparent and easily explainable.

BNs are a popular method for decision support in medicine [1, 110] perhaps because they can be built using a combination of data and knowledge elicited from domain experts or literature [150, 54, 31, 190]. The use of knowledge is an advantage when there is a lack of sufficient good-quality data, but essential when the application requires a causally coherent model. BNs are widely used in medical diagnosis because they can express medical knowledge and uncertainty [190] and deal with incomplete data [219].

RA is a chronic inflammatory disease causing swollen and painful joints. Since diagnostic delays may occur due to inadequate evidences at the onset of IA [77], rheumatologists have proposed diagnostic criteria, such as the American College of Rheumatology (ACR) 1987 [9] and its revision jointly with the European League Against Rheumatism (EULAR)[6]-known as the 2010 criteria. As introduced in Subsection 4.2.1, the 2010 criteria have four domains: joint involvement, serology results, acute-phase reactants, and duration of symptoms. Joint involvement refers to the count of swollen and tender joints including large and small joints and considering the symmetrical or asymmetrical pattern of the involved joints [6]. Serology results entail two antibodies involved in the development of RA: rheumatoid factor (RhF) and cyclic citrullinated peptide (CCP). Acute phase reactants consider any abnormality of two blood markers of erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP). Duration of symptoms is a patient's self-reported duration of signs or symptoms of synovitis and is major if equals or exceeds 6 weeks [6]. The 2010 criteria are mainly used by our experts at Barts Rheumatology Department for diagnosing RA; however they consider the recommendations of the National Institute for Health and Care Excellence (NICE) for RA diagnosis including X-ray scanning of hand and feet joints [138], as mentioned in Subsection 4.2.1. Additionally, clinicians at Barts Rheumatology Department use ultra-sound scanning of hand and feet joints since it can be helpful to diagnose patients with negative antibody CCP, but inflamed joints (see Section 4.3.1). Although the 2010 criteria for diagnosing RA have improved the early diagnosis compared to the 1987 criteria, there is still a need to improve the accuracy of RA diagnosis [77].

We collaborate with our rheumatology experts to create a model of care pathway for the diagnosis of RA, as shown in Figure 4.5. This pathway represents the patient history, clinical examination, ultrasound scan, and blood test results as activity nodes of the diagnosis process that lead to make a decision on the diagnosis of RA. We also analyse the PEAC data on risk factors, comorbidities, signs, symptoms, and a basic intervention in Subsections

4.4.2 and 4.4.3. These data match with the activities and decision represented in the pathway for diagnosis, except for the ultrasound scan that is not available for all cases of the PEAC study. The matching data and activities lead us to the selected decision support point of 'RA Diagnosis?' of the diagnosis pathway, as we introduce in Section 4.5.

In this chapter, we describe a BN model for differentiating RA from other forms of IA using a combination of experts' knowledge and data. The knowledge is acquired from the pathway for diagnosis of RA and close collaboration with our main expert. We obtain the required data from the PEAC dataset and for a small number of exceptions, we consulted both experts and relevant medical literature. Using the knowledge and data, we specify the variables to build the structure of BN model and parameterise the variables. The BN model enables us to differentiate RA cases from other IA cases by giving observations to the risk factors, comorbidities, signs, symptoms, and serology results.

5.2 Description of Variables for Diagnosis of Rheumatoid Arthritis

We interview the experts to create a model of care pathway for diagnosis of RA, as shown in Figure 4.5. The pathway consists of activity nodes and decision points represented by rectangular and diamond shapes, respectively. Considering the concepts mentioned in the pathway and the available dataset, we identified the categories of variables needed in the BN model for the diagnosis of RA, namely 1) personal information, risk factors, and comorbidity variables; 2) disease manifestations; 3) pathogenesis mechanism; 4) diagnosis; and 5) intervention. The variables of these categories are described in the following parts.

5.2.1 Personal Information, Risk Factors, and Comorbidities

Once a referral is received, the rheumatologist should collect the patient history consisting of their personal, risk, and comorbidity variables. Personal variables include age and sex. The records of age are continuous in the PEAC dataset. The main expert defines two thresholds for discretising age data into three groups of Young ($45 \leq$), 'Middle Age' ($45 <$ and $60 \leq$), and Older ($60 <$). We use the binary data of sex (Female and Male), as recorded in the original dataset. To combine personal variables, we define a synthetic variable called 'Personal Risks' with three states of Low, Medium, and High. The parameterisation of this variable is explained in Subsection 5.3.3.

Literature review on hormonal pathogenesis of RA indicates the effect of pregnancy or postpartum [200] and early menopause (before 45 years old) [163]. We add two variables of

'Early Menopause' and 'Pregnancy or Postpartum', as personal variables specific to female patients. 'Early Menopause' has two states of No and Yes. 'Pregnancy or Postpartum' has four states of Pregnancy, '0-3 months postpartum', '4-12 months postpartum', and Outside, i.e., outside pregnancy or 12 months postpartum.

We consider three demographic factors that may associate with the diagnosis of RA: ethnicity, occupation, and education. Exposure to industrial dusts is believed to be a risk factor of RA [38]. So, we classify occupations into three states of Industrial, Non-industrial, and Unemployed. Ethnicity has an association with RA diagnosis [162]. Taken from NHS's ethnicity categories [128], we defined four states for Ethnicity variable: 'Asian or Asian British', 'Black or Black British', 'Mixed and Others', and White. We also consider education factor with two states of School and University. To combine the demographic factors, we define a synthetic variable called 'Demographic Risks' with three states of Low, Medium, and High. Details of parameterisation of this variable is presented in Subsection 5.3.3.

Moderate consumption of alcohol can be a preventive factor against RA; however, none or high consumption of alcohol pose risk of RA. Alcohol consumption is measured with a unit called 'alcohol unit', as explained in Subsection 4.4.2. We consider four states of None, Low (< 2.625), Medium ($2.625 \leq$ and < 26.25), and High ($26.25 \leq$). Smoking is a major environmental risk factor playing a role in development of RA [223, 151, 92]. Smoking is commonly measured with the pack-year, as explained in Subsection 4.4.2. We define three states for smoking: None ($= 0$), Low (< 10 and > 0), and High ($10 \leq$). Body mass index (BMI) has association with RA as elaborated in [58]. We consider the well-known categories of BMI: Underweight (< 18.5), Normal ($18.5 \leq$ and < 25), Overweight ($25 \leq$ and < 30), and Obese ($30 \leq$). We create a synthetic variable called 'Lifestyle Risks' for combining lifestyle variables of BMI, Alcohol, and Smoking. It has three states of Low, Medium, and High, and its parameters are defined by a set of rules given by our experts.

Relevant medical background risks consist of family history of RA, represented with 'RA FHx', and family history of other IA, represented with 'Other IA FHx'. Each variable has four states of None, 'RA Threat', 'Other IA Threat', and 'RA and Other IA Threat'. We combine these two variables using a synthetic variable called 'IA FHx' with the same four states as the combined variables. We also create a synthetic variable called 'Comorbidity Background' with the same four states, again. Its parameters are specified by our experts indicating whether a patient has an autoimmune comorbidity. For simplification, we define another synthetic variable called 'Medical Background Risks' to aggregate the two synthetic variables of 'IA FHx' and 'Comorbidity Background'. It has the same four states: None, 'RA Threat', 'Other IA Threat', and 'RA and Other IA Threat'.

Table 5.1 Summary of personal information, risk factors, and comorbidity variables for diagnosis of RA

Category	Variable Name	States	Source	Type	Description
Personal	Age	Young($45 \leq$), Middle Age($45 <$ & $60 \leq$), Older($60 <$)	Data	Evidence	Person's age
	Sex	Female, Male	Data	Evidence	Person's sex
	Early Menopause	No, Yes	Literature	Absent	Women patients who experienced early menopause
	Pregnancy or Postpartum	Pregnancy, 0-3 months postpartum, 4-12 months postpartum, Outside	Literature	Absent	Women individuals being pregnant or in postpartum period
	Personal Risks	Low, Medium, High	Expert	Synthetic	Combination of age and sex
Demographic	Ethnicity	Asian or Asian British, Black or Black British, Mixed and Others, White	Data	Evidence	Person's ethnicity
	Education	School, University	Data	Evidence	Person's education level
	Occupation	Industrial, Non-industrial, Unemployed	Literature	Evidence	Person's occupation
	Demographic Risks	Low, Medium, High	Expert	Synthetic	Combination of ethnicity, education, and occupation using rules provided by experts
Lifestyle	Alcohol	None($= 0$), Low($0 <$ & < 2.625), Medium($2.625 \leq$ & < 26.25), High($26.25 \leq$)	Experts	Evidence	Duration of disease since being diagnosed
	Smoking	None($= 0$), Low($0 <$ & < 10), High($10 \leq$)	Experts	Evidence	Person experiencing a flare
	BMI	Underweight(< 18), Normal($18 \leq$ & $25 <$), Overweight($25 \leq$ & $30 <$), Obese($30 \leq$)	Experts	Evidence	Current disease activity of the chronic condition appeared in short-term
	Lifestyle Risks	Low, Medium, High	Expert	Synthetic	Combination of lifestyle risks of alcohol, smoking, and BMI using rules provided by experts
Medical background	RA FHx	No, Yes	Data	Evidence	Family history of RA
	Other IA FHx	No, Yes	Data	Evidence	Family history of other IA, i.e. PsA, monoarthritis, and UA
	IA FHx	None, RA Threat, Other IA Threat, RA and Other IA Threat	Experts	Synthetic	Family history of IA, or family history of RA and other IA
	Comorbidity background	None, RA Threat, Other IA Threat, RA and Other IA Threat	Experts	Synthetic	Threat of comorbidity background on RA or other IA or both
	Medical Background Risks	None, RA Threat, Other IA Threat, RA and Other IA Threat	Experts	Synthetic	Threat of comorbidity background on RA or other IA or both
Comorbidities	OA	Yes, No	Data	Evidence	Person having OA
	TAD	Yes, No	Data	Evidence	Person having TAD
	Skin Psoriasis	Yes, No	Data	Evidence	Person having skin psoriasis
	Other IA	Yes, No	Data	Evidence	Person having Other IA psoriasis
	CTD	Yes, No	Data	Evidence	Person having CTD psoriasis
	CA	Yes, No	Data	Evidence	Person having CA psoriasis
	Alternative Explanation of Symptoms	Yes, No	Expert	Synthetic	Alternative explanation of symptoms due to having OA, skin psoriasis, other IA, CTD, and CA
	Alternative Disease Effect on CRP and ESR	Yes, No	Expert	Synthetic	Alternative disease effect on CRP and ESR due to having CTD and CA

Relevant comorbidities are either autoimmune ones or those with similar symptoms and serology results as RA. Any autoimmune comorbidity tends to increase the risk of developing another autoimmune disease [209]; therefore, we consider thyroid autoimmune disease (TAD) and skin psoriasis as the relevant autoimmune comorbidities. We wished to include further autoimmune comorbidities such as type I diabetes, but ignored them since none of them were present in the final version of the PEAC dataset. The two variables of TAD and 'Skin Psoriasis' have two states of Yes and No. The two variables of TAD and 'Skin Psoriasis' combine into the aforementioned 'Comorbidity Background'.

We consider variables for other comorbidities which are involved in causing similar symptoms and serology results as RA. OA (osteoarthritis), 'Other IA' (including RA, psoriatic arthritis (PsA), spondyloarthropathy (SA), and undifferentiated arthritis (UA), CTD (connective tissue disease), and CA (crystal arthropathy). We define a synthetic variable called 'Alternative Explanation of Symptoms' to combine OA, 'Skin Psoriasis', CTD, and CA. This variable is associated with all symptom variables which are described in Subsection 5.2.2. To represent the effect of CTD and CA on the serology results of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), we add a synthetic variable called 'Alternative Disease Effect on CRP and ESR'. ESR and CRP are both described in Subsection 5.2.2. Above-mentioned synthetic variables have two states of Yes and No.

Table 5.1 summarises the description of personal information, risk factors, and comorbidity variables.

5.2.2 Disease Manifestation

From the care pathway node of 'Perform Clinical Examination' (Figure 4.5), we derived disease manifestations, namely, signs, symptoms, and serology results. Two main signs of IA diseases are the count of tender joints (TJC) and the count of swollen joints (SJC) among 28 joints, as explained in Subsection 4.4.3. We define a variable called TJC with five states of None ($= 0$), Few ($1-5$) Some ($6-10$), Many ($11-16$), and 'Too many' ($17-28$). Similarly, we define a variable called SJC with five states of None ($= 0$), Few ($1-3$) Some ($4-7$), Many ($8-12$), and 'Too many' ($13-28$). Another sign is the symmetrical pattern of swollen joints that is known to be observed in RA patients, not other IA diseases. We consider a variable called SSJ (symmetrical swollen joints) with two states of Yes and No.

The most important symptom in diagnosing of RA is morning stiffness for more than 30 minutes. We consider a variable called Stiffness with binary states of Yes and No. Other symptoms are general malaise, measured by global health (GH) scores, pain, and fatigue, which patients are asked to indicate or provide a number between 0 and 100, where 0 indicates the best and 100 the worst. We define three variables called GH, Pain, and Fatigue with

Low (≤ 33.33), Medium ($33.33 <$ and ≤ 66.66), and High ($66.66 <$) states. An additional symptom measured in clinical practice is the health assessment questionnaire (HAQ). It estimates patient's disability and mobility [76] and scales patients in a range between 0 and 3. We discretise HAQ into four categories of None ($= 0$), Mild ($0 <$ and ≤ 1), Moderate ($1 <$ and ≤ 2), and Severe ($2 <$).

Table 5.2 Summary of disease manifestation variables associated with RA

Category	Variable Name	States	Source	Type	Description
Signs	TJC	None($= 0$), Few($1-5$), Some($6-10$), Many($11-16$), Too many($17-28$)	Data	Evidence	Count of tender joints
	SJC	None($= 0$), Few($1-3$), Some($4-7$), Many($8-12$), Too many($13-28$)	Data	Evidence	Count of swollen joints
	SSJ	No, Yes	Data	Evidence	Person having symmetrical pattern of swollen joints
Symptoms	Stiffness	No, Yes	Data	Evidence	Person experiencing morning stiffness for more than 30 minutes
	GH	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Evidence	Person's self-assessed global health represented in a range between 0 (best) and 100 (worst)
	Fatigue	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Evidence	Person feeling fatigue represented in a range between 0 (best) and 100 (worst)
	Pain	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Evidence	Person feeling pain represented in a range between 0 (best) and 100 (worst)
	HAQ	None ($= 0$), Mild ($0 <$ & ≤ 1), Moderate ($1 <$ & ≤ 2), Severe ($2 <$)	Data	Evidence	Person's disability measurement quantified between 0 and 3 representing the lowest and the highest disability respectively
Serology	ESR	Normal (< 20), Moderate($20 \leq$ and ≤ 35), High(> 35 and ≤ 50), Very high(> 50 and ≤ 75), Extreme(> 75)	Data	Evidence	ESR measured in a blood test
	CRP	Normal (< 5), Moderate($5 \leq$ and ≤ 15), High(> 15 and ≤ 30), Very high(> 30 and ≤ 50), Extreme(> 50)	Data	Evidence	CRP measured in a blood test

Serology results include measurements of two antibodies called cyclic citrullinated peptide (CCP) and rheumatoid factor (RhF). CCP and RhF involve in pathogenesis mechanism that is described in detail in Subsection 5.2.3. We use three common categories to discretise CCP data: Negative (≤ 10), 'Low positive' (> 10 and ≤ 30), and 'High positive' (> 30). Similarly, we categorise RhF data into: Negative (≤ 20), 'Low positive' (> 20 and ≤ 60), and 'High positive' (> 60). In clinical examination, two inflammation markers are also measured:

ESR and CRP. We consider medically meaningful thresholds to categorise ESR and CRP data provided by our main expert. This leads to five states for ESR: Normal (< 20), Moderate ($20 \leq$ and ≤ 35), High (> 35 and ≤ 50), 'Very high' (> 50 and ≤ 75), and Extreme (> 75), and CRP is categorised into: Normal (< 5), Moderate ($5 \leq$ and ≤ 15), High (> 15 and ≤ 30), 'Very high'(> 30 and ≤ 50), and Extreme (> 50). The summary of disease manifestation variables is presented in Table 5.2.

The pathway node of 'Ultrasound (if available)' represents the examination of joints by ultrasound scan (Figure 4.5). It enables the rheumatologists to have a clearer image of the joints, but it may not be available in all clinics; therefore we skip it.

After collecting risk factors, comorbidities, and manifestations of disease (including all blood test results), rheumatologists may suspect of RA. Finally, they make a decision on diagnosing a person with RA or another IA.

5.2.3 Pathogenesis

Pathogenesis mechanism of RA is not present in the diagnosis pathway, but our experts describe it as a complex and multifactorial interplay of genetic and environmental factors, known as the gene-environment pathogenesis mechanism. Our model considers genetics and its impact on the development of pathogenic antibodies of CCP and RhF through a serological pathogenesis jointly with smoking as a lifestyle factor. CCP is associated with HLA-DRB1 gene (known as the "shared epitope") and smoking [160] and RhF is associated with HLA-DRB1 gene, PTPN22 gene, and smoking [151, 14]. Two known antibodies of CCP and RhF are measured by blood test and are implicated in RA pathogenesis through a gene-environment interaction. Serostatus is a synthetic variable that we consider to combine CCP and RhF, and our experts define a set of rules to quantify it (see Subsection 5.3.3).

Table 5.3 Summary of pathogenesis variables associated with RA development

Category	Variable Name	States	Source	Type	Description
Serological pathogenesis	CCP	Negative (≤ 10), Low positive (> 10 and ≤ 30), High positive (> 30)	Data	Evidence	Antibody CCP measured in a blood test
	RhF	Negative (≤ 20), Low positive (> 20 and ≤ 60), High positive (> 60)	Data	Evidence	Antibody RhF measured in a blood test
	Serostatus	Negative, Low positive, High positive	Expert	Synthetic	Serostatus combining RhF and CCP
Genetic pathogenesis	HLA-DRB1 Gene	None, Heterozygotes, Homozygotes	Literature	Absent	Patient's HLA-DRB1 gene
	PTPN22 Gene	None, Heterozygotes, Homozygotes	Literature	Absent	Patient's PTPN22 gene
Hormonal pathogenesis	FSH Fluctuation	None, Low, High, Inapplicable	Expert	Latent	Effect of fluctuation of female sex hormones on development of RA

We showed the higher percentage of females being diagnosed with RA that is revealed in the PEAC data (Table 4.3) and we touched on the females' personal information in Subsection 5.2.1. It is believed that the fluctuation of female sex hormones (FSH) play a role in developing RA, mainly during postpartum [200] and early menopause [163], which have been observed to be classical times for the onset of RA in women [8].

Our model includes a latent variable called 'FSH Fluctuation' to represent the hormonal pathogenesis of RA. It is associated with two personal variables: 'Early Menopause' and 'Pregnancy or Postpartum'. It has four states of None, Low, High, and Inapplicable. The last state excludes male patients and any other impossible states (e.g., a menopaused female being pregnant) from the hormonal pathogenesis process. There is no available data for 'FSH Fluctuation' in the datasets; therefore we use our experts' judgement to parameterise it, as explained in Subsection 5.3.3. The summary of pathogenesis variables is presented in Table 5.3.

5.2.4 Intervention

Intervention refers to a possible prescription of steroids by a a practitioner or a doctor in primary care to alleviate symptoms and reduce inflammation, as described in Section 4.4.5. It can reduce the severity of disease manifestation including serology results (ESR and CRP) and symptoms (stiffness, GH, pain, fatigue, and HAQ). We consider a variable called Steroid with binary states of Yes and No. Table 5.4 shows a summary of the intervention variable.

Table 5.4 Summary of intervention variable associated with RA development

Category	Variable Name	States	Source	Type	Description
Intervention	Steroid	No, Yes	Data	Evidence	Prescribed steroid by practitioners prior to diagnosis

5.2.5 Diagnosis

Diagnosis is the target variable and has two states of RA and 'Other IA'. RA refers to anyone diagnosed with RA in the baseline and other IA encompasses those diagnosed with psoriatic arthritis (PsA), monoarthritis (MonoA), undifferentiated arthritis (UA), and other sorts of IA, as recorded in the PEAC dataset (see Subsection 4.4.1).

Table 5.5 Summary of diagnosis variable

Category	Variable Name	States	Source	Type	Description
Diagnosis	Diagnosis	RA, Other IA	Data	-	Prediction of diagnosis being RA or other IA

5.3 Bayesian Network Model for Diagnosis of Rheumatoid Arthritis

5.3.1 Data

We use a subset of the PEAC dataset with 373 cases that are diagnosed with RA (226 cases), UA (79 cases), PsA (49 cases), MonoA (12 cases), and others (7 cases). The dataset consists of 30 variables covering patients' personal information, lifestyle demographic, medical background, comorbidities, disease manifestation, diagnosis, and a basic intervention. Some variables of the dataset have outlier values, beyond possible ranges, and some of them contain missing values. The extreme outliers were found in the age and smoking data and were removed. Inaccurate records, e.g., '<11' or '<5' in RhF or CCP data, were replaced with random values, as described in Subsection 4.4.3. The data of some variables are discretised using medically meaningful thresholds proposed by our rheumatology experts. After these processes, we removed 13 cases which had four or more missing values out of 30. In the 360 remaining cases, only 4% of data were missing. The diagnoses of the remaining cases were RA (215 cases), UA (77 cases), PsA (49 cases), monoarthritis (12 cases), and others (7 cases). Since our goal was to classify RA from other arthritic diseases, we grouped all conditions into RA and 'other IA', which contains UA, PsA, monoarthritis, and others.

5.3.2 Structure

We use the extracted variables to create the structure of BN model for diagnosis of RA, as displayed in Figure 5.1 We used the AgenaRisk software to build the model [3]. Ovals represent the variables that we described in Section 5.2. There are thirteen areas in the model called fragments containing ovals that are representing variables. Fragments indicate the category of each variable as were mentioned in Tables 5.1, 5.2, 5.3, 5.4, and 5.5.

The Diagnosis variable is shown with a white solid oval in the middle of the model. The variables within the Demographic, Personal, Lifestyle, and 'Medical Background' fragments associate with diagnosis of RA corresponding with the risk factor idiom (Figure 2.10c) that we introduced in Section 2.5.1. The comorbidities (except TAD) that can cause similar symptoms and increase serology results fit the common symptomology idiom (Figure 2.10e). TAD and 'Skin Psoriasis' are two autoimmune diseases that can increase the risk of development of another autoimmune disease such as RA; thus they behave like risk factors, as described in Subsection 5.2.1. The variables in the hormonal, serological, and genetic pathogenesis mechanisms bridge the risk factors and diagnosis of disease that makes them similar to the structure of the pathogenesis idiom shown in Figure 2.10d. The variables within the Signs,

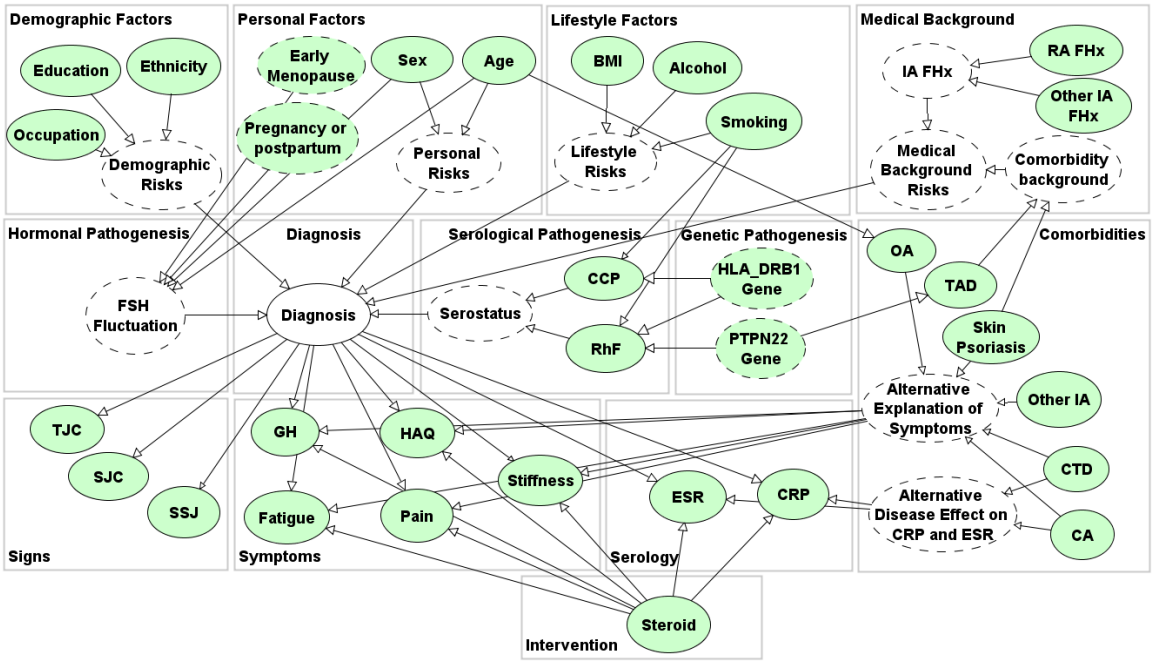


Fig. 5.1 BN model for RA diagnosis built from experts' knowledge.

Symptoms, and Serology fragments manifest the outcomes of the disease and correspond to the manifestation idiom (Figure 2.10a). Steroid is an intervention in an initial stage and is similar to the treatment idiom displayed in Figure 2.10g.

Ovals coloured in green express the evidence variables, i.e., the variables that an observation can be given to. Dashed ovals are either synthetic variables that we defined to simplify the model, and latent or absent variables that are not recorded in the datasets and we added them to enrich the model, despite their unavailability in the PEAC dataset. There are nine synthetic variables in the model. Five of them, namely, 'Demographic Risks', 'Personal Risks', 'Lifestyle Risks', 'Medical Background Risks', and Serostatus, combine their parent variables and reduce the chances of overfitting of the model by reducing the number of incoming dependencies into their child variable, Diagnosis variable. In the 'Medical Background Risks' fragment, the 'IA FHx' and 'Comorbidity Background' variables combine their parent variables and collectively influence their children variables. Similarly, two synthetic variables in the Comorbidities fragment, 'Alternative Explanation of Symptoms' and 'Alternative Explanation of CRP and ESR', combine their parent variables.

These variables do not exist in the dataset, however, our experts defined their values by providing a set of rules. To parameterise the latent variables of the hormonal and genetic factors, experts defined rules based on the established literature.

5.3.3 Parameterisation

Solid green ovals in Figure 5.1 imply the evidence variables that we have available data to parameterise. For this, we use a dataset with discrete values, as described in Subsection 5.3.1. There are four dashed green ovals in the model representing absent variables, which there is no available data for them. We exploit the information published in [224], [200], [205], and [14] to parameterise the node probability tables (NPTs) of 'Early Menopause', 'Pregnancy or Postpartum', 'HLA-DRB1 Gene', and 'PTPN22 Gene', respectively.

Synthetic variables and a latent variable are depicted in white dashed ovals in Figure 5.1. To parameterise these variables, we ask our experts to specify a set of rules for each variable. We prepared the personal variables with initial dependencies between them in the form of risk factor idioms, and added multiple rules, such as: <*IF Age is Young and Sex is Male, THEN Personal Risks is Low*>. We presented them to the experts and asked them to approve the rules; otherwise reject them and provide modified ones. We informed the experts that we will find a consensus between their rules and that we take the senior expert's rules into account in case of any conflict between the rules given by three experts. Figures B.1, B.2, B.3, B.4, B.5, and B.6 demonstrate some shots of the rules approved or modified by the main expert.

After translating the rules into parameters of NPTs, we come across a high percentage of missing values (e.g. 36% for Serostatus and 45% for 'Personal Risks'). To reduce the percentage of missing values of the synthetic variables, we suggested a set of new rules and asked our main expert to approve them or otherwise reject them and modify them, if possible. Our idea was to improve the performance of BN by dealing with missing values more effectively, including reducing their number in the synthetic variables. We also wanted to improve the dependencies of hormonal factors and update the rules for 'FSH Fluctuation'. Figure B.7, B.8, B.9, and B.10 show the shots of the secondary rules given by our main expert. Secondary rules reduced the missing values significantly (e.g. 4% for Serostatus and 1% for 'Personal Risks'); however, that prevented the expectation-maximisation (EM) algorithm from learning the missing values properly and led to poor performance of the model. Therefore, we ignored the secondary rules to protect the performance of the model, although we kept only the update of hormonal factors and secondary rules for 'FSH Fluctuation'.

We train the NPTs of the rest of the model, mainly the Diagnosis variable, using the available discretised dataset. We apply the EM algorithm (described in Subsection 2.5.2) in the AgenaRisk software. The synthetic variables have missing values that can be learnt by incorporating knowledge using Equation 2.7 with the relative weight of expert knowledge and available data obtained by Equation 2.8. The value of r is given as the percentage of

knowledge incorporation to the AgenaRisk. We use 100 iterations and 0.001 threshold for the convergence of EM.

5.4 Results and Evaluation

We exploited the discretised dataset to learn the structure of a BN as an alternative to the causal BN model built from experts' knowledge. We use the structural EM (SEM) algorithm in the bnlearn package that is able to deal with missing data [186]. SEM uses the Hill-Climbing method for scoring candidate models and the Maximum Likelihood for parameter learning. Details about SEM is available in Subsection 2.5.1.

We obtain a learnt BN model as displayed in Figure 5.2. The model contains 61 dependencies between 30 nodes. It is entirely learnt from the data and there is no direct intervention by experts in designing the model. However, we made many processes (e.g., cleaning and discretising) on the data along with the main expert to prepare the dataset for structure and/or parameter learning. The model has neither synthetic variables, nor latent and synthetic variables, since it is purely built from data.

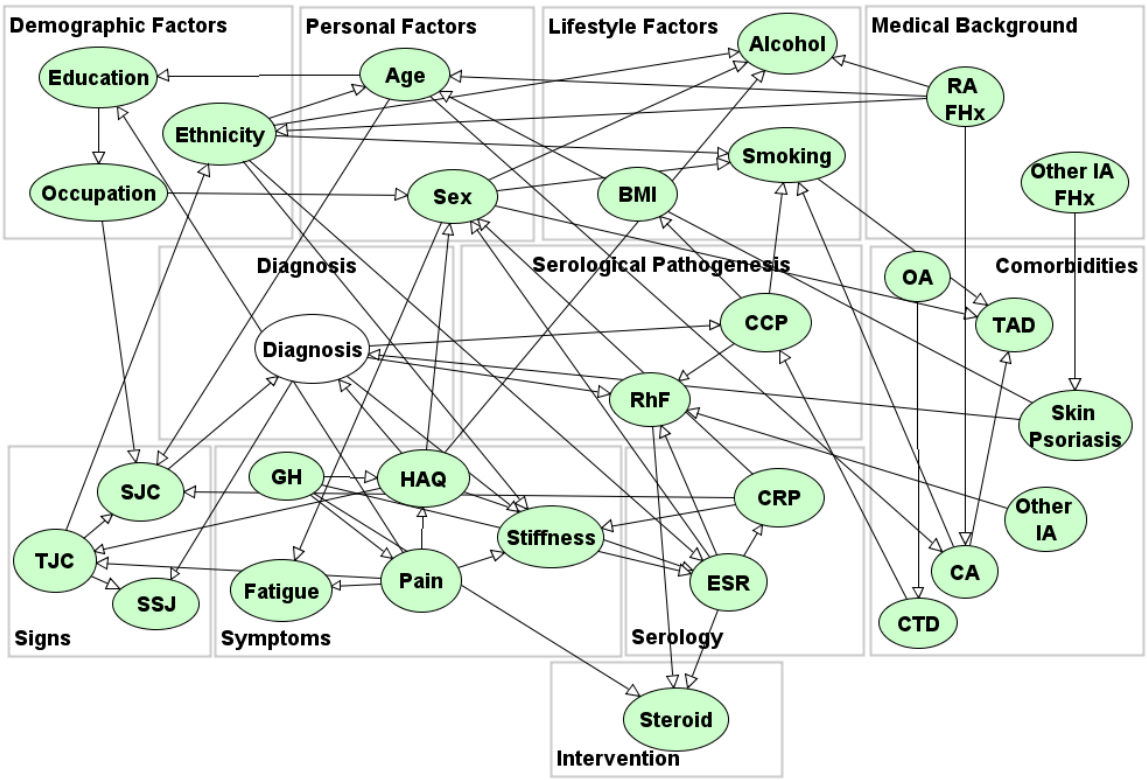


Fig. 5.2 Learnt BN model for RA diagnosis built using SEM algorithm.

In this model, it is medically meaningful that the Diagnosis variable is linked to the SSJ and Stiffness variables. It also makes medical sense that Diagnosis variable is connected in the reverse direction to CCP, RhF, SJC, HAQ, and 'Skin Psoriasis' which is a component of 'Comorbidity Background' of the knowledge-based model. The reverse connections from ESR and GH to Steroid and from CCP to Smoking represent a medically meaningful association between those variables.

5.4.1 Cross-Validation

We used the results of a 10-fold cross-validation to compare the performance of the models in terms of their discrimination and accuracy. Discrimination measures if the model is able to differentiate between two states of the diagnosis, i.e., RA and 'Other IA', and the accuracy investigates whether the predicted outcomes are actually close to the recorded outcomes. For discrimination comparison, we compare the receiver operating characteristics (ROC) curves of each model, sensitivity, and specificity values. The area under the ROC curve (AUROC) of the knowledge-based BN is 0.84 with 95% confidence intervals (CI) (0.79-0.88), but the AUROC of the learnt BN is 0.71 with 95% CI (0.66-0.77). Figure 5.3 displays the ROC curves of the knowledge-based BN model (black line) and the data-driven model (grey line). At 90% sensitivity, the knowledge-based BN shows 63% specificity and the learnt BN results 22% specificity.

We use the Brier score (BS) and the Brier skill score (BSS) to evaluate the accuracy of the models. BS is the measurement of the mean squared difference between the predicted results and the real ones. Its value can be between 0 and 1, where 0 refers to the perfect prediction and 1 is the worst outcome. BSS is the improvement of the prediction relative to a reference prediction, which is usually the average probability of the prediction event recorded in the real data. Its values can be between minus infinity and 1, where 1 indicates a perfect prediction and lesser or negative values are worse outcomes. The BS of the knowledge-based BN model is 0.17, but that of the learnt model is 0.23. The BSS of the knowledge-based model is 0.30, whereas that of the learnt model is 0.04. The evaluation of two models in terms of discrimination and accuracy is summarised in Table 5.6.

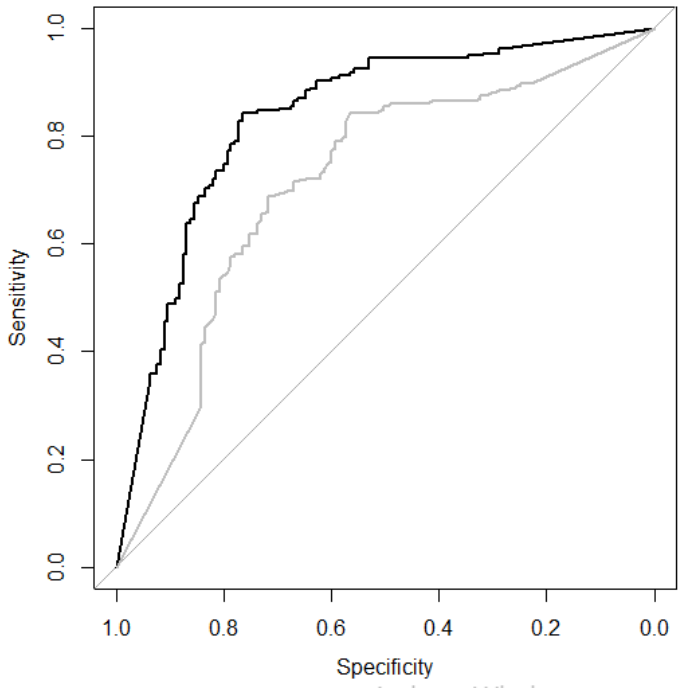


Fig. 5.3 ROC curves of knowledge-based BN (black) and BN model learnt with SEM (grey).

Table 5.6 Comparison of diagnosis models performance

	Knowledge-based BN	Learnt BN
AUROC	0.84	0.71
Specificity (at 90% sensitivity)	0.63	0.22
Sensitivity (at 80% sensitivity)	0.77	0.57
BS	0.17	0.23
BSS	0.30	0.04

We evaluate the performance of the models with a range of thresholds from 0.0 to 0.9. Four metrics of performance, including accuracy, precision, recall, and F1-score, are shown Figure 5.4 that elaborate the performance of the knowledge-based model. Similarly, Figure 5.5 display the performance evaluation of the learnt model with the four metrics of accuracy, precision, recall, and F1-score.

As an example, the confusion matrices of the knowledge-based and learnt models for 0.6 or 60% threshold are shown in Tables 5.7 and 5.8, respectively. Considering the proportion of 215 cases of RA out of 360 records, we assumed the threshold probability of diagnosis to separate predicted RA cases from 'Other IA' to be 60%. We can see a better prediction accuracy of RA compared to 'Other IA' in both models. The accuracy of RA and 'Other IA'

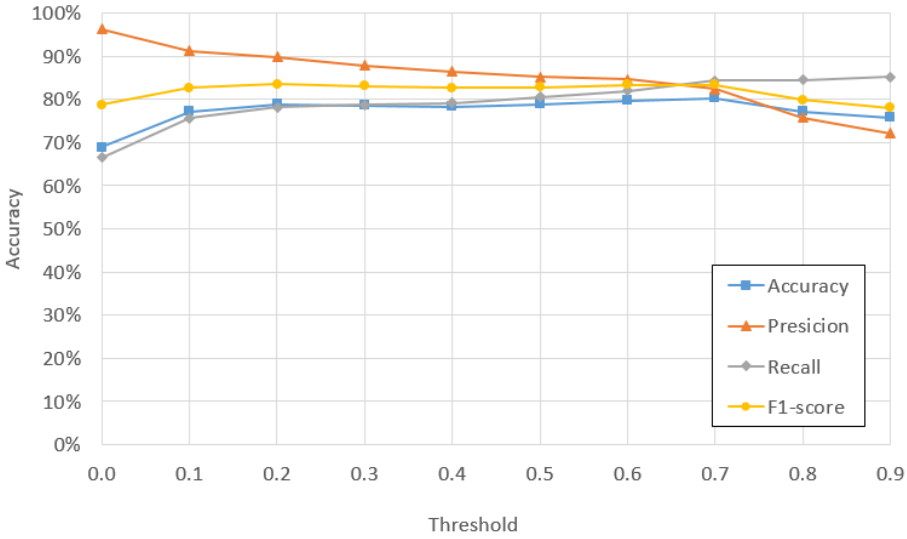


Fig. 5.4 Performance evaluation of the knowledge-based BN model with a range of thresholds from 0.0 to 0.9.

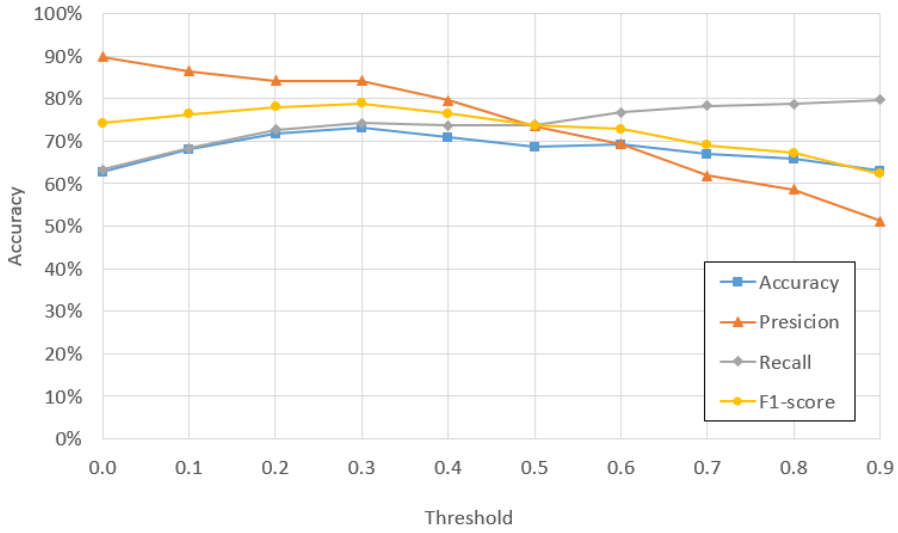


Fig. 5.5 Performance evaluation of the learnt BN model with a range of thresholds from 0.0 to 0.9.

using knowledge-based BN model are 85% and 72%, respectively. Learnt BN model achieves 69% accuracy in classifying both RA and 'Other IA' categories. The overall accuracy of the confusion matrices of the knowledge-based and learnt BN models is 80% and 69%, respectively.

Table 5.7 Confusion matrix of knowledge-based BN model.

Diagnosis	RA	Other IA
$60\% \leq$ Prediction	182	40
Prediction $< 60\%$	33	105

Table 5.8 Confusion matrix of learnt BN model.

Diagnosis	RA	Other IA
$60\% \leq$ Prediction	149	45
Prediction $< 60\%$	66	100

Using another threshold of 0.7 or 70% alters the differentiation and creates two new confusion matrices shown in Tables 5.9 and 5.10. The accuracy of RA and 'Other IA' using knowledge-based BN model is 82% and 77%, respectively, whereas the accuracy of RA and 'Other IA' using the learnt BN model is 62% and 74%, respectively. The overall accuracy of the confusion matrices of the knowledge-based and learnt BN models is 80% and 67%, respectively. These figures reveal that changing the threshold to 70% does not improve the overall accuracy of either model's results, although the alternative threshold increases the accuracy of the 'Other IA' and decreases that of the RA in the results of both models.

Table 5.9 Alternative confusion matrix of knowledge-based BN model.

Diagnosis	RA	Other IA
$70\% \leq$ Prediction	177	33
Prediction $< 70\%$	38	112

Table 5.10 Alternative confusion matrix of learnt BN model.

Diagnosis	RA	Other IA
$70\% \leq$ Prediction	133	37
Prediction $< 70\%$	82	108

The accuracy of 'Other IA' in both models is relatively lower. We believe that the lower accuracy may result from the relative lack of known risk factors of the 'Other IA' conditions. PsA, one of the other IA conditions, has only two major risk factors in our models: 'Other IA FHx' and 'Skin Psoriasis' comorbidity. For another condition, MonoA, 'Other IA FHx'

is the only risk factor. UA, the other main IA condition, refers to those arthritic cases that cannot be differentiated; therefore UA cases have no specific risk factors involved in the models. We further compared the accuracy of classifying other IA diseases separately. Table 5.11 indicates the number of correctly classified cases of UA (out of 77), PsA (out of 49), MonoA (out of 12), and others (out of 7) using both models.

Table 5.11 Prediction accuracy of other IA separately

Other IA conditions	Knowledge-based BN	Learnt BN
UA	51/77(66%)	57/77(74%)
PsA	39/49(80%)	32/49(65%)
MonoA	11/12(92%)	8/12(67%)
Others	4/7(57%)	3/7(43%)

To gain deeper insight into the predicted outcomes and the real diagnosis, we divide the prediction probabilities into five bins and investigate the number of the accurate cases and their percentages. Table 5.12 shows the accurate cases and their percentages in each bin for both knowledge-based and data-driven models.

Table 5.12 Accuracy of RA prediction in 5 bins

RA prediction bins	Knowledge-based BN		Learnt BN	
	Cases	Accuracy	Cases	Accuracy
$0.8 \leq P \leq 1.0$	193/360	163/193(84%)	160/360	126/160(79%)
$0.6 \leq P < 0.8$	29/360	19/29(66%)	34/360	23/34(68%)
$0.4 \leq P < 0.6$	13/360	9/13(69%)	38/360	16/38(42%)
$0.2 \leq P < 0.4$	12/360	5/12(63%)	17/360	7/17(41%)
$0.0 \leq P < 0.2$	113/360	91/113(81%)	111/360	77/111(69%)

The discrimination and accuracy results show that the overall performance of the knowledge-based BN model is better than that of the learnt model. Although the knowledge-based model is sparser than the learnt model, both of them are susceptible to overfitting. The Diagnosis variable in the knowledge-based model and two variables of 'RA FHx' and Pain in the learnt model have at least four parent variables. These variables can potentially lead the corresponding models to overfit.

5.4.2 Evaluation with Scenarios of Dummy Patients

Our rheumatology experts designed 20 scenarios of dummy patients with RA and other IA diseases. We use these scenarios to evaluate the performance of the BN model qualitatively. For designing the scenarios, we prepared a table of observations (including risk factors, signs, and symptoms) and another table for experts' expected diagnosis considering the observations. The observations and expected outcome of dummy patient 1 is presented in Table B.1.

We enter the observation of each dummy patient in the BN models and compared the results of each model with the experts' expectation, as displayed in Table 5.13. We applied 60% threshold to determine the diagnosis category from predicted probabilities of the two models, since RA patients comprise 60% of PEAC dataset (explained in Subsection 5.4.1). However, we used 50% threshold to categorise experts' expected diagnosis because their expectations were not based on any specific population or a proportion of RA patients in that population. The comparison between expectations and predictions using knowledge-based BN and learnt BN indicates 12 (60%) and 10 (50%) matching scenarios, respectively.

Table 5.13 Scenario-based evaluation of knowledge-based BN and learnt BN.

Scenario	Expectation	Prediction	
		Knowledge-based BN	Learnt BN
1	RA	Other IA	RA
2	RA	RA	RA
3	Other IA	Other IA	Other IA
4	RA	Other IA	Other IA
5	RA	RA	Other IA
6	RA	Other IA	Other IA
7	RA	RA	Other IA
8	RA	Other IA	Other IA
9	Other IA	Other IA	Other IA
10	RA	RA	RA
11	Other IA	Other IA	Other IA
12	RA	Other IA	Other IA
13	RA	RA	Other IA
14	RA	Other IA	Other IA
15	RA	RA	Other IA
16	RA	Other IA	Other IA
17	Other IA	RA	Other IA
18	RA	RA	RA
19	Other IA	Other IA	Other IA
20	RA	RA	RA

5.4.3 Review of Inaccurate Cases

We investigated the inaccurate cases of the knowledge-based model. There are 40 inaccurate RA predictions and 33 inaccurate predictions of 'Other IA', considering the 60% threshold of differentiation. Of 40 inaccurate RA cases, 26 are diagnosed with UA, and 6 and 4 cases (all UAs) have no recorded CCP and RhF, respectively. This is 15% and 10% missing values, whereas CCP and RhF have respectively 12% and 5% missing values in the dataset (out of 360). Considering the role of CCP and RhF in diagnosing RA, their absence increases the chance of inaccurate prediction. Less cases with inaccurate prediction of 'Other IA' have missing values in CCP and RhF, i.e., 4 and 2, respectively. This is expected since other IA cases are not associated with CCP and RhF.

We also noticed that 30% of the accurately predicted cases were prescribed a steroid, whereas a slightly higher percentage (33%) of the inaccurately predicted ones were prescribed steroid. This can reveal the Steroid variable's role in causing a potential overfitting in the NPTs of its child variables such as Fatigue and Pain.

5.4.4 Explanation of One Case of Rheumatoid Arthritis

Explaining a BN model by doing reasoning increases the trustworthiness of the model [100] and can lead to achieve a useful decision support. To explain the model, we pick one RA case to enter its records as observations into the BN model and do reasoning.

The recorded diagnosis of this specific case is RA. By giving the available evidences of risk factors, comorbidities, and manifestations, we achieve an 96% probability of getting a diagnosis with RA using the knowledge-based model, whereas the data-driven model obtains 41%. The first posterior probability of diagnosis is greater than the overall ratio of RA cases (60%) and therefore the knowledge-based model correctly predicts. On the other hand, 41% is less than the ratio and the learnt model fails to predict correctly.

We consider the diagnosis probability while removing the evidences one-by-one to check if variables support the diagnosis or not. In the knowledge-based model, all variables update diagnosis probability favourably, except BMI, Smoking, OA, TAD, CCP, RhF, ESR, Pain, SJC and TJC. Posterior probability of Diagnosis is insensitive to the observation given to 'Other IA', CTD, CA, and Steroid.

In the data-driven model, Ethnicity, Smoking, RA FHx, TAD, CA, ESR, SJC, TJC, and Steroid do not support Diagnosis. Diagnosis is insensitive to any observation to OA and Other IA variables. However, any observations to other variables of the learnt model are in favour of diagnosing RA.

CCP, RhF, ESR, Pain, SJC and TJC from the knowledge-based model and ESR, SJC, and TJC from the learnt BN unexpectedly do not support Diagnosis, perhaps because this patient has mild signs, symptoms, and serology results and has no antibodies - also called seronegative.

5.4.5 Reasoning with Absent Variables

We selected five critical cases that were diagnosed with UA or RA, as shown in Table 5.14. We can see a substantial difference between the posterior probabilities of RA, before and after giving observation to the absent variables. We gave Yes to 'Early Menopause' in Cases 1 and 5. We observed pregnancy for Cases 2 and 4, and we gave '0-3 months postpartum' observation for Case 3 in the 'Pregnancy or postpartum' variable. These observations update the probability of diagnosis in favour of RA. This proves the positive effect of having access to women's factors such as menopause and pregnancy in diagnosing RA.

Table 5.14 Reasoning with knowledge-based BN model on absent variables: 'Early Menopause' and 'Pregnancy or Postpartum'.

Case	Diagnosis	Posterior Probability of Diagnosis: RA	
		Absent Variable Unobserved	Absent Variable Observed
1	UA	61%	78%
2	RA	56%	60%
3	UA	48%	55%
4	UA	46%	65%
5	UA	63%	74%

We analyse the sensitivity of the BN model by a Tornado graph, given Diagnosis variable as the outcome of the model and the evidence variables and the two absent variables (Early Menopause' and 'Pregnancy or Postpartum') as the inputs. This graph shows the effect of each input variable in updating the probabilities of the outcome variable, Diagnosis. Figure 5.6 displays the Tornado graph for RA state of Diagnosis variable. The baseline is on 0.6 and we can see the effect of inputs in order. As expected, TJC, SSJ, SJC, CCP, and RhF are all on top of the graph. 'Pregnancy or postpartum' and 'Early Menopause' are at the bottom of the diagram; however, they demonstrate a higher effect on RA diagnosis than other routinely collected variables such as Alcohol and various comorbidities. This establishes the importance of collecting women's factors in clinical practice, especially because they comprise a higher percentage of women being diagnosed with RA, which is also the case in the PEAC study (see Table 4.1).

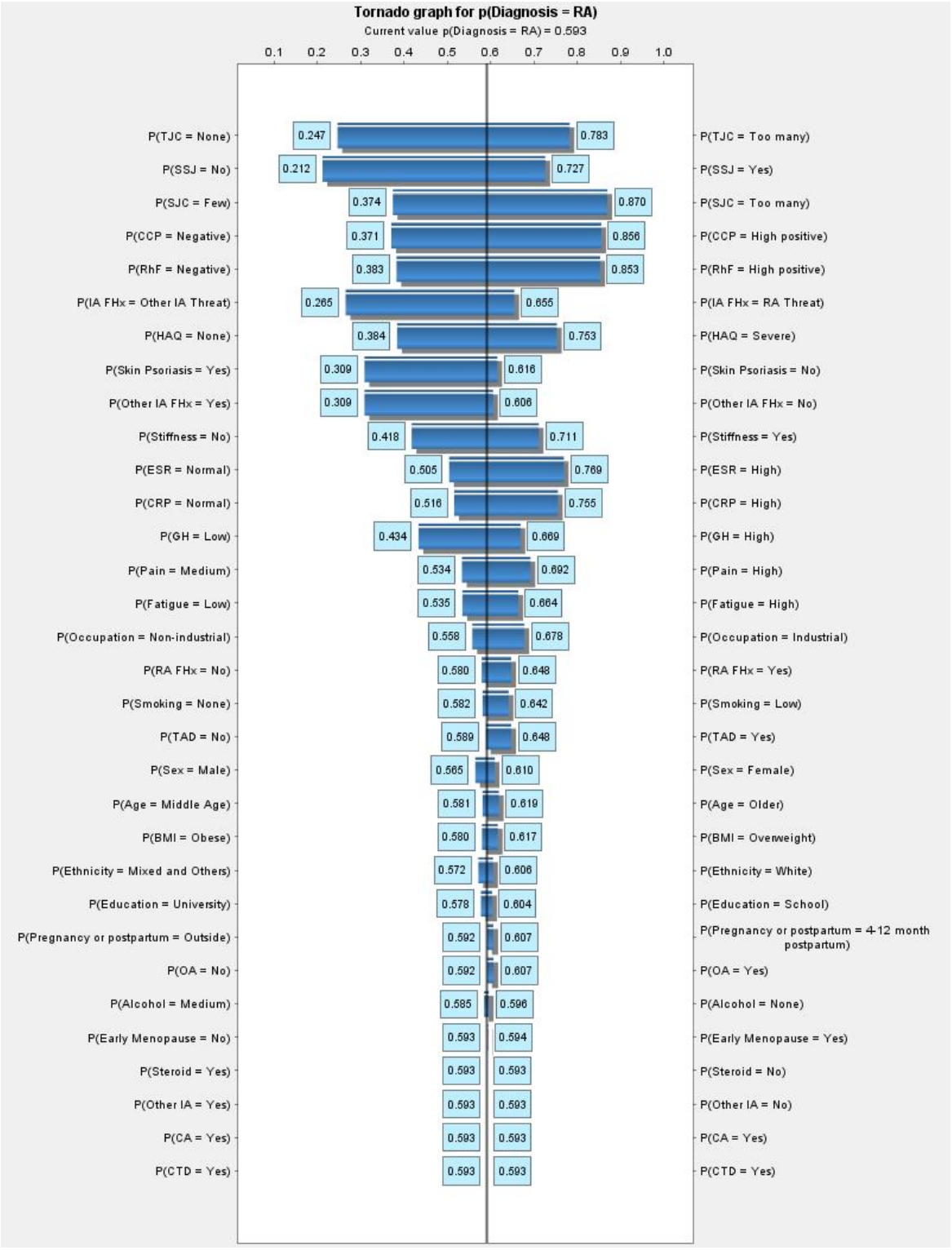


Fig. 5.6 Tornado graph of RA state of Diagnosis variable of BN model for diagnosis of RA with inputs of evidence variables and two absent variables of 'Early Menopause' and 'Pregnancy or postpartum'.

5.5 Summary

In this chapter, we proposed a BN model for RA diagnosis corresponding to the Objective 1. We call it the knowledge-based BN model since it is built mainly based on experts' knowledge. We show how we employed the available data, elicited knowledge from experts, used a care pathway, and learnt from the literature to develop the BN model. The model covers a wide range of variables including personal information, risk factors, pathogenesis mechanisms, relevant comorbidities, signs, symptoms, serology results to differentiate the RA cases from other IA ones. The pathogenesis mechanisms include a gene-environment and a hormonal process. The gene-environment pathogenesis considers the joint effect of smoking and two genes on the development of RA, and the hormonal process focuses on the role of female sex hormones on developing RA.

In this chapter, data has played an important role in the model development. Consequently, a machine learning approach is a viable alternative to the combined knowledge and learning approach that we have used. To investigate this, the performance of the knowledge-based model is compared with that of a model learnt entirely from data called 'learnt model'. The knowledge-based model outperforms the learnt model in multiple comparisons. We also use a set of dummy patient data to contrast the performance of the two models, which indicated more accurate results being achieved by the knowledge-based model. We have not compared the learnt BN with other machine learning methods that could be applied to our dataset. Suitable methods include both a logistic regression and a random forest, since both assign probabilities to outcomes in the same manner as a BN. Although it is possible that the performance might improve with a different ML technique, we know of no reason that suggests a substantial improvement is likely. It is more likely that until a much larger dataset allows the use of methods such as deep neural networks, the performance would be similar. We believe that the advantage of the BN lies more in its relative transparency to review.

To create the diagnostic BN models, we deal with a limited number of cases in the dataset to train the models which can potentially lead to a risk of overfitting. We also had to evaluate the models quantitatively with the same limited data. Although our experts were dedicated and closely collaborated with us, we acknowledge the fact that it could be fairer to have access to more dedicated experts. Another limitation was about qualitative evaluation of the model. We used a set of scenarios that were designed by our clinical experts, whom were also involved in creating the structure of the model. Thus, it could be fairer to validate the model with a group of independent experts.

Chapter 6

Building Medical Bayesian Networks from Care Pathways and Knowledge Graphs

In Chapter 5, we showed how we built a Bayesian network (BN) model - called 'knowledge-based' model - for the diagnosis of Rheumatoid Arthritis (RA). The name of this BN model refers to the medical knowledge elicited from experts, learnt from clinical guidelines and established literature. This BN model addressed the selected decision support point of 'RA Diagnosis?' (described in Section 4.5). In this chapter, we propose two step-by-step methods to create the structure of BN models for medical diagnosis using the knowledge represented in a model of care pathway and by creating knowledge graphs from the knowledge stored in pre-existing medical ontologies (Objective 2). This can help us to (partially) automate the process of building BN models that can be helpful in situations that BN modellers might not always have access to a dedicated domain expert.

In Section 6.1, we introduce the care pathways, knowledge graphs, and medical ontologies, followed by an analysis of care pathways and an explanation on creating knowledge graphs from medical ontologies in Section 6.2. In Section 6.3, we explains how to create knowledge graphs using the knowledge retrieved from the SNOMED CT and UMLS, and then we propose two methods for each ontology's knowledge to be translated into a structure of a BN model using two sets of translation rules. Section 6.4 is devoted to a discussion about the proposed methods for a case study of RA diagnosis and the resulting BN models.

6.1 Introduction

Clinical care pathways or carempas express the activities, decisions, and their sequence to manage a medical condition. These pathways are graphical representations of the sequence of patient care activities required in managing a particular condition [39, 118, 177, 178], and this can be an effective representation of knowledge that facilitates knowledge transfer [118, 117]. Care pathways have been designed in different shapes; therefore, some researchers (e.g. [119]) have tried to standardise the structure of care pathways. As presented in Section 4.3, we have developed four models of pathways - including a model of pathway for diagnosis of RA - using [119]'s standardisation.

Ontologies are a formal and explicit description of a set of concepts in a domain of knowledge. They can provide a common language to share knowledge in a domain including vocabulary, classification, taxonomy, relations (hierarchies and constraints), and domain axioms [68]. See Section 3.1 for further details about ontologies. Knowledge graphs are the same as ontologies with a difference that they acquire and integrate the information of ontologies [52, 13]. As described in Subsection 3.5.1, knowledge graphs are hierarchical networks consisting of concepts or classes, their properties, attributes, and instances. The connections between concepts of a knowledge graph are semantic relations such as is-a that represents the relation between subclass and superclass concepts.

BNs are directed acyclic graphs with a set of random variables and dependencies between them [60]. The nodes of the graph correspond to the variables of BNs and the arcs correspond to the conditional dependencies, as described in Section 2.1. BNs are commonly made from experts' knowledge and data (Section 2.5). BN models are classically built from experts' knowledge that can be time-consuming and costly. In some cases, the BN modellers may not have access to domain experts and therefore they cannot easily create a model from experts' knowledge. Here, we show how to use the activities and decision nodes of a care pathway and the concepts of a knowledge graph to build the structure of a BN model. The care pathway provides initial knowledge about the clinical process and contains medically validated keywords related to it. We propose to use these keywords to retrieve further knowledge from pre-existing medical ontologies, namely, SNOMED CT and UMLS (described in Section 3.4). We then translate the retrieved knowledge, including medical concepts and their properties, into the variables and dependencies of a BN model. Our aim is to partially automate the process of building the structure of BNs.

6.2 Analysing Care Pathways and Creating Knowledge Graphs from Medical Ontologies

This section provides a generic explanation of how we build knowledge graphs starting from care pathways and then using medical ontologies.

Care pathways help medical and decision experts to create a common understanding of the care process. We need to select which node(s) of the care pathway is targeted to get support. This node corresponds to the selected decision support points that we presented in Section 4.5.

By searching the name of the selected node(s), we retrieve all properties and related concepts of the selected node(s) from the ontologies, i.e., SNOMED CT and UMLS. The properties include semantic tags, semantic groups, semantic types, related concepts, and the relationships between the searched concepts and the retrieved ones. These properties may differ from one ontology to another. The medical terms of other nodes of pathway are also searched and their properties are retrieved.

We obtain the required properties and concepts to build a knowledge graph by removing synonym concepts and refining other concepts and their properties. We take the hierarchy of concepts from the original ontology, i.e., the semantic tags in SNOMED CT and the semantic types in UMLS. In this step, we can create the structure of knowledge graph based on the hierarchy of concepts, and bind them by their relationships or attributes.

Based on the associational or causal nature of these relationships or attributes, we categorise the relationships into two groups: 'from-to' and 'to-from'. For instance, consider the following concepts and their relationships in the form of triples:

- Concept1 *is risk factor of* Concept2
- Concept1 *is outcome of* Concept2

The first one represents an association from Concept1 to Concept2; however, the second example shows an association from Concept2 to Concept1.

These classifications of relationships are practically the translation rules as they convert the relationships of knowledge graphs into associational or causal connections. For instance, above concepts and relationships get translated as follows:

- Concept1 *is risk factor of* Concept2 means Concept1 *is associated to* Concept2
- Concept1 *is outcome of* Concept2 means Concept2 *is associated to* Concept1

The *Is associated to* refers to the dependency between concepts with a direction that leads us to create small fragments of the BN model - also called idioms (Subsection 2.5.1) - and finally build the entire BN structure with associational or causal relationships between its variables.

6.3 Case Study: Diagnosis of Rheumatoid Arthritis

First, we create a model of care pathway for the diagnosis of RA. We have already presented a model of care pathway for RA diagnosis in Figure 4.5 that was drawn based on rheumatology experts' knowledge and existing guidelines in the UK, as described in Section 4.3.1. Here, we simplify that model of care pathway for RA diagnosis to make it easier for this case study. Figure 6.1 shows a simplified version of the care pathway for the diagnosis of RA.

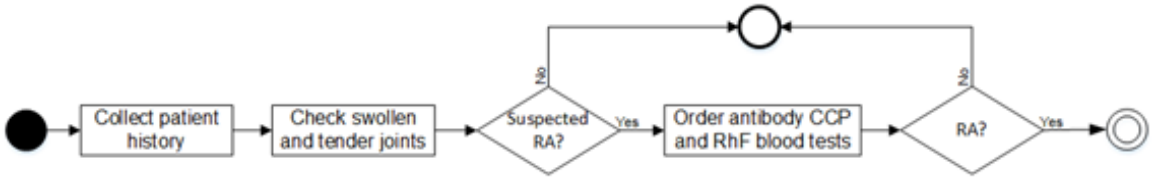


Fig. 6.1 A simplified model of care pathway for diagnosis of RA.

According to our rheumatology experts, the simplified care pathway has one selected decision support point: RA?, which is equal to 'rheumatoid arthritis?'. The activity nodes (represented with rectangles) have the following medical terms: 'patient history', 'swollen joints', 'tender joints', 'antibody CCP', and RhF or 'rheumatoid factor'. Using two medical ontologies of SNOMED CT and UMLS, we gather the concepts related to 'rheumatoid arthritis' as well as the other medical terms, namely, 'patient history', 'swollen joints', 'tender joints', 'antibody CCP', and RhF or 'rheumatoid factor'. The following subsections explain how we retrieve knowledge, refine it, create knowledge graphs, and translate them into candidate BN models.

6.3.1 Systematized Nomenclature of Medicine-Clinical Terms

We propose a step-by-step method to retrieve knowledge from SNOMED CT, refining processes, creating a usable knowledge graph, and finally translate the knowledge graph into a candidate BN model.

Step 1: Knowledge retrieval

Searching 'rheumatoid arthritis' in SNOMED CT matches 262 concepts in English, which contain 300 concepts from 6 hierarchies, namely, disorder, organism, procedure, regime/therapy, situation, and substance. Since the semantic tags disambiguate various concepts matched by searching a term, 'rheumatoid arthritis' matches 113 concepts from 9 semantic tags: assessment scale, disorder, finding, observable entity, organism, procedure, regime/therapy, situation, and substance. We consider the matched concepts with the semantic tags, and investigate their expressions from the inferred concept definition, which includes all non-redundant relationships. This gives us the consequence concepts related to the 'rheumatoid arthritis' from 9 hierarchies (or semantic tags) of the actual, body structure, disorder, morphological abnormality, observable entity, person, procedure, qualifier value, and substance.

By searching the medical terms of other pathway nodes including 'rheumatoid arthritis history', 'swollen joint count', 'tender joint count', 'antibody CCP', and 'rheumatoid factor', we retrieved the related concepts, their relationship attributes, and semantic tags. Any repeated retrieval concepts are removed. Some retrieval sets (i.e., a concept with a semantic tag and usually relationship attributes) have no relationship attribute, which we suggest defining an attribute for them by adding a verb to their semantic tag. For instance, the semantic tag of assessment scale becomes is assessment scale. The new attributes involve the concepts without attributes in the knowledge graph. We create the knowledge graph using the retrieved concepts and semantic tags or hierarchies.

Step 2: Refining knowledge

Once the retrieval is finished, we can start refining the concepts in two stages: (1) refining based on semantic tags and (2) refining based on relationship attributes. We have removed those with the regime/therapy, organism, observable entity, substance, and procedure semantic tags because of their inaction in diagnosis of RA, but the concepts with the following semantic tags are necessary for diagnosis of RA: assessment scale, disorder, finding, and situation.

At this stage, 154 remaining retrieval sets should be refined based on their relationship attributes considering semantic tags. Concepts with the semantic tags of assessment scale, finding, observable entity, situation, and substance are all needed, unless they have unnecessary attributes, including finding context, finding site, has interpretation, interprets, method, is physical object, subject relationship context, and temporal context. By deleting them, 11 concepts remain which only 7 of them hold a relationship attribute and the rest have none.

Among the concepts with the semantic tag of disorder, the 'rheumatoid arthritis' concept has three attributes: *associated morphology*, *pathological process*, and *finding site*. The first attribute is needed, but the other two are less important and we remove them. There are two synonyms of 'rheumatoid arthritis' - 'Rheumatoid Arthritis' and RhA. We remove

both. There are also many concepts with the disorder semantic tag and they repeat the same attributes and expressions as 'rheumatoid arthritis', e.g., 'rheumatoid arthritis of wrist' or 'rheumatoid arthritis of foot'. Their difference with 'rheumatoid arthritis' is that they have a distinct *finding site* attribute which specifies the joint affected by RA. These concepts are less important in diagnosis of RA; so that they can be removed, unless they have a *due to* or *causative agent* relationship attribute.

There are some sub-classes of 'rheumatoid arthritis' including the 'juvenile arthritis', 'seropositive rheumatoid arthritis', 'seronegative arthritis', 'acute polyarticular juvenile rheumatoid arthritis', and 'uveitis-rheumatoid arthritis syndrome'. We remove all of them, unless they hold an attribute of *due to* or *causative agent*, as mentioned before. Those with the *associated morphology* attributes can also be kept, but we have removed them at this point.

After all refining processes of the concepts with the disorder semantic tag, the 'rheumatoid arthritis' concept should be remained, plus any concept with the *due to* or *causative agent* attributes, because of their causal or associational importance. Finally, we search for the synonyms and keep only one of them, if they are related to an identical concept. In case we removed a concept that was retrieved by searching the medical terms of the care pathway nodes except that in the selected decision support node (except 'rheumatoid arthritis' in this case study), we should recover the concept and its semantic tag. Here, we recovered the 'rheumatoid factor' and the 'anti-cyclic citrullinated peptide antibody positive' concepts. In the end of this step, all attributes are checked and updated.

Step 3: Finalising knowledge graph

Among the remaining retrieval sets, those with the 'rheumatoid arthritis' term , in the concept itself or the consequence concept, should be kept as they are. In other retrieval sets, if the concept is from an observable entity, assessment scale, finding, or disorder hierarchy and the consequence concept is also a concept from the same hierarchy, then that retrieval set loses the consequence concept and gets 'rheumatoid arthritis' instead. In case the consequence concept is blank, the 'rheumatoid arthritis' is given as the consequence concept. We obtain the final knowledge graph by connecting the concepts to the consequence concepts using the new relationships. Table 6.1 shows the final list of concepts, their attributes, and the related concepts that are the expressions from the Inferred Concept Definition of SNOMED CT.

Step 4: Translating knowledge graph into Bayesian network

Both semantic relations and causal-associational connections are directed. Their difference is that the causal-associational connections are either the 'causes/associates with' or 'is

caused/associated by', whereas there can be various semantic relations between concepts. Here, we convert the semantic relations of the knowledge graph into causal-associational connections required for the BN model using a set of translation rules. These rules are not applied to the 'is observable entity' semantic tags because it involves many trivial concepts. Rules also do not translate the hierarchy of the knowledge graph, i.e., the 'has subclass' hierarchy is not translated. The translation rules of the final knowledge graph for RA diagnosis using the retrieved knowledge from SNOMED CT are as follows:

- If semantic relationship is *associated finding*, causal-associational connection is *causes/associates with*.
- If semantic relationship is *associated morphology*, causal-associational connection is *causes/associates with*.
- If semantic relationship is *due to*, causal-associational connection is *is caused/associated by*.
- If semantic relationship is *is assessment scale of*, causal-associational connection is *is caused/associated by*.
- If semantic relationship is *is finding of*, causal-associational connection is *is caused/associated by*.

Table 6.1 Refined concepts of rheumatoid arthritis and their relationships retrieved from SNOMED CT.

Concept	Updated Attribute	Related Concepts
Rheumatoid arthritis	Associated morphology	Rheumatic inflammation
Erosion of joint surface co-occurrent and due to rheumatoid arthritis	Due to	Rheumatoid arthritis
Deformity of hand due to rheumatoid arthritis	Due to	Rheumatoid arthritis
Disease activity score in rheumatoid arthritis	is assessment scale	Rheumatoid arthritis
Disease activity score 28 joint in rheumatoid arthritis	is assessment scale	Rheumatoid arthritis
Deformity of foot due to rheumatoid arthritis	Associated morphology	Deformity
Deformity of foot due to rheumatoid arthritis	Due to	Rheumatoid arthritis
Deformity of wrist due to rheumatoid arthritis	Associated morphology	Deformity
Deformity of wrist due to rheumatoid arthritis	Due to	Rheumatoid arthritis
Family history: Rheumatoid arthritis	Associated finding	Rheumatoid arthritis
History of rheumatoid arthritis	Associated finding	Rheumatoid arthritis
Swollen joint count	is assessment scale	Rheumatoid arthritis
Tender joint count	is assessment scale	Rheumatoid arthritis
Rheumatoid factor	is substance	Rheumatoid arthritis
Anti-cyclic citrullinated peptide antibody positive arthritis	is finding	Rheumatoid

- If semantic relationship is *is substance of*, causal-associational connection is *causes/associates with*.

Figure 6.2 shows the candidate BN model for RA diagnosis translated from the final knowledge graph that was retrieved from the SNOMED CT. This BN is built in AgenaRisk10 software [3].

Fig. 6.2 Candidate BN model for RA diagnosis built based on knowledge retrieved from SNOMED CT.

6.3.2 Unified Medical Language System

We explain the results of knowledge retrieval from UMLS metathesaurus browser and our proposed step-by-step method to process the results, create usable knowledge graphs, and ultimately translate the final knowledge graph into the structure of a candidate BN model.

Step 1: Knowledge retrieval

We find 86 atoms containing 1674 relationships by searching the selected decision support point from the simplified care pathway, 'rheumatoid arthritis', in the metathesaurus browser. Each atom may have one or multiple 'relations' which consist of the relationships (RELs), Root Source Abbreviations (RSABs), strings, concept unique identifiers (CUIs), and in most

cases the relationship attributes (RELAs). Searching other terms of care pathway, namely 'rheumatoid arthritis history', 'rheumatoid factor', 'antibody CCP', 'swollen joint count', 'tender joint count', respectively matching 4, 15, 9, 3, and 3 atoms with 56, 395, 264, 85, and 85 RELAs, respectively. There is no exact match of 'rheumatoid arthritis history' and 'antibody CCP', so that we considered the atoms of all search results in these two cases in order to prevent any missing knowledge.

Step 2: Refining knowledge

In this step, we perform two stages of refining: (1) based on RELAs and (2) based on the semantic type among the selected RELA. As the first stage, we look for relationships with RELAs since they enable us to process them as well as reduce the number of concepts. Once we removed the concepts without RELAs, 54, 1213, 206, 28, 26, and 26 relationships remain from the atoms of 'rheumatoid arthritis history', 'rheumatoid arthritis, 'rheumatoid factor', 'antibody CCP', 'swollen joint count', and 'tender joint count'. Since the care pathway is for diagnosis, we do not need the concepts with the following RELAs: *answer to*, *associated with*, *associated with malfunction of gene product*, *classified as*, *classifies*, *clinically similar*, *clinician form of*, *component of*, *concept in subset*, *concept in subset*, *concept in subset*, *ddx, disease has normal tissue origin, disease has primary anatomic site, diseasehas associated anatomic site, do not code with, due to, expanded form of, focus of, fragments for synonyms of, gene product encoded by gene, gene product has biochemical function, gene product has chemical classification, gene product plays role in biological process, has alias, has associated finding, has cdrh parent, has clinician form, has component, has consumer friendly form, has contraindicated drug, has definitional manifestation, has expanded form, has expanded form, has finding site, has finding site, has interpretation, has location, has method, has nichd parent, has nichd parent, has pathological process, has permuted term, has specialty, has specimen, has subject relationship context, has translation, has translation, interprets, inverse isa, inverse was a, isa, manifestation of, mapped from, mapped to, may be associated disease of disease, may be prevented by, may be treated by, measured by, measured component of, member of, occurs before, permuted term of, phenotype of, possibly equivalent to, primary mapped from, primary mapped to, same as, used for*, and *was a*.

We are interested in diagnosis-related RELAs including *associated finding of, causative agent of, cause of, clinically associated with, co-occurrence with, disease has associated gene, has associated finding, has associated morphology*, and *has causative agent*.

In the second stage of refining, among the interesting RELAs, we refine the relations based on the semantic type of the string. The relations with RELAs of *associated finding of, cause of, co-occurs with, disease has associated gene, has associated finding*, and *has*

associated morphology are removed, regardless of their semantic type. However, those with the RELAs of *clinically associated with*, are kept if they belong to the Sign or Symptom or Pathologic Function semantic types.

Applying the two stages of refining on the relations retrieved by searching the 'rheumatoid arthritis' provides us with 35 relations that 17 of them are repeated. With respect to the relations of the other terms, 1 relation remains for the 'rheumatoid factor' and 5 relations for the 'rheumatoid arthritis history' which 3 of them are repeated relation. The 'Antibody CCP', 'swollen joint count', and 'tender joint count' lost all their relations, but since they are explicitly mentioned in the care pathway, we recovered them and their semantic types and assign the 'rheumatoid arthritis' as their related concept. We also define a new RELA for

Table 6.2 Atoms of rheumatoid arthritis and their properties retrieved from UMLS

Concept	RELA	Consequence concept	Semantic type
Rheumatoid arthritis	co-occurs with	X RAY ABNORMAL	Finding
Rheumatoid arthritis	clinically associated with	FEVER	Sign or Symptom
Rheumatoid arthritis	clinically associated with	HEMARTHROSIS	Pathologic Function
Rheumatoid arthritis	co-occurs with	CONNECTIVE TISSUE DISEASE MIXED	Disease or Syndrome
Rheumatoid arthritis	co-occurs with	POLYMYOSITIS	Disease or Syndrome
Rheumatoid arthritis	disease has associated gene	MIF Gene	Gene or Genome
Rheumatoid arthritis	disease has associated gene	CCL2 Gene	Gene or Genome
Rheumatoid arthritis	disease has associated gene	PTPN22 Gene	Gene or Genome
Rheumatoid arthritis	Disease has associated gene	IL17A Gene	Gene or Genome
Rheumatoid arthritis	has causative agent	Rheumatoid factor	Immunologic Factor
Rheumatoid arthritis	associated finding of	FH: Rheumatoid arthritis	Finding
Rheumatoid arthritis	associated finding of	H/O: rheumatoid arthritis	Finding
Rheumatoid arthritis	cause of	Deformity of wrist due to rheumatoid arthritis	Finding
Rheumatoid arthritis	cause of	Deformity of foot due to rheumatoid arthritis	Anatomical Abnormality
Rheumatoid arthritis	cause of	Deformity of hand due to rheumatoid arthritis	Finding
Rheumatoid arthritis	cause of	Erosion of joint surface co-occurrent and due to rheumatoid arthritis	Disease or Syndrome
Rheumatoid arthritis	has associated morphology	Inflammation	Pathologic Function
Rheumatoid arthritis	has associated morphology	Rheumatic inflammation	Pathologic Function
Rheumatoid Factor	causative agent of	Rheumatoid arthritis	Disease or Syndrome
Anti-cyclic citrullinated peptide antibody positive	is laboratory or test result of	Rheumatoid arthritis	Laboratory or Test Result
Swollen joint count	is intellectual product of	Rheumatoid arthritis	Intellectual Product
Tender joint count	is intellectual product of	Rheumatoid arthritis	Intellectual Product
FH: Rheumatoid arthritis	has associated finding	Rheumatoid arthritis	Finding
H/O: rheumatoid arthritis	has associated finding	Rheumatoid arthritis	Finding

them by adding a verb and a preposition before and after their semantic type. Here, we have the following new RELAs:

- The laboratory or Test Result becomes *is laboratory or test result of*.
- The intellectual Product becomes *is intellectual product of*.

Table 6.2 summarises all search terms, RELAs, refined concepts, and their semantic types after two stages of refining and defining new RELAs. In this table, the concept column contains the terms that we searched and the concepts in the consequence concept column are related to the searched concept. The concepts and consequence concepts as well as the RELAs of this table imply the knowledge graph based on the concept and properties that we searched from UMLS and refined.

Step 3 – Translating knowledge graph into Bayesian network model
Using RELAs, we translated the semantic links between the concepts of the knowledge graph into the causal or associational connections required for a BN model. We investigated the causal or associational meaning of each RELA and developed the following rules:

- If RELA is *associated finding of*, causal-assiociational connection is *is caused/associated by*
- If RELA is *causative agent of*, causal-assiociational connection is *causes/associates with*
- If RELA is *cause of*, causal-assiociational connection is *causes/associates with*
- If RELA is *clinically associated with*, causal-assiociational connection is *causes/associates with*
- If RELA is *co-occurs with*, causal-assiociational connection is *causes/associates with*
- If RELA is *disease has associated gene*, causal-assiociational connection is *is caused/associated by*
- If RELA is *has associated finding*, causal-assiociational connection is *causes/associates with*
- If RELA is *has associated morphology*, causal-assiociational connection is *causes/associates with*

- If RELA is *has associated morphology*, causal-assiociational connection is *causes/associates with*
- If RELA is *has causative agent*, causal-assiociational connection is *is caused/associated by*
- If RELA is *is intellectual product of*, causal-assiociational connection is *is caused/associated by*
- If RELA is *is laboratory or test result of*, causal-assiociational connection is *is caused/associated by*

We translate the RELAs of Table 6.2 into causal-associational connections using the above-mentioned translation rules. This converts the knowledge graph into a candidate BN model for RA diagnosis by translating the RELAs, as displayed in Figure 6.3. This BN is built in AgenaRisk10 software [3].

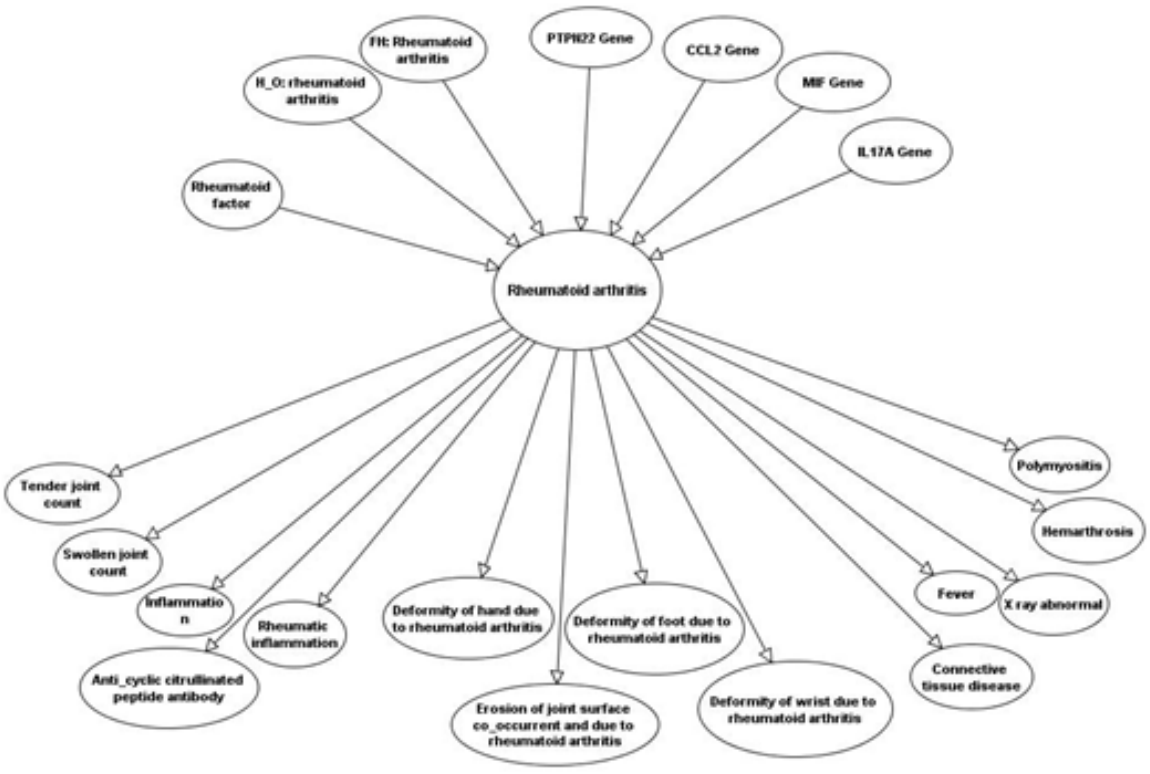


Fig. 6.3 Candidate BN model for RA diagnosis built based on UMLS knowledge.

6.4 Discussion

The proposed step-by-step methods to convert knowledge graphs for RA diagnosis into BN models are the initial attempt to automate the process of building BN structure using care pathways and pre-existing ontologies. According to our main expert, the structure of both candidate models is medically meaningful, although some variables in either model are rather less important to be involved in a diagnosis model. In the candidate BN model from SNOMED CT (Figure 6.2), there are two variables of 'Disease activity score in Rheumatoid Arthritis' and 'Disease activity score 28 joints in RA' that both refer to disease activity, although the latter can be a variation of the former. This shows the difficulty to refine the retrieved knowledge because we still have many concepts (or variables) with close definitions, despite the multiple refining stages that we performed in Section 6.3.1.

We think that the candidate BNs may not be able to detect early RA, but they can identify established RA. This is because these candidate BNs have multiple signs and symptoms, e.g., different variables related to deformity in the candidate BN model made from UMLS. Signs like deformity or erosion of joints may be developed in an established RA, not an early RA. On the other hand, we can see that the candidate BNs have a relatively smaller number of risk factors and lack a pathogenesis mechanism like what we made in the knowledge-based BN model in Section 5.3. Although the candidate BN model from UMLS includes four genes, any inter-connection between them and other risk factors is missing. This can be improved by retrieving the concepts for RA, i.e., all descendants and ancestors of the matched concepts need to be investigated to find inter-connections. Having a richer layer of risk factors and pathogenesis mechanisms can ultimately improve the ability of the candidate BNs in diagnosing early RA.

Future work can add transitive or functional relations to the knowledge graphs and define translation rules for them to improve the outcome of BN models. We can also parameterise the candidate BN models with the dataset that we used to parameterise the knowledge-based BN model (Subsection 5.3.1). Then, we can compare the performance of the candidate BN models with that of the knowledge-based and also with that of the learnt model (described in Section 5.4).

6.5 Summary

In this chapter, we propose two methods to build BN models using a care pathway and the retrieved knowledge from pre-existing medical ontologies of SNOMED CT and UMLS, corresponding to the Objective 2. In these methods, we first create a care pathway based

on experts' knowledge that represents the main medical terms and the decision support that they want to achieve. We use this care pathway as the basis to create knowledge graphs. For this, we search the medical terms and the decision support point in the ontologies and we retrieve the concepts and their properties. We refine the retrieved information by removing the redundant concepts and transforming some properties into useful ones. This results in a knowledge graph that represents the concepts and the semantic relations between them.

We define a set of rules to translate the semantic relations into causal or associational connections based on the causation or association meaning of the semantic relations. We specify the concepts of the knowledge graph as the variables of the BN model and link them with the translated connections. This leads us to build a meaningful structure for BN models.

Chapter 7

Building Dynamic Bayesian Network Model for Self-Management of Rheumatoid Arthritis: Appointment Scheduling

In Chapter 5, we presented a Bayesian network (BN) model for the diagnosis of Rheumatoid Arthritis (RA). After diagnosis, the RA disease needs to be managed until the end of patients' lives. The current chapter shows how to build a Dynamic BN (DBN) for self-management of RA using available data and by acquiring knowledge from experts, clinical guidelines, and literature. We suggest using this DBN model to support clinicians in appointment scheduling, as the first stage of decision support for self-management of RA. The DBN model for self-management includes signs, symptoms, and serology results, and estimates the disease state and the occurrence of flares. These estimations lead to an advice variable on the scheduling of follow-up appointments. We specify the components of the model from the available data, experts' knowledge, guidelines' information, and literature, corresponding to the Objective 3. To evaluate the performance of the DBN model in providing monitoring advice, we use a set of interpolated data and scenarios of dummy patients to do reasoning with the model (Objective 3).

The outline of this chapter is as follows: Section 7.1 gives and introduction, recapping from Chapter 4 the current care of RA and introducing the potential of better decision support to improve care. Section 7.2 covers a description of the relevant variables for self-management classified as evidence variables, latent variables, and an advice variable. Section 7.3 presents an explanation of the data for building the model, the structure of DBN model for self-management, followed by an explanation of the parameter learning of DBN.

Section 7.4 describes the DBN model for self-management of RA as a decision support tool, elaborating its inputs, the dynamics within the model, and the prediction of disease state and an appropriate advice for appointment scheduling. In Section 7.5, we present the reasoning and evaluate the model using the interpolated the Pathobiology of Early Arthritis (PEAC) data and a set of scenarios of dummy patient data. Finally, a summary of the chapter is presented in Section 7.6.

7.1 Introduction

As described in Chapter 4, the monitoring of chronic diseases has been classically done directly by clinicians while patients attend the clinic at fixed and infrequent intervals, e.g., every six weeks for people with RA. For monitoring of RA, clinicians follow the existing clinical guidelines. Based on these guidelines and by eliciting the knowledge of our experts (introduced in Section 4.1), two models of care pathways were created for the initial and ongoing management of RA, as displayed in Figures 4.6 and 4.7. These models of care pathway consider the guidelines of the National Institute for Health and Care Excellence (NICE) for the management of RA that recommend monitoring the disease monthly, until a low and stable disease activity is achieved [138]. Monthly monitoring may be applicable in the first follow-up for patients registered with the Barts and the London NHS Trust where our rheumatology experts are based or other clinics or hospitals in the UK. However, regular monthly monitoring is impractical and the interval of appointments stretches to 3-monthly or longer. This leads to undermonitoring of people with RA and clinicians being overbooked even though they only provide fixed and infrequent monitoring, ignoring any possible issues arose by the shortage of staff at the National Health Service (NHS).

Considering the models of care pathways for the management of RA (Section 4.3) and the analysis of available data (Section 4.4), we selected multiple decision support points, as presented in Table 4.6. The decision support points of 'Book Follow-up Appointment' and 'New Appointment?' can refer to the aforementioned issues of monitoring at fixed and infrequent intervals. Therefore, we suggest creating decision support with a DBN model for self-management of RA to assist with appointment scheduling. This can realise the continuous monitoring without actually attending unnecessary follow-up visits. It can also reduce the burden of monitoring patients who achieve a low and stable disease activity without them attending monitoring visits as frequently as 3-monthly. We decided to exclude medications at this stage and only focus on disease state estimation and appointment scheduling. In Chapter 8, we extend the DBN model for self-management by adding medication-related variables.

7.2 DBN Variables for Self Management

We specify the required variables to build a DBN model for self-management of RA. As shown in Figure 4.6, the pathway for initial management of RA represents the measurement of disease activity in the beginning of the care, and then clinicians make a decision on the next follow-up appointment. In the pathway for ongoing management of RA (Figure 4.7), there are three entries: the regular clinic visit by a rheumatologist, the review of test results by a clinician, and when a patient feels flare. Clinicians check the manifestation factors (signs, symptoms, and blood test results) and calculate their Disease Activity Score 28 (DAS28) to categorise them into high, moderate and low, and remission, as clinicians usually do in clinical practice.

Knowledge elicitation and data analysis lead us to specify the signs, symptoms, and blood test results.We classify the variables into three groups: evidence variables, latent variables, and advice variable.

7.2.1 Evidence Variables

Evidence variables include signs, serology results, and symptoms. The main signs of RA are tender joints count (TJC) and swollen joints count (SJC). The swelling of joints can be easily observed, but tenderness is claimed by patients. We define five states for TJC: None ($=0$), Few ($1-5$), Some ($6-10$), Many ($11-16$), and 'Too many' ($17-28$). We consider five states for SJC: None ($=0$), Few ($1-3$), Some ($4-7$), Many ($8-12$), and 'Too many' ($13-28$). The difference between the thresholds for TJC and SJC implies the importance of the observed SJC compared to the claimed TJC.

Two common serology results to monitor RA are the erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). We consider medically meaningful thresholds to categorise ESR and CRP data provided by our main expert. ESR is categorised into: Normal (<20), Moderate($20 \leq$ and ≤ 35), High(>35 and ≤ 50), Very high(>50 and ≤ 75), and Extreme(>75), and CRP is categorised into: Normal (<5), Moderate($5 \leq$ and ≤ 15), High(>15 and ≤ 30), Very high(>30 and ≤ 50), and Extreme(>50).

There are five symptoms commonly collected in clinical practice to monitor RA, namely, stiffness, global health (GH), fatigue, pain, and health assessment questionnaire (HAQ). The most important symptom of RA is the morning stiffness for more than 30 minutes. We consider a variable called Stiffness with binary states of Yes and No. GH refers to the patient's general health, Fatigue is their feeling of tiredness, and Pain is their feeling of pain, all quantified between 0 to 100 representing the best and the worst values, respectively. We

categorised GH, Fatigue, and Pain into three groups of Low ($0 <$ and ≤ 33.33), Medium ($33.33 <$ and ≤ 66.66), and High ($66.66 <$ and ≤ 100).

HAQ is a measurement of disability quantified between 0 to 3 [76]. It is a combination of twenty questions regarding mobility such as the ability to move hands, arms, and running errands. We consider a variable called HAQ with four states of None ($= 0$), Mild (≤ 1), Moderate (> 1 and ≤ 2), and Severe (> 2).

Table7.1 shows a summary of RA monitoring factors and their characteristics.

Table 7.1 Summary of evidence variables for self-management of RA

Category	Variable	States	Source	Description
Sign	TJC	None($=0$), Few($1-5$), Some($6-10$), Many($11-16$), Too many($17-28$)	Data	Number of tender joints examined by pressing between two fingers
	SJC	None($=0$), Few($1-3$), Some($4-7$), Many($8-12$), Too many($13-28$)	Data	Number of swollen joints observed
Symptom	Stiffness	No, Yes	Data	Patient experiencing morning stiffness for more than 30 minutes
	GH	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Patient's general health quantified between 0 and 100 representing the best and the least health, respectively
	Fatigue	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Patient's feeling of fatigue quantified between 0 and 100 representing the lowest and the highest fatigue, respectively
	Pain	Low ($0 <$ & ≤ 66.66), Medium ($33.33 <$ & ≤ 66.66), High ($66.66 <$ & ≤ 100)	Data	Patient's feeling of pain quantified between 0 and 100 representing the lowest and the highest pain, respectively
	HAQ	None ($=0$), Mild ($0 <$ & ≤ 1), Moderate ($1 <$ & ≤ 2), Severe ($2 <$)	Data	Patient's disability measurement quantified between 0 and 3 representing the lowest and the highest disability respectively
Serology	ESR	Normal (< 20), Moderate($20 \leq$ and ≤ 35), High(> 35 and ≤ 50), Very high(> 50 and ≤ 75), Extreme(> 75)	Data	ESR measured in blood test
	CRP	Normal (< 5), Moderate($5 \leq$ and ≤ 15), High(> 15 and ≤ 30), Very high(> 30 and ≤ 50), Extreme(> 50)	Data	CRP measured in blood test

7.2.2 Latent Variables

Latent variables are the underlying factors of disease progression associated with multiple clinical and laboratory measurements. In RA, disease activity is a major underlying factor required to judge the severity of disease. Rheumatologists proposed various measurements to quantify disease activity. One of the first attempts was presented in [211], which suggested an index called Disease Activity Score (DAS) combining the Ritchie's index (count of tender joints), the count of swollen joints, acute-reactant phase, and general health. DAS was later modified for only 28 joints and was called DAS28 that combines TJC, SJC, acute-reactant phase (ESR or CRP), and GH [166]. In addition to single indices like DAS and DAS28, the

American College of Rheumatology (ACR) suggested seven criteria to specify improvement in RA severity including 20% improvement of tender and swollen joints counts and 20% improvement of at least 3 or the following measures: pain, patient and physician global assessments, self-assessed physical disability, and acute-phase reactant [57]. Also, the European League Against Rheumatism (EULAR) proposed a set of criteria, called EULAR response criteria, to determine the improvement of RA based on DAS28 change and DAS28 at present [215].

Considering the above measures and various criteria for measuring RA disease activity, we define a variable called 'Disease State'. The 'Disease State' has three states defined based on DAS28's common thresholds: Low (< 2.6), Moderate ($2.6 <$ and < 5.1), and High ($5.1 \leq$). The Low state refers to remission and the common threshold of $DAS28 = 2.6$ to specify remission. We define a variable called 'Disease State Change' with three states of Increasing, Steady, and Decreasing. This variable is a comparison between the disease state variable of the present time with that in the previous time, scaling increase, steadiness, or decrease of the disease state. This helps to capture the trend of disease state by combining the change of disease states in the two times and the trend of disease state in the previous time. For this, we define a variable called 'Disease Progression Trend' with three states of Deteriorating, Stable, and Improving. The change of disease state and the trend of disease progression are two distinct concepts. The former just compares two consecutive disease states, whereas the latter combines that comparison with a long-term change of disease progression that is embedded in the previous disease progression trend.

A sudden rise of disease state is known as a flare-up or flare. However, there is a lack of consensus among either rheumatologists or patients on the exact definition of flares. An international research project called the RA Flare Definition Working Group, organised by the Outcome Measures in Rheumatology Clinical Trials (OMERACT), targets to standardise the definition of flare and establish the characteristics and the impact of flares. The OMERACT defines flare as "any worsening of disease activity that would, if persistent, lead to the initiation or change of therapy" [17]. Recently, the OMERACT has identified the features of flares – including fatigue, stiffness, symptom persistence, systemic features, and participation – in addition to the existing core set of RA features [16]. Unlike OMERACT, some other researchers have tried to define the flare based on 'objective' signs of disease. In [213], two criteria are proposed to determine a flare: 1.2 increase in DAS28 or more than 0.6 increase in DAS28, if DAS28 is already greater than 3.2.

We adopt the flare definition of [213] to create a variable called 'Flare Occurrence' and associate it with the TJC sign and the symptoms of RA. These associations represent the 'subjective' nature of flares as a self-reported state by RA patients. We consider three states of

None, Minor, and Major for the 'Flare Occurrence' and we define them by extending [213]'s suggestion. To specify these states, we compare the change of DAS28 in two consecutive weeks. The None state is in two situations: 1) < 0.1 change of DAS28, if DAS28 is < 5.1 and 2) $0 \leq$ change of DAS28, if DAS28 is $5.1 \leq$. We define a Minor state to be: 1) $0.1 \leq$ and < 1.2 change of DAS28, if DAS28 is < 3.2, 2) $0.1 \leq$ and < 0.6 change of DAS28, if DAS28 is $3.2 \leq$ and < 5.1, and 3) $0 \leq$ and < 0.1 change of DAS28, if DAS28 is $5.1 \leq$. A Major state of flare is when: 1) $1.2 \leq$ change of DAS28, if DAS28 is < 3.2, 2) $0.6 \leq$ change of DAS28, if DAS28 is $3.2 \leq$ and < 5.1, and 3) $0.1 \leq$ change of DAS28, if DAS28 is $5.1 \leq$.

In addition to the occurrence of flare, we consider two features of flares: frequency and duration. In [87], it is found that 26% of cases had a $<$ 3-day flare during 3 months, whereas only 7% of cases experienced a $>$ 3-day flare at the same length of time. An analysis of flare and disease activity showed that patients with a higher disease activity tend to have a longer flare and more intense symptoms [26], as also implied in the thresholds proposed by [213]. We define a variable called 'Flare Duration' with three states of None, Short, and Long, to capture no or a brief ($\leq$ 3-day) flares, short length (3 $<$-day up to 1-week) flares, or long ($<$ 1-week) flares, respectively. We also define the 'Flare Frequency' variable with three states: Rare, Some, and Many. To specify the states of 'Flare Frequency', we give a weight of 2 to the Major flares and a weight of 1 for Minor flares, and then sum up the weighted scores of flares of 4 consecutive times to calculate the weighted score of flares. We define the Rare state to be ≤ 1, Some state to be $2 \leq$ and ≤ 5, and Many state to be $6 \leq$.

We create a variable called 'Overall Flare' representing the collective severity of flares by combining three variables of 'Flare Occurrence i', 'Flare Duration i', and 'Flare Frequency i'. It has three states: Severe, Mild, and None. A trend variable, called 'Overall Flare Trend

Table 7.2 Summary of latent variables for self-management of RA.

Category	Variable	States	Source	Description
Disease	Disease State	Low, Moderate, High	Data	Patient's disease state
	Disease State Change	Decreasing, Steady, Increasing	Expert	Change of disease state in the current time compared to that of the previous time
	Disease Progression Trend	Deteriorating, Stable, Improving	Expert	Trend of disease progression combining the change of disease states in two consecutive times and the trend of disease state of the previous time
Flare	Flare Occurrence	None, Minor,Major	Data	Occurrence of a Major or Minor or no flare
	Flare Duration	None, Short, Long	Literature	The duration of a flare
	Flare Frequency	Rare, Some, Many	Data	Frequency of flare occurrence during 4 weeks
	Overall Flare	None, Mild, Severe	Expert	Patient's overall flare combining flare occurrence, duration, and frequency
	Overall Flare Change	Decreasing, Steady, Increasing	Expert	Change of overall flare in the current time compared to that of the previous time
	Overall Flare Trend	Deteriorating, Stable, Improving	Expert	Overall flare trend representing the trend of overall flare by comparing two consecutive overall flares
	Overall Disease Control	Good, Alarm, Poor	Expert	Disease being overall under control, on alarm, or unmanaged, combining disease state and overall flare

i', measures the change of 'Overall Flare i' between the current time and the previous time, which has three states: Deteriorating, Stable, and Improving.

Since we intend to predict whether the disease is managed or not, we aggregate 'Disease State i' and 'Overall Flare i' in a variable called 'Overall Disease Control i'. It has three states of Good, Alarm, and Poor, which refer to the managed, uncertain, and unmanaged states of the disease, respectively.

Table 7.2 shows a summary of the latent variables and their characteristics.

7.2.3 Advice Variable

We consider an advice variable called 'Appointment Scheduling Advice' with two states: 'Standard monitoring' and 'Enhanced monitoring'. The 'Standard monitoring' advice indicates that a patient's disease is managed well and can attend a standard clinic visit (e.g., 6-monthly). It implies that a patient can self-manage their disease by the next standard monitoring visit. In contrast, the 'Enhanced monitoring' advice recommends attending an appointment earlier than the standard monitoring to investigate and improve the severity of the disease. The advice variable is associated with the overall disease control, the progression trend of disease, and the trend of flares. Table 7.3 shows a summary of the advice variable.

Table 7.3 Summary of the advice variable for self-management of RA: appointment scheduling.

Category	Variable	States	Source	Description
Advice	Appointment Scheduling Advice	Standard monitoring, Enhanced monitoring	Data	Advice on appointment scheduling for monitoring based on overall disease control, disease progression, and overall flare trend

7.3 DBN Model for Self Management of Rheumatoid Arthritis: Appointment Scheduling

DBNs are a temporal extension of BN models comprising identical BN models repeated in discrete time slices. The BN model of each time is connected to the BN model of other time(s) with temporal dependencies. DBNs are described in detail in Subsection 2.3.1. Here, we propose a DBN model for self-management of RA containing the variables described in Section 7.2. We create the structure of DBN model based on expert's knowledge and established literature. For the parameterisation of DBN model, we prepared a dataset from the available data, and we used expert's judgement and information published in the medical literature, if the data were unavailable.

7.3.1 Structure

DBNs consist of repeated BN models in multiple discrete time slices forming a stochastic process in each time slice (see Subsection 2.3). Based on this definition, first we create the structure of a BN model for one time slice of the DBN model. Then, we transform it into a temporal structure by adding temporal dependencies between the BNs. We depict the structure of the DBN model with the simplified notation, as expressed in Figure 2.3.1. In this notation, the repeated BN models are rolled-up and represented as one BN model in the temporal plate.

The proposed DBN model for self-management is shown in Figure 7.1. We used the AgenaRisk software [3] to create the DBN model. It represents a DBN model with initial conditions or time slice, a temporal plate, and terminal conditions. Variables are represented with ovals and all dependencies are indicated with arcs, but temporal dependencies have a number 1 on the arc.

The temporal plate is located in the middle of the model and represents the 'rolled-up' BN models. The variables of the temporal plate are either evidence variables represented with

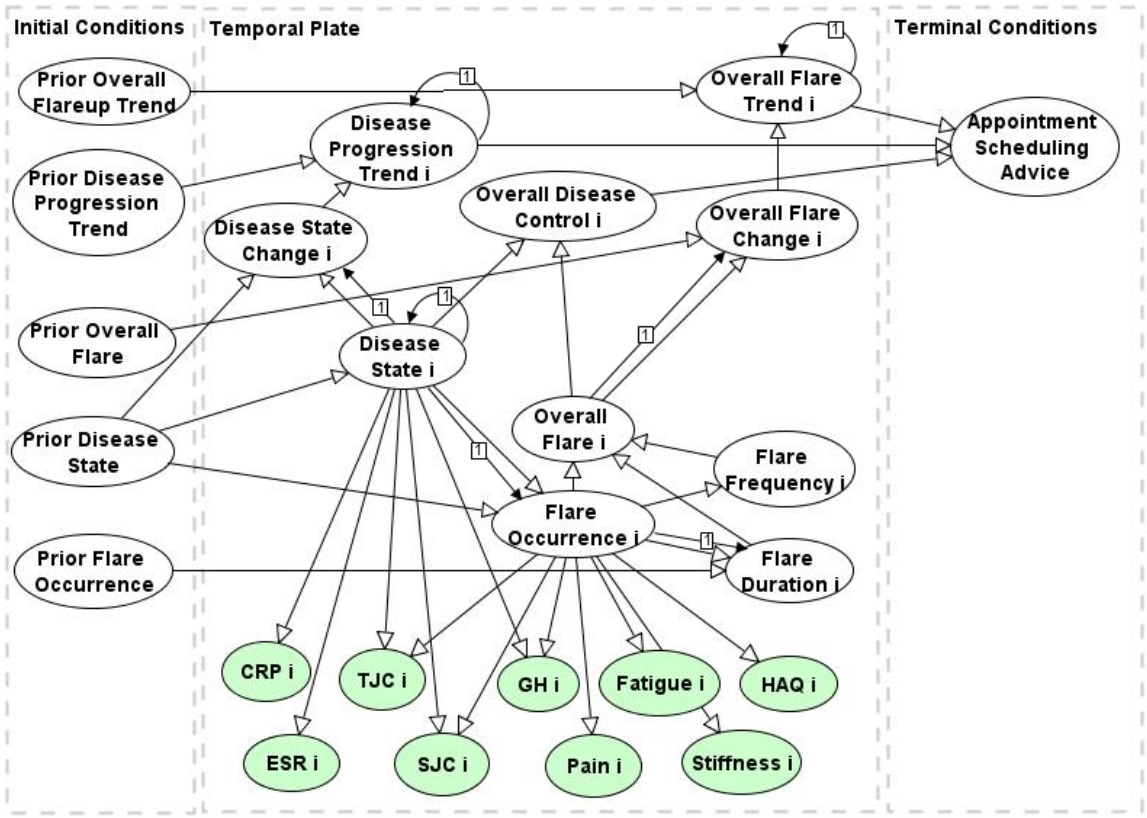


Fig. 7.1 DBN model for self-management of RA for appointment scheduling.

green ovals, or latent variables represented with white ovals. The variables in the temporal plate have an index *i*, representing the time slice that they belong to. Each index has an underscore in Figure 7.2, which is the necessity of the software; however, we remove the underscore in the manuscript. To build the BN of the temporal plate, we investigate the causal and associational relations between the evidence and latent variables (described in Subsection 7.2). The main latent variable is 'Disease State i' that is associated with 'CRP i', 'ESR i', 'TJC i', 'SJC i', and 'GH i' variables that correspond to the manifestation idiom (Figure 2.10a). This is adapted from the DAS28 and its components. To capture the trend of disease state, we consider the variable 'Disease State Change i' to scale the increase, steadiness, or decrease of disease state of the current time compared to the previous time. The change of disease state associates with the 'Disease Progression Trend i' variable.

The 'Flare Occurrence i' is linked to the 'Disease State i' and 'Disease State i-1' variables of two consecutive times, estimating a worsening of disease state. It is associated with the symptoms of RA, namely, 'GH i', 'Fatigue i', 'Pain i', 'Stiffness i', and 'HAQ i', as well as two sign variables of 'TJC i' and 'SJC i' that patients are involved in their measurement. Dependencies from the 'Flare Occurrence i' variable to the relevant signs or symptoms match the manifestation idiom (Figure 2.10a). We also consider the 'Flare Frequency i' and 'Flare Duration i', and combine them with the 'Flare Occurrence i' in the synthetic variable of 'Overall Flare i'. Similar to the disease state, we capture the trend of overall flare by two variables of 'Overall Flare Change i' and 'Overall Flare Trend i'.

The initial condition area, shown on the left side of the model, comprises the prior variables of the latent variables of the temporal plate that are involved in the temporal process.

The terminal condition, shown on the right side of the model contains the 'Appointment Scheduling Advice' variable that is associated with 'Overall Disease Control i', 'Disease Progression Trend i', and 'Overall Flare Change i' variables.

7.3.2 Data

We use a subset of the dataset collected in the PEAC study, as described in Section 4.4. The PEAC dataset consists of personal factors, risk factors, comorbidities, clinical record of signs, symptoms, and serology results (described in Subsections 4.4.2 4.4.3). The PEAC data are gathered in the baseline, 1st follow-up (3-month), 2nd follow-up (6-month), 3rd follow-up (9-month), and 4th follow-up (12-month) visits, which respectively contain 215, 181, 176, 157, and 146 RA cases.

For the parameterisation of the DBN model, we choose the 1st and 2nd follow-up visits of the PEAC data since they contain the highest number of cases in two consecutive

times, excluding the baseline data. The exclusion of the baseline data is because people get diagnosed in baseline and receive their first medications that result in a sharp fluctuation of their clinical records and symptoms. We skip this fluctuation by excluding the baseline data from parameter learning. As mentioned before, the 3rd or 4th follow-up visits of the PEAC data contain progressively less data and using them could lead to a poor parameterisation.

The dataset consists of signs (TJC and SJC), symptoms (stiffness, GH, fatigue, pain, and HAQ), serology results (ESR and CRP), and disease state (DAS28). It also includes the flare occurrence and the frequency of flares calculated as described in Subsections 7.2.2 and thoroughly explained in Subsection 4.4.4. By interpolating the data, we obtain a more frequent dataset that resembles a dataset collected to build and evaluate a self-management. The interpolation of the quarterly collected data is an easy way of creating more frequent (weekly) data required for evaluating the DBN model for self-management. For interpolation, we merge the subsets of the PEAC data of the first and second follow-up visits and then interpolate the values between the recorded data of the first and second follow-up visits. Merging the two subsets increases the percentage of missing values to 10%, mainly because the number of cases in the two subsets are not equal. We only keep the cases (rows) with at least 50% missing values; otherwise we drop them. This results in 163 remaining cases (13 cases dropped) and 4% missing values in the entire dataset.

The dataset that would be ideally frequent to evaluate the performance of the DBN model is not available. Even the BioT study (introduced in Subsection 4.4.1), that studies the effect of a mobile app to remotely manage disease activity and medication tapering, has no such data recorded. We suggest generating an appropriate data by interpolating a subset of the PEAC dataset, as described in Subsection 4.4.4. The interpolation of PEAC dataset expanded it from quarterly to weekly, and therefore increased its frequency from 5 times (quarterly) to 49 (weekly), comprising of 5 records and 44 interpolations. This is an appropriate frequency to do reasoning with the DBN model for self-management of RA.

7.3.3 Parameterisation

We used data, experts' knowledge, and information published in the literature to parameterise the DBN model. The evidence, latent, and advice variables of the DBN model are parameterised as follows:

- **Evidence Variables.** The available data is a subset of the PEAC dataset (described in Subsection 7.3.2), consisting of all evidence variables, including signs (TJC and SJC), symptoms (stiffness, GH, fatigue, pain, and HAQ), and serology results (ESR and CRP), disease state, flare, and flare frequency. These are all discretised data using

medically meaningful thresholds defined by our main expert (explained in Subsections 7.2.1 and 7.2.2).

- **Latent Variables.** We exploited the information published in [87] and modified it for parameterising the 'Flare Duration i' variable. For parameterisation of the 'Disease Progression Trend i' and 'Overall Flare Trend i' variables, we used the ranked nodes with Truncated Normal distribution and the weighted average function between their parent variables, e.g., the weighted average function between 'Prior Disease State' and 'Disease State 1' to train the parameters of 'Disease State Change 1'. The Truncated Normal distribution as suggested by [60] facilitates parameterising using expert's judgement. We elicited the weights for each ranked node from our main expert.

- **Advice Variable.** The advice variable for appointment scheduling is associated with 'Overall Disease Control i', 'Overall Flare Trend i', and 'Disease Progression i' variables. We derived the parameters of the advice variable from our main expert.

To parameterise the DBN model, we use the two-column method to train the parameters. In this method, we apply the expectation-maximisation (EM) algorithm (explained in Subsection 2.5.2) using a subset of available data of two consecutive times of the PEAC data, as mentioned in Subsection 7.3.2. We use the EM algorithm with 100 iterations and 0.001 convergence threshold in the AgenaRisk software [3].

7.4 Using the DBN for Decision-Support

The DBN model for self-management of RA tracks the disease state of patients and can support clinicians to decide on monitoring appointments. The inputs of the model are the patient's observations and the outputs are the estimations of disease state, flare, and their trends, leading to a final advice on appointment scheduling.

We consider the interval between entering inputs to be weekly, i.e., each time slice of the DBN represents one week. The weekly interval is because of the average interval between taking the common RA medications. As described in Section 8.2, medicines should be taken in various intervals - some medicines called conventional disease-modifying antirheumatic drugs (csDMARDs) should be taken daily (e.g., sulfasalazine or SSZ) or weekly (e.g., methotrexate or MTX) and some other medicines called the targeted synthetic or biological disease-modifying antirheumatic drug (ts/bDMARDs) should be taken fortnightly (e.g., CTZ and RTX) or monthly (e.g., tocilizumab or TCZ). Table 4.4 indicates that the majority of persons in the PEAC study are prescribed csDMARDs, and ts/bDMARDs medicines are

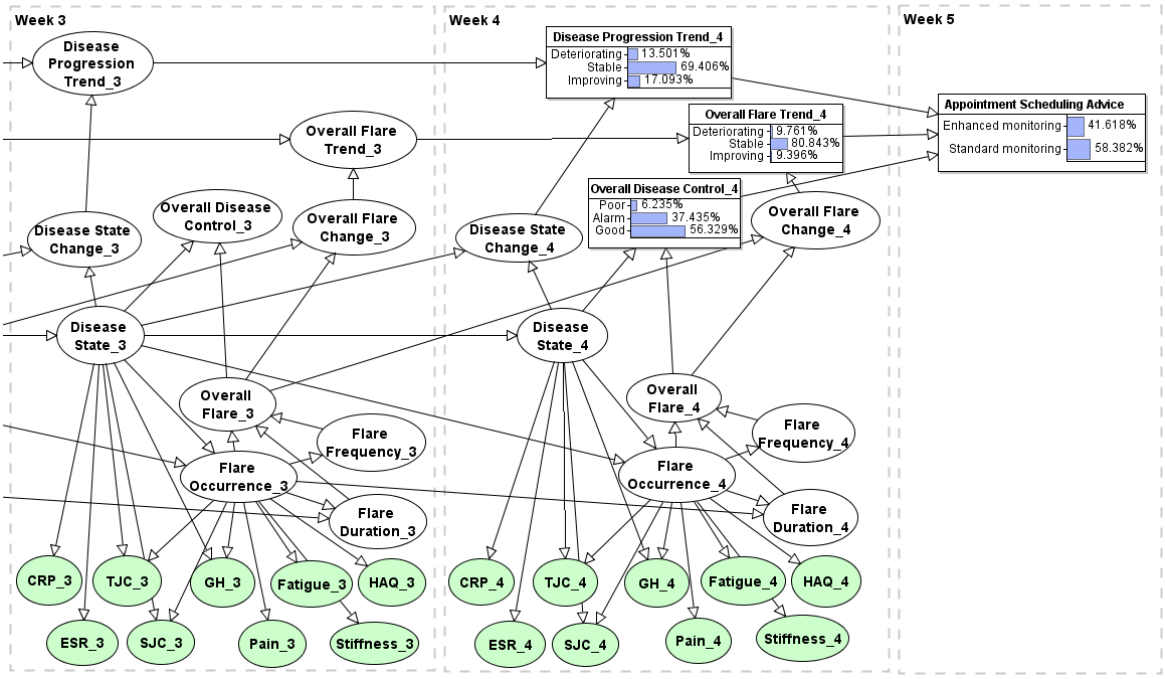


Fig. 7.2 Unrolled DBN model for self-management of RA.

given rarely. In the BioT data, all persons receive one form of ts/bDMARDs, mainly ETN, TCZ, or CTZ, as shown in Table 4.5. We consider a time slice of the DBN model to represent one week, because it is an average of the intervals of RA medications in the PEAC and BioT studies and also due to the fact that the decision support for self-management of RA is designed to be ultimately used for medication management, mainly biological medicines or bDMARDs (explained in the beginning of Chapter 1 and in Section 4.1).

Figure 7.2 displays the unrolled DBN model for self-management of RA. As an example, we assume a patient started providing their observations four weeks before, i.e., they are in Week 4 currently. In the following parts, we explain the inputs of the model, the dynamics of disease state and overall flare, and the prediction of disease state and the advice on monitoring.

7.4.1 Inputs

The input observations would be given to the evidence variables represented with green ovals in the bottom left of Figure 7.2. The inputs entail signs (TJC and SJC), symptoms (Stiffness, GH, Fatigue, and Pain), and serology results (ESR and CRP), described in Subsection 7.2.1 and also represented with green ovals in Figure 7.1. Our proposal is that the signs and symptoms would be observed by the patient and entered into the model. We also propose

that the serology results would be regularly measured by a local clinic every two weeks and would be given to the model through an electronic health record system.

7.4.2 Dynamics of Disease State and Overall Flare

Giving observations to the inputs of the DBN updates the probabilities of the latent and advice variables (described in Sections 7.2.2 and 7.2.3). The updated probabilities of 'Disease State 4' influence the probabilities of 'Disease State Change 4' and 'Disease Progression Trend 4'. The difference between the posterior probabilities of 'Disease State 4' and those in 'Disease State 3' updates the posterior probability of 'Flare Occurrence 4'. If the probabilities of High or Moderate states of 'Disease State 4' variable rise substantially compared to those in the previous week, a minor or major flare could occur. Figure 7.3 displays a fragment of DBN model expressing the probabilities of 'Disease State 4' and 'Flare Occurrence 4' variables, and excluding the outgoing dependencies. As shown in Figure 7.2, the overall disease control is most likely on a good state, representing a higher likelihood of having low disease state and no overall flare. The 'Flare Duration 4' considers the occurrence of flares in Weeks 3 and 4 and increases the probabilities of the Long or Short states, if two flares occur over two consecutive weeks or, in other words, the flare of the previous week continues to the current week. The 'Flare Frequency 4' appraises the number of flares that occurred before. As shown in Figure 7.3, the probability of the Rare state of flare frequency is 99%. This means that the person has most likely experienced none or only one minor flare within four weeks, based on the definition of the flare frequency variable (see Subsection 7.2.2). The combination of variables for flare occurrence, duration, and frequency updates the probabilities of 'Overall Flare 4' variable, which indicates 89% probability for the None state implying the lack of a severe or mild flare.

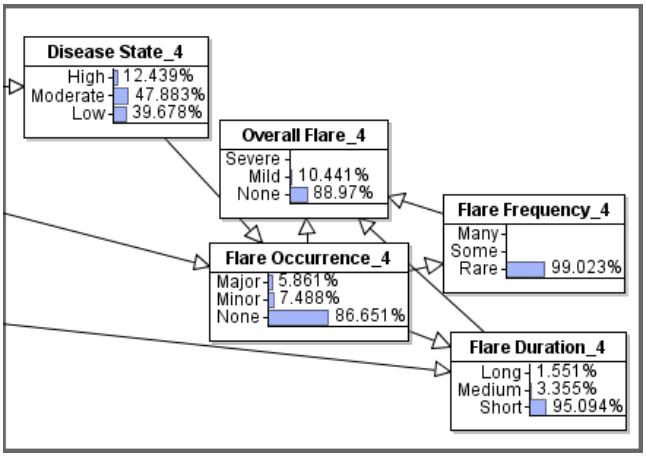


Fig. 7.3 Fragment of 'Disease State 4' and 'Overall Flare 4' and its parent variables.

We defined the change and trend of overall flare similar to the change and progression of disease state variables in each time. The 'Overall Flare Change 4' estimates the difference between the 'Overall Flare 4' and 'Overall Flare 3'. The 'Overall Flare Change 4' and 'Overall Flare 3' calculate the posterior probabilities of 'Overall Flare Trend 4'.

We define disease control to be a combination of disease state and overall flare. As shown in Figure 7.2, the 'Overall Disease Control 4' combines 'Disease State 4' and 'Overall Flare 4' with almost 56% likelihood of being in a good disease control.

7.4.3 Prediction of Advice on Appointment Scheduling

Three variables of 'Overall Disease Control 4', 'Overall Flare Trend 4', and 'Disease Progression Trend 4' lead to the 'Appointment Scheduling Advice'. It gives a recommendation on appointment scheduling, whether the person can self-manage, or whether they need enhanced monitoring. In Figure 7.2, the probabilities of 'Overall Disease Control 4' indicate that the person is most probably in a good disease control since the Good state has the highest probability (56%), and the probabilities of 'Overall Flare Trend 4', and 'Disease Progression Trend 4' imply that the disease is most likely stable as the Stable states of the two variables are the greatest probability (81% and 69%). Accordingly, the probabilities of the advice variable recommend continuing with the standard monitoring by the week after (Week 5). In other words, the person has self-managed their disease in Week 4 successfully and they are expected to do so for another week.

7.5 Evaluation

In this section, we evaluate the performance of the DBN model for self-management of RA. For this, we do reasoning by giving observations to the evidence variables (or inputs) and update the probabilities of latent and advice variables. We use the sliding window to limit the length of DBN model, as explained in Section 2.4. The sliding window moves forward when the time goes ahead, as depicted in Figures 2.9a and 2.9b. In those figures, we assumed the sliding window to contain two time slices and we depicted the move of the sliding window.

To reason with the DBN model for self-management of RA, we apply the sliding window approach. First, we define the length of the window by analysing the DBN model with different window lengths in the temporal plate: 2, 3, 4, 5, and 6. For this, we picked one case from the PEAC dataset and gave the interpolated data of that specific case to the DBN models with sliding window lengths of 2, 3, 4, 5, and 6. Since the 'Disease State' is the main latent variable, we compare the probabilities of its Low state with the same probabilities in another

model with a longer sliding window. We expect a longer window to increase the probability of the Low state. The Low state is our desire because the management of RA focuses on reducing the disease state and sustaining it. We stop lengthening the sliding window, once a longer window does not result in a greater probability for the Low state. We found out that a sliding window containing 5 time slices gets the continuous increase in the probabilities of Low states; therefore, it is an appropriate length of the sliding window. Figure 7.4 indicates the total difference of the probabilities of the Low state in DBN models with sliding windows of 2, 3, 4, 5, and 6 length. The orange, gray, and red colours represent the difference between a sliding window containing 2 and 3, 3 and 4, and 4 and 5, respectively. These lines are all above the 0 line, meaning they have positive differences. However, the difference between the probability of Low state in DBNs with two sliding windows containing 5 and 6 time slices (shown with the blue line) is negative, i.e., adding an extra time slice has reduced the probability of the Low state. Therefore, we conclude that a DBN with a sliding window containing 5 time slices has a suitable length.

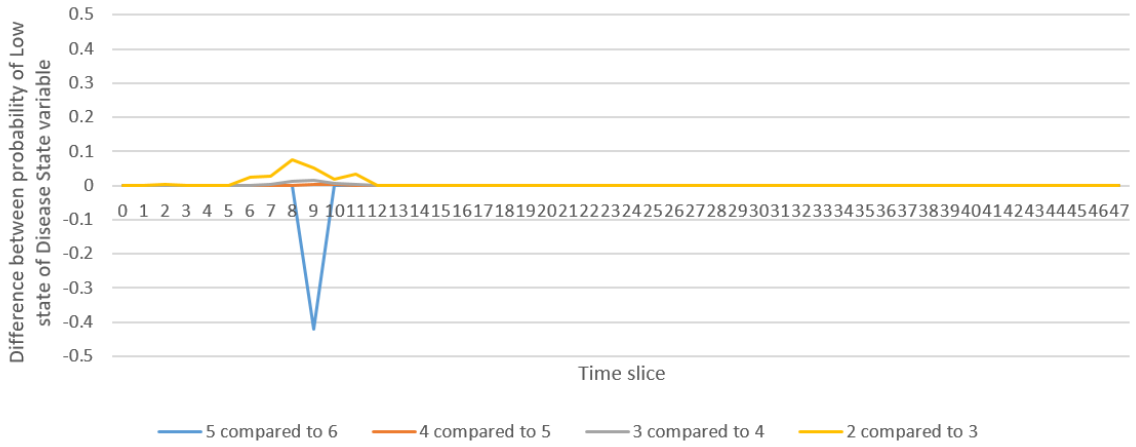


Fig. 7.4 The total difference of the probabilities of the Low state of the 'Disease State' variables in the initial DBN model for self-management of RA with sliding windows of 2, 3, 4, 5, and 6 length.

7.5.1 Evaluation with PEAC Data

We evaluate the performance of DBN model using interpolated PEAC data, as mentioned in Section 7.3 and described in Subsection 4.4.4. For this, we discretise the subsets of the PEAC data comprising of signs (TJC and SJC), symptoms (Stiffness, GH, Fatigue, Pain, and HAQ), serology results (ESR and CRP), disease state, flare occurrence, and flare frequency. By interpolation, we expand the dataset from quarterly to weekly. To form the dataset for evaluation, we merge the datasets that are interpolated separately. We drop any case with

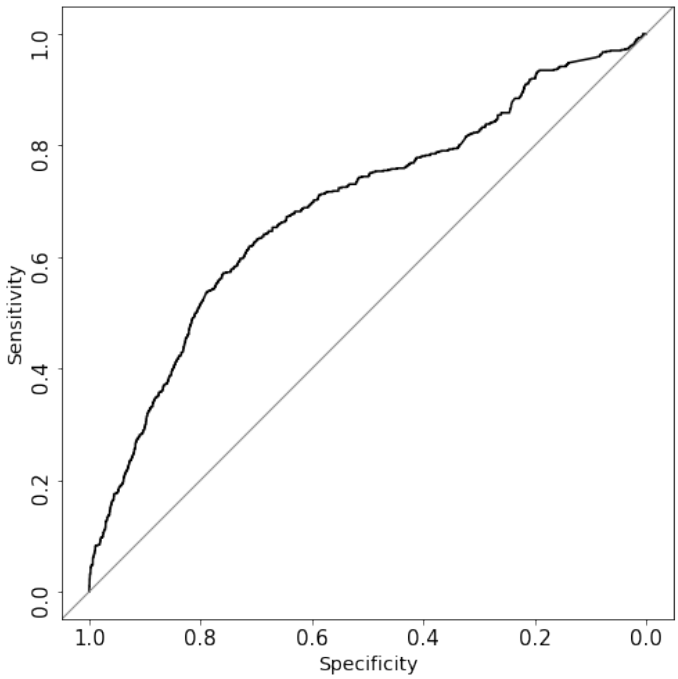


Fig. 7.5 ROC curve of DBN model for self-management of RA for appointment advice.

greater than 50% missing values. Finally, we obtain a dataset with 119 cases containing 5% of missing values.

We analyse the performance of the DBN model for self-management in terms of its discrimination by drawing a receiver operating characteristics (ROC) curve. Figure 7.5 shows the ROC curve of DBN model for self-management of RA for appointment advice. The area under ROC curve (AUROC) for this model is 0.69.

Since there is no gold standard (e.g., recorded advice on appointments) in the original dataset to compare the predicted advice with, we assume that the patient would need an enhanced monitoring appointment when they experience a major flare. Thus, we compare the probabilities of the advice variable with the major flare occurrence.

Figure 7.6 shows the analysis of performance of the DBN model for self-management of RA for appointment advice with a range of thresholds from 0.0 to 0.9. This figure includes four metrics of accuracy, precision, recall, and F1-score that express the prediction performance when data are imbalanced. The intersection point of precision and recall (also called sensitivity) is known to indicate the balanced threshold to have a high precision and recall at the same time. As shown in Figure 7.6, the balanced threshold is unusually less than 0.1 than can happen due to a small proportion of the enhanced monitoring cases compared to the standard monitoring ones (160 versus 4517) in the original dataset. Therefore, a small

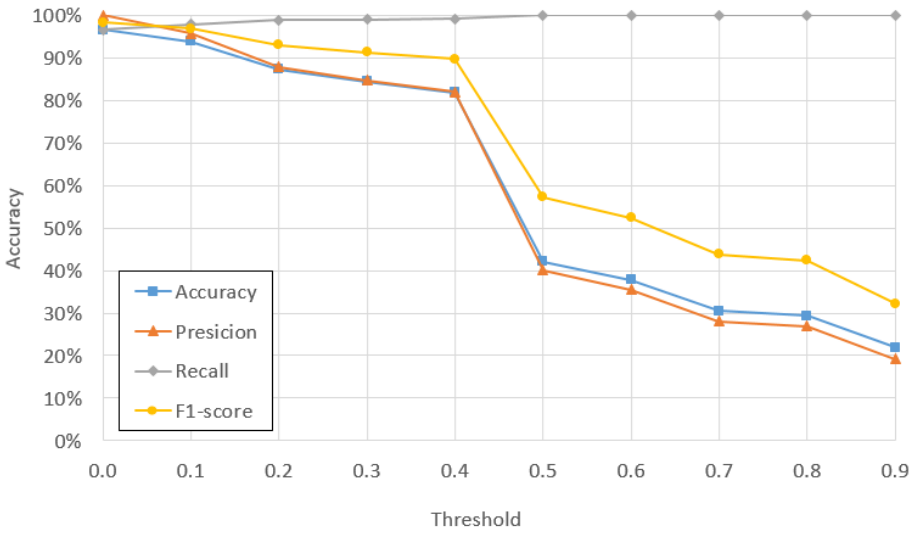


Fig. 7.6 Performance analysis of the DBN model for self-management of RA for appointment advice with a range of thresholds from 0.0 to 0.9.

percentage of the predicted values correspond to the enhanced monitoring cases and that would decrease the balanced threshold to below 0.1.

A pairwise comparison between the probabilities of the advice and the (interpolated) data of flare occurrence in each week shows 82% accuracy considering a threshold of 40% to differentiate the standard monitoring and enhanced monitoring states. Table 7.4 demonstrates the confusion matrix of this evaluation. The accuracy of 'Standard monitoring' is 82% and the accuracy of 'Enhanced monitoring' is 78%, which indicates a relatively close accuracy of the two states.

Table 7.4 Confusion matrix of DBN model for self-management of RA for appointment advice with threshold of 40%.

Appointment Advice	Standard monitoring (no major flare)	Enhanced monitoring (major flare)
$40\% \leq$ Prediction	3700	35
Prediction $< 40\%$	817	125

Choosing another smaller threshold of 20% increases the overall accuracy up to 85%, as presented in Table 7.5. This results in an increase of the accuracy of 'Standard monitoring' advice (85%), but the accuracy of 'Enhanced monitoring' decreases to 75%.

Table 7.5 Confusion matrix of DBN model for self-management of RA for appointment advice with threshold of 30%.

Appointment Advice	Standard monitoring (no major flare)	Enhanced monitoring (major flare)
$30\% \leq$ Prediction	3828	40
Prediction $< 30\%$	689	120

As shown in Figure 7.7, we visualise the probabilities of appointment advice of six cases with histograms and highlight the threshold of 0.6 probability with a blue line in the middle. The darker plum colours represent the probability of 'Enhanced monitoring' and the lighter plum colour shows the probability of 'Standard monitoring'. For each case, as well as the bar chart of the probabilities of the advice variable, we also show a line plot of the interpolated DAS28 data. The red and green dashed lines respectively represent 5.1 and 2.6 values, indicating the thresholds of the high DAS28 and remission, respectively.

In Figure 7.7a, the histogram of appointment advice demonstrates a higher probability for 'Enhanced monitoring', which corresponds to the high DAS28 in the first couple of weeks. Later, the probability of standard monitoring increases to 58% (below the blue line), and we can see that the DAS28 becomes moderate (between the red and green lines). The 'Enhanced monitoring' is advised again after Week 32 since a flare occurs and DAS28 also rises above the red line again and levels off.

Figure 7.7b advises an 'Enhanced monitoring' for three weeks in the beginning, that is because DAS28 is in a high state. Once DAS28 diminishes to moderate and close to remission states, the probability of standard monitoring increases, except in two cases in Weeks 15 and 30 that DAS28 slightly increases and that rises the probability of 'Enhanced monitoring' above the blue line.

In Figure 7.7c, the 'Enhanced monitoring' is recommended in the beginning, but the standard monitoring is advised after Week 11. At the same time, the person enters into remission and stays there almost stably. We can see a slight rise of DAS28 after Week 36 which increases the probability of 'Enhanced monitoring', but still the standard monitoring is the dominant advice.

A high DAS28 is shown in Figure 7.7d that decreases initially but then rises slightly and then drops towards a moderate state. The slight rise of DAS28 is known as a flare in high DAS28 states, which increases the probability of 'Enhanced monitoring' beyond 60% after Week 30 until a substantial decrease in DAS28.

Figure 7.7e displays a fluctuating monitoring advice and DAS28. In the first weeks, DAS28 is high but 'Enhanced monitoring' is not recommended. Once DAS28 is diminished and stabilised, the standard monitoring becomes the dominant advice. The 'Enhanced

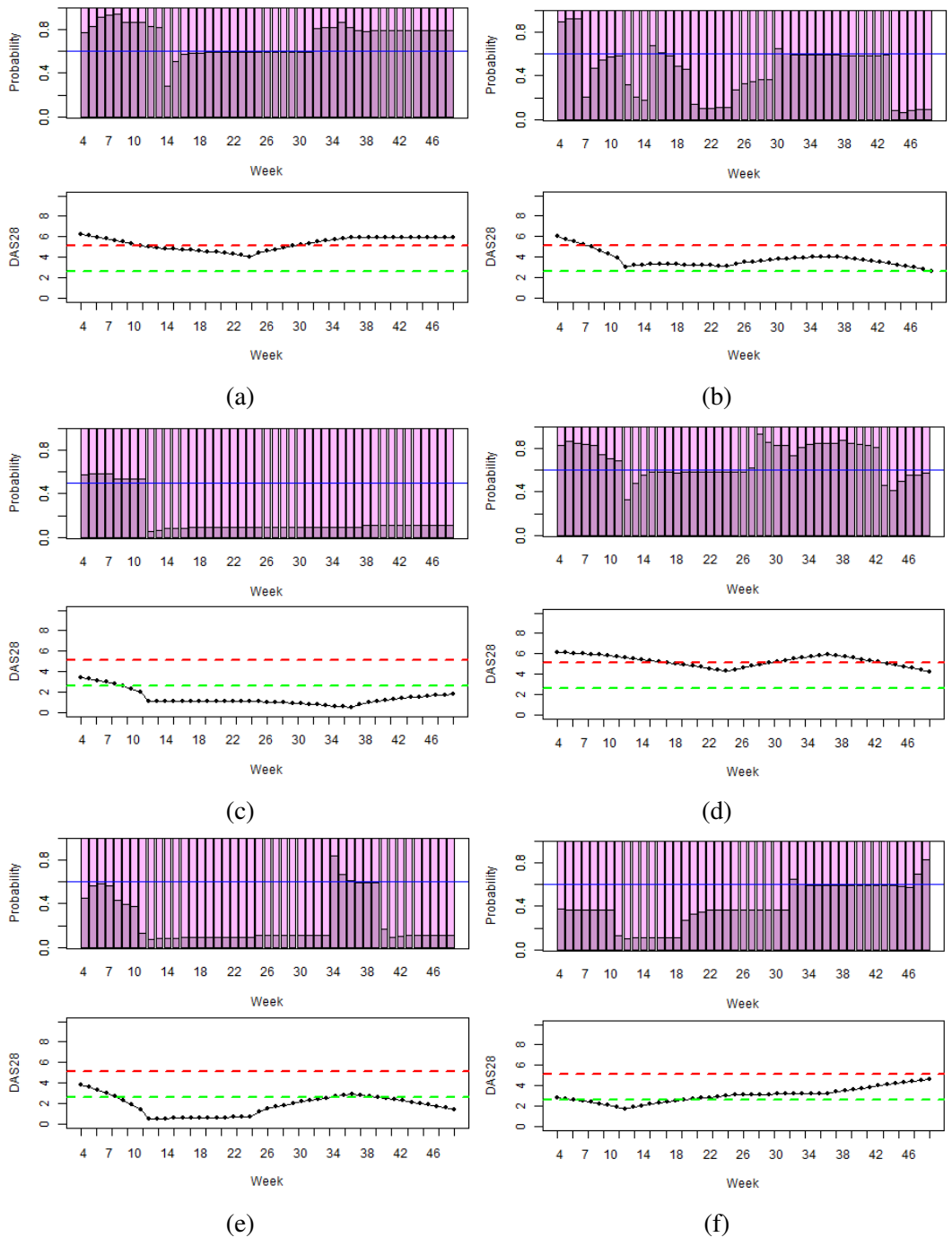


Fig. 7.7 Probability of appointment advice corresponding to the interpolated DAS28 data.

monitoring' is recommended in Weeks 34, and afterwards until Week 39, and we can see that a rise of DAS28 or a flare is experienced almost at the same time, between Weeks 25 and 38.

A case with a mild disease activity in the beginning is demonstrated in Figure 7.7f. Standard monitoring is advised up to a point that DAS28 grows beyond the remission

threshold. Later, DAS28 gradually rises and causes at least one flare. This is projected on the rise of 'Enhanced monitoring' probability after Week 32.

7.5.2 Evaluation with Scenarios of Dummy Patients

Since the ideal data to evaluate the DBN model does not exist, we asked our experts to create a set of scenarios of dummy patients. We evaluate the DBN model for self-management of RA using these scenarios. As explained in Subsection 5.4.2, we evaluated the diagnostic model with 20 scenarios of dummy patients. Of them, 15 scenarios were expected to have RA as indicated in Table 5.13. We considered at least four follow-up visits after the diagnosis visit. Each follow-up visit has a table of observations (including signs, symptoms, and serology results) and another table for experts' expected disease activity and monitoring interval. As an example, the observations and expected outcome of the dummy patient 1 in follow-up visit 1 is presented in Table B.2.

The design of the scenarios is based on the current practice, i.e., dummy patients would be monitored in fixed interval appointments, except in the beginning of their treatment or in the case of flares. However, we need more frequent data of weekly. Like PEAC data, we interpolated the data of each dummy patient to expand their granularity and create a weekly dataset matching the DBN model (see Section 7.4 for details). We also used the DBN model with the same size of the sliding window (containing 5 time slices), as described in the beginning of this section. We use the API of the AgenaRisk Developer [3] to give the inputs of the model (observations) and receive the outcomes on monitoring. The 15 dummy patients have at least five follow-up records comprising a total of 89 appointment scheduling cases, either 'Standard monitoring' or 'Enhanced monitoring'. Of 89, 23 have an enhanced appointment (4, 6, or 8 weeks apart) scheduled and 66 are standard (13, 17, or 26 weeks apart).

For example, we differentiate the probabilities of the standard and 'Enhanced monitoring' with a threshold of 0.5 or 50%. As shown in Table 7.6, the overall accuracy is 60%. The 'Standard monitoring' and 'Enhanced monitoring' obtain 48% and 91%, respectively.

Table 7.6 Confusion matrix of DBN model for self-management of RA for appointment advice using dummy patient scenarios.

Prediction of advice for	True decisions	
appointment scheduling	Standard monitoring (no major flare)	Enhanced monitoring (major flare)
$50\% \leq$ Prediction	32	2
Prediction $< 50\%$	34	21

Using an alternative threshold of 0.4 or 40% to differentiate the standard and enhanced monitoring predictions, the confusion matrix becomes as shown in Table 7.7. The overall accuracy rises to 78% and the accuracy of 'Standard monitoring' and 'Enhanced monitoring' obtain 83% and 56%, respectively.

Table 7.7 Alternative confusion matrix of DBN model for self-management of RA for appointment scheduling advice using dummy patient scenarios.

Prediction of advice for appointment scheduling	True decisions	
	Standard monitoring (no major flare)	Enhanced monitoring (major flare)
40% $\leq$ Prediction	55	9
Prediction $<$ 40%	11	14

Table 7.8 Accurately predicted monitoring advice separated based on the follow-up visit numbers.

Follow-up visit	Prediction of advice for appointment scheduling	
	Standard monitoring	Enhanced monitoring
1	1/4(25%)	10/11(91%)
2	7/11(64%)	2/4(50%)
3	10/11(91%)	1/4(25%)
4	12/12(100%)	1/3(33%)
5	12/13(92%)	0/2(0%)
6	13/13(100%)	0/1(0%)

Table 7.8 expresses the accurately predicted monitoring advice separated based on the follow-up visit numbers of 1 to 6. The table shows how the current parameters perform better in the Follow-up 4 onward, compared to those in the first three follow-ups. We use threshold of 40% to differentiate the predictions of the standard and enhance monitoring since it has a slightly greater overall accuracy than the threshold of 50%. We see a decline of the accuracy of the 'Enhanced monitoring' and a rise of the accuracy of the 'Standard monitoring' as follow-up visits move forward, although the number of cases are limited and difficult to judge.

7.6 Summary

This chapter shows how to build an initial DBN model for self-management of RA that corresponds to the Objective 3 of this thesis. The model includes various variables of signs,

symptoms, and serology results that estimate the disease state and the occurrence of flares. In the model, we consider two features of flares: flare duration and flare frequency. The estimation of disease state and flares allows us to predict if the disease is under control or not. The trend of disease state and flare as well as the prediction of disease control lead us to a monitoring advice whether standard monitoring is enough or enhanced monitoring is needed. The former refers to the self-management of disease by the patient and a need for standard monitoring (e.g., every 6 months), whereas the latter advice implies more frequent or urgent investigation. To parameterise the model, we used a subset of the PEAC dataset to train the parameters and where unavailable, we extracted from the data or elicited expert's knowledge, or used information published in the literature.

We explained how the DBN model for self-management of RA operates as a decision support system for clinicians and patients. It required active involvement of patients to provide their observations of signs and symptoms, and serology results to be given through an electronic health system. Since the frequent data for reasoning and evaluating of the model does not exist, we interpolated the available data and expanded its granularity. Without having access to a gold standard to evaluate the performance of the model, we used proxy records and interpolated values to compare the predictions with and evaluate the performance of the model. We also used a set of dummy patient data to further evaluate the model. In both evaluation approaches, we obtained promising results in predicting the monitoring advice.

The contribution self-management modelling was limited with a small data set that is less granular than what we needed. Although we interpolate the data to expand it (see Subsections 7.3.2 and 4.4.4), we train and reason with the same limited number of cases in the data set. Another limitation was the small number of dedicated experts who were involved in the creating the structure of BN model and model evaluation with scenarios.

The DBN model for self-management of RA is an initial model that excludes any variables related to medication and provides advice on monitoring scheduling. In the next chapter, we extend the DBN model by adding medications and adverse events that can provide advice on medication review.

Chapter 8

Building Dynamic Bayesian Network Models for Self-Management of Rheumatoid Arthritis: Medication Review

In Chapter 7, we demonstrated how to use medical knowledge and data to build a dynamic Bayesian network (DBN) model for self-management of Rheumatoid Arthritis (RA) to give advice on appointment scheduling. The current chapter builds on Chapter 7, covering the treatment of RA as a part of the self-management of RA, leading to support clinicians in reviewing medications. In this chapter, we show how to employ the available data, experts' knowledge, and clinical guidelines to create a DBN-based decision support for self-management of RA, which corresponds to the Objective 3. We then use an interpolated dataset and the data of a set of dummy patients to do reasoning with the DBN model for self-management of RA and evaluate the performance of DBN model in generating advice for medication review (Objective 3). The interpolated data and the dummy patient data sets are the same as those in Chapter 7, but both of them consist of medications and adverse medication events (AMEs).

The outline of this chapter is as follows: Section 8.1 introduces the management of RA and the DBN model for self-management of RA, aiming to provide advice on medication review. In Section 8.2, we explain the guidelines for the medications of RA, including a description of conventional disease-modifying antirheumatic drugs (csDMARDs), targeted synthetic and biological disease-modifying antirheumatic drugs (ts/bDMARDs), and steroids. Section 8.3 describes the treatment strategies for RA, treatment regimens for RA, and decision-making on the treatment of RA. In Section 8.4, we describe the variables for the

treatment of RA, including medications, AMEs, and advice on medication review. Section 8.5 explains the DBN model for self-management of RA, focusing on medication review. This section consists of an explanation of the data, the structure of DBN, and the parameterisation of DBN. A description of the DBN-based decision support for self-management of RA is presented in Section 8.6. It demonstrates how a DBN model can be used to support clinicians in medication review. In Section 8.7, we do reasoning with the DBN model and so evaluate its performance. Finally, we summarise the chapter in Section 8.8.

8.1 Introduction

The monitoring of RA is to track the disease state and the effects or side-effects of the medications. These medications need to be reviewed from time to time to make sure about their efficacy and the lack of adverse events. The National Institute for Health and Care Excellence (NICE) recommends a pathway for the management of RA [140] based the treat-to-target strategy (as shown in Figure 4.2) and provides multiple guidelines, describing the medications for RA, their potential effects, and contraindications, e.g., [137, 135, 139]. Two models of care pathways for the initial management and ongoing management of RA (shown in Figures 4.6 and 4.7) were created based on the clinical guidelines and the knowledge elicited from our experts.

By reviewing these models of pathways and analysing the available data (see Section 4.4), we selected four decision support points, namely, 'Book Follow-up Appointment', 'New Appointment?', 'Change Meds?', and 'Review Blood Test Frequency', as listed in Table 4.6. In Chapter 7, we built a DBN model for self-management of RA to recommend appropriate advice on appointment scheduling, addressing the first two selected decision support points, i.e., 'Book Follow-up Appointment' and 'New Appointment?'. In the current chapter, we show how we extend the DBN model for self-management of RA built in Chapter 7, considering all above-mentioned selected decision support points. The extended DBN model contains medication and AME variables to estimate the disease state and flare. The main medications prescribed to manage RA are called disease-modifying anti-rheumatic drugs (DMARDs) that are categorised into conventional, targeted synthetic, and biological. Steroids are also prescribed to quickly reduce the severity of disease. These medications need to be prescribed in a way that leads to a low disease state or remission, but do not cause AMEs, including organ damage, symptomatic side-effects, or infections.

The aim of this DBN model is to generate advice on medication review, whether the same regimen should be continued, or whether changes of medication or medication dosage are needed. Advice on medication review implies advice on appointment scheduling of the DBN

model in Chapter 7, since the change of medication or medication dosage would occur in an enhanced monitoring appointment.

8.2 Medications for Rheumatoid Arthritis

DMARDs are the main medications for adult persons with RA. As we briefly described in Subsection 4.4.5, DMARDs can be classified into three groups of csDMARDs, targeted synthetic DMARDs (tsDMARDs), and biological DMARDs (bsDMARDs). Steroids or glucocorticoids are also prescribed to adults with RA. In the following subsection, we describe the csDMARDs, targeted synthetic and biological DMARDs (ts/bDMARDs), and steroids.

8.2.1 Conventional Disease-Modifying Anti-Rheumatic Drugs

csDMARDs are synthetic drugs to modify RA disease, rather than to alleviate its symptoms [138]. It is prescribed to the person with inflammatory arthritic diseases including RA. The onset of csDMARDs is slow and may take 2 to 3 months to respond [138]. csDMARDs are suitable for first-line treatment of RA. Three main csDMARDs are Methotrexate (MTX), Sulfasalazine (SSZ), and Hydroxychloroquine (HCQ), which were used the Pathobiology of Early Arthritis Cohort (PEAC) study and are most commonly prescribed in practice, and other csDMARDs – leflunomide (LEF) and azathioprine (AZA) – are rarely offered (see Subsection 4.4.5). Of csDMARDs, HCQ is less effective and recommended for those with mild disease or palindromic disease - an inflammatory arthritis that causes attacks of joint pain and swelling like RA [138].

csDMARDs of MTX, SSZ, and HCQ include almost 95% of the prescribed DMARDs, according to our main rheumatology expert. MTX comes in three forms of tablets, drinkable liquid, or pre-filled injection pens or syringes. MTX should be taken weekly, its doses vary between 5mg to 25mg, and it is usually advised to be escalated by 2.5mg. SSZ is usually tablets of 500mg which should be taken daily. It is usually prescribed between 500mg to 2000mg. Like SSZ, HCQ is taken daily, but its tablets are 200mg and may escalate only up to 400mg.

Clinical practice and experience indicate the incompatibility of csDMARDs with specific comorbidities. For example, if a patient has a lung disease, MTX is not a safe prescription. In the case of family planning, SSZ needs to be stopped and replaced. Different csDMARDs and their incompatibilities and possible side-effects express the complexity of decision-

making for RA treatment with csDMARDs that is reflected in building a decision support for self-management of RA with treatment factors involved in it.

8.2.2 Targeted Synthetic and Biological Disease-Modifying Anti-Rheumatic Drugs

tsDMARDs are synthesised drugs developed to target a particular molecular structure [91]. The main tsDMARDs are tofacitinib, fostamatinib, baricitinib, and apremilast, which are offered for RA, whereas imatinib or ibrutinib agents are not focused primarily on rheumatic diseases [91]. Of them, the NICE guideline for treatment of RA [140] recommends tofacitinib and baricitinib to be prescribed with MTX to treat active RA patients who have not responded adequately to a combination of csDMARDs, if the disease activity is high ($DAS28 > 5.1$). These drugs can be prescribed without MTX, if it is contraindicated or the person cannot tolerate it. These two drugs should be continued if a moderate response using the European League Against Rheumatism (EULAR) criteria was achieved after 6 months of therapy onset, otherwise they can be withdrawn [140].

bDMARDs or biological DMARDs are single-dose drugs and need to be injected. Some bDMARDs are called first-line, namely, Adalimumab (ADA), Etanercept (ETN), Certolizumab (CTZ), Golimumab (GMB), and Tocilizumab (TCZ). The NICE guideline [140] recommends to prescribe these bDMARDs with MTX for the treatment of those with RA disease on high DAS28 who have not responded to intensive therapy with a combination of csDMARDs. The bDMARDs, except GMB, cannot be used as monotherapy for those who cannot take MTX due to contraindication or intolerance [140]. All these bDMARDs, except TCZ, are anti-TNF (Tumour Necrosis Factor-alpha) which reduces the inflammation. Rituximab (RTX) is not a first-line bDMARD, meaning it cannot be prescribed before other bDMARDs are prescribed. Therefore, RTX can be prescribed after the prescription of at least one first-line bDMARD.

The main bDMARDs, their characteristics, contraindications, and common AMEs are listed as follows:

- ADA is administered at 40mg dose fortnightly via subcutaneous injection. In monotherapy of ADA, it may be increased to 40 mg every week, if the patient experiences a decrease in response. ADA is frequently offered and is the most successful bDMARD. ADA is contraindicated in people with moderate to severe heart failure, those with active tuberculosis or those with other severe or opportunistic infections [135], and patients with cancer due to anti-TNF. ADA's common adverse events are injection-site reactions and infections [135].

- ETN is administered by subcutaneous injection at 25mg dose twice weekly [135]. Its common adverse events are injection-site reactions, infections, and allergic reactions. Clinical knowledge indicates that ETN may be more effective, if infections are foreseen. ETN is contraindicated in people with sepsis or risk of sepsis, those with other active infections [135], and current or previous cancer patients. Common adverse effects of ETN are injection-site reactions, infections and allergic reactions [135].

- CTZ is administered subcutaneously as initial 400mg doses at 0, 2, and 4 weeks, followed by maintenance doses of 200mg every 2 weeks. Alternatively, it can be administered 400mg every 4 weeks, if clinical response is confirmed [135, 137]. CTZ is contraindicated in people with active tuberculosis or other severe infections, those with moderate or severe heart failure [135], and current or previous cancer patients because of anti-TNF. CTZ's most common adverse events in clinical trials are bacterial and viral infections [135].

- GMB is administered subcutaneously as a 50mg dose every month on the same day each month [135]. A dose of 100mg may be considered for persons more than 100kg weight, if the disease has an inadequate clinical response after 3–4 doses. According to our main expert, GMB is effective in patients with high BMI. Contraindications of GMB are in people with active tuberculosis or other severe infections, those with moderate or severe heart failure [135], and people with current or history of cancer because of anti-TNF factor. The most commonly reported adverse reactions of GMB are upper respiratory tract infections [135].

- TCZ is administered as an intravenous infusion, given over 1 hour and its recommended dosage is 8mg/kg, given once every 4 weeks [136]. TCZ is recommended for seronegative patients, but it can be offered to seropositive patients as well. It is not recommended for people with more than 100 kg weight since the doses exceed 800mg per infusion. TCZ is contraindicated in people with active, severe infections. Its common adverse reactions are upper respiratory tract infections, nasopharyngitis, headache, hypertension and increased alanine transaminase [136].

- RTX includes two 1000mg intravenous infusions given 2 weeks apart, and it should be given at intervals of no less than 16 weeks [135]. RTX is known to be effective in seropositive patients. RTX is contraindicated in people with severe heart failure or severe, uncontrolled cardiac disease, and those with active, severe infections. RTX's common adverse effects are infusion reactions and infections [135].

- Abatacept (ABA) is administered as a 30-minute intravenous infusion, and it should be repeated at week 2, week 4 and every 4 weeks thereafter, i.e. 14 infusions are needed in the first year and 13 infusions in the next years [135]. ABA's common adverse effects are infections of sepsis and pneumonia. ABA is contraindicated in people with severe, uncontrolled infections, and opportunistic infections [135]. ABA is the least frequently offered bDMARD.

Various ts/bDMARDs, their expected effects and side-effects, as well as their contraindications reveal the complexity of making a decision for RA treatment and the complexity of providing decision support to review the medications for RA.

8.2.3 Steroids

Since all DMARDs have a slow onset action and their response may not be seen in 2 to 3 months, clinicians are recommended to offer bridging treatments [138, 139]. Bridging treatment refers to using glucocorticoids or steroids that are anti-inflammatory drugs offered to relieve joint pain and swelling immediately [138] – within a week. Steroids are recommended for a short period of time to improve symptoms and manage flares while waiting for the new DMARD to become effective. Steroids are highly toxic, and their long-term use can cause adverse reactions. They may be offered for long-term use to adults with established RA, if the long-term complications of them are discussed and all other drugs including tsDMARDs and bDMARDs were offered before [138].

The main steroids for RA treatment in the PEAC study are depo-medrone and prednisone, as mentioned in Subsection 4.4.5. Steroids are categorised into three groups of oral, intramuscular (IM) (including depo-medrone), and intra-articular (IA) [140]. These types of steroids are offered based on the type of joint inflammation and pain as well as patient's preference. Their possible side-effects need also to be considered since they are known as very toxic drugs. These different factors express how complex it is to prescribe effective drugs and keep patients safe from side-effects as well.

8.3 Strategies, Regimens, and Decisions for Treatment of Rheumatoid Arthritis

To learn about the treatment of RA, we have reviewed the NICE guidelines to investigate the strategies for the treatment of RA. We have also elicited knowledge from our experts and merged the extracted knowledge with strategies for the treatment of RA. These led us

to devise treatment regimens and decision-making on regimens (see Section 4.4.5). In the following parts, we describe the strategies, regimens, and decision-making on regimens.

8.3.1 Treatment Strategies

The pathway of NICE guideline for the treatment of RA (displayed in Figure 4.2) suggests that adults with RA receive treatment based on the treat-to-target strategy. This strategy aims to treat active RA to achieve a target of remission or low disease activity, if remission cannot be achieved.

The treat-to-target strategy applies tight control, e.g., monthly follow-up visits and respective treatment adjustments, until reaching the target [138]. People who the treat-to-target strategy is applied on them should be checked to measure the C-reactive protein (CRP) and disease state monthly in specialist care till the target is achieved [140]. Multiple csDMARDs and ts/bDMARDs with different mechanisms of action and specific order may involve in achieving the target.

The initial treatment of RA includes a monotherapy strategy that is the first-line treatment with a single csDMARD therapy [140]. The strategy of sequential monotherapy is a treatment commencing with a single csDMARD which is replaced with a different single csDMARD, in the case of inadequate response [139]. The strategy of step-up therapy is additional DMARDs added to the DMARD monotherapy when disease is not adequately controlled [139]. Evidence supports the priority of step-up strategy over sequential monotherapy, but economic analysis supports the reverse priority [138]. The strategy of parallel combination therapy is two or more DMARDs commenced at the same time without a step-down strategy [139]. If DMARDs become adequate, the strategy of step-down therapy can be applied, which is tapering and stopping at least one drug, once the disease was adequately controlled during the treatment with two or more DMARDs [139]. The NICE guideline [138] recommends considering the step-down strategy only after a person has maintained the treatment target for at least one year without using steroids.

If csDMARDs failed to lead to an adequate response, the first-line bDMARDs would be offered considering the risk factors, comorbidities, and preferences of the person. In the case of having an inadequate response or intolerance to first-line bDMARDs, RTX can be offered, if the person has no contraindications to RTX and has tolerance to RTX. An adequate response refers to an improvement in DAS28 of 1.2 points or more 6 months after initiation of therapy [140]. If RTX is not suitable and other bDMARDs have not had adequate responses, then steroids or other therapies can be used to control the disease.

8.3.2 Treatment Regimens

We process long-term treatment strategies to define treatment regimens as plans or courses of action for the treatment of RA in short-term. Each regimen includes one medication or a combination of multiple medications. We found out the most common regimens for RA treatment are: single csDMARD, combined csDMARDs, single bDMARDs, and combined csDMARDs and bDMARDs. Each regimen may have steroids or not depending on the patient experiencing a flare or their preference.

- The single csDMARD regimen refers to the treatment with one single csDMARD, mainly MTX, SSZ, and HCQ. This regimen is usually applied in the beginning of treatment as an initial treatment. All regimens start with a lower dose in the beginning and prescribed an incremental escalation in the upcoming weeks.
- The combined csDMARDs is a common regimen which is a combination of csDMARDs, mainly MTX with SSZ, MTX with HCQ, or SSZ with HCQ. It is recommended to offer combined csDMARDs after a monotherapy of csDMARDs, but in some cases combined csDMARDs may be offered from the beginning of treatment to reduce the disease activity. A similar regimen is applied to those who get prescribed one type of csDMARD in the baseline, but they may receive a prescription of combined csDMARDs, in the case of having no side-effects or other issues before.
- The regimen of combined csDMARDs and bDMARDs is the recommendation of the NICE guidelines to prescribe bDMARDs with MTX or other csDMARDs, single or combined. This aggressive regimen can control aggressive disease activity. It starts with a first-line bDMARD and may be advanced to bDMARDs like RTX in the next steps.
- The single bDMARD regimen may be offered, if the person cannot receive csDMARDs combined with bDMARDs. The prescription of a single bDMARD is a regimen that is relatively rare in the first year after diagnosis.

Treatment regimens for RA provide us with information on decision-making for the treatment of RA. Treatment regimens help us to define the medication variables and their states, as described in Section 8.4.

8.3.3 Treatment Decision-Making

Treatment strategies and treatment regimens help us to define the treatment decisions for RA. We consider three main decisions for the treatment regimen: (1) prescribing, (2) tapering,

and (3) keeping the same regimen. Prescribing decision starts with medication selection. csDMARDs are prescribed in the beginning of treatment and may need to be offered with steroids. In the case of csDMARDs, clinicians need to make a decision on dose, dose escalation steps, and the maximum dose. These detailed decisions should be made based on the guidelines, but clinicians may consider their expertise in applying these guidelines.

If disease activity is high and severe adverse events have happened, then clinicians may not go back to the previous drugs. In the case of bDMARDs, they cannot prescribe previously given (ineffective or unsuccessful) bDMARDs. According to our main expert, the main comorbidities to consider for prescribing or tapering are infections, namely, tuberculosis (TB), human immunodeficiency virus (HIV), Chickenpox, hepatitis B and C.

Tapering decisions are made when a person is prescribed multiple medications and their disease activity is managed. Corresponding to the step-down strategy, decision-making on tapering may optimistically happen one year after the onset of the first therapy. The first decision on tapering is whether a person is eligible for tapering or not. Being on remission, having a stable disease progression, and having no adverse events associated with tapering eligibility. Tapering may be applied to those on a regimen containing bDMARDs, but persons who are on a regimen with csDMARDs may be tapered as well. In the case of tapering of csDMARDs, tapering decisions include tapering dose steps and the minimum dose, whereas bDMARDs exempt these decisions since they are mostly single-dose (except GMB).

Similar to the prescription decisions, clinicians need to make a multicriteria decision for tapering by considering disease activity, disease progression trend, and adverse events. In the case of prescribing bDMARDs, they should consider the criterion of expensiveness of bDMARDs. If a patient has a long list of failed prescriptions, they may have a severe disease or their tolerance to drugs is low. These cases are usually difficult to taper.

The treatment decision of keeping the same regimen is recommended for patients who have an active disease. This decision can be changed to a tapering or escalation decision if the clinicians reach a conclusion that the regimen is not effective to lead patients to remission or at least lower disease activity.

An additional decision for the treatment of RA is in the case that AMEs emerge. As mentioned before, AMEs can be of three main categories of organ damage, symptomatic side-effects, or infection. Severe AMEs need immediate stop of medication, but moderate ones can be dealt with by reducing the medication dosage. Moderate and stable AMEs can be ignored.

Treatment decisions lead us to define the states of the advice variable for medication review, as we later describe in Section 8.4. Decisions also help to build the structure and parameterise the DBN model that will be presented in Section 8.5.

8.4 Description of Variables for Treatment of Rheumatoid Arthritis

We specified the variables for self-management of RA for monitoring scheduling in Section 7.2. In this section, we consider the same evidence and latent variables presented in Subsections 7.2.1 and 7.2.2 and we add a set of new evidence and latent variables, and we introduce a new advice variable related to medication review.

8.4.1 Evidence Variables

The evidence variables include the evidence variables of the initial DBN model for self-management (described in Section 7.2.1) as well as the variables for medication and adverse events. Medication variables consist of three variables: ts/bDMARDs, csDMARDs, and Steroids.

We consider the ts/bDMARDs variable to have three states of None, 'Extended interval', and 'Standard interval'. The None state means no ts/bDMARDs were taken or prescribed. Since almost all ts/bDMARDs are single-dose drugs (except for GMB), we define the 'Extended interval' as a tapered dose. It refers to waiting longer than the standard break between two injections of ts/bDMARDs, which the extension usually means 50% of the standard break between two injections. The 'Standard interval' refers to the standard break between two injections of each ts/bDMARDs.

The csDMARDs variable represents the common csDMARDs of MTX, SSZ, and HCQ. It has three states of None, 'Low dose' and 'High dose'. The None state means taking or being prescribed no csDMARDs. The 'Low dose' refers to the low dose of MTX (5mg,

Table 8.1 Summary of evidence variables of medication and AME for treatment of RA.

Category	Variable	States	Source	Description
Medication	ts/bDMARDs	None, Extended interval, Standard interval	Expert	Targeted synthetic and biological DMARDs including ADA, ETN, CTZ, TCZ, and RTX prescribed to be taken weekly
	csDMARDs	None, Low dose, High dose	Data	csDMARDs of MTX, SSZ, and HCQ prescribed to be taken weekly
	Steroids	No, Yes	Data	Steroids or glucocorticoids prescribed to manage high disease activity or flare cases in short-term
AME	Organ Damage	None, Mild/Moderate, Severe	Data	Adverse events of damaging organs such as liver, lungs, and blood which can be detected by getting tests
	Symptomatic Side-Effects	None, Mild/Moderate, Severe	Data	Adverse events of systematic side-effect which can be detected and reported by patients
	Infection	None, Mild/Moderate, Severe	Data	Contracting an infectious disease

7.5mg, 10mg, 12.5mg, or 15mg) or the low dose of SSZ (500mg, 1000mg, or 1500mg) or any dose of HCQ (200mg or 400mg). The 'High dose' state of csDMARDs variable refers to high MTX (17.5mg, 20mg, 22.5mg, or 25mg) or high SSZ (2000mg or more).

The Steroids variable represents the glucocorticoids. It has two states of Yes and No meaning steroids were prescribed and taken or not, respectively. Steroids help to manage the high disease activity or flare quickly within a week, whereas ts/bDMARDs or csDMARDs become effective in about 3 months.

Steroids, ts/bDMARDs, and csDMARDs have toxicity and can develop different adverse events in RA patients. Some adverse events can damage organs, which are usually asymptomatic but can also be symptomatic sometimes. For example, liver injury, white blood cell (WBC) reduction, lung issues, and ocular side-effects. Another category of adverse events is symptomatic, such as hair loss or headache. In addition, ts/bDMARDs and csDMARDs can lead to increase the risk of getting serious infectious diseases [167]. We specify three variables called 'Organ Damage', 'Symptomatic Side-Effect', and Infection to represent the above-mentioned adverse events each of which has three states: None, Mild/Moderate, and Severe.

The 'Organ Damage' variable covers adverse events that are not easily detectable by patients and need to be tested to be detected. The 'Organ Damage' is None in the case of having no damaging adverse events, mainly if the alanine aminotransferase (ALT) enzyme of the liver is less than 35. It is Mild/Moderate, if the ALT is equal or greater than 35 and less than 100, the alkaline phosphatase (ALP) enzyme of liver changes, the count of lymphocytes WBC reduces, neutrophils WBC stays between normal (none) and 1, any chest infection of lungs that does not require hospital admission, and asymptomatic ocular side-effects detectable by an eye test. The Severe state of 'Organ Damage' includes: liver ALT being more than 100, neutrophils WBC being less than 1 unit, any breathlessness or coughing which requires hospital admission, the reduction in lung function tests or recurrent chest infections, and symptomatic ocular side-effects detectable by an eye test. In all cases above, the aspartate aminotransferase (AST) is considered combined with ALT.

The 'Symptomatic Side-Effects' variable refers to adverse events that are reported by persons that do not necessarily require a test to detect them. The 'Symptomatic Side-Effects' is normally on the None state. The main case of Mild/Moderate state is the skin irritation due to drug injection. The Severe state of 'Symptomatic Side-Effects' includes headache, which is mainly caused by SSZ, skin reaction as a severe rash, or any blood problems.

The Mild/Moderate state of the Infection variable refers to urinary infection, chest infection, and recurrent infections. A Severe state of the Infection variable can prevent

a patient from being able to take ts/bDMARDs, and the most prevalent severe case is osteomyelitis.

The Severe state of adverse events means the drugs must be stopped, immediately investigated, and new drugs prescribed. The Mild/Moderate state means the drug dose needs to reduce and the result should be observed carefully. In the case of the Mild/Moderate Infection, patients may not be able to use ts/bDMARDs. If a patient is taking antibiotics for their infection, they are not able to take ts/bDMARDs. After antibiotics, they need a week of time to be able to take ts/bDMARDs.

A summary of the evidence variables of medication and AME is presented in Table 8.1.

8.4.2 Latent Variables

Latent variables of the self-management of RA entail seven latent variables of the initial self-management model of RA, namely, 'Disease State', 'Disease Progression Trend', 'Flare Occurrence', 'Flare Duration', 'Flare Frequency', 'Overall Flare', and 'Overall Flare Trend', that we described in Subsection 7.2.2.

We add five new variables and amend one existing one. The five new variables are: AME, 'AME Change', 'AME Trend', 'Tolerance for DMARDs', and 'Overall Disease Progression Trend'. The AME variable represents any adverse event caused by RA medications that can be manifested as an organ damage, a symptomatic side-effect, or an infection. It has three states of None, Mild/Moderate, and Severe. We define a variable called 'AME Change' that compares AME variables of its time slice with those in the previous time slice to scale increase, steadiness, or decrease of the AMEs. This helps to capture the trend of AMEs by

Table 8.2 Summary of latent variables of medication and AME for self-management of RA.

Category	Variable	States	Source	Description
Adverse Events	AME	None, Mild/Moderate, Severe	Data	Adverse events caused by the medications manifests as organ damage, symptomatic side-effects, and infection
	AME Change	Decreasing, Steady, Increasing	Expert	Change of AME in the current time compared to that of the previous time
	AME Trend	Deteriorating, Stable, Improving	Expert	The trend of AMEs combining the change of AME in two consecutive times and the trend of AME of the previous time
	Tolerance for DMARDs	Low, Medium, High	Expert	Tolerance for taking csDMARDs and ts/bDMARDs without having a mild, moderate, or severe AME and having stable (or improving) trend of AME
Disease	Overall Disease Progression Trend	Deteriorating, Stable, Improving	Expert	Combining the trend of disease progression and the trend of overall flare
	Overall Disease Control	Good, Alarm, Poor	Expert	Disease being overall under control, on alarm, or unmanaged, combining tolerance for DMARDs, disease state, and overall flare

combining the change of AME in two time slices and the trend of AME in the previous time slice. For this, we define a variable called 'AME Trend' with three states of Deteriorating, Stable, and Improving. Similar to the change of the disease state and its trend (contrasted in Subsection 7.2.2), the change of AME and the trend of AME are two distinct concepts - the former compares AMEs in two consecutive states of the AME, but the latter combines that comparison with a long-term change of AMEs embedded in the AME trend of the previous time slice.

According to our main expert, different people have different tolerance for csDMARDs and ts/bDMARDs, especially to the latter DMARDs due to its greater toxicity and possible adverse events compared to the former category of DMARDs. To express this concept, we define a latent variable called 'Tolerance for DMARDs' with three states of Low, Medium, and High. This latent variable associates with the evidence variables of csDMARDs and ts/bDMARDs and two latent variables of AME and 'AME Trend'.

We define the 'Overall Disease Progression Trend' variable to combine the two trend variables of 'Disease Progression Trend' and 'Overall Flare Trend'. It has three states of Deteriorating, Stable, and Improving.

We alter the definition of the 'Overall Disease Control' variable from what it was in Subsection 7.2.2. We keep the same state of Good, Alarm, and Poor, but we redefine it as a combination of 'Tolerance for DMARDs', 'Disease State', and 'Overall Flare'.

Table 8.2 summarises the new latent variables for the self-management of RA involving medications and adverse events.

8.4.3 Advice Variable

Based on the treatment decisions for RA (described in Subsection 8.3.3), we define an advice variable called 'Medication Review Advice' with four states: Escalation, 'Same regimen', Tapering, and 'Medication change'.

- Escalation advice suggests increasing the dosage of medications to suppress the active disease or deteriorating disease.

Table 8.3 Summary of the advice variable for self-management of RA: medication review.

Category	Variable	States	Source	Description
Advice	Medication Review Advice	Escalation, Same regimen Tapering Medication change	Expert	Advice on medication review based on the tolerance for DMARDs, AME trend, the trend of overall disease progression, and overall disease control

- Tapering advice recommends reducing the dose of medication or stop them since the person's disease is in remission and stays stable.
- 'Same regimen' advice suggests continuing the same medications and dosages, implying them being effective and having no severe AME or mild or moderate but stable AME.
- 'Medication change' advice refers to the cases that medications have caused severe adverse events or medications have been ineffective; therefore, the change of medications is needed to suppress the disease.

The advice variable is associated with the tolerance for DMARDs, the trend of AME, the trend of overall disease progression, and overall disease control. Table 8.3 shows a summary of the advice variable.

8.5 DBN Model for Self-Management of Rheumatoid Arthritis: Medication Review

The structure of DBNs represents the variables and their causal or associational dependencies, and there are underlying parameters for each variable. In this section, we describe the structure of the DBN model for self-management of RA involving medications, the data required to train the parameters, and the parameterisation of the DBN model.

8.5.1 Structure

We extend the structure of the initial DBN model (Figure 7.1). For this, we add medication, AME, and medication review advice variables that were described in Section 8.4. We follow the notation of the simplified DBN model as presented in Subsection 2.3.1.

The structure of the extended DBN model for self-management of RA is shown in Figure 8.1. We used the AgenaRisk software [3] to create the structure of the DBN. There are three areas in the model: initial conditions, temporal plate, and terminal conditions. The temporal plate represents the structure of a BN model that repeats over time, i.e., the BN models in the temporal plate are rolled-up. The variables are represented with ovals. The green ovals are the evidence variables (described in Subsection 8.4.1) that are expected to receive an observation, and the white ones show the latent variables (described in Subsection 8.4.2). The arcs between the variables within the temporal plate represent intra-time-slice dependencies and those with a square and number 1 are the inter-time-slice or temporal dependencies. The

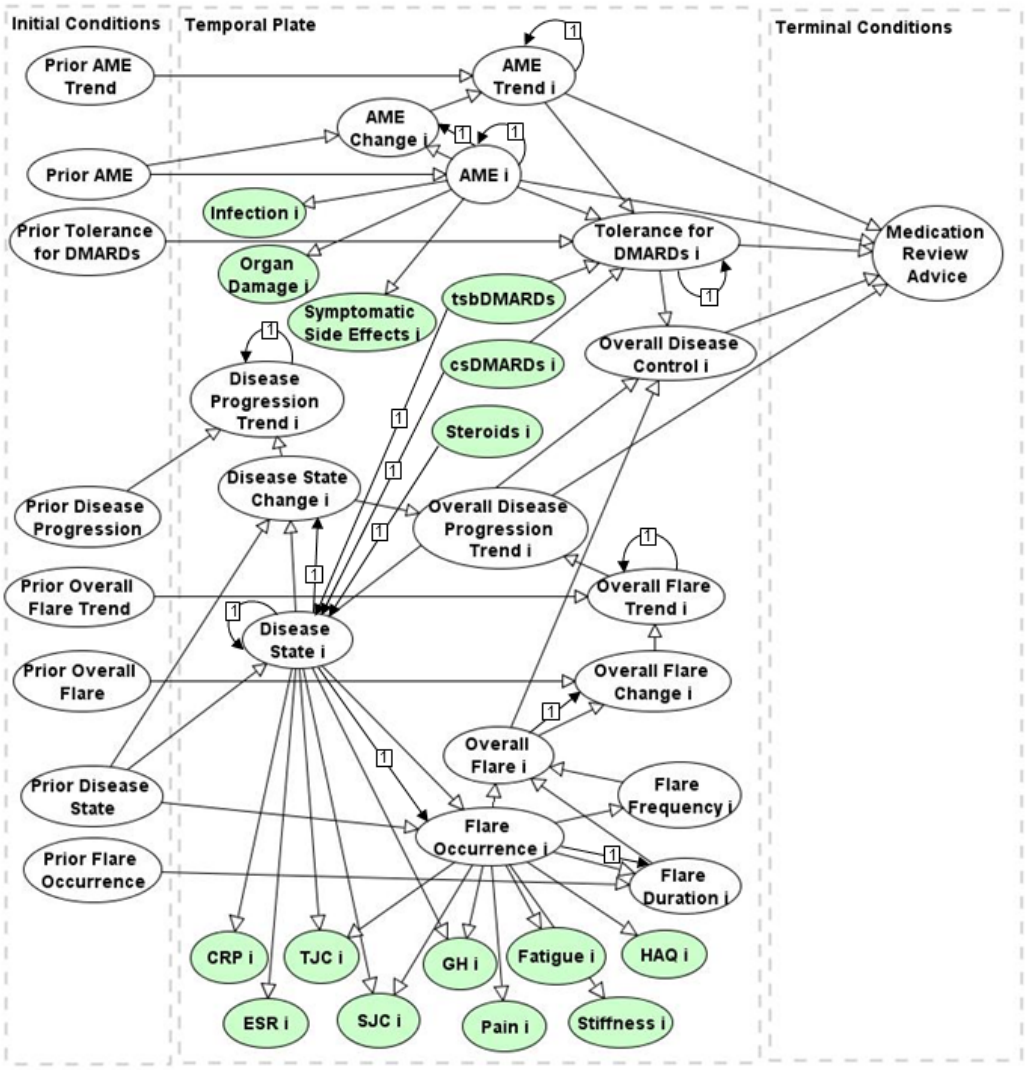


Fig. 8.1 DBN model for self-management of RA for medication review.

variables in the temporal plate have an index *i*, representing the time slice that they belong to. Each index has an underscore in Figure 8.1, which is the necessity of the software; however, we remove the underscore in the manuscript.

This DBN model considers three medication variables as interventional factors, namely, 'ts/bDMARDs i'[1], 'csDMARDs i', and 'Steroids i'. These variables are connected to 'Disease State i' variable as suggested by the treatment idiom (Figure 2.10g). ts/bDMARDs and csDMARDs are prescribed to be taken at the current time, but they become effective in the future knowing the length of time for these medications to show effects (mentioned in Subsection 8.2.3 and 4.3.2). We are aware of the effectiveness of the csDMARDs and

[1]The name of this variable has no slash in the figure since the software does not accept this character.

ts/bDMARDs is rather long-term (e.g., 3 months); however, we did not want to create a complicated DBN model with temporal dependencies across the time slices, i.e., a higher order Markov model.

The 'AME i' variable is associated with three variables of 'Organ Damage i', 'Symptomatic Side-Effects i', and 'Infection i', like the manifestation idiom (Figure 2.10a). The 'AME Change i' variable compares the AMEs of the current time slice with those of the previous time slice and determines whether AMEs are decreased, steady, or increased. The 'AME Change i' is combined with the trend of AME in the previous time slice to specify the 'AME Trend i' variable of the current time slice. The AME, 'AME Trend i', and two variables for DMARD medications ('csDMARDs i' and 'ts/bDMARDs i') are combined to create the 'Tolerance for DMARDs' variable. This variable and the 'Disease State i' and 'Overall Flare i' variables of a time slice lead to estimate whether the disease is under control or not, represented with the 'Overall Disease Control i' variable. Another combination variable is the 'Overall Disease Progression Trend i' variable that aggregates the trend of the disease state and overall flare and gives a unified measure of the disease trend. The terminal conditions include the advice variable for medication review (described in Subsection 8.4.3).

The initial conditions show the prior variables for the latent variables in the temporal plate that have at least one temporal dependency. These variables are needed for reasoning with the DBN model, as will be explained in Section 8.7.

8.5.2 Data

We use a subset of the dataset collected in the PEAC study, as described in Section 4.4. The data are the same as those for the initial DBN model (Subsection 7.3.2), but they include records of medications and AMEs (described in Subsection 4.4.5).

Like in Subsection 7.3.2, we select the data of the first and the second follow-up visits of the PEAC study to parameterise the DBN model. These two follow-up visits contain the highest number of cases excluding the baseline data. The final dataset for parameter learning of the extended DBN model consists of the same variables (e.g., TJC and SJC) as explained in Subsection 7.3.2, plus medications (csDMARDs, ts/bDMARDs, and Steroids), and adverse events (organ damages, symptomatic side-effects, and infections). The dataset has 163 cases and contains 4% missing values.

8.5.3 Parameterisation

To parameterise the DBN model, we use the dataset with 163 cases that is originally collected in the PEAC study, as described in Subsection 8.5.2. The evidence, latent, and advice variables of the DBN model are parameterised as follows:

- **Evidence Variables.** The dataset is a subset of the first and second follow-up visits of the PEAC study containing 36 variables in total for the evidence variables (except ts/bDMARDs), disease state, flare occurrence, flare frequency, and AME. The ts/bDMARD medications are not prescribed in the early visits of the PEAC study as the guidelines suggest starting the prescription of csDMARDs first (see table 4.4). We assume that all cases that had a prescription record of a ts/bDMARD, those with a referral to receive ts/bDMARDs, and those with a refusal of taking ts/bDMARDs are actually given a ts/bDMARD (3 cases in the first follow-up and 25 cases in the second follow-up as shown in Table 4.4). We assumed 1% of patients' ts/bDMARD may get tapered since 3 in 32 cases of the BioT data received a tapering decision (see Figure 4.23).

- **Latent Variables.** As mentioned in Subsection 7.3.3, the parameters of the 'Flare Duration i' are taken from a literature [87], and we defined the ranked variables weighted by our main expert to parameterise the following variables: 'Overall Flare i' , 'Overall Flare Trend i' , and 'Disease Progression Trend i'. We apply the same approach to parameterise the 'Overall Disease Progression Trend i', 'AME Trend i', 'Tolerance for DMARDs i', and 'Overall Disease Control i'. For this, we defined them as the ranked nodes with a Truncated Normal distribution (as suggested by [60]) and a weighted average function to aggregate the parameters of their parent variables. The weights for each variable are defined by the main expert.

- **Advice Variable.** We specify the parameters of the 'Medication Review Advice' based on the knowledge elicited from the main expert and expressed in the care pathway for ongoing management of RA (Figure 4.7). The parameters are defined in a way that a severe AME leads to a recommendation of medication change. A mild/moderate AME needs requires a reduction of medication dose; therefore, the medication is recommended to be changed. A mild/moderate AME that remains stable is negligible. Any increase of disease state requires to be prevented by medication escalation for those who can tolerate an escalation of DMARDs. Those who cannot tolerate extra DMARDs, will not be recommended to escalate. Tapering is applied to those who meet the following criteria: good overall disease control, stable overall disease progression

trend, without AMEs, and with a stable AME trend. The rest of the cases continue the same regimen that has been given.

8.6 Using the DBN for Decision-Support

The DBN model for self-management of RA supports decision-making by a clinician reviewing the medications of RA patients. The inputs of the model are the observations (variables shown with green ovals in Figure 8.1) of persons entered to the model at least once a week. We decided the interval of entering observations to be weekly since one week is the average length between taking medications, as described in Section 7.4.

Figure 8.2 demonstrates the unrolled structure of DBN model in Weeks 4 and 5 and the advice on medication on the right hand side. The probability distribution of the five parent variables of the advice variable is visible. Assume Week 5 is the current time. Based on the visible probabilities, the disease is under control and the overall progression trend is stable. The lack of adverse events and its stable trend reveals a medium state of tolerance for DMARDs.

8.6.1 Inputs

The inputs of DBN model are the evidence variables of the model, namely, signs (TJC and SJC), symptoms (Stiffness, GH, Fatigue, Pain, and HAQ), serology results (ESR and CRP), medications (csDMARDs, ts/bDMARDs, and Steroids), and AME (Organ Damage, Symptomatic Side-Effects, and Infection). These inputs are represented with green ovals in Figure 8.2. We assume that the serology is measured once in two weeks in a local clinic.

8.6.2 Dynamics of Adverse Medication Events and Tolerance for DMARDs

Since ts/bDMARDs and csDMARDs are toxic medications that could cause adverse events. Some people may be tolerant for the toxicity of these medications, but some others may not. The latter group are usually more difficult to find an appropriate medication for. Figure 8.3 shows a fragment of the DBN model containing the ‘Tolerance for DMARDs i’ variable and its parent variables. It expresses the dependencies within the variables of the fragment and the incoming dependencies into them; however, it hides the outgoing dependencies from these variables to those outside the fragment. The None state of ‘Infection i’, ‘Organ Damage i’, ‘Tolerance for DMARDs i’ are associated with the AME. The ‘AME Change i’ variable measures the changes of AMEs in the current time compared with the previous one. ‘AME

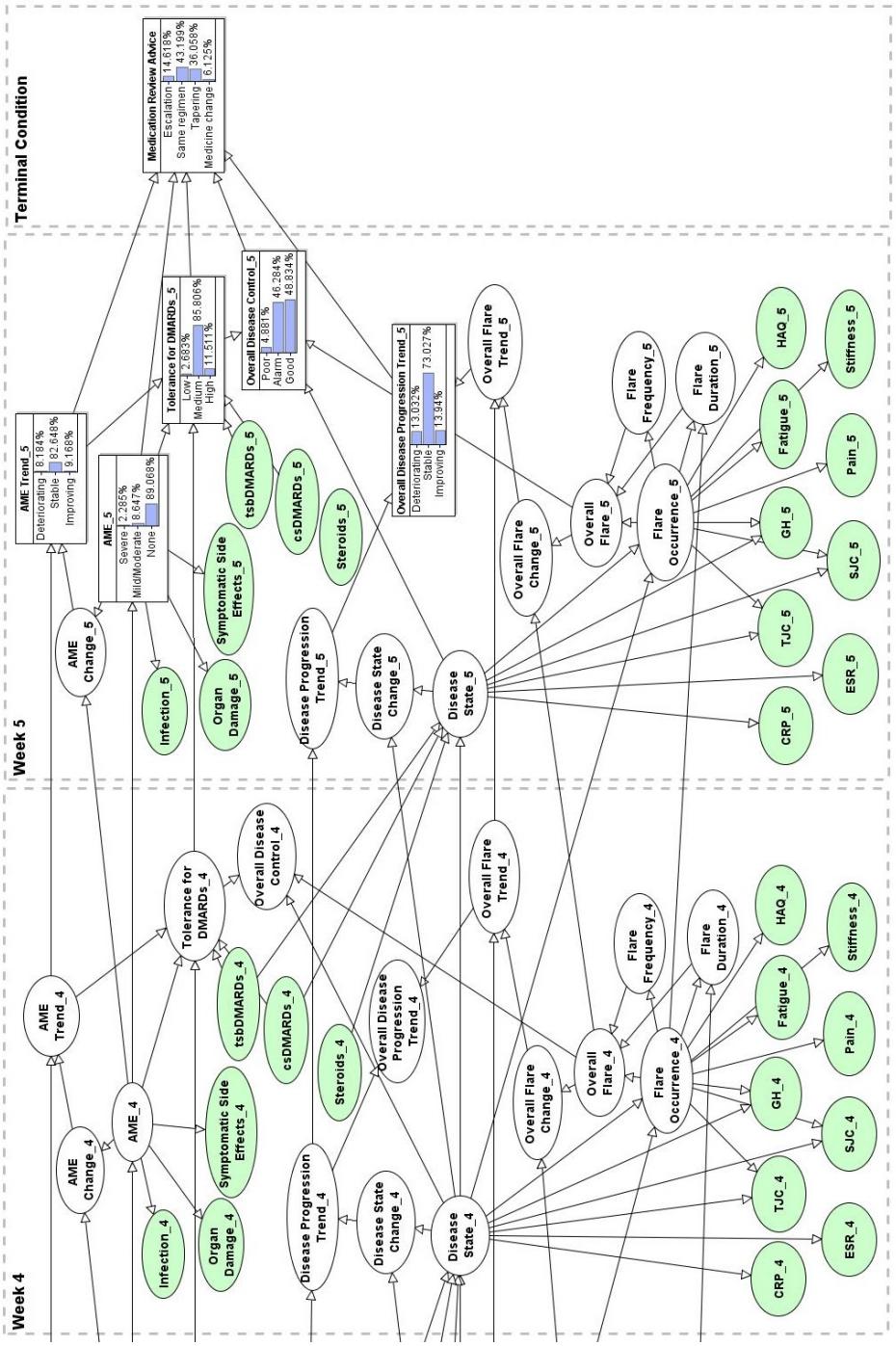


Fig. 8.2 Unrolled DBN model for self-management of RA for medication review.

Trend i' is linked to the trend of AME in the previous time and captures the trend of AME beyond the two times (current and previous).

The AME, 'AME Trend i', csDMARDs, and ts/bDMARDs variables combine and form the 'Tolerance for DMARDs i'. A tolerant person for DMARDs is the one that takes a higher dosage of csDMARDs or ts/bDMARDs and manifests no AME and their trend of AME stays stable or improves. In contrast, an intolerant person to DMARDs is who takes a lower dosage of csDMARDs or ts/bDMARDs and shows mild, moderate, or severe AMEs.

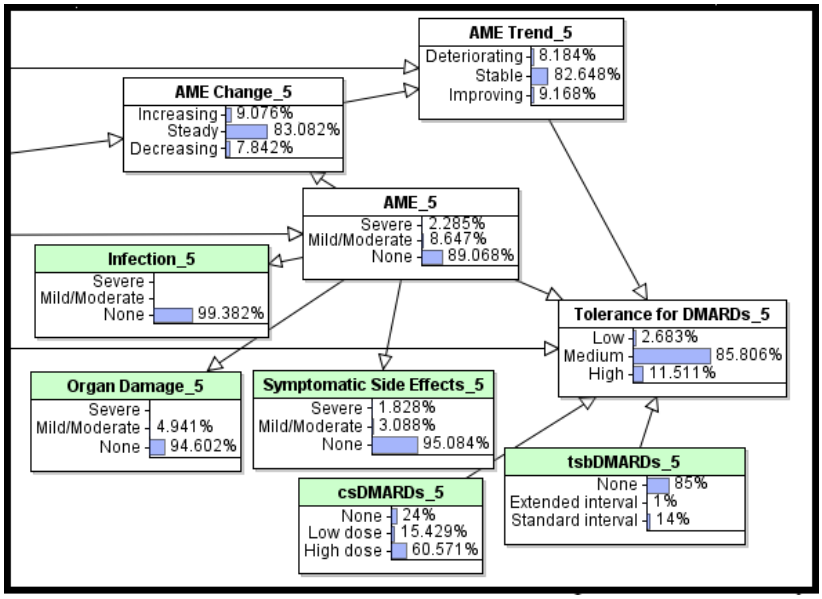


Fig. 8.3 Fragment of 'Tolerance for DMARDs' variable and its parent variables in Week 5.

8.6.3 Dynamics of Disease State and Overall Flare

Figure 8.4 depicts a fragment of DBN model including the 'Overall Disease Control 5' variable and its parent variables. It shows the dependencies within the variables of the fragment and the incoming ones into these variables.

Once the inputs are given to the model in the current time (Week 5), the posterior probabilities of 'Disease State 5' are calculated and the comparison of the 'Disease State 5' in the current time and the previous one updates the probabilities of 'Flare Occurrence 5'. The updated probabilities of 'Flare Occurrence 5' leads to calculate the posterior probabilities of 'Flare Duration 5', 'Flare Frequency 5', and finally 'Overall Flare 5'.

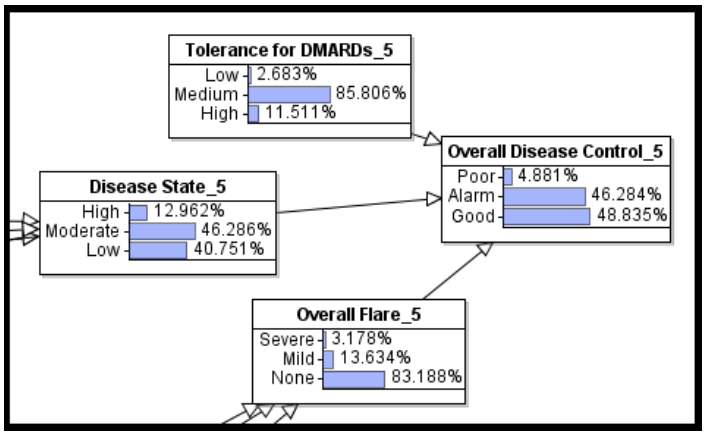


Fig. 8.4 Fragment of 'Overall Disease Control' variable and its parent variables.

8.6.4 Prediction of Advice on Medication Review

The advice variable has four recommendations: Escalation, 'Same regimen', Tapering, and 'Medication change'. Like a real prescribing decision-making in a clinic, the advice variable on medication review involves multiple variables, namely, AME, 'AME Trend', 'Overall Disease Control', 'Overall Disease Progression Trend', and 'Tolerance for DMARDs', and combines them. If a severe AME is observed, it advises to change the medication. The same advice is given when a mild or moderate AME happens in a deteriorating trend; however, if the mild or moderate AME stays steady or gets improved, the advice will not be 'Medication change'. If the 'Overall Disease Control' becomes good (disease is under control) and the 'Overall Disease Progression Trend' stays stable, the person is on remission and their medication can be tapered. Tapering is not recommended to those whom their disease is under control and the overall progression trend is improving, since improvement of the disease implies instability and therefore ineligibility for tapering. Anyone with a deteriorating trend is advised to escalate the dose of their medication to prevent deterioration, but if their tolerance for DMARDs is low, they will not receive the Escalation advice. For any circumstances except the above ones, the person is advised to keep the same regimen.

8.7 Evaluation

We evaluate the performance of the DBN model in providing advice for medication review. For this, we enter observations into the inputs (explained in Subsection 8.6.1) and we reason on the advice variable. As described in 2.4, we specify a sliding window to limit the length of the DBN chain since it can grow rapidly. We follow the same approach as we did for

the initial DBN model. We found out that a window length of 6, i.e., a DBN containing six BN models within its temporal plate, has the lowest increase of probability of the Low state. Figure 8.5 shows the pairwise difference between the probability of Low state in DBN models with window lengths of 2, 3, 4, 5, and 6. Although the total difference between the probability of Low states in a DBN model with a sliding window that contains 4 BNs compared to that with 3 BNs is negative, adding further BNs (5 and 6) compensates the negative change. Therefore, we do not choose a DBN with a sliding window that contains 3 BNs.

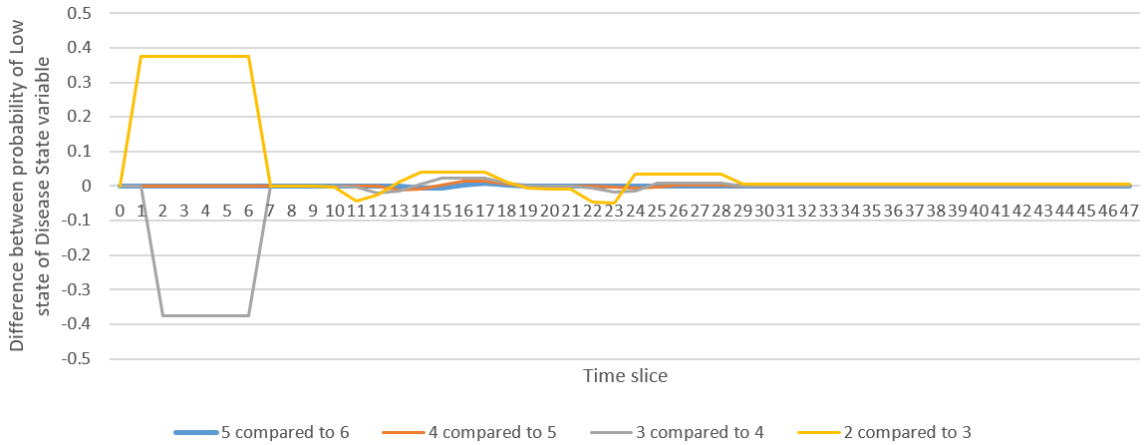


Fig. 8.5 The total difference of the probabilities of the Low state of the 'Disease State' variables in the extended DBN model for self-management of RA with sliding windows of 2, 3, 4, 5, and 6 length.

8.7.1 Evaluation with PEAC Data

We use the interpolated PEAC data (described in 4.4.4) to evaluate the performance of DBN model. For this, we enter observations for TJC, SJC, GH, Fatigue, Pain, HAQ, csDMARDs, Steroids, csDMARDs, ts/bDMARDs, Organ Damage, Symptomatic Side-effects, and Infection. The posterior probabilities of the latent variables led to an advice for each case. A comparison between the advice and medication decisions in each follow-up visit of the PEAC study (presented in Figure 4.22) help us to evaluate the performance of the model. In Figure 4.22, there is a fifth decision called Reduction, indicating to the situation that a mild or moderate AME happened (explained in Section 4.4.5). Here, we merge it with the 'Medication change' decision, since the two categories together comprise a small number of cases and also because both decisions refer to an adverse event.

Tables 8.4, 8.5, 8.6, and 8.7 show the confusion matrix of the advice provided by the model and the decisions extracted from the medication decisions in the first, second, third,

and forth follow-up visits of PEAC study. The decisions in the first, second, third, and fourth visits contain 35, 3, 12, and 2 missing values (out of 119); therefore, we cannot compare the obtained advice with those missing cases. Having at least one missing value in the dataset (e.g., a medication being prescribed in the baseline) leads to a missing decision. The advice for medication change is the 100% in the four tables, whereas escalation is the poorest. The advice for continuing the same regimen and tapering gains an average accuracy about 50% (except in one case). We can see that the model cannot easily differentiate the same regimen from tapering. This may be caused by the insensitivity of the parent variables of the medication review advice variable to capture the uncertainty between the two advice states. We can increase the number of states of 'Overall Disease Control' and 'Overall Disease Progression Trend' in order to increase the sensitivity of the advice variable.

Table 8.4 Confusion matrix of DBN model for self-management of RA for medication review advice - first follow-up visit.

Prediction of advice	True decisions			
for medication review	Escalation	Same regimen	Tapering	Medication change
Escalation	0/36(0%)	0/23(0%)	0/10(0%)	0/9(0%)
Same regimen	26/36(72%)	12/23(52%)	8/10(80%)	0/9(0%)
Tapering	10/36(28%)	10/23(43%)	2/10(20%)	0/9(0%)
Medication change	0/36(0%)	1/23(4%)	0/10(0%)	9/9(100%)

Table 8.5 Confusion matrix of DBN model for self-management of RA for medication review advice - second follow-up visit.

Prediction of advice	True decisions			
for medication review	Escalation	Same regimen	Tapering	Medication change
Escalation	0/8(0%)	0/97(0%)	0/10(0%)	0/1(0%)
Same regimen	7/8(88%)	49/97(51%)	5/10(50%)	0/1(0%)
Tapering	1/8(12%)	45/97(46%)	5/10(50%)	0/1(0%)
Medication change	0/8(0%)	3/97(3%)	0/10(0%)	1/1(100%)

Table 8.6 Confusion matrix of DBN model for self-management of RA for medication review advice - third follow-up visit.

Prediction of advice for medication review	True decisions			
	Escalation	Same regimen	Tapering	Medication change
Escalation	0/9(0%)	0/90(0%)	1/6(17%)	0/5(0%)
Same regimen	6/9(67%)	55/90(61%)	2/6(33%)	0/5(0%)
Tapering	3/9(33%)	35/90(39%)	3/6(50%)	0/5(0%)
Medication change	0/9(0%)	0/90(0%)	0/6(0%)	5/5(100%)

Table 8.7 Confusion matrix of DBN model for self-management of RA for medication review advice - fourth follow-up visit.

Prediction of advice for medication review	True decisions			
	Escalation	Same regimen	Tapering	Medication change
Escalation	0/12(0%)	2/91(2%)	0/8(0%)	0/6(0%)
Same regimen	5/12(42%)	47/91(52%)	7/8(88%)	0/6(0%)
Tapering	7/12(58%)	42/91(46%)	1/8(12%)	0/6(0%)
Medication change	0/12(0%)	0/91(0%)	0/8(0%)	6/6(100%)

Table 8.8 shows the aggregated confusion matrix of DBN model for self-management of RA for medication review advice that contains the predictions and true decisions of the first, second, third, and fourth follow-up visits together.

Table 8.8 Aggregated confusion matrix of DBN model for self-management of RA for medication review advice.

Prediction of advice for medication review	True decisions			
	Escalation	Same regimen	Tapering	Medication change
Escalation	0/65(0%)	2/301(1%)	1/34(3%)	0/21(0%)
Same regimen	44/65(68%)	163/301(54%)	22/34(65%)	0/21(0%)
Tapering	21/65(32%)	132/301(44%)	11/34(32%)	0/21(0%)
Medication change	0/65(0%)	4/301(1%)	0/34(0%)	21/21(100%)

To analyse the performance of DBN model in terms of discrimination, we aggregated the two states of ‘Same regimen’ and ‘Medication change’ and aggregated the other two states of Escalation and tapering. This way we obtain binary states that we could easily demonstrate the ability of the DBN to discriminate the states. We draw a receiver operating

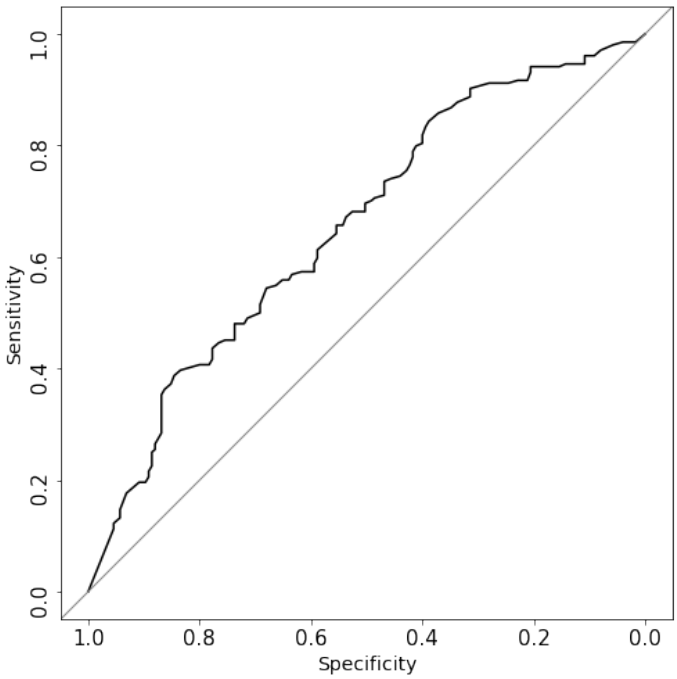


Fig. 8.6 ROC curve of DBN model for self-management of RA for medication review.

characteristics (ROC) curve for the DBN model as shown in Figure 8.6. The area under ROC curve (AUROC) for this model is 0.66.

Figure 8.7 shows the analysis of performance of the DBN model for self-management of RA for medication review with a range of thresholds from 0.0 to 0.9. This figure includes

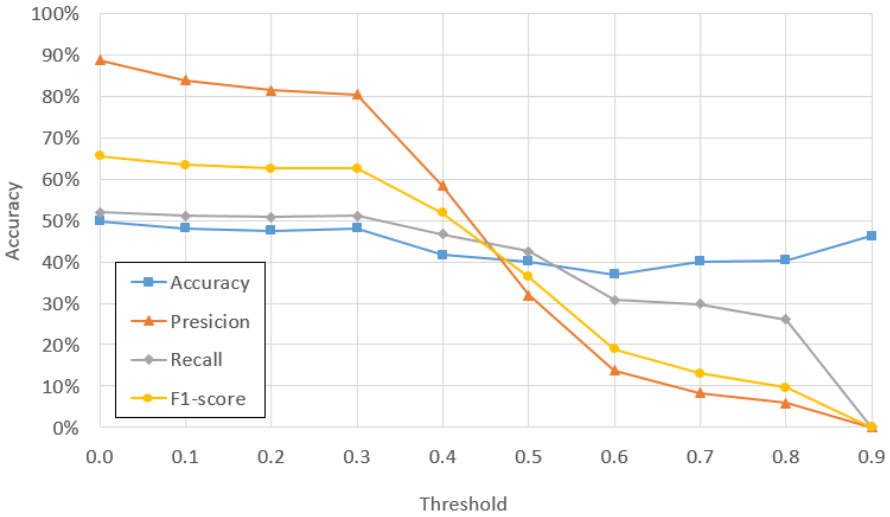


Fig. 8.7 Performance analysis of the DBN model for self-management of RA for medication review with a range of thresholds from 0.0 to 0.9.

four metrics of accuracy, precision, recall, and F1-score which demonstrate the performance of prediction models when data are imbalance.

8.7.2 Evaluation with Scenarios of Dummy Patients

We do reasoning with the data of a set of dummy patient scenarios to evaluate the DBN model for self-management of RA. We had used the same scenarios in Subsections 5.4.2 and 7.5.2 to evaluate the diagnostic BN model and the initial DBN model for self-management, respectively. Scenarios of dummy patients for evaluating the extended DBN model are the same as those we used for evaluating the initial DBN model (described in Subsection 7.5.2). These scenarios include observation data of medications and AMEs, and also expert's expected decisions for medication. As explained in Subsection 7.5.2, our experts designed these scenarios considering the current clinical practice that patients are monitored in fixed intervals. Therefore, we expand the frequency of medication and AMEs, like what we did for the observation data of signs, symptoms, and serology results (described in Subsection 7.5.2). The observations and expert's expected outcome for the dummy patient 1 in follow-up visit 1 are presented in Table B.2.

Table 8.9 shows the merged confusion matrix of all 15 scenarios. The percentage of accurate predictions in the diagonal of the matrix is close to the average of those in the first, second, third, and forth follow-up visits presented in Tables 8.4, 8.5, 8.6, and 8.7, respectively. Among the percentages on the diagonal, 'Medication change' and 'Same regimen' gain 83% and 58% accurate predictions, whereas the accuracy of Tapering and Escalation are poor. A large percentage of Escalation decisions are predicted as 'Same regimen' (95%) and similarly 42% of the 'Same regiment' decisions are predicted as Tapering. This shows a confusion between the Escalation and 'Same regimen' and between the 'Same regimen' and Tapering in the parameters of the 'Medication Review Advice'.

Table 8.9 Confusion matrix of DBN model for self-management of RA for medication review advice - dummy patient scenarios.

Prediction of advice for medication review	True decisions			
	Escalation	Same regimen	Tapering	Medication change
Escalation	1/20(5%)	0/45(0%)	0/3(0%)	0/6(0%)
Same regimen	19/20(95%)	26/45(58%)	3/3(100%)	1/6(17%)
Tapering	0/20(0%)	19/45(42%)	0/3(0%)	0/6(0%)
Medication change	0/20(0%)	0/45(0%)	0/3(0%)	5/6(83%)

8.8 Summary

This chapter extends the initial DBN model for self-management of RA (described in Chapter 7), covering the treatment of RA. It corresponds to the Objective 3 of this thesis. In this chapter, we first introduced common medications for RA, including csDMARDs, ts/bDMARDs, and steroids, and we explained the strategies, regimens, and decision-making for the treatment of RA. We then described the variables related to medications and AMEs that we added to extend the initial DBN model. We also define a variable that estimates the tolerance for DMARDs and how that variable along with AMEs, the trend of AMEs, overall disease control, and overall disease progression trend lead to generate advice on medication review. The advice entails four recommendations to clinicians: escalate medications, keep the same regimen, taper, or change medications. These recommendations are extracted from the regimens and decision-making for the treatment of RA, suggested by the clinical guidelines.

We explained an unrolled DBN model to illustrate its operation as a decision support. We elaborated the inputs of the DBN-based decision support, the dynamics of AMEs and tolerance for DMARDs, the dynamics of disease state and overall flare, and how the model generates advice on medication.

To evaluate the performance of DBN model, we created an interpolated dataset to be able to do reasoning with DBN model. From the original dataset, we extracted a set of medication decisions made by the clinicians. We entered the interpolated data as observations into the DBN model for self-management of RA, and compared the predicted advice on medication review with the medication decisions extracted. We also used a dataset of dummy patients to further evaluate the model.

As mentioned in Section 7.6, the contribution to self-management modelling suffers from the limited data set with small number of cases and less granularity that our expectations. The study is also limited to a small number of experts involved in both model development and its evaluation with scenarios. The results of both evaluation approaches are helpful to understand the strengths and weaknesses of the model and they indicate the need for further improvements of the model. Improvements should focus on the parameterisation of the advice variable for medication review. Evaluation with real data collected from the patients can help to better evaluate the model. Improvement and further evaluations can help us to fully achieve the Objective 3.

Chapter 9

Building a Bayesian Network Model for Personalised Care for Living with Rheumatoid Arthritis

Chapters 7 and 8 proposed two dynamic Bayesian network (DBN) models for self-management of Rheumatoid Arthritis (RA), which predict the state of disease and flare to provide advice on self-management on appointment scheduling (Section 7.4) and medication review (Section 8.6). This chapter presents the steps of building a recommender system (RS) to provide personalised advice for living with RA. The RS includes a Bayesian networks (BNs) model and a rule-based system. The BN estimates disease activity, flare, quality of life (QoL), and QoL components, and the rule-based system provides personalised advice to adjust lifestyle and improve living with RA, rather than providing a set of advice for self-management and treatment merely based on disease state and flare. It aims to show the steps of building and evaluating a Bayesian network (BN) model and a rule-based system to recommend personalised advice for living with RA using knowledge gathered from experts, people with RA, and literature. We investigate how to create a BN model to estimate QoL using knowledge and without having data, and whether this BN model can help to provide useful advice to improve the experience of living with chronic diseases, corresponding to the Objective 4.

In Section 9.1, we introduce the personalised care for living with RA. Section 9.2 covers a description of the variables required for personalised care. The questionnaires and other sources of data collection are presented in Section 9.3. In Section 9.4, we explain the proposed BN model for personalised care including the structure of the model and its parameterisation. In Section 9.5, we describe a rule-based system to generate a set of advice and modify them to find the appropriate advice for each person. Finally, Section 9.6 presents a scenario

analysis and a sensitivity analysis to evaluate the BN model and rule-based system. Finally, we summarise the chapter in Section 9.7.

9.1 Introduction

People with chronic diseases need lifelong care and their lives can get improved by personalising care skills. In the case of RA, people receive some leaflets and the address of the National Rheumatoid Arthritis Society (NRAS) website to access personalised care prepared as leaflets, online RA community, relevant mobile apps, or publications [147]. As we described in Subsection 4.2.3, the leaflets of NRAS introduce RA and explain a wide range of information about living with RA, such as emotional aspects, work with RA, benefits and support, lifestyle, and practical help. This information is publicly available and has a structured presentation, but it includes various contents of generic advice for all sorts of patients, e.g. see [145]. Clinicians may also provide people with other leaflets prepared by the NHS [126] and the Arthritis Research UK [11], which contain a description of RA and advice for living with RA. In addition, clinicians would hand a set of links to the NHS web pages to learn about RA [132] and how to live with it [131]. This advice is necessary to be given to patients, yet they are not targeted and may not be as useful as expected by their organisers.

Personalised care refers to having choices and control to select care planning and delivery [133]. Therefore, personalised care for living with RA is not limited to the provision of information and sources of information prepared by NRAS or NHS. To achieve personalised care, the advice needs to be personalised and targeted to people to learn how to live with RA.

One way to provide personalised care to patients is to identify each person, estimate their QoL, and recommend them a set of personalised advice based on their QoL. To establish these processes, we created a care pathway for personalised care of RA based on the knowledge learnt from experts, interviews with patients, and literature, as described in Subsection 4.3.4. Analysing the activity and decision nodes of the care pathway leads us to select a decision support point called ‘Personalised Advice Given to Patient’, as described in Section 4.5. This point indicates the target of our decision support for personalised care for living with RA. The other nodes of the pathway help us to specify the variables of a BN model. The BN model can estimate three components of QoL, namely, participation, empowerment, and independence, that correspond to the social, psychological, and physical aspects of QoL, respectively. These estimations allow the rule-based system to provide personalised advice. Our aim is to give useful advice to balance disease activity and lifestyle choices. This model

and system can be mainly used by patients through a user interface, and clinicians can also use it to monitor patients' disease activity remotely.

9.2 Description of Variables for Personalised Care for Living with Rheumatoid Arthritis

We specify the variables of personalised care for living with RA from the activities and decision nodes of the proposed care pathway (Figure 4.8). The variables are supported by experts' knowledge, coded interviews with the members of a Patient and Public Involvement (PPI) group, or literature. We classify the variables into personal and environmental, disease manifestation and disease activity, characteristics of QoL, lifestyle choices, and advice.

9.2.1 Personal and Environmental Variables

We derive personal and environmental variables from the node 'Patient Enters Personal Information'. Our experts (introduced in Section 4.1) suggested including Age, Sex, and body mass index (BMI) as the relevant personal factors. We consider three states for Age: Young (< 40), Middle age ($40 \leq$ and < 65), and Older ($65 \leq$). Sex is considered as binary of Female and Male. For BMI, we use the common categories: Underweight (< 18.5), Normal ($18.5 \leq$ and < 25), Overweight ($25 \leq$ and < 30), and Obese ($30 \leq$).

Environmental factors play a role in the feeling of pain and mood, and physical activity of people with RA [48, 201, 208]. Here, we consider temperature, humidity, air pressure, and wind as suggested by [171]. We define four states for the Temperature variable: Low ($12°C <$), Medium ($\leq 12°C$ and $18°C <$), High ($19°C \leq$ and $< 30°C$), and Very high ($30°C \leq$). We categorise Humidity and 'Air Pressure' based on Millibar (MB) into High ($\leq 1050MB$), Medium ($970MB \leq$ & $1050MB <$), and Low ($\leq 970MB\%$). Like 'Air Pressure', we define the Humidity variable with three states of Low ($\leq 40\%$), Medium ($40\% \leq$ & $60\% <$), and High ($\leq 60\%$). We specify the Wind variable as a binary of Yes and No, as suggested by [27]. To combine the environmental factors, we define a latent variable called 'Weather Condition' with four states of Cold, Cool, Warm, and Hot.

Some people may be sensitive to a particular environmental situation, but some others may not. Therefore, we define one variable called 'Weather Sensitivity' and we consider three states of None, Some, and High respectively for insensitive, some sensitive, and highly sensitive people to the environmental factors. This variable and the 'Weather Condition' variable combine and form a variable called 'Weather Effect' that has three states of Unpleasant,

Ineffective, and Pleasant. The variables enable us to personalise the environmental factors based on each person's sensitivity.

Table 9.1 summarises the personal and environmental factors.

Table 9.1 Summary of personal and environmental factors.

Type	Variable	States	Source	Description
Personal Factors	Age	Young (< 40), Middle age ($40 \leq$ & < 65), Older ($65 \leq$)	Experts	Person's age
	Sex	Female, Male	Experts	Person's sex
	BMI	Underweight (< 18.5), Normal ($18.5 \leq$ & < 25), Overweight ($25 \leq$ & < 30), Obese ($30 \leq$)	Experts	Person's BMI
Environmental Factors	Temperature	Low($\leq 30°C$), Medium($12°C \leq$ & $18°C <$), High ($19°C \leq$ & $< 30°C$), Very high ($\leq 30°C$)	Literature	Temperature where the person lives
	Humidity	Low($\leq 40\%$), Medium ($40\% \leq$ & $60\% <$), High($\leq 60\%$)	Literature	Humidity where the person lives
	Wind	Yes, No	Literature	Wind where the person lives
	Air Pressure	Low($\leq 970MB\%$), Medium ($970MB \leq$ & $1050MB <$), High($\leq 1050MB$)	Literature	Air pressure where the person lives
	Weather Condition	Cold, Cool, Warm, Hot	Experts	Person's feeling of the weather
	Weather Sensitivity	None, Some, High	Expert	Person being sensitive to weather condition
	Weather Effect	Unpleasant, Ineffective, Pleasant	Expert	Combination of the weather condition and sensitivity to weather

9.2.2 Disease Manifestation and Disease Activity Variables

The pathway node 'Patient Enters Observations' refers to the manifestations of disease that can be used to estimate disease activity. Our experts specify five main manifestation factors – namely, pain, fatigue, morning stiffness, mobility, and deformation.

Pain and Fatigue variables represent two common manifestations of chronic diseases [51]. These two variables have three categories of None, Some, and Severe. 'Morning Stiffness' of joints is another common manifestation of chronic diseases such as RA [199]. According to our rheumatology experts, morning stiffness is classified into three groups based on its length: 'More than 2 hours', 'Between 30 minutes and 2 hours', 'Less than 30 minutes'. These groups refer to no stiffness or any stiffness lasting for less than or equal to 30 minutes, more than 30 minutes and less than or equal to 2 hours, and more than 2 hours.

RA may result in disability and restriction of mobility [116] which is usually measured by the health assessment questionnaire (HAQ) [63]. HAQ ranges from 0 to 3 with a higher scale

representing a greater level of disability. It is categorised into Mild (< 1), Moderate ($1 \leq$ and < 2), and Severe ($2 \leq$). Here, we consider a variable called Mobility and we translate HAQ's categories into three states of Low, Average, and Good. The joints of some RA patients may develop deformation in long-term. We add a variable called 'Deformation Effect' with three states of None, Medium, and High.

We specify four variables related to disease activity – namely, 'Disease Duration', Flare, 'Current Disease Activity', and 'Overall Disease Activity'. 'Disease Duration' is the duration from diagnosis to the present time. It has three categories: 'less than 1 year', 'between 2 years to 5 years', and 'more than 5 years'. Flare variable is known as the worsening of symptoms or an increase of disease activity. We classify it into three states: None, Mild, and Severe. 'Current Disease Activity' variable estimates the activity of disease in short-term, whereas the 'Overall Disease Activity' keeps track of the accumulation of disease activity during long-term. Both variables have four states of Remission, Low, Moderate, and High, similar to the categories of the Disease Activity Score 28 (DAS28). Table 9.2 summarises the disease manifestation and disease activity variables.

Table 9.2 Summary of disease manifestation and disease activity variables.

Type	Variable	States	Source	Description
Disease Manifestations	Fatigue	None, Some, Severe	Literature	Person feeling of fatigue
	Arthritis Pain	None, Some, Severe	Literature	Person feeling pain
	Morning Stiffness	More than 2 hours, Between 30 minutes and 2 hours, Less than 30 minutes	Literature	Feeling stiffness in joints in the mornings
	Mobility	Low, Average, Good	Literature	Mobility of a person
	Deformation Effect	None, Medium, High	Literature	Person having deformed joints
Disease Activity	Disease Duration	Less than 1 year, Between 2 and 5 years, More than 5 years	Experts	Duration of disease since being diagnosed
	Flare	None, Mild, Severe	Experts	Person having a flare
	Current Disease Activity	Remission, Low, Moderate, High	Experts	Current disease activity appeared in short-term
	Overall Disease Activity	Remission, Low, Moderate, High	Experts	Overall disease activity accumulated in long-term

9.2.3 Characteristics of Quality of Life and Lifestyle Choices

We derive three categories of variables – social, psychological, and physical - from three pathway nodes of 'Social Aspect Assessment', 'Psychological Aspect Assessment', and 'Physical Aspect Assessment'. Social variables include 'Financial Status', 'Work Capacity', Belonging, and 'Social Activity'. Financial status is believed to have an association with QoL of patients with a history of cancer [196], and our main expert finds it important for

patients with arthritic diseases. 'Financial Status' variable refers to a person's perception of their financial status, which we specify it in three categories of Bad, Average, and Good. 'Work Capacity' variable is the amount of work that a person can do. We consider it to have three states of Low, Medium, and High. A patient's level of socialising with their friends is called 'Social Activity'. Belonging variable indicates the feeling of belonging to family or friends, environment, and community. Both 'Social Activity' and Belonging have three states of Low, Medium, and High. Among these, 'Financial Status' is a characteristic of QoL; however 'Work Capacity', Belonging, and 'Social Activity' are lifestyle choices. We combine all variables of social aspect in one synthetic variable called Participation, which has three states of Low, Medium, and High. Participation refers to a patient's QoL experience in terms of being part of a social community, participating in different projects at home, at work, or in the community [115].

For the psychological aspect, we define three variables – Mood, 'Level of Tension', and 'Medication Adherence'. Mood has three states of Depressed, Numb, or Happy, which are the modified states of [51]'s suggestion. The 'Level of Tension' refers to the response to the emotional pressure or tension that a person suffers for a (prolonged) period of time in which an individual perceives they have little or no control. 'Medication Adherence' variable indicates a person's level of adherence to the medication recommendations from the care provider [221]. The lack of adherence to medication can cause severer disease, lesser functional ability, and lower QoL [89]. The Mood variable is a characteristic of QoL that cannot be changed by the patient; however, we consider the level of tension and medication adherence as lifestyle choices. We aggregate all psychological variables and define a synthetic variable called Empowerment with three states of Low, Medium, and High. Empowerment implies a patient's experience of taking charge of their life and using different resources, such as positive thinking, to manage fatigue, pain, and physical functioning [115].

'Sleep Quality', Self-care, and 'Physical Activity' variables cover the physical aspect of QoL. 'Sleep Quality' variable is the quality level of sleep categorised into Interrupted, Normal, and Deep [51, 86]. Self-care points out the state of being able to self-care [131]. 'Physical Activity' variable is the level of physical activity of a patient. Both Self-care and 'Physical Activity' have three states of Low, Medium, and High. We classify 'Sleep Quality' and Self-care as the characteristics of QoL, and we consider 'Physical Activity' to be a lifestyle choice that can be changed by the patient. We define a synthetic variable called Independence to combine physical variables, and we determine three states of Low, Medium, and High. Independence refers to a patient's experience of QoL when they are free and independent in choosing and managing their daily activities, living at home, at work, or in their leisure time [115].

In addition to the three aspects of QoL, we consider the acceptance of disease as a key element that influences QoL. We define the 'Disease Acceptance' variable as the level of disease being accepted by a patient associated with gaining knowledge about the disease during time [97]. This knowledge is represented by the 'Disease Knowledge' variable, which refers to a patient's experience and knowledge about their disease, earned by experiencing the disease and studying about it [29]. Both 'Disease Acceptance' and 'Disease Knowledge' have three states of Low, Medium, and High. We classify the 'Disease Acceptance' as a characteristic of QoL, and the 'Disease Knowledge' as a lifestyle choice.

We consider a variable called QoL that combines its components – participation, empowerment, and independence. Therefore, we associate the QoL variable with the Participation, Empowerment, and Independence and we specify three states of Low, Medium, and High for it.

A summary of social, psychological, physical, and disease acceptance factors is presented in Table 9.3.

Table 9.3 Summary of characteristics of QoL and lifestyle choices.

Type	Variable	States	Source	Description
Characteristic of QoL	Financial Status	Bad, Average, Good	Literature	Financial status of a person
	Participation	Low, Medium, High	Literature	Participating in family and society
	Mood	Depressed, Numb, Happy	Literature	Person's mood
	Empowerment	Low, Medium, High	Literature	Being able tomanage physical functioning
	Sleep Quality	Interrupted, Normal, Deep	Literature	Person's quality of sleep
	Self-care	Low, Medium, High	Literature	Ability to look after self
	Independence	Low, Medium, High	Literature	Being free and independent
	Disease Acceptance	Low, Medium, High	Literature	Level of disease being accepted
	QoL	Low, Medium, High	Literature	Person's QoL combining independence, empowerment, and participation factors
Lifestyle Choices	Work Capacity	Low, Medium, High	Literature	Patient's capacity of working
	Belonging	Low, Medium, High	Literature	Level of belonging to family members and friends
	Social Activity	Low, Medium, High	Literature	Socially being active
	Level of Tension	Low, Medium, High	Literature	Level of tension
	Medication Adherence	Low, Medium, High	Experts, Literature	Person's level of adherence to medication
	Disease Knowledge	Low, Medium, High	Literature	Person's knowledge about their disease

9.2.4 Advice Variable

We define a variable called 'Advice Priority' that is taken from the pathway node of 'Personalised Advice Given to Patient'. It estimates a priority of advice category that is most needed

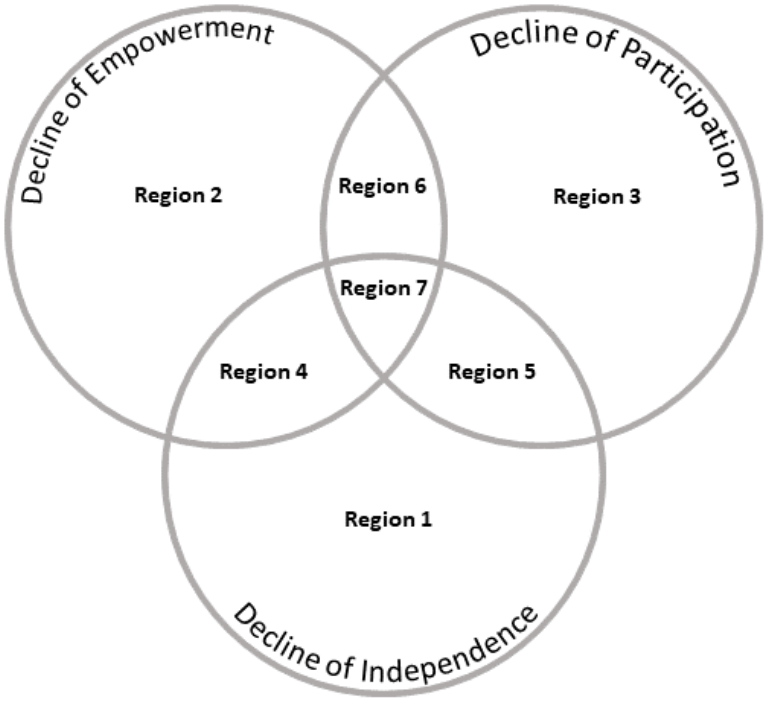


Fig. 9.1 QoL components change and advice

to improve QoL. The states of the 'Advice Priority' variable are based on the Low state being estimated for one or a combination of Empowerment, Independence, and Participation; therefore, we define seven states for the 'Advice Priority' variable: Independence, Empowerment, Participation, 'Independence & empowerment', 'Independence & participation', 'Empowerment & participation', and 'Independence & empowerment & participation'.

The specifications of these states are depicted as the regions of a Venn diagram for declining QoL components, as shown in Figure 9.1. This diagram represents three main regions of decline of empowerment, decline of participation, and decline of independence. Seven regions of this diagram correspond to the seven states of the 'Advice Priority' variable.

Table 9.4 Summary of advice variable.

Type	Variable	States	Source	Description
Advice	Advice Priority	Independence, Empowerment, Participation, Independence & empowerment, Independence & participation, Empowerment & participation, Independence & empowerment & participation	Literature, expert	Priority of advice to improve QoL

In Section 9.5, we will use these regions and their estimated priorities to generate personalised advice based on [115]'s suggestion.

Table 9.4 summarises the advice variable for personalised care.

9.3 Data Collection Questions and Information Retrieval

We need to collect data from patients for the variables of the BN model for personalised care (described in Section 9.2). These data are the observations that we feed to the BN model to estimate QoL and 'Advice Priority'. We retrieve the weather data (Temperature, Humidity, Wind, and 'Air Pressure') from an online source. To collect personal information, we ask one set of questions. Then, we ask two sets of questions, namely, initial questions and secondary questions, to collect observations for the rest of the variables from the following types: disease manifestation, characteristics of QoL, and lifestyle choices (listed in Tables 9.2 and 9.3). The secondary questions gather further information when the person who is not satisfied with the items. We process the answers to these questions to prepare the observations for the BN model.

We adapt the secondary questions from widely used questionnaires including [5, 222, 76, 122, 79]. These questions are originally designed to be answered using either a four-level or a five-level Likert scale, which are quantified using 1 to 4 or 1 to 5 numbers, respectively. Since there are multiple questions corresponding to each variable, we suggest applying the factor analysis method to reduce the dimension of the received data. These are the back-end processes that are not directly related to this chapter.

9.3.1 Personal, Disease Duration, and Environmental Variables

We gather data of the patient's age, sex, and BMI. The duration of the disease is also asked to patients. As mentioned before, we retrieve the weather data from an online source.

A summary of data collection questions for personal information, disease duration, and sensitivity to weather conditions is presented in Table C.1. The choices of the initial or secondary questions for 'Weather Sensitivity' are then translated into the states of 'Weather Sensitivity', as presented in Table 9.2. A summary of the initial and secondary questions to collect data of personal information and 'Weather Sensitivity' is presented in Table C.1

9.3.2 Disease Manifestations

We collect patients' observations about their disease manifestation variables - Fatigue, Pain, 'Morning Stiffness', Mobility, and 'Deformation Effect' - by asking a set of questions for

each variable. The initial information about disease manifestations is gathered by asking the initial questions. The initial questions have three options: 'Not satisfied', 'Somewhat satisfied', and Satisfied. If Satisfied is chosen, we translate it into None state of Fatigue, as presented in Subsection 9.2.2. In the case of 'Not satisfied' and 'Somewhat satisfied', secondary questions are asked which scale patients' feeling of fatigue. These questions have five options – Never, 'Almost never', Sometimes, 'Very Often', and Always, which are quantified as 1, 2, 3, 4, and 5, respectively. Table C.2 shows the primary and secondary questions for each manifestation factor.

9.3.3 Characteristics of Quality of Life

The observation about the characteristics of QoL, namely, 'Sleep Quality', Self-care, Mood, and 'Financial Status', is initially gathered by asking primary questions, which have three options of 'Not satisfied', 'Somewhat satisfied', and Satisfied. If Satisfied is chosen, we consider it as a desired choice being opted. However, in the case of 'Not satisfied' and 'Somewhat satisfied', secondary questions are asked, which scale the unsatisfactory level. The secondary questions have five options of Never, 'Almost never', Sometimes, 'Very Often', and Always, which are quantified as 1, 2, 3, 4, and 5, respectively. Table C.3 shows the initial and secondary questions for QoL characteristics.

9.3.4 Lifestyle Choices

We collect patients' lifestyle observations including 'Medication Adherence', 'Physical Activity', 'Level of Tension', 'Work Capacity', 'Social Activity', Belonging, and 'Disease Knowledge' by asking initial questions to measure the satisfaction level regarding each lifestyle choice. These questions have three options: 'Not satisfied', 'Somewhat satisfied', and Satisfied. If Satisfied is chosen, we translate it as the desired state, as presented in Subsection 9.2.3. In the case of 'Not satisfied' and 'Somewhat satisfied' being chosen, the secondary questions are asked to scale the dissatisfaction with lifestyle choices. These questions have five options – Never, 'Almost never', Sometimes, 'Very Often', and Always, which are quantified as 1, 2, 3, 4, and 5, respectively. A summary of the initial and secondary questions to collect data of lifestyle choices is presented in Table C.4.

9.4 BN Model for Personalised Care for Living with Rheumatoid Arthritis

In this section, we present a BN model for personalised care of RA. In the following subsections, we explain the structure of the BN and its parameterisation using expert's knowledge.

9.4.1 Structure

We described the variables of the BN model for personalised care for living with RA in Section 9.2, which are taken from coded interviews with the members of a PPI group, and from the literature. We classify the BN variables into seven groups: personal factors, environmental factors, disease activity, disease manifestation, QoL characteristics, lifestyle choices, and advice.

We specify the causal or associational dependencies between the variables of personalised care for living with RA by reviewing the literature and eliciting knowledge from our experts. To specify dependencies, we employ the medical idioms as the building blocks of BNs with medical applications, as described in Subsection 2.5.1. The 'Overall Disease Activity' variable - an accumulation of disease activity during long-term - associates with Age, Sex, BMI, and 'Disease Duration' that corresponds with the risk factor idiom (Figure 2.10c). Since a persistent high overall disease activity can develop disability and deformation, we connect the 'Overall Disease Activity' to the Mobility and 'Deformation Effect' variables. The 'Overall Disease Activity' variable is also associated with flares which cause pain, morning stiffness, and fatigue. The link between Flare and 'Overall Disease Activity' variables is like the complication idiom (Figure 2.10i). Flare is affected by the medication adherence which can trigger or prevent the occurrence of a flare. The 'Medication Adherence' variable is a risk factor to the Flare variable, fitting the risk factor idiom (Figure 2.10c). The Flare variable associates with 'Arthritis Pain', 'Morning Stiffness', and Fatigue, like the manifestation idiom (Figure 2.10a). The 'Weather Effect' is an indirect combination of four environmental factors, namely, temperature, humidity, air pressure, and wind, aggregated with the 'Weather Sensitivity' variable. The 'Weather Effect' variable acts as a risk factor for 'Arthritis Pain', Mood, and 'Physical Activity' variables, as suggested by the risk factor idiom (Figure 2.10c).

The Flare and 'Overall Disease Activity' lead to a short-term disease activity called the 'Current Disease Activity' which is a disease activity in short-term. Thus, Flare and 'Overall Disease Activity' variables associate with 'Current Disease Activity', matching the complication idiom (Figure 2.10i). A high current disease activity happens when a patient

experiences a flare while having a high overall disease activity. Current disease activity influences patient's QoL - the higher the current disease activity, the lower the QoL.

QoL is affected by patient's disease acceptance since those accepted their disease can manage it better than others and can sustain their QoL [97]. In addition, the 'Disease Acceptance' variable is associated with 'Disease Duration' and 'Disease Knowledge'.

QoL is a consequence of its components: Participation, Independence, and Empowerment. Participation of a patient is caused by 'Financial Status' as a QoL characteristic and three lifestyle choices – 'Work Capacity', Belonging, and 'Social Activity'. Participation is associated with Mobility, which is a consequence of the 'Overall Disease Activity'. The association between Participation and Mobility refers to the fact that the participation of a disabled patient can be restricted. Independence variable is associated with the 'Sleep Quality' and Self-care variables as the characteristics of QoL. Independence variable is also linked to the 'Morning Stiffness', which is a consequence of Flare. Empowerment variable is

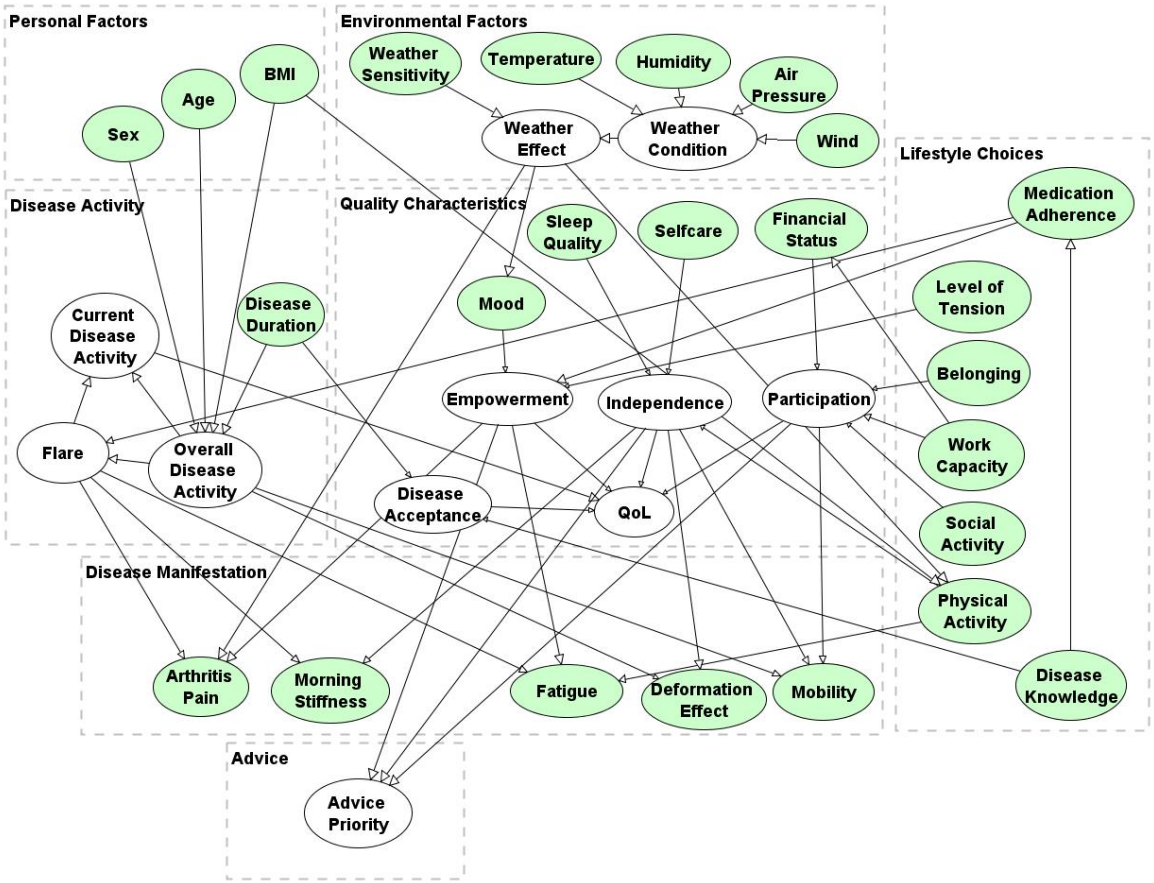


Fig. 9.2 BN model for personalised care for living with RA.

associated with the Mood, a QoL characteristic, and two lifestyle choices of the 'Level of Tension' and 'Medication Adherence', which is associated with 'Disease Knowledge'.

The BN model for personalised care for living with RA is shown in Figure 9.2. The evidence variables or the inputs of the model are displayed by green ovals. These variables receive observations that are collected through multiple questions (described in Section 9.3). The white ovals represent the synthetic and latent variables or the outputs of the model. We intend to estimate the probabilities of these variables for advice generation. The 'Weather Condition' is a synthetic variable that combines its parent variables and simplifies the model. The other white ovals represent the outputs of the model, namely Flare, 'Current Disease Activity', 'Overall Disease Activity', 'Disease Acceptance', Independence, Participation, Empowerment, QoL, and 'Advice Priority'. Flare variable has three states: None, Mild, and Severe. The 'Current Disease Activity' and 'Overall Disease Activity' have four states: Remission, Low, Moderate, and High. The rest of the output variables of the QoL characteristics have three states of Low, Medium, and High. 'Advice Priority' variable combines the Empowerment, Independence, and Participation variables and estimates the priority of the seven advice categories (as described in Subsection 9.2.4).

9.4.2 Parameterisation

We used expert's knowledge to specify the probabilities as there is no data available. The probability elicitation is simplified using the 'ranked nodes' a Truncated Normal distribution as defined in [60]. These nodes allow us to specify the parameters of a variable using the weighted mean of the parameters of their parent variables. We applied the ranked nodes to specify the parameters of synthetic and latent variables (represented with white ovals in Figure 9.2) and our main expert determined a set of weights for the following latent variables. For the 'Weather Condition' variable, we applied a weight of 2 for the Temperature and 'Air Pressure' variables since our expert believes these two play a major role in the feeling of the weather condition for people with RA. To parameterise the Participation variable, we defined the weights of 2.5 and 1.5 respectively for 'Financial Status' and 'Work Capacity' variables, since the two factors have crucial effects on a person's social aspects. 'Disease Acceptance' variable is associated with 'Disease Duration' and 'Disease Knowledge', where the latter is believed to have a double effect on the acceptance of the disease; therefore, we gave a weight of 2 to the 'Disease Knowledge'. Since a flare can affect the short-term disease activity, we defined a weight of 2 for the Flare variable in the parameterising of 'Current Disease Activity' variable. We applied no weights greater than one for the rest of the latent variables including 'Weather Effect', Flare, 'Overall Disease Activity', Independence, Empowerment, and QoL.

The evidence variables (represented with green ovals) with no parent variables, e.g., Sex, Temperature, or 'Social Activity', have equal probabilities for their states. For 'Medication Adherence', 'Financial Status', and 'Physical Activity', we have defined no ranked nodes, nor their node probability tables (NPTs) have non-equal probabilities for the parent nodes. We know that defining a NPT with equal probabilities for the states of the parent variables has no effect on the child variables. For all variables in the Disease Manifestation fragment, namely, 'Arthritis Pain', 'Morning Stiffness', Fatigue, 'Deformation Effect', and Mobility, we specify ranked nodes with parameters calculated by the weighted mean method on the parameters of their parent variables. We applied equal weights of 1 for all of them as our main expert found no difference between the effect of their parent variables.

For the parameterisation of 'Advice Priority' variable, we use expert's opinion to specify the NPT. Our main expert gave the priority of each combination of the parents of 'Advice Priority', i.e., Independence, Empowerment, and Participation. Each combination receives a priority based on the regions of Figure 9.1 and the priorities of each combination for the regions are originally based on [115]'s suggestion, as we mentioned in Subsection 9.2.4. Table 9.5 presents the priorities of each region numbered from 1 to 7 in any combination of the state of independence, empowerment, and participation. Our main expert has given the

Table 9.5 Priority of regions based on combination of states of independence, empowerment, and participation

Independence	Empowerment	Participation	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Region 7
Low	Low	Low	5	6	7	2	3	4	**1**
		Medium	3	4	7	**1**	5	6	2
		High	2	3	7	**1**	4	5	6
	Medium	Low	2	7	3	4	**1**	5	2
		Medium	**1**	6	7	2	3	4	5
		High	**1**	5	7	2	3	4	6
	High	Low	3	7	4	5	**1**	6	2
		Medium	**1**	7	4	5	2	6	3
		High	**1**	6	7	2	3	5	4
Medium	Low	Low	7	2	3	4	5	**1**	6
		Medium	6	**1**	7	2	5	3	4
		High	5	**1**	7	2	4	6	3
	Medium	Low	6	7	**1**	4	2	3	5
		Medium	5	6	7	2	3	4	**1**
		High	2	3	7	**1**	4	5	6
	High	Low	6	7	**1**	5	2	4	3
		Medium	2	7	3	4	**1**	5	6
		High	**1**	6	7	2	3	4	5
High	Low	Low	7	2	3	4	5	**1**	6
		Medium	7	**1**	4	5	6	2	3
		High	6	**1**	7	2	5	3	4
	Medium	Low	7	4	**1**	6	5	2	3
		Medium	7	2	3	4	5	**1**	6
		High	6	**1**	7	2	5	3	4
	High	Low	6	7	**1**	5	2	3	4
		Medium	6	7	**1**	5	2	3	4
		High	5	6	7	2	3	4	**1**

priorities in a way that the Low or Medium states of QoL components get improved and the High states of QoL components stabilise. The priority 1 in each combination (row) is in a bold font to highlight the region that we should focus on to give advice, which we describe in detail in Section 9.5. To transform the table of priority into a NPT for the 'Advice Priority' variable, we first reversed the priorities (e.g., 7 to 1 and 1 to 7), then we normalised the table.

9.5 Rule-based System for Personalised Care for Living with Rheumatoid Arthritis

We propose a rule-based mechanism to generate a set of advice and then modify them to obtain personalised advice. The following subsections present the rules for advice generation and the rules for advice modification.

9.5.1 Advice Generation

The rule-based system generates a set of advice based on the priority of advice that was estimated in the 'Advice Priority' variable of the BN model (described in Section 9.4) after receiving observations for the evidence variables. In this stage, we suggest (only) selecting the state of the 'Advice Priority' with the highest posterior probability. The selected state - that can be Independence, Empowerment, Participation, 'Independence & empowerment', 'Independence & participation', 'Empowerment & participation', or 'Independence & empowerment & participation' - is mapped to a prioritised region - that can be Region 1, Region 2, Region 3, Region 4, Region 6, or Region 7, respectively corresponding

Table 9.6 Prioritised regions with the NRAS advice.

Prioritised Region	NRAS Advice
Region 1	Help on improving self-care
	Help on improving physical activities
Region 2	Help on medication adherence
	Help on psychological distress and well-being
Region 3	Help with meeting others
Region 4	Fatigue matters
	Help with sleep
	Help with pain and stiffness
	Help with disease knowledge and acceptance
	Help with flares
Region 5	Help with role activities
Region 6	Help with emotions and relationships
Region 7	Help on improving lifestyle

to one of the selected states. The mapped prioritised regions lead to one or multiple advice that are listed in Table 9.6. The prioritised regions correspond to the priority 1 in each row of Table 9.5 that highlights the priority of each combination of QoL components. The advice of Table 9.6 is taken from the NRAS website and has a hyperlink that directs to the NRAS advice. Each advice contains appropriate recommendations for improvement of declined independence, empowerment, or participation, or for stabilising the already improved ones.

After generating advice based on the state of QoL components, we add further advice, if an undesired state of the evidence variable is observed. This advice intends to improve the evidence variables specifically, without taking the QoL components into consideration. The advice are originally from the NRAS. Table 9.7 displays the evidence variables, their undesired state, and an advice from the NRAS. Each advice has a hyperlink that directs to a web page of the NRAS.

Table 9.7 Further advice for evidence variables with undesired state being observed.

Type	Evidence Variable	Undesired State	NRAS Advice
Personal Factor	BMI	Underweighted, Obese	Dietary recommendations
QoL Characteristics	Financial Status	Bad	Financial benefits and personal independence
	Self-care	Low	Self-care recommendations
	Mood	Depressed	Depression recommendations
	Sleep Quality	Low	Sleep matters
Lifestyle Choices	Work Capacity	Low	Work recommendations
	Social Activity	Low	Recommendations on social activity
	Physical Activity	Low	Importance of exercise
	Belonging	Low	Hobbies
	Level of Tension	High	Relaxation during flares
	Medication Adherence	Low	Importance of medication adherence
	Disease Knowledge	Low	Knowing about RA
Disease Manifestations	Arthritis Pain	Severe	Managing pain
	Morning Stiffness	More than 2 hours	Managing stiffness in flares
	Fatigue	Severe	Fatigue matters
	Deformation Effect	High	Joint protection
	Mobility	Low	Mobility and exercise recommendations

Table 9.8 presents further advice for Flare and 'Weather Effect' latent variables, if their undesired states are estimated to have the highest posterior probability compared to the other states. The advice is taken from NRAS and each of them has a hyperlink that directs to a web page of NRAS.

Table 9.8 Further advice for latent variables with undesired state being estimated.

Type	Latent variable	Undesired State	NRAS Advice
Disease Activity	Flare	Severe	Help on improving lifestyle
Environmental Factors	Weather Effect	Unpleasant	Coping with weather conditions

9.5.2 Advice Modification

The set of advice collected in the advice generation (Subsection 9.5.1) needs to be reviewed and modified, if needed. The modification removes any advice that contradicts with the observations given by patients; therefore, provides us with personalised advice. For this, we go through the collected advice and remove any advice that targets an evidence variable which is already having a desired state, according to the observations, or any advice that targets a latent variable which is estimated to have the highest posterior probability for their desired state. As an example, if a patient observes a low sleep quality, high self-care, and low physical activity, they will have a medium independence. Assume this patient has high participation and empowerment and desired states of other evidence variables are observed, then the advice list will include the following advice: 'help with improving self-care' and 'help on improving physical activities' because Region 1 is prioritised (see Table 9.6) as well as two further advice: 'sleep matters' and 'importance of exercise' for having Low states

Table 9.9 Modification of redundant advice for evidence variables with desired state being observed.

Type	Evidence Variable	Desired State	No NRAS Advice
Personal Factor	BMI	Normal	Dietary recommendations
QoL Characteristics	Financial Status	Good	-
	Self-care	High	Help on improving self-care
	Mood	Happy	*Help on psychological distress and well-being
	Sleep Quality	High	Help with sleep
Lifestyle Choices	Work Capacity	High	Help with role activities
	Social Activity	High	Help with role activities
	Physical Activity	High	*Help on improving physical activities
	Belonging	High	Help with emotions and relationships
	Level of Tension	Low	*Help on psychological distress and well-being
	Medication Adherence	High	Help on medication adherence
	Disease Knowledge	High	Help with disease knowledge and acceptance
Disease Manifestations	Arthritis Pain	None	*Help with pain and stiffness
	Morning Stiffness	Less than 30 minutes	*Help with pain and stiffness
	Fatigue	None	Fatigue matters
	Deformation Effect	None	-
	Mobility	High	*Help on improving physical activities

being observed for sleep quality and physical activity (see Table 9.7). The first advice, 'help with improving self-care', is in the list despite self-care being observed as high. Therefore, this advice is redundant and needs to be removed.

Table 9.9 shows the NRAS advice that should be removed if the desired state of evidence variables are observed. The NRAS advice of Table 9.9 are all presented in Table 9.6. The stared advice in Table 9.9 are repeated in another row of the same table, e.g. 'Help on psychological distress and well-being' mentioned for Mood and 'Level of Tension'. This advice should be removed only if the desired states of both evidence variables are being observed at the same time. Otherwise, advice is needed and can be given to the patient.

Table 9.10 shows the NRAS advice that should be removed if the desired state of the Flare and 'Weather Effect' latent variables are estimated to have the highest posterior probability compared with the other states.

Table 9.10 Modification of redundant advice for latent variables with desired state being estimated.

Type	Latent variable	Desired State	No NRAS Advice
Disease Activity	Flare	None	Help on improving lifestyle
Environmental Factors	Weather Effect	Pleasant	Coping with weather conditions

9.6 Evaluation

As stated in [216], the ground-truth is often unknown in medicine and we need to use experts' judgement instead. Our study on personalised care for living with RA lacks available data to compare the prediction of the BN model with and evaluate its performance. Therefore, we ask our experts to provide their knowledge and allow us to evaluate BN performance in two ways: scenario analysis and sensitivity analysis.

9.6.1 Scenario Analysis

We evaluate the model using a set of scenarios. For this, we use a set of interviews from a formal semi-structured interview study, called AtTRA, about life with RA as a basis to initially validate the model from a patients' perspective. We develop four scenarios directly from the text of the interviews by coding them and attaining the evidences and expected state of the output variables in a blind way. The BN model receives evidence for each scenario and estimates the output variables (explained in Section 9.2). The evidences extracted from AtTRA interviews for scenario number 1, 2, 3, and 4 in Tables C.5, C.6, C.7, and C.8, respectively.

The output variables include nine variables: Independence, Empowerment, Participation, QoL, 'Disease Acceptance', 'Overall Disease Activity', 'Current Disease Activity', Flare, and 'Advice Priority', which are represented by green ovals in Figure 9.2.

On the other hand, we entered the evidences as inputs to the BN model and obtained the outputs of BN. We compare the expected states and BN estimations to evaluate the model. One issue is that the expected outputs are just one state of the output variable, however, the estimations of BN model are posterior probabilities of the output variables.

The main expert's expected outcomes and the estimation of outcome variables in the BN model are shown in Table 9.11. The two expectations with two states are expected by our main expert since he expects either of the states. We take both cases as the expected states. There is one estimation ('Disease Acceptance' for Scenario 2) with two estimated posterior probabilities of 50%. Table 9.12 shows a comparison between the outcomes of the model and the main expert's expectations.

Table 9.11 BN outputs and expected states of Flare, 'Current Disease Activity', 'Overall Disease Activity', and 'Disease Acceptance'.

Scenario	Flare		Current Disease Activity		Overall Disease Activity		Disease Acceptance	
	Estimation	Expectation	Estimation	Expectation	Estimation	Expectation	Estimation	Expectation
1	None	None	Remission	Low or Remission	Remission	Low or Remission	High	High
2	Mild	None	Low	Remission	Low	Remission	Medium or High	High
3	Mild	Mild	Low	Low	Low	Low	Low	Medium
4	Mild	Severe	Low	Moderate	Low	Moderate	High	High

Table 9.12 BN outputs and expected states of independence, participation, empowerment, and QoL.

Scenario	Independence		Participation		Empowerment		QoL	
	Estimation	Expectation	Estimation	Expectation	Estimation	Expectation	Estimation	Expectation
1	Medium	Low	Medium	Low	Medium	Low	Medium	Medium
2	High	Low	High	-	Medium	-	Medium	Low
3	Medium	High	High	High	Medium	Medium	Medium	Medium
4	Medium	Low	Low	Low	Medium	Medium	Medium	Low

Table 9.13 shows the advice of the model and compares them with the main expert's expectations. The expected advice is based on the expected QoL components given by the main expert.

Table 9.13 BN advice and expected states of advice priority.

Scenario	Advice 1		Advice 2	
	Estimation	Expectation	Estimation	Expectation
1	Independence & empowerment & participation	Independence & empowerment & participation	Independence & empowerment	Independence & empowerment
2	Empowerment	Empowerment	Empowerment & participation	Empowerment & participation
3	Independence & empowerment	Empowerment	Independence	Independence & empowerment
4	Participation	Independence & participation	Empowerment & participation	Independence & empowerment & participation

9.6.2 Sensitivity Analysis

Using Tornado graphs, we analyse the output of the BN model given the inputs. We consider the 'Advice Priority' variable as the outcome of the BN model and we select the 25 evidence variables (represented with green ovals in Figure 9.2) as the inputs of the BN model. The Tornado graph for the 'Independence & empowerment & participation' state of the 'Advice Priority' variable is displayed in Figure 9.1. Given the median values of the inputs, the posterior probability of the state becomes 0.177, which is the baseline of the graph. The horizontal bars show the change of the posterior probability of 'Independence & empowerment & participation' state, given the most desired state and the medium desired state of each input, e.g., Good and Average states of the Mobility variable, respectively.

The inputs with the greatest effect on the output variable are on the top of the figure and those with the least effect are on the bottom. According to Figure 9.3, the Mobility has the highest effect on the 'Advice Priority', followed by the 'Deformation Effect', 'Morning Stiffness', 'Financial Status', 'Arthritis Pain', Fatigue, and so on. Our main expert approves the order of the variables to be as he expected.

The Tornado graph for other states of the 'Advice Priority' variable are shown in Figures C.1, C.2, C.3, C.4, C.5, and C.6.

9.7 Summary

This chapter shows how we build a RS to provide personalised advice for living with chronic diseases with a BN model in the heart of this system, that corresponds to the Objective 4

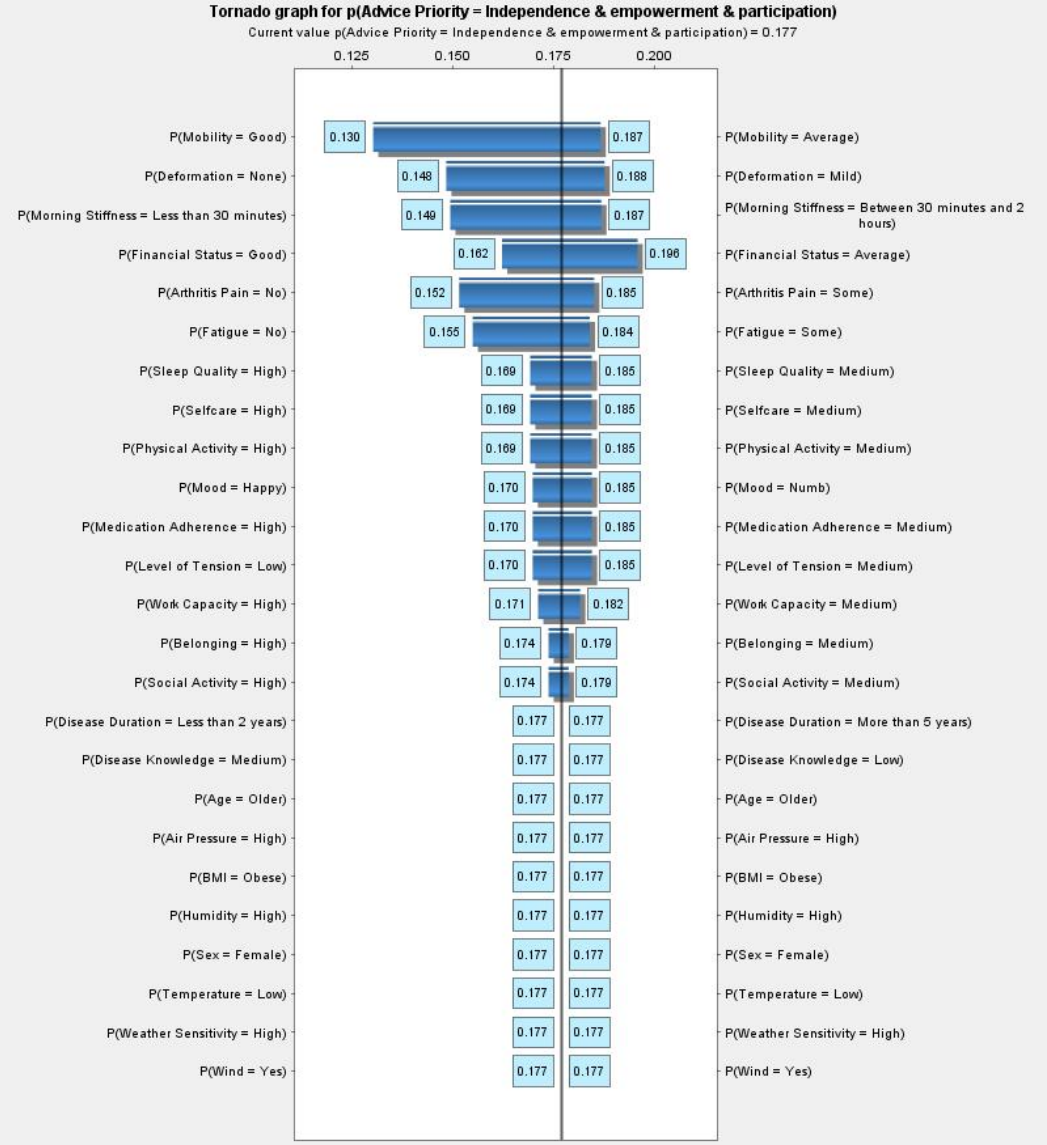


Fig. 9.3 Tornado graph of the 'Independence & empowerment & participation' state of the 'Advice Priority' variable of BN model for personalised care for living with RA.

of this thesis. The BN model estimates the disease activity, flare, and leads us to estimate the QoL components - empowerment, participation, and independence - and QoL itself. A rule-based system then operates the estimation of BN model and generates personalised advice.

In the beginning, we describe the variables required for personalised care for living with RA. These variables are specified using a model of care pathway that illustrates the steps of the proposed RS, coded interviews with patients, and a literature review. We then explain how we can collect information and observations of the patients by presenting a set of questions

to ask from patients. We specify the parameters of the BN model using ranked nodes derived from experts' knowledge.

We formed an advice graph by adopting [115]'s categorisation of advice which is based on the state of QoL components. The advice categories contain a set of advice links taken from NRAS's website.

Once we entered patient's information and observations in the BN model, it provides us with the estimation of empowerment, participation, and independence. These estimations lead us to a map the person to one or multiple advice from the advice graph. This way we tackle the challenge of providing personalised advice for living with RA. It helps us to estimate the disease activity, flare, and QoL components and recommend patients useful advice to manage their QoL.

In this chapter, we deal with unavailability of any quantitative data for model development and evaluation. We have access to a limited number of dedicated experts in creating the structure of the model and quantifying it. The evaluation lacks a larger number of scenarios that could reflect the performance of the BN model and so the RS.

Chapter 10

Summary and Future Directions

In this thesis, we addressed some of the challenges of building Bayesian network (BN) models for diagnosis, self-management, and personalised care of Rheumatoid Arthritis (RA). In Chapters 5 and 6, we described how we build BN models for medical diagnosis and we showed how we created BN models for the diagnosis of RA using experts' knowledge and retrieved knowledge from the pre-existing medical ontologies. In Chapters 7 and 8, we explained how to use data and knowledge to build dynamic BN (DBN) models for the self-management of RA. In Chapter 9, we proposed a BN model and a rule-based system to provide personalised advice for living with RA.

In Section 10.1, we summarise the contributions of this thesis considering the objectives of the thesis that was presented in Section 1.2. We also explain the future directions in Section 10.2.

10.1 Contributions

The overall objective of this thesis is that we can use probabilistic models to provide decision support for a chronic disease such as RA, covering diagnosis, self-management, and personalised care. This overall objective entails four research objectives. Objectives 1 and 2 focus on using data and/or knowledge to create diagnostic BN models. Objective 3 addresses the building of DBN models for self-management. Finally, Objective 4 is about building a BN model and a rule-based system for personalised care. The following parts show how we complete the objectives of this thesis:

Objective 1: Show how data and knowledge elicited from experts, literature, and clinical guidelines were used to build BN models for efficient prediction of medical diagnosis.

We contribute to this objective by showing how to build a BN model for the diagnosis of RA. Specifically:

- We reviewed the literature on BN models and clinical decision support (CDS) developed for the diagnosis of RA, and we did not find any BN models developed for the diagnosis of RA, as described in Subsections 2.6.1 and 2.7.1. To create the structure of the BN model for diagnosing RA, we showed how the knowledge of experts, literature, and clinical guidelines were used (Subsection 5.3.2).

- We explained how the data collected in a cohort study, experts' judgement, and the information published in the literature were used to parameterise the BN model (Subsection 5.3.3). We used various sources of parameters because the a complete dataset that contains the required records was not available or perhaps does not exist.

- We compare the results of the BN model learnt from knowledge with another BN model learnt entirely from data, indicating the efficiency of diagnosis predicted by the proposed BN model (Subsection 5.4).

Objective 2: Investigate how the knowledge of medical care processes is captured from care pathways and retrieved from pre-existing medical ontologies and translated into candidate BN models for medical diagnosis.
To contribute to this objective, we show how to build BN models for medical diagnosis using the knowledge represented in the models of care pathway and the knowledge retrieved from the medical ontologies (Chapter 6). Specifically:

- We did a literature review to find the state-of-the-art of building BNs from ontologies and came across a gap in developing diagnostic BNs from medical ontologies, as presented in Subsection 3.5.2. For building BNs from ontologies, we first investigated care pathways, knowledge graphs, and BNs by contrasting their properties in Section 3.5. We further analysed care pathways and explained how care pathways can be used to create knowledge graphs from pre-existing medical ontologies in Section 6.2. We find out that care pathways can represent initial knowledge and this knowledge can lead us to retrieve detailed knowledge from medical ontologies.

- In Section 6.3, we presented a simple care pathway for the diagnosis of RA. Then, we described a step-by-step method to build the structure of a BN model from the simple care pathway and by retrieving further knowledge from Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) ontology (Subsection 6.3.1) and Unified Medical Language System (UMLS) ontology (Subsection 6.3.2). The steps of each

method encompass multiple steps including knowledge retrieval, refining knowledge, and finally translating the relationships of the knowledge graph into causal or association dependencies using a set of rules on the meaning of the relationships. By converting the concepts of the knowledge graph into the variables of the BN model and adding the dependencies, we were able to build the candidate BN models for diagnostic application.

- We built two candidate BN models for the diagnosis of RA using knowledge graphs made from knowledge retrieved from SNOMED CT and UMLS. These models represent the main variables required for RA diagnosis; therefore, our aim of partial automation of BN modelling is achieved. However, the models lack variables related to the pathogenesis mechanism of RA and may not be useful diagnostic models compared to other models built from experts' knowledge and/or available data.

Objective 3: Show how data, elicited knowledge from experts, clinical guidelines, and literature are used to build DBN models for self-management of chronic diseases and how these DBN models are used to support clinicians in monitoring and medication review.

We contribute to the objective in two stages. The first stage is an initial DBN model for self-management of RA, described in Section 7.3. The second stage is an extension of the initial DBN model with the treatment-related variables added to the DBN model, as described in Section 8.5. We show how to do reasoning with both DBN models and compare the recommendations of the model on monitoring and medication with the expected recommendations. Specifically:

- We reviewed the literature on BN models and CDS for the treatment and self-management of RA. We did not find any studies applying BNs as CDS for the treatment and self-management of RA, as we explained in Subsections 2.6.2 and 2.7.1. To build a DBN model, we specify the variables of DBN model for self-management of RA from expert's knowledge and guidelines, as we described in Sections 7.2 and 8.4.

- Subsections 7.3.1 and 8.5.1 explained the steps we took to determine the dependencies between the variables of the models to create a meaningful structure for self-management of RA.

- Subsections 7.3.3 and 8.5.3 elaborated how the available data, expert's judgement, and the information from the literature helped us to specify the parameters of the DBN model for self-management of RA.

- Sections 7.4 and 8.6 elucidated the DBN models for self-management of RA as decision support systems. Subsection 7.4.2 expressed the dynamics of the predicted variables, mainly disease state and flare. Subsection 7.4.3 demonstrates how we estimate whether a disease is under control or not and how the two trend variables for disease state and flare help us to determine if the disease is stable or not. The subsection presents the advice as a combination of disease control and the two trend variables. Subsection 8.6.3 elaborated the dynamics of disease control and tolerance for commonly prescribed drugs for RA, known as disease-modifying antirheumatic drugs, which involves a person's medications taken and the adverse medication events (AMEs) that they may experience. The estimation of tolerance for toxic drugs can support clinicians to stratify patients based on their tolerance for drugs. Subsection 8.6.4 explained the combination of the AMEs and their trend, the tolerance for disease-modifying antirheumatic drugs (DMARDs), disease control, and an overall progression trend of the disease, which leads to provide a recommendation on medication review.

- We presented the outcomes of reasoning with the DBN models for self-management of RA using a set of interpolated data in Subsections 7.5.1 and 8.7.1. For these, we used a set of interpolated data (described in Subsections 4.4.4) and a set of expanded medication data (described in Subsection 4.4.5). While a gold standard was not available or does not exist, we considered the gold standards for appointment scheduling and medication review to be the occurrence of major flares (described in Subsection 4.4.4) and a set of medication decisions extracted from medication records using deterministic rules (described in Subsection 4.4.5), respectively. Comparing the major flare data and the advice on appointment scheduling revealed promising results of the initial DBN's performance. The comparison between medication decisions with the extended DBN's predicted advice on medication review indicates the ability of the DBN model to give advice to clinicians on medication and support them in decision-making for prescribing medications.

- We demonstrated the reasoning with DBN models using a set of scenarios of dummy patients in Subsections 7.5.2 and 8.7.2. The results evaluate the performance of DBN models and indicate their strengths and weaknesses.

Objective 4: Investigate how to use the knowledge gathered from experts and literature and to build a BN model to predict the quality of life and how to recommend personalised advice to improve the experience of living with chronic diseases.

To contribute to this objective, we created a BN model and a rule-based system to generate a set of advice based on BN's prediction of advice priority, and then modify the advice generated to personalise the advice. Specifically:

- We did a literature review to obtain the state-of-the-art of building BN models and CDS for providing personalised care for living with chronic diseases. We noticed the gap of BN-based CDS for personalised care for living with, as we explained in Subsection 2.6.2. To build the BN model for personalised care, we first created a pathway for personalised care (described in Subsection 4.3.4) to specify the variables of the BN model for personalised care, as presented in Subsection 9.2. In Subsection 9.4.1, we presented how the dependencies between the variables are determined and formed the structure of the BN model for personalised care. The BN model represents the disease manifestation variables and lifestyle choices on the two sides of the model, aiming to estimate the disease activity, flare, and the components of quality of life (QoL), namely, independence, empowerment, and participation, corresponding to the physical, psychological, and social aspects of QoL.

- We showed the parameterisation of the BN model from the expert's judgement in Subsection 9.4.2. We demonstrated how the expert considers [115]'s advice and provides a set of parameters to define the priorities of the advice variable, if the BN estimation indicates the decline of one or multiple components of QoL.

- In Section 9.5, we elaborated the steps of generating advice based on the advice priority estimated by the BN, and then the modification of the generated advice to remove the possibly redundant advice and add targeted advice based on personal information, disease manifestation, or lifestyle choices.

- We presented a brief evaluation using a set of scenarios and sensitivity analysis in Section 9.6. The outcomes indicate promising results that the proposed BN and rule-based system for personalised advice can improve the experience of living with RA.

10.2 Future Directions

Future directions can focus on further validation of the models to improve their performance and fulfil the potential benefits of the proposed decision support. Another future direction can include the extension of the proposed models. The future directions corresponding to each research objective are as follows:

Objective 1: Show how data and knowledge elicited from experts, literature, and clinical guidelines were used to build BN models for efficient prediction of medical diagnosis.
The BN model for the diagnosis of RA can be improved and extended. Specifically:

- The knowledge-based BN model for RA diagnosis outperforms the learnt BN model (Section 5.4). The knowledge-based model can be further evaluated, mainly to determine the possible negative influence of the missing antibody values on its parameter learning. In addition, we have noticed that the Steroid variable complicates the conditional probabilities of the knowledge-based model (described in Subsection 5.4.3). Removal of the Steroid variable may have a positive effect on the accuracy of prediction. Using more states for signs, symptoms, and serology results is expected to improve the performance. We can also use narrower intervals for critical values for discretising the data of some variables, particularly the tender joints count and swollen joints count.

- The BN model for the diagnosis of RA includes a gene-environment pathogenesis mechanism supported by the established literature (explained in Subsection 5.2.3). Although the mechanism represents interactions between two genes and an environmental factor, the variables of the two genes in the proposed model can not involve in reasoning since the flow of information would be blocked by their child variables (see Figure 5.1). Further investigation of the gene-environment may bring the genes in the reasoning process and can prove their potential effect in the diagnosis of RA.

Objective 2: Investigate how the knowledge of medical care processes is captured from care pathways and retrieved from pre-existing medical ontologies and translated into candidate BN models for medical diagnosis.
Candidate BN models for diagnostic application can be investigated and the proposed step-by-step methods to build candidate BNs can be extended. Specifically:

- In the future, the candidate BN models for medical diagnostic that we built from medical ontologies can be parameterised using available data and their predicting performance can be evaluated. Their performance can also be compared with that of other similar BN models that are built from experts' knowledge and/or available data.

- Proposed methods to build BNs from ontologies can be extended to have another refining mechanism that prevents any duplication of variable and trivial variables. One or multiple additional steps can also be outlined to specify the pathogenesis mechanisms. For example, the method for BN modelling using UMLS ontology

(Subsection 6.3.2) finds incorporating genes in RA development; however, they are not interconnected with other variables such as rheumatoid factor or any other variable that they may have a relationship with. This can be investigated and perhaps added to the proposed methods.

Objective 3: Show how data, elicited knowledge from experts, clinical guidelines, and literature are used to build DBN models for self-management of chronic diseases and how these DBN models are used to support clinicians in monitoring and medication review.

Proposed DBN models for self-management can be improved and extended. Specifically:

- Although the results of evaluating the initial DBN model for self-management of RA are promising, the performance of the extended DBN model is not as we expected. This can be improved by eliciting the knowledge of a group of rheumatology experts and finding their consensus on the parameters of the advice variable for medication review. The states of other latent variables such as 'Overall Disease Control' can be updated, so that they would be able to estimate the disease control more accurately. As suggested in the future direction of the diagnostic model, using more states for signs, symptoms, and serology results and narrower discretisation for some variables are expected to improve the performance of DBN models as well.

- The next step in improving the DBN model for self-management of RA might include investigating if the suitable data is available in the electronic health records or whether running a study to learn if the decision support is practical in clinical use. The latter would also include gathering data.

- The DBN model for self-management of RA can be extended into another time-based BN model that can operate with irregular time intervals. We expect that clinicians and patients would need a model with specifications that are closer to real-world use.

- Since clinicians consider some risk factors, mainly body mass index, serostatus, comorbidities, and medical background, in prescribing DMARDs, they can be added to the DBN model for self-management of RA and get involved in generating medication advice.

- Another extension of DBN models for self-management of RA is to add the various medications commonly prescribed for RA, their dosage, and detailed adverse events in the model. This can give a higher accuracy to the model in predicting the disease state and the effect of medications on it; therefore, the prediction of advice on medication

can be improved. This can ultimately provide a detailed recommendation for each medication, although the parameters of the recommendations need to be defined precisely by consulting a group of experts and finding their consensus.

Objective 4: Investigate how to use the knowledge gathered from experts and literature and to build a BN model to predict the quality of life and how to recommend personalised advice to improve the experience of living with chronic diseases.
The proposed recommender system, including the BN model and rule-based system, that provide personalised advice for living with chronic diseases can be evaluated and extended. Specifically:

- The proposed BN model and rule-based system for personalised advice were achieved to provide the expected results, based on our brief evaluation. Although there is no available data, the performance of the BN and the rule-based system can be further evaluated and the model or rules can be updated accordingly.

- An extension of the recommender system can include investigating whether running a study to learn if personalised advice is helpful in practice and if it is useful for patients to manage their disease. This extension would include gathering data from patients.

References

[1] Abu-Hanna, A. and Lucas, P. J. (2001). Prognostic models in medicine. *Methods of information in medicine*, 40(01):1–5.

[2] Adel, T. and de Campos, C. (2017). Learning bayesian networks with incomplete data by augmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

[3] Agena (2018). Agenarisk. https://www.agenarisk.com/ (Accessed on 16/10/2020).

[4] Ahire, S. B. and Khanuja, H. K. (2015). A personalized framework for health care recommendation. In *2015 International Conference on Computing Communication Control and Automation*, pages 442–445. IEEE.

[5] AIMS2 (1990). Arthritis impact measurement scales 2. http://www.ser.es/wp-content/uploads/2016/07/AIMS2-ORIGINAL.pdf (Accessed on 22/10/2020).

[6] Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham III, C. O., Birnbaum, N. S., Burmester, G. R., Bykerk, V. P., Cohen, M. D., et al. (2010). 2010 rheumatoid arthritis classification criteria: an american college of rheumatology/european league against rheumatism collaborative initiative. *Arthritis & rheumatism*, 62(9):2569–2581.

[7] Alharbi, R. F., Berri, J., and El-Masri, S. (2015). Ontology based clinical decision support system for diabetes diagnostic. In *2015 Science and Information Conference (SAI)*, pages 597–602. IEEE.

[8] Alpízar-Rodríguez, D., Pluchino, N., Canny, G., Gabay, C., and Finckh, A. (2017). The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatology*, 56(8):1254–1263.

[9] Arnett, F. C., Edworthy, S. M., Bloch, D. A., Mcshane, D. J., Fries, J. F., Cooper, N. S., Healey, L. A., Kaplan, S. R., Liang, M. H., Luthra, H. S., et al. (1988). The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 31(3):315–324.

[10] Arp, R., Smith, B., and Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Mit Press.

[11] Arthritis Research UK (2014). Rheumatoid arthritis. https://www.arthritisresearchuk.org/arthritis-information/conditions/rheumatoid-arthritis.aspx (Accessed on 02/04/2020).

[12] Arthritis Research UK (2018). Rheumatoid arthritis. https://www.arthritisresearchuk.org/arthritis-information/conditions/rheumatoid-arthritis.aspx (Accessed on 15/03/2018).

[13] Auer, S., Kovtun, V., Prinz, M., Kasprzik, A., Stocker, M., and Vidal, M. E. (2018). Towards a knowledge graph for science. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*, pages 1–6.

[14] Begovich, A. B., Carlton, V. E., Honigberg, L. A., Schrodi, S. J., Chokkalingam, A. P., Alexander, H. C., Ardlie, K. G., Huang, Q., Smith, A. M., Spoerke, J. M., et al. (2004). A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (ptpn22) is associated with rheumatoid arthritis. *The American Journal of Human Genetics*, 75(2):330–337.

[15] Bellazzi, R., Larizza, C., and Riva, A. (1998). Temporal abstractions for interpreting diabetic patients monitoring data. *Intelligent Data Analysis*, 2(1-4):97–122.

[16] Bingham, C. O., Alten, R., Bartlett, S. J., Bykerk, V. P., Brooks, P. M., Choy, E., Christensen, R., Furst, D. E., Hewlett, S. E., Leong, A., et al. (2011). Identifying preliminary domains to detect and measure rheumatoid arthritis flares: report of the omeract 10 ra flare workshop. *The Journal of Rheumatology*, 38(8):1751–1758.

[17] Bingham, C. O., Pohl, C., Woodworth, T. G., Hewlett, S. E., May, J. E., Rahman, M. U., Witter, J. P., Furst, D. E., Strand, C. V., Boers, M., et al. (2009). Developing a standardized definition for disease "flare" in rheumatoid arthritis (omeract 9 special interest group). *The Journal of Rheumatology*, 36(10):2335–2341.

[18] Bodenreider, O. (2004). The unified medical language system (umls): integrating biomedical terminology. *Nucleic acids research*, 32(suppl_1):D267–D270.

[19] Boyen, X. and Koller, D. (2013). Tractable inference for complex stochastic processes. *arXiv preprint arXiv:1301.7362*.

[20] Briggs, F., Ramsay, P., Madden, E., Norris, J., Holers, V., Mikuls, T. R., Sokka, T., Seldin, M. F., Gregersen, P., Criswell, L., et al. (2010). Supervised machine learning and logistic regression identifies novel epistatic risk factors with ptpn22 for rheumatoid arthritis. *Genes & Immunity*, 11(3):199–208.

[21] Bucci, G., Sandrucci, V., and Vicario, E. (2011). Ontologies and bayesian networks in medical diagnosis. In *2011 44th Hawaii International Conference on System Sciences*, pages 1–8. IEEE.

[22] Buckley, L., Ware, E., Kreher, G., Wiater, L., Mehta, J., and Burnham, J. M. (2020). Outcome monitoring and clinical decision support in polyarticular juvenile idiopathic arthritis. *The Journal of rheumatology*, 47(2):273–281.

[23] Bueno, M. L., Hommersom, A., Lucas, P. J., and Janzing, J. (2019). A probabilistic framework for predicting disease dynamics: A case study of psychotic depression. *Journal of biomedical informatics*, 95:103232.

[24] Burns, B. and Morrison, C. T. (2003). Temporal abstraction in bayesian networks. In *AAAI Spring Symposium, Palo Alto, CA*, volume 21.

[25] Bykerk, V. P., Bingham, C. O., Choy, E. H., Lin, D., Alten, R., Christensen, R., Furst, D. E., Hewlett, S., Leong, A., March, L., et al. (2016). Identifying flares in rheumatoid arthritis: reliability and construct validation of the omeract ra flare core domain set. *RMD open*, 2(1):e000225.

[26] Bykerk, V. P., Shadick, N., Frits, M., Bingham, C. O., Jeffery, I., Iannaccone, C., Weinblatt, M., and Solomon, D. H. (2014). Flares in rheumatoid arthritis: frequency and management. a report from the brass registry. *The Journal of rheumatology*, 41(2):227–234.

[27] Cedar Lake Ventures Inc. (2020). Average weather in london. "https://weatherspark.com/y/45062/Average-Weather-in-London-United-Kingdom-Year-Round#Sections-Temperature (Accessed on 04/06/2020).

[28] Çelik Ertuğrul, D. and Elçi, A. (2020). A survey on semanticized and personalized health recommender systems. *Expert Systems*, 37(4):e12519.

[29] Chabowski, M., Polański, J., Jankowska-Polanska, B., Lomper, K., Janczak, D., and Rosinczuk, J. (2017). The acceptance of illness, the intensity of pain and the quality of life in patients with lung cancer. *Journal of thoracic disease*, 9(9):2952.

[30] Chang, Y.-S., Fan, C.-T., Lo, W.-T., Hung, W.-C., and Yuan, S.-M. (2015). Mobile cloud-based depression diagnosis using an ontology and a bayesian network. *Future Generation Computer Systems*, 43:87–98.

[31] Chen, S. H. and Pollino, C. A. (2012). Good practice in bayesian network modelling. *Environmental Modelling & Software*, 37:134–145.

[32] Chen, X., Jia, S., and Xiang, Y. (2020). A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications*, 141:112948.

[33] Chi, Y.-L., Chen, T.-Y., and Tsai, W.-T. (2015). A chronic disease dietary consultation system using owl-based ontologies and semantic rules. *Journal of biomedical informatics*, 53:208–219.

[34] Community, T. S. (2020). scipy.interpolate.interp1d. https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.interp1d.html (Accessed on 04/03/2021).

[35] Constantinou, A. C., Fenton, N., Marsh, W., and Radlinski, L. (2016). From complex questionnaire and interviewing data to intelligent bayesian network models for medical decision support. *Artificial intelligence in medicine*, 67:75–93.

[36] Constantinou, A. C., Freestone, M., Marsh, W., and Coid, J. (2015). Causal inference for violence risk management and decision support in forensic psychiatry. *Decision Support Systems*, 80:42–55.

[37] Cooper, G. F. (1990). The computational complexity of probabilistic inference using bayesian belief networks. *Artificial intelligence*, 42(2-3):393–405.

[38] Cooper, G. S. (2008). Occupational exposures and risk of rheumatoid arthritis: continued advances and opportunities for research.

[39] Cordon, D. B. (1996). Critical pathways: a road to institutionalizing pain management. *Journal of pain and symptom management*, 11(4):252–259.

[40] Costa, P. C. and Laskey, K. B. (2006). Multi-entity bayesian networks without multi-tears.

[41] Da Costa, P. C. G., Laskey, K. B., and Laskey, K. J. (2006). Pr-owl: A bayesian ontology language for the semantic web. In *Uncertainty Reasoning for the Semantic Web I*, pages 88–107. Springer.

[42] Dean, T. and Kanazawa, K. (1989). A model for reasoning about persistence and causation. *Computational intelligence*, 5(2):142–150.

[43] Dellaert, F. (2002). The expectation maximization algorithm. Technical report, Georgia Institute of Technology.

[44] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

[45] Devitt, A., Danev, B., and Matusikova, K. (2006). Constructing bayesian networks automatically using ontologies.

[46] Ding, Z. and Peng, Y. (2004). A probabilistic extension to ontology language owl. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, pages 10–pp. IEEE.

[47] Ding, Z., Peng, Y., and Pan, R. (2006). Bayesowl: Uncertainty modeling in semantic web ontologies. In *Soft computing in ontologies and semantic web*, pages 3–29. Springer.

[48] Dixon, W. G., Beukenhorst, A. L., Yimer, B. B., Cook, L., Gasparrini, A., El-Hay, T., Hellman, B., James, B., Vicedo-Cabrera, A. M., Maclure, M., et al. (2019). How the weather affects the pain of citizen scientists using a smartphone app. *NPJ digital medicine*, 2(1):1–9.

[49] Do, C. B. and Batzoglou, S. (2008). What is the expectation maximization algorithm? *Nature biotechnology*, 26(8):897–899.

[50] Drinkaware (2019). Unit calorie calculator. https://www.drinkaware.co.uk/understand-your-drinking/unit-calculator#main (Accessed on 16/01/2020).

[51] Druce, K. L., Cordingley, L., Short, V., Moore, S., Hellman, B., James, B., Lunt, M., Kyle, S. D., Dixon, W. G., and McBeth, J. (2018). Quality of life, sleep and rheumatoid arthritis (quasar): a protocol for a prospective uk mhealth study to investigate the relationship between sleep and quality of life in adults with rheumatoid arthritis. *BMJ open*, 8(1).

[52] Ehrlinger, L. and Wöß, W. (2016). Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCCESS)*, 48:1–4.

[53] Engelbrecht, R. et al. (2005). Ontology driven construction of a knowledgebase for bayesian decision models based on umls. In *Connecting Medical Informatics and Bioinformatics: Proceedings of MIE2005: the XIXth International Congress of the European Federation for Medical Informatics*, volume 116, page 223. IOS Press.

[54] Estabragh, Z. S., Kashani, M. M. R., Moghaddam, F. J., Sari, S., and Oskooyee, K. S. (2011). Bayesian network model for diagnosis of social anxiety disorder. In *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, pages 639–640. IEEE.

[55] Fahmi, A., MacBrayne, A., Kyrimi, E., McLachlan, S., Humby, F., Marsh, W., and Pitzalis, C. (2020). Causal bayesian networks for medical diagnosis: A case study in rheumatoid arthritis. Technical report, EasyChair.

[56] Farmer, N. (2014). An update and further testing of a knowledge-based diagnostic clinical decision support system for musculoskeletal disorders of the shoulder for use in a primary care setting. *Journal of evaluation in clinical practice*, 20(5):589–595.

[57] Felson, D. T., Anderson, J. J., Boers, M., Bombardier, C., Furst, D., Goldsmith, C., Katz, L. M., Lightfoot Jr, R., Paulus, H., Strand, V., et al. (1995). American college of rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 38(6):727–735.

[58] Feng, X., Xu, X., Shi, Y., Liu, X., Liu, H., Hou, H., Ji, L., Li, Y., Wang, W., Wang, Y., et al. (2019). Body mass index and the risk of rheumatoid arthritis: an updated dose-response meta-analysis. *BioMed research international*, 2019.

[59] Fenton, N. and Neil, M. (2010). Comparing risks of alternative medical diagnosis using bayesian arguments. *Journal of Biomedical Informatics*, 43(4):485–495.

[60] Fenton, N. and Neil, M. (2012). *Risk assessment and decision analysis with Bayesian networks*. Crc Press.

[61] Flores, M. J., Nicholson, A. E., Brunskill, A., Korb, K. B., and Mascaro, S. (2011). Incorporating expert knowledge when learning bayesian network structure: a medical case study. *Artificial intelligence in medicine*, 53(3):181–204.

[62] Friedman, N., Ninio, M., Pe'er, I., and Pupko, T. (2002). A structural em algorithm for phylogenetic inference. *Journal of Computational Biology*, 9(2):331–353.

[63] Fries, J. F. (1982). The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *j Rheumatol*, 9:789–793.

[64] Galan, S. F., Aguado, F., Dıez, F., and Mira, J. (2002). Nasonet, modeling the spread of nasopharyngeal cancer with networks of probabilistic events in discrete time. *Artificial Intelligence in Medicine*, 25(3):247–264.

[65] Galán, S. F. and Díez, F. J. (2002). Networks of probabilistic events in discrete time. *International Journal of Approximate Reasoning*, 30(3):181–202.

[66] Gámez, J. A., Mateo, J. L., and Puerta, J. M. (2011). Learning bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Mining and Knowledge Discovery*, 22(1):106–148.

[67] Garcia-Zapirain, B., Garcia-Chimeno, Y., and Rogers, H. (2015). Machine learning techniques for automatic classification of patients with fibromyalgia and arthritis. *Int. J. Comput. Trends Technol*, 25(3):149–152.

[68] Gaševic, D., Djuric, D., and Devedžic, V. (2009). *Model driven engineering and ontology development*. Springer Science & Business Media.

[69] Gatti, E., Luciani, D., and Stella, F. (2012). A continuous time bayesian network model for cardiogenic heart failure. *Flexible Services and Manufacturing Journal*, 24(4):496–515.

[70] Gómez, J., Oviedo, B., and Zhuma, E. (2016). Patient monitoring system based on internet of things. *Procedia Computer Science*, 83:90–97.

[71] Gómez-Pérez, A. and Corcho, O. (2002). Ontology languages for the semantic web. *IEEE Intelligent systems*, 17(1):54–60.

[72] Gosselt, H. R., Verhoeven, M., Bulatović-Ćalasan, M., Welsing, P. M., de Rotte, M. C., Hazes, J. M., Lafeber, F. P., Hoogendoorn, M., and de Jonge, R. (2021). Complex machine-learning algorithms and multivariable logistic regression on par in the prediction of insufficient clinical response to methotrexate in rheumatoid arthritis. *Journal of Personalized Medicine*, 11(1):44.

[73] Gottheil, S., Khemani, E., Copley, K., Keeney, M., Kinney, J., Chin-Yee, I., and Gob, A. (2016). Reducing inappropriate esr testing with computerized clinical decision support. *BMJ Open Quality*, 5(1):u211376–w4582.

[74] Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5-6):907–928.

[75] Guan, Y., Zhang, H., Quang, D., Wang, Z., Parker, S. C., Pappas, D. A., Kremer, J. M., and Zhu, F. (2019). Machine learning to predict anti–tumor necrosis factor drug responses of rheumatoid arthritis patients by integrating clinical and genetic markers. *Arthritis & Rheumatology*, 71(12):1987–1996.

[76] HAQ-DI (2003). Health assessment questionnaire- disability index. https://integrationacademy.ahrq.gov/sites/default/files/HAQ-DI_0.pdf (Accessed 28/04/2020).

[77] Heidari, B. (2011). Rheumatoid arthritis: Early diagnosis and treatment outcomes. *Caspian journal of internal medicine*, 2(1):161.

[78] Helsper, E. M. and Van Der Gaag, L. C. (2002). Building bayesian networks through ontologies. In *ECAI*, volume 2002, page 15th.

[79] Hennell, S., Brownsell, C., and Dawson, J. (2004). Development, validation and use of a patient knowledge questionnaire (pkq) for patients with early rheumatoid arthritis. *Rheumatology*, 43(4):467–471.

[80] Honka, A. M., van Gils, M. J., and Pärkkä, J. (2011). A personalized approach for predicting the effect of aerobic exercise on blood pressure using a fuzzy inference system. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 8299–8302. IEEE.

[81] Horridge, M (2009). A practical guide to building owl ontologies using protege 4 and co-ode tools. http://mowlpower.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_2.pdf (Accessed on 31/08/2019).

[82] Horridge, M (2017). Snomed ct concepts. https://confluence.ihtsdotools.org/display/DOCTIG/3.1.+Components (Accessed on 19/08/2019).

[83] Hu, B., Hu, B., Wan, J., Dennis, M., Chen, H.-H., Li, L., and Zhou, Q. (2010). Ontology-based ubiquitous monitoring and treatment against depression. *Wireless communications and mobile computing*, 10(10):1303–1319.

[84] Hu, H., Elkus, A., and Kerschberg, L. (2016). A personal health recommender system incorporating personal health records, modular ontologies, and crowd-sourced data. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 1027–1033. IEEE.

[85] Hügle, M., Omoumi, P., van Laar, J. M., Boedecker, J., and Hügle, T. (2020). Applied machine learning and artificial intelligence in rheumatology. *Rheumatology advances in practice*, 4(1):rkaa005.

[86] IQWiG (2016). What is "normal" sleep? https://www.ncbi.nlm.nih.gov/books/NBK279322/ (Accessed on 22/10/2020).

[87] Jacquemin, C., Molto, A., Servy, H., Sellam, J., Foltz, V., Gandjbakhch, F., Hudry, C., Mitrovic, S., Granger, B., Fautrel, B., et al. (2017). Flares assessed weekly in patients with rheumatoid arthritis or axial spondyloarthritis and relationship with physical activity measured using a connected activity tracker: a 3-month study. *RMD open*, 3(1).

[88] Janssen, K. M., de Smit, M. J., Brouwer, E., de Kok, F. A., Kraan, J., Altenburg, J., Verheul, M. K., Trouw, L. A., van Winkelhoff, A. J., Vissink, A., et al. (2015). Rheumatoid arthritis–associated autoantibodies in non–rheumatoid arthritis patients with mucosal inflammation: a case–control study. *Arthritis research & therapy*, 17(1):174.

[89] Jimmy, B. and Jose, J. (2011). Patient medication adherence: measures in daily practice. *Oman medical journal*, 26(3):155.

[90] Jones, G., Nash, P., and Hall, S. (2017). Advances in rheumatoid arthritis. *Medical Journal of Australia*, 206(5):221–224.

[91] Josef S., S., van der Heijde, D., Machold, K. P., Aletaha, D., and Landewé, R. (2014). Proposal for a new nomenclature of disease-modifying antirheumatic drugs. *Annals of the Rheumatic Diseases*, 73:3–5.

[92] Karlson, E. W., Lee, I.-M., Cook, N. R., Manson, J. E., Buring, J. E., and Hennekens, C. H. (1999). A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 42(5):910–917.

[93] Kay, J. and Upchurch, K. S. (2012). Acr/eular 2010 rheumatoid arthritis classification criteria. *Rheumatology*, 51(suppl_6):vi5–vi9.

[94] Ker, J., Wang, L., Rao, J., and Lim, T. (2017). Deep learning applications in medical image analysis. *Ieee Access*, 6:9375–9389.

[95] Kong, G., Xu, D.-L., and Yang, J.-B. (2008). Clinical decision support systems: a review on knowledge representation and inference under uncertainties. *International Journal of Computational Intelligence Systems*, 1(2):159–167.

[96] Korb, K. B. and Nicholson, A. E. (2010). *Bayesian artificial intelligence*. CRC press.

[97] Kostova, Z., Caiata-Zufferey, M., and Schulz, P. J. (2014). The process of acceptance among rheumatoid arthritis patients in switzerland: A qualitative study. *Pain Research and Management*, 19(2):61–68.

[98] Kyrimi, E. (2019). *Bayesian Networks for Clinical Decision Making: Support, Assurance, Trust*. PhD thesis, Queen Mary University of London.

[99] Kyrimi, E., McLachlan, S., Dube, K., Neves, M. R., Fahmi, A., and Fenton, N. (2020a). A comprehensive scoping review of bayesian networks in healthcare: Past, present and future. *arXiv preprint arXiv:2002.08627*.

[100] Kyrimi, E., Mossadegh, S., Tai, N., and Marsh, W. (2020b). An incremental explanation of inference in bayesian networks for increasing model trustworthiness and supporting clinical decision making. *Artificial Intelligence in Medicine*, 103:101812.

[101] Kyrimi, E., Neves, M. R., McLachlan, S., Neil, M., Marsh, W., and Fenton, N. (2020c). Medical idioms for clinical bayesian network development. *Journal of Biomedical Informatics*, 108:103495.

[102] Lacave, C. and Díez, F. J. (2002). A review of explanation methods for bayesian networks. *The Knowledge Engineering Review*, 17(2):107–127.

[103] Larik, A. S. and Haider, S. (2010). Efforts to blend ontology with bayesian networks: An overview. In *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, volume 2, pages V2–598. IEEE.

[104] Lee, S., Eun, Y., Kim, H., Cha, H.-S., Koh, E.-M., and Lee, J. (2020). Machine learning to predict early tnf inhibitor users in patients with ankylosing spondylitis. *Scientific reports*, 10(1):1–9.

[105] Lezcano-Valverde, J. M., Salazar, F., León, L., Toledano, E., Jover, J. A., Fernandez-Gutierrez, B., Soudah, E., González-Álvaro, I., Abasolo, L., and Rodriguez-Rodriguez, L. (2017). Development and validation of a multivariate predictive model for rheumatoid arthritis mortality using a machine learning approach. *Scientific reports*, 7(1):1–10.

[106] Liao, K. P., Alfredsson, L., and Karlson, E. W. (2009). Environmental influences on risk for rheumatoid arthritis. *Current opinion in rheumatology*, 21(3):279.

[107] Liu, M., Hommersom, A., van der Heijden, M., and Lucas, P. J. (2017). Hybrid time bayesian networks. *International Journal of Approximate Reasoning*, 80:460–474.

[108] Liu, M., Stella, F., Hommersom, A., Lucas, P. J., Boer, L., and Bischoff, E. (2019). A comparison between discrete and continuous time bayesian networks in learning from clinical time series data with irregularity. *Artificial intelligence in medicine*, 95:104–117.

[109] Lötsch, J., Alfredsson, L., and Lampa, J. (2020). Machine-learning–based knowledge discovery in rheumatoid arthritis–related registry data to identify predictors of persistent pain. *Pain*, 161(1):114–126.

[110] Lucas, P. J., Van der Gaag, L. C., and Abu-Hanna, A. (2004). Bayesian networks in biomedicine and health-care. *Artificial intelligence in medicine*, 30(3):201–214.

[111] MacBrayne, A. (2019). Ra case study: Update. https://pambayesian.org/wp-content/uploads/2019/02/Amy-MacBrayne-PAMBAYESIAN-RA-6th-Feb-2019.pdf (Accessed on 01/02/2021).

[112] MacBrayne, A. C. B., Pott, J., Petrovic, V., Pitzalis, C., and Humby, F. (2020). P214 preliminary results: driving improvements in disease outcomes for rheumatoid arthritis patients using remote disease activity monitoring via smartphone app. *Rheumatology*, 59(Supplement_2):keaa111–209.

[113] Mahmoud, N. and Elbeh, H. (2016). Irs-t2d: Individualize recommendation system for type2 diabetes medication based on ontology and swrl. In *Proceedings of the 10th International Conference on Informatics and Systems*, pages 203–209.

[114] Mäkinen, H., Kautiainen, H., Hannonen, P., and Sokka, T. (2005). Is das28 an appropriate tool to assess remission in rheumatoid arthritis? *Annals of the Rheumatic Diseases*, 64(10):1410–1413.

[115] Malm, K., Bergman, S., Andersson, M. L., Bremander, A., and Larsson, I. (2017). Quality of life in patients with established rheumatoid arthritis: A phenomenographic study. *SAGE open medicine*, 5:2050312117713647.

[116] Markusse, I. M., Dirven, L., Gerards, A. H., van Groenendael, J. H., Ronday, H. K., Kerstens, P. J., Lems, W. F., Huizinga, T. W., and Allaart, C. F. (2015). Disease flares in rheumatoid arthritis are associated with joint damage progression and disability: 10-year results from the best study. *Arthritis research & therapy*, 17(1):232.

[117] McLachlan, S., Dube, K., and Gallagher, T. (2016). Using the caremap with health incidents statistics for generating the realistic synthetic electronic healthcare record. In *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 439–448. IEEE.

[118] McLachlan, S., Kyrimi, E., Dube, K., Hitman, G., Simmonds, J., and Fenton, N. (2020). Towards standardisation of evidence-based clinical care process specifications. *Health Informatics Journal*, page 1460458220906069.

[119] McLachlan, S., Kyrimi, E., FENTON, N., Dube, K., et al. (2019). Clinical caremap development: How can caremaps standardise care when they are not standardised?

[120] Moreira, L. B. and Namen, A. A. (2018). A hybrid data mining model for diagnosis of patients with clinical suspicion of dementia. *Computer methods and programs in biomedicine*, 165:139–149.

[121] Moreira, M. W., Rodrigues, J. J., Oliveira, A. M., Ramos, R. F., and Saleem, K. (2016). A preeclampsia diagnosis approach using bayesian networks. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–5. IEEE.

[122] Morisky, D. E., Green, L. W., and Levine, D. M. (1986). Concurrent and predictive validity of a self-reported measure of medication adherence. *Medical care*, pages 67–74.

[123] Murphy, K. P. (2002). Dynamic bayesian networks: representation, inference and learning.

[124] Myasoedova, E., De Thurah, A., Erpelding, M.-L., Schneeberger, E. E., Maribo, T., Citera, G., Davis, J. M., Matteson, E. L., Crowson, C. S., Fautrel, B., et al. (2020). Definition and construct validation of clinically relevant cutoffs on the flare assessment in rheumatoid arthritis (flare-ra) questionnaire. In *Seminars in arthritis and rheumatism*, volume 50, pages 261–265. Elsevier.

[125] Neil, M., Fenton, N., and Nielson, L. (2000). Building large-scale bayesian networks. *The Knowledge Engineering Review*, 15(3):257–284.

[126] NHS (2015). Swollen joints, stiffness, pain? these are the signs of rheumatoid arthritis. https://www.nhs.uk/ra/Documents/Rheumatoid_arthritis_easy-print_leaflet.pdf (Accessed on 04/05/2020).

[127] NHS (2016). Arthritis. https://www.nhs.uk/Conditions/Arthritis/ (Accessed on 16/10/2020).

[128] NHS (2017). Ethnicity. https://data.developer.nhs.uk/specifications/NHS-CDA-eDischarge/Vocabulary/Ethnicity.html (Accessed 05/01/2021).

[129] NHS (2018). Alcohol units. https://www.nhs.uk/live-well/alcohol-support/calculating-alcohol-units/ (Accessed on 13/01/2020).

[130] NHS (2019a). Ethnic category code. https://www.datadictionary.nhs.uk/data_dictionary/attributes/e/end/ethnic_category_code_de.asp/ (Accessed on 04/09/2019).

[131] NHS (2019b). Living with. https://www.nhs.uk/conditions/rheumatoid-arthritis/living-with/ (Accessed on 03/04/2020).

[132] NHS (2019c). Rheumatoid arthritis. https://www.nhs.uk/conditions/rheumatoid-arthritis/ (Accessed on 04/05/2020).

[133] NHS (2020). What is personalised care? https://www.england.nhs.uk/personalisedcare/what-is-personalised-care/ (Accessed 02/04/2020).

[134] NHS Health and Social Care Information Centre (2019). Snomed ct a user guide for mental health. file:///C:/Users/Ali/Downloads/User_Guide_for_Mental_Health_V1.0.pdf (Accessed on 12/02/2021).

[135] NICE (2010). Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tnf inhibitor. https://www.nice.org.uk/guidance/ta195 (Accessed 01/01/2021).

[136] NICE (2012). Tocilizumab for the treatment of rheumatoid arthritis. https://www.nice.org.uk/guidance/ta247/resources/tocilizumab-for-the-treatment-of-rheumatoid-arthritis-pdf-82600436098501 (Accessed 02/04/2021).

[137] NICE (2016). Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with dmards or after conventional dmards only have failed. https://www.nice.org.uk/guidance/ta375 (Accessed 01/01/2021).

[138] NICE (2018a). Rheumatoid arthritis in adults: management. https://www.nice.org.uk/guidance/ng100 (Accessed 01/01/2021).

[139] NICE (2018b). Tocilizumab for the treatment of rheumatoid arthritis. https://www.nice.org.uk/guidance/ng100/evidence/evidence-review-g-analgesics-pdf-4903172324 (Accessed 01/01/2021).

[140] NICE (2020). Drug treatment for rheumatoid arthritis. https://pathways.nice.org.uk/pathways/rheumatoid-arthritis/drug-treatment-for-rheumatoid-arthritis.pdf (Accessed 01/01/2021).

[141] Nicholson, A., Boneh, T., Wilkin, T., Stacey, K., Sonenberg, L., and Steinle, V. (2013). A case study in knowledge discovery and elicitation in an intelligent tutoring application. *arXiv preprint arXiv:1301.2297*.

[142] Nicholson, A. E. and Brady, J. M. (1994). Dynamic belief networks for discrete monitoring. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(11):1593–1610.

[143] Nicholson, A. E. and Flores, M. J. (2011). Combining state and transition models with dynamic bayesian networks. *Ecological Modelling*, 222(3):555–566.

[144] Noy, N. F., McGuinness, D. L., et al. (2001). Ontology development 101: A guide to creating your first ontology.

[145] NRAS (2018). Living better with ra. https://nras.org.uk/product/living-better-with-ra/?attribute_pa_file-type=download (Accessed on 10/03/2021).

[146] NRAS (2020a). The das28 score. https://nras.org.uk/resource/the-das28-score/ (Accessed on 10/03/2021).

[147] NRAS (2020b). Help for you. https://www.nras.org.uk/help-for-you (Accessed on 04/05/2020).

[148] Orphanou, K., Keravnou, E., and Moutiris, J. (2012). Integration of temporal abstraction and dynamic bayesian networks in clinical systems. a preliminary approach. In *2012 Imperial College Computing Student Workshop*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[149] Orphanou, K., Stassopoulou, A., and Keravnou, E. (2014). Temporal abstraction and temporal bayesian networks in clinical domains: A survey. *Artificial intelligence in medicine*, 60(3):133–149.

[150] Orphanou, K., Stassopoulou, A., and Keravnou, E. (2015). Dbn-extended: a dynamic bayesian network model extended with temporal abstractions for coronary heart disease prognosis. *IEEE journal of biomedical and health informatics*, 20(3):944–952.

[151] Padyukov, L., Silva, C., Stolt, P., Alfredsson, L., and Klareskog, L. (2004). A gene–environment interaction between smoking and shared epitope genes in hla–dr provides a high risk of seropositive rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 50(10):3085–3092.

[152] PamBayesian (2018). Pambayesian: Patient managed decision-support using bayesian networks. https://pambayesian.org/overview/ (Accessed 15/01/2021).

[153] Pan, R., Ding, Z., Yu, Y., and Peng, Y. (2005). A bayesian network approach to ontology mapping. In *International Semantic Web Conference*, pages 563–577. Springer.

[154] Panella, M., Marchisio, S., and Di Stanislao, F. (2003). Reducing clinical variations with clinical pathways: do pathways work? *International Journal for Quality in Health Care*, 15(6):509–521.

[155] Parks, C. G., Meyer, A., Freeman, L. E. B., Hofmann, J. N., and Sandler, D. P. (2019). Farming tasks and the development of rheumatoid arthritis in the agricultural health study. *Occupational and environmental medicine*, 76(4):243–249.

[156] Paydar, K., Kalhori, S. R. N., Akbarian, M., and Sheikhtaheri, A. (2017). A clinical decision support system for prediction of pregnancy outcome in pregnant women with systemic lupus erythematosus. *International journal of medical informatics*, 97:239–246.

[157] PEAC (2018). PEAC. http://gtr.rcuk.ac.uk/projects?ref=G0800648 (Accessed on 22/02/2018).

[158] Pearl, J. (2014). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier.

[159] Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3):54–60.

[160] Pedersen, M., Jacobsen, S., Klarlund, M., Pedersen, B. V., Wiik, A., Wohlfahrt, J., and Frisch, M. (2006). Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis research & therapy*, 8(4):R133.

[161] Peelen, L., De Keizer, N. F., De Jonge, E., Bosman, R.-J., Abu-Hanna, A., and Peek, N. (2010). Using hierarchical dynamic bayesian networks to investigate dynamics of organ failure in patients in the intensive care unit. *Journal of biomedical informatics*, 43(2):273–286.

[162] Peschken, C. A. and Esdaile, J. M. (1999). Rheumatic diseases in north america's indigenous peoples. In *Seminars in arthritis and rheumatism*, volume 28, pages 368–391. Elsevier.

[163] Pikwer, M., Bergström, U., Nilsson, J.-Å., Jacobsson, L., and Turesson, C. (2012). Early menopause is an independent predictor of rheumatoid arthritis. *Annals of the rheumatic diseases*, 71(3):378–381.

[164] Pollino, C. A., Woodberry, O., Nicholson, A., Korb, K., and Hart, B. T. (2007). Parameterisation and evaluation of a bayesian network for use in an ecological risk assessment. *Environmental Modelling & Software*, 22(8):1140–1152.

[165] Pradhan, M., Provan, G., Middleton, B., and Henrion, M. (1994). Knowledge engineering for large belief networks. In *Uncertainty Proceedings 1994*, pages 484–490. Elsevier.

[166] Prevoo, M., Van'T Hof, M., Kuper, H., Van Leeuwen, M., Van De Putte, L., and Van Riel, P. (1995). Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 38(1):44–48.

[167] Quartuccio, L., Zabotti, A., Del Zotto, S., Zanier, L., De Vita, S., and Valent, F. (2019). Risk of serious infection among patients receiving biologics for chronic inflammatory diseases: usefulness of administrative data. *Journal of Advanced Research*, 15:87–93.

[168] Queen Mary University London (2020). Biot das app 2. https://apps.apple.com/lk/app/biot-das-app-2/id1432609979 (Accessed 15/01/2021).

[169] Ramati, M. (2010). *Irregular-time Markov Models*. Citeseer.

[170] Ramati, M. and Shahar, Y. (2012). Irregular-time bayesian networks. *arXiv preprint arXiv:1203.3510*.

[171] Rebhan, M. (2017). Towards a systems approach for chronic diseases, based on health state modeling. *F1000Research*, 6.

[172] Research, U. and Innovation (2017). Pambayesian: Patient managed decision-support using bayesian networks. https://gtr.ukri.org/projects?ref=EP%2FP009964%2F1#/tabOverview (Accessed 15/01/2021).

[173] Roos, J., Gavin, G., and Bonnevay, S. (2017). A dynamic bayesian network approach to forecast short-term urban rail passenger flows with incomplete data. *Transportation research procedia*, 26:53–61.

[174] Ropero, R. F., Nicholson, A. E., Aguilera, P. A., and Rumí, R. (2018). Learning and inference methodologies for hybrid dynamic bayesian networks: a case study for a water reservoir system in andalusia, spain. *Stochastic Environmental Research and Risk Assessment*, 32(11):3117–3135.

[175] Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

[176] Sa-Ngamuang, C., Haddawy, P., Luvira, V., Piyaphanee, W., Iamsirithaworn, S., and Lawpoolsri, S. (2018). Accuracy of dengue clinical diagnosis with and without ns1 antigen rapid test: Comparison between human and bayesian network model decision. *PLoS neglected tropical diseases*, 12(6):e0006573.

[177] Sackman, J. E. and Citrin, L. (2014). Cracking down on cost outliers. *Healthcare financial management*, 68(3):58–63.

[178] Saint-Jacques, H., Burroughs, V. J., Watkowska, J., Valcarcel, M., Moreno, P., and Maw, M. (2005). Acute coronary syndrome critical pathway: Chest pain caremap: A qualitative research study—provider-level intervention. *Critical pathways in cardiology*, 4(3):145–160.

[179] Salmeron, J. L., Rahimi, S. A., Navali, A. M., and Sadeghpour, A. (2017). Medical diagnosis of rheumatoid arthritis using data driven pso–fcm with scarce datasets. *Neurocomputing*, 232:104–112.

[180] Salunke, A. B. and Kasar, S. L. (2015). Personalized recommendation system for medical assistance using hybrid filtering. *International Journal of Computer Applications*, 975:8887.

[181] Sandilya, S. and Rao, R. B. (2004). Continuous-time bayesian modeling of clinical data. In *Proceedings of the 2004 SIAM International Conference on Data Mining*, pages 512–516. SIAM.

[182] Scanagatta, M., Corani, G., De Campos, C. P., and Zaffalon, M. (2018). Approximate structure learning for large bayesian networks. *Machine Learning*, 107(8):1209–1227.

[183] Scanagatta, M., Salmerón, A., and Stella, F. (2019). A survey on bayesian network structure learning from data. *Progress in Artificial Intelligence*, 8(4):425–439.

[184] Schneider, M. and Krüger, K. (2013). Rheumatoid arthritis—early diagnosis and disease management. *Deutsches Ärzteblatt International*, 110(27-28):477.

[185] Scutari, M. (2014). Bayesian network constraint-based structure learning algorithms: Parallel and optimised implementations in the bnlearn r package. *arXiv preprint arXiv:1406.7648*.

[186] Scutari, M. (2020). Structure learning from missing data. https://www.bnlearn.com/documentation/man/structural.em.html (Accessed on 16/10/2020).

[187] Scutari, M. (2021). Structure learning algorithms. https://www.bnlearn.com/documentation/man/structure.learning.html (Accessed on 23/02/2021).

[188] Scutari, M., Vitolo, C., and Tucker, A. (2019). Learning bayesian networks from big data with greedy search: computational complexity and efficient implementation. *Statistics and Computing*, 29(5):1095–1108.

[189] Seixas, A. A., Henclewood, D. A., Williams, S. K., Jagannathan, R., Ramos, A., Zizi, F., and Jean-Louis, G. (2018). Sleep duration and physical activity profiles associated with self-reported stroke in the united states: application of bayesian belief network modeling techniques. *Frontiers in neurology*, 9:534.

[190] Seixas, F. L., Zadrozny, B., Laks, J., Conci, A., and Saade, D. C. M. (2014). A bayesian network decision model for supporting the diagnosis of dementia, alzheimer's disease and mild cognitive impairment. *Computers in biology and medicine*, 51:140–158.

[191] Serwylo, P. E. (2010). *Eliciting Bayesian networks via online surveys: a new approach to knowledge elicitation*. PhD thesis, Monash University.

[192] Sesen, M. B., Nicholson, A. E., Banares-Alcantara, R., Kadir, T., and Brady, M. (2013). Bayesian networks for clinical decision support in lung cancer care. *PloS one*, 8(12):e82349.

[193] Sesen, M. B., Peake, M. D., Banares-Alcantara, R., Tse, D., Kadir, T., Stanley, R., Gleeson, F., and Brady, M. (2014). Lung cancer assistant: a hybrid clinical decision support application for lung cancer care. *Journal of The Royal Society Interface*, 11(98):20140534.

[194] Shachter, R. D. (1986). Evaluating influence diagrams. *Operations research*, 34(6):871–882.

[195] Shanmugam, S. and Preethi, J. (2019). Improved feature selection and classification for rheumatoid arthritis disease using weighted decision tree approach (react). *The Journal of Supercomputing*, 75(8):5507–5519.

[196] Shao, Z., Zhu, T., Zhang, P., Wen, Q., Li, D., and Wang, S. (2017). Association of financial status and the quality of life in chinese women with recurrent ovarian cancer. *Health and quality of life outcomes*, 15(1):1–8.

[197] Shi, B. and Weninger, T. (2018). Open-world knowledge graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

[198] Shiezadeh, Z., Sajedi, H., and Aflakie, E. (2015). Diagnosis of rheumatoid arthritis using an ensemble learning approach. *Comput. Sci. Inf. Technol.(CS & IT)*, 5(15):139–148.

[199] Sierakowski, S. and Cutolo, M. (2011). Morning symptoms in rheumatoid arthritis: a defining characteristic and marker of active disease. *Scandinavian Journal of Rheumatology*, 40(sup125):1–5.

[200] Silman, A., Kay, A., and Brennan, P. (1992). Timing of pregnancy in relation to the onset of rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 35(2):152–155.

[201] Smedslund, G., Mowinckel, P., Heiberg, T., Kvien, T. K., and Hagen, K. B. (2009). Does the weather really matter? a cohort study of influences of weather and solar conditions on daily variations of joint pain in patients with rheumatoid arthritis. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 61(9):1243–1247.

[202] SNOMED International (2019). Domain specific modeling. https://confluence.ihtsdotools.org/display/DOCEG/Domain+Specific+Modeling (Accessed on 01/09/2021).

[203] Somnay, Y. R., Craven, M., McCoy, K. L., Carty, S. E., Wang, T. S., Greenberg, C. C., and Schneider, D. F. (2017). Improving diagnostic recognition of primary hyperparathyroidism with machine learning. *Surgery*, 161(4):1113–1121.

[204] Spirtes, P., Glymour, C. N., Scheines, R., and Heckerman, D. (2000). *Causation, prediction, and search*. MIT press.

[205] Stastny, P. (1978). Association of the b-cell alloantigen drw4 with rheumatoid arthritis. *New England journal of medicine*, 298(16):869–871.

[206] Subramaniyaswamy, V., Manogaran, G., Logesh, R., Vijayakumar, V., Chilamkurti, N., Malathi, D., and Senthilselvan, N. (2019). An ontology-driven personalized food recommendation in iot-based healthcare system. *The Journal of Supercomputing*, 75(6):3184–3216.

[207] Šváb, O. and Svátek, V. (2006). Combining ontology mapping methods using bayesian networks. *Ontology Matching*, page 206.

[208] Timmermans, E. J., Van Der Pas, S., Schaap, L. A., Sánchez-Martínez, M., Zambon, S., Peter, R., Pedersen, N. L., Dennison, E. M., Denkinger, M., Castell, M. V., et al. (2014). Self-perceived weather sensitivity and joint pain in older people with osteoarthritis in six european countries: results from the european project on osteoarthritis (eposa). *BMC Musculoskeletal Disorders*, 15(1):66.

[209] Tizaoui, K., Kim, S. H., Jeong, G. H., Kronbichler, A., Lee, K. S., Lee, K. H., and Shin, J. I. (2019). Association of ptpn22 1858c/t polymorphism with autoimmune diseases: a systematic review and bayesian approach. *Journal of clinical medicine*, 8(3):347.

[210] UMLS (2021). Unified medical language system (umls). https://www.nlm.nih.gov/research/umls/index.html (Accessed on 09/05/2021).

[211] Van der Heijde, D., van't Hof, M. A., Van Riel, P., Theunisse, L., Lubberts, E. W., van Leeuwen, M. A., van Rijswijk, M. H., and Van de Putte, L. (1990). Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Annals of the rheumatic diseases*, 49(11):916–920.

[212] Van der Heijden, M., Velikova, M., and Lucas, P. J. (2014). Learning bayesian networks for clinical time series analysis. *Journal of biomedical informatics*, 48:94–105.

[213] van der Maas, A., Lie, E., Christensen, R., Choy, E., de Man, Y. A., van Riel, P., Woodworth, T., and den Broeder, A. A. (2013). Construct and criterion validity of several proposed das28-based rheumatoid arthritis flare criteria: an omeract cohort validation study. *Annals of the rheumatic diseases*, 72(11):1800–1805.

[214] Van Gerven, M. A., Taal, B. G., and Lucas, P. J. (2008). Dynamic bayesian networks as prognostic models for clinical patient management. *Journal of biomedical informatics*, 41(4):515–529.

[215] van Gestel, A. M., Haagsma, C. J., and van Riel, P. L. (1998). Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 41(10):1845–1850.

[216] Velikova, M., van Scheltinga, J. T., Lucas, P. J., and Spaanderman, M. (2014). Exploiting causal functional relationships in bayesian network modelling for personalised healthcare. *International Journal of Approximate Reasoning*, 55(1):59–73.

[217] Verma, T. and Pearl, J. (1991). *Equivalence and synthesis of causal models*. UCLA, Computer Science Department.

[218] W3 (2015). Semantic web. https://www.w3.org/standards/semanticweb/ (Accessed on 22/08/2019).

[219] Wang, X.-H., Zheng, B., Good, W. F., King, J. L., and Chang, Y.-H. (1999). Computer-assisted diagnosis of breast cancer using a data-driven bayesian belief network. *International journal of medical informatics*, 54(2):115–126.

[220] Weiss, J. C., Natarajan, S., Peissig, P. L., McCarty, C. A., and Page, D. (2012). Machine learning for personalized medicine: predicting primary myocardial infarction from electronic health records. *Ai Magazine*, 33(4):33–33.

[221] WHO (2003). Adherence to long-term therapies: Evidence for action. https://www.who.int/chp/knowledge/publications/adherence_report/en/ (Accessed on 22/10/2020).

[222] WHOQOL-BREF (1996). Introduction, administration, scoring and generic version of the assessment. https://www.who.int/mental_health/media/en/76.pdf (Accessed 22/10/2020).

[223] Wolfe, F. (2000). The effect of smoking on clinical, laboratory, and radiographic status in rheumatoid arthritis. *The Journal of rheumatology*, 27(3):630–637.

[224] Wong, L. E., Huang, W.-T., Pope, J. E., Haraoui, B., Boire, G., Thorne, J. C., Hitchon, C. A., Tin, D., Keystone, E. C., Bykerk, V. P., et al. (2015). Effect of age at menopause on disease presentation in early rheumatoid arthritis: results from the canadian early arthritis cohort. *Arthritis care & research*, 67(5):616–623.

[225] Yang, Y. and Calmet, J. (2005). Ontobayes: An ontology-driven uncertainty model. In *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, volume 1, pages 457–463. IEEE.

[226] Yet, B., Perkins, Z., Fenton, N., Tai, N., and Marsh, W. (2014a). Not just data: A method for improving prediction with knowledge. *Journal of biomedical informatics*, 48:28–37.

[227] Yet, B., Perkins, Z. B., Rasmussen, T. E., Tai, N. R., and Marsh, D. W. R. (2014b). Combining data and meta-analysis to build bayesian networks for clinical decision support. *Journal of biomedical informatics*, 52:373–385.

[228] Yet, B., Perkins, Z. B., Tai, N. R., and Marsh, D. W. R. (2017). Clinical evidence framework for bayesian networks. *Knowledge and Information Systems*, 50(1):117–143.

[229] Yin, Z., Zhao, Y., Lu, X., and Duan, H. (2015). A hybrid intelligent diagnosis approach for quick screening of alzheimer's disease based on multiple neuropsychological rating scales. *Computational and mathematical methods in medicine*, 2015.

[230] Zhang, X., Hu, B., Ma, X., Moore, P., and Chen, J. (2014). Ontology driven decision support for the diagnosis of mild cognitive impairment. *Computer methods and programs in biomedicine*, 113(3):781–791.

Appendix A

Analysis, Interpolation, and Expansion of Data of Pathobiology of Early Arthritis Cohort

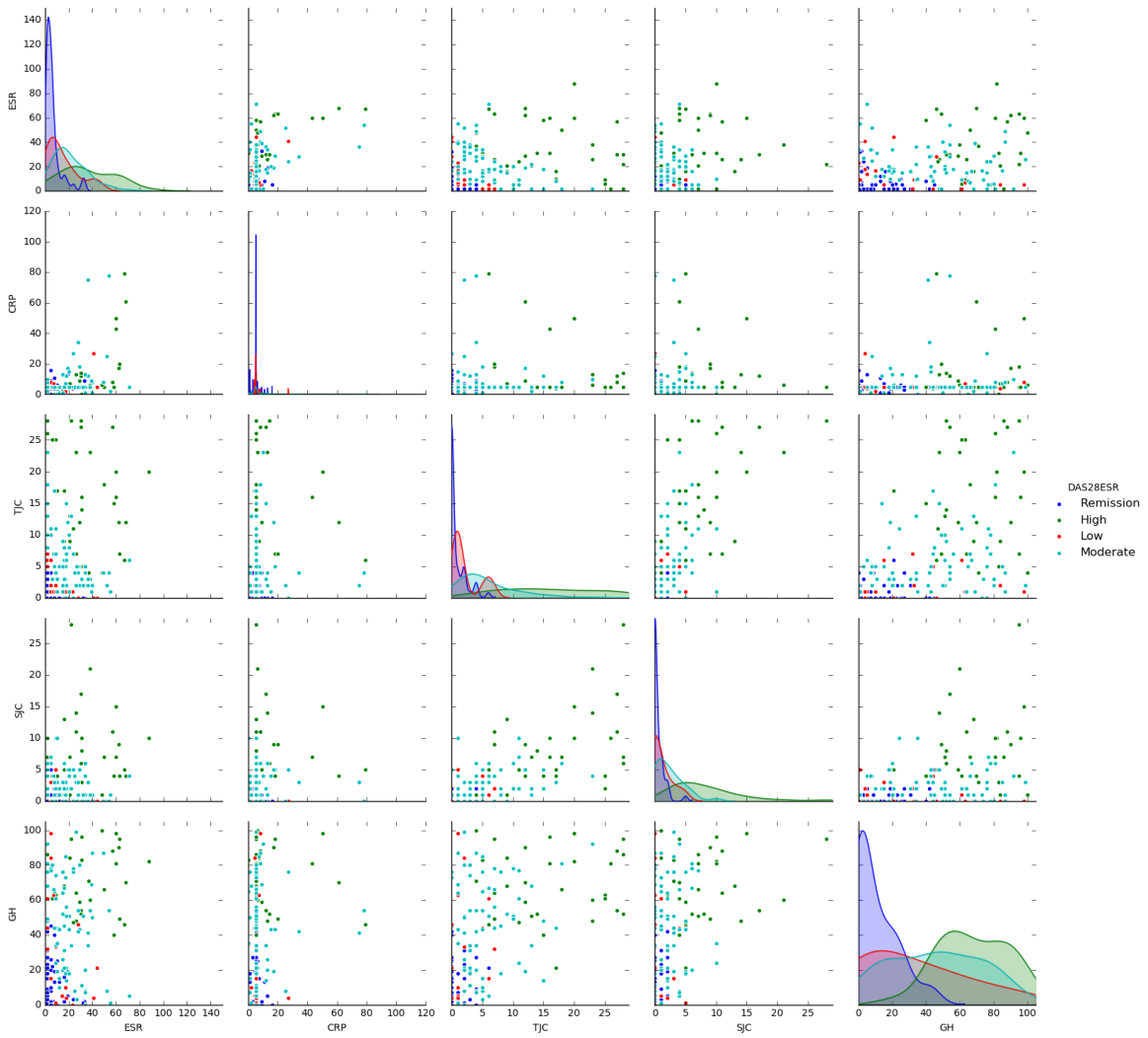


Fig. A.1 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

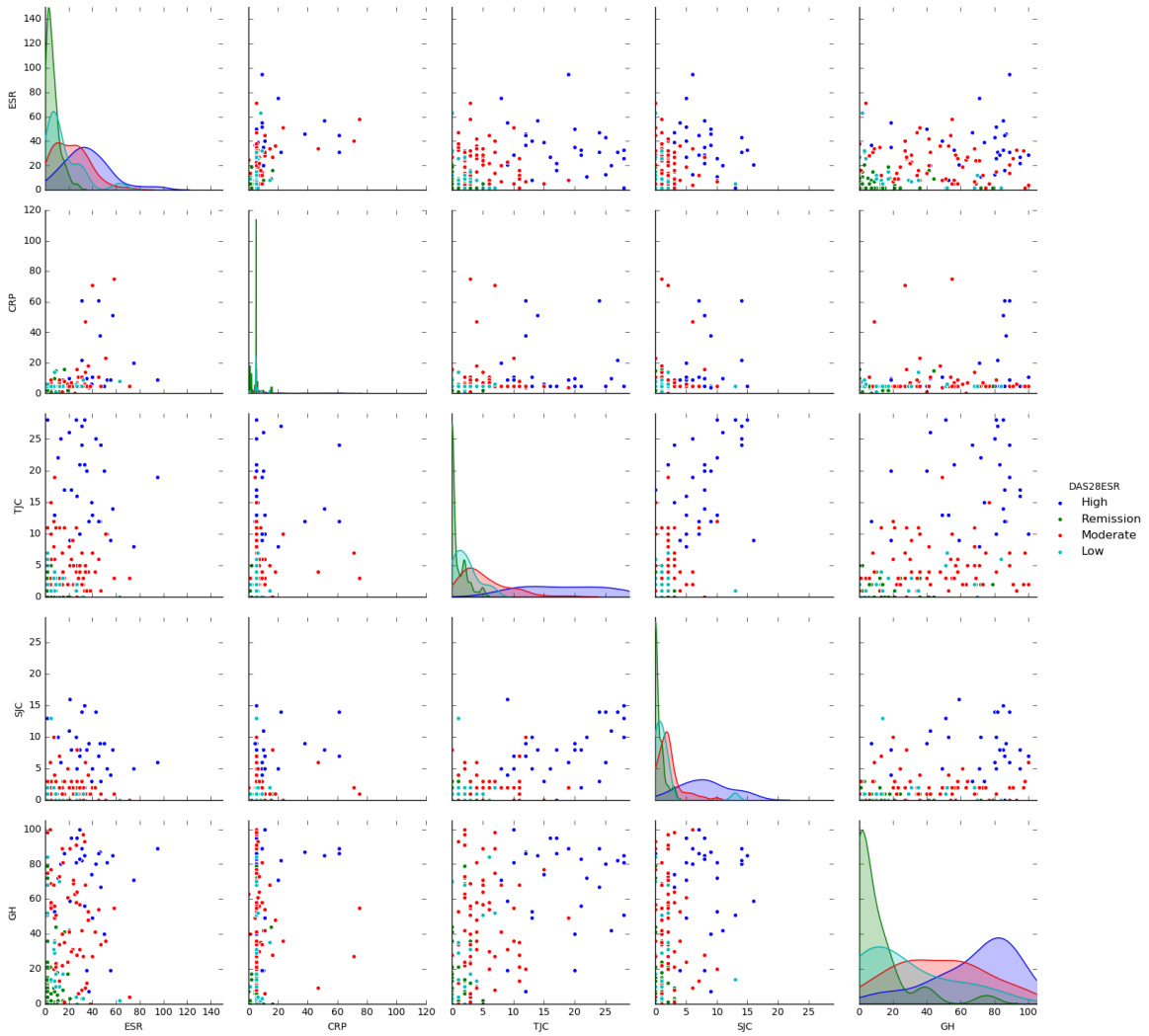


Fig. A.2 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

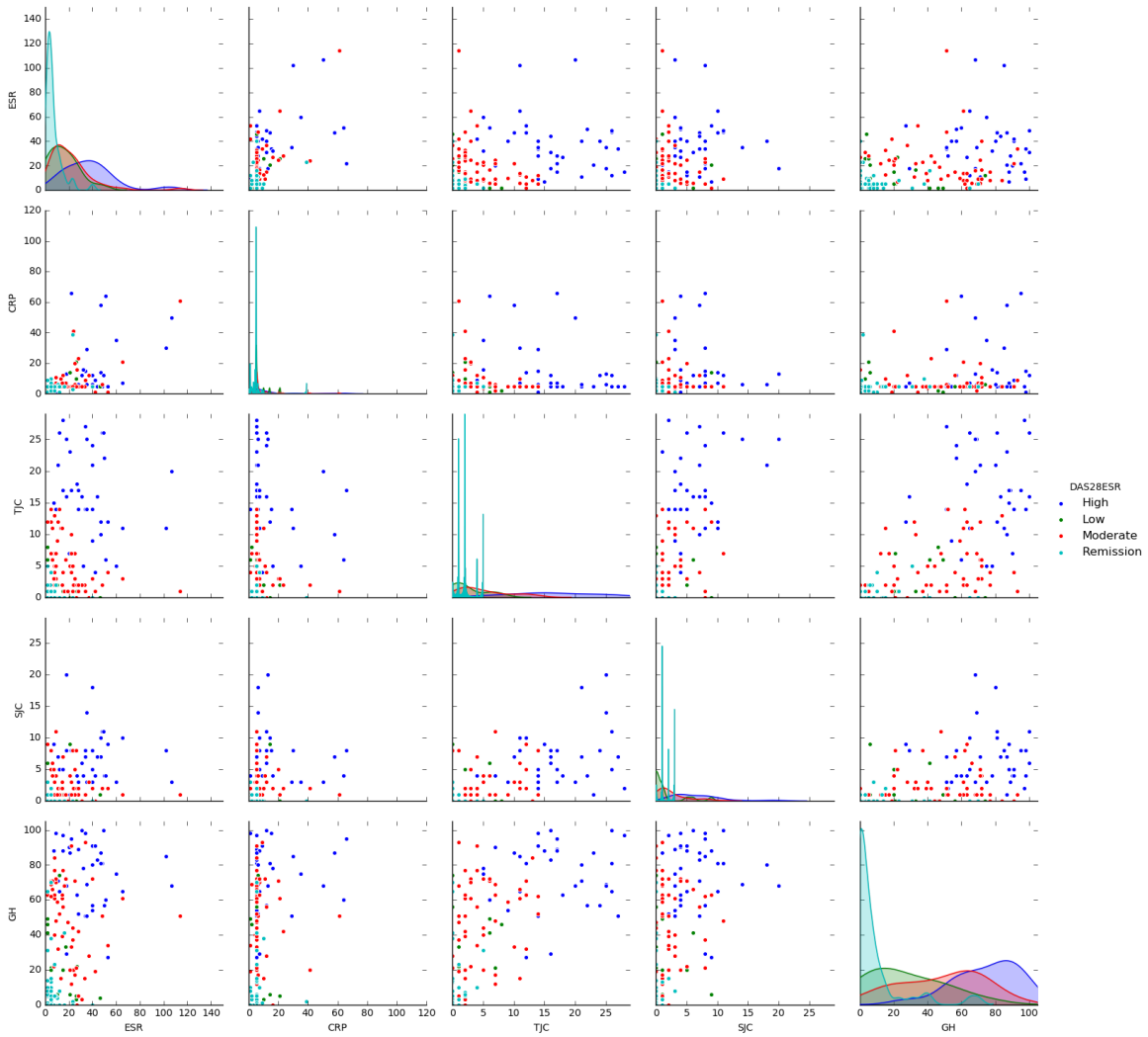


Fig. A.3 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.

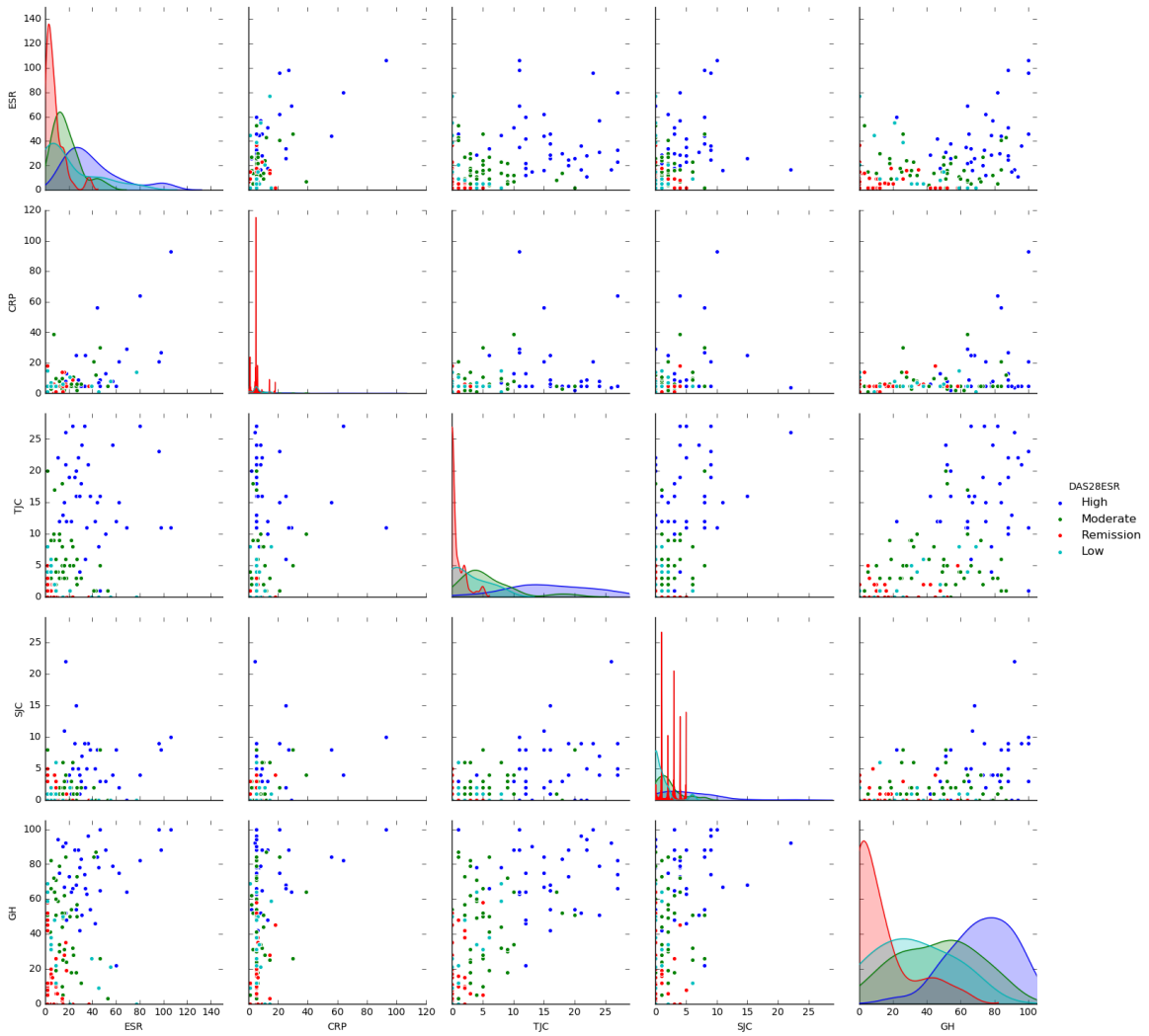
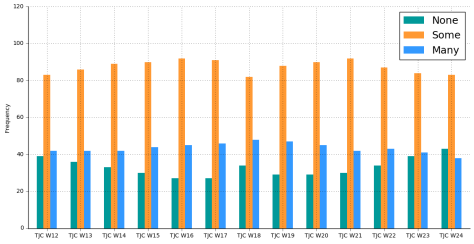
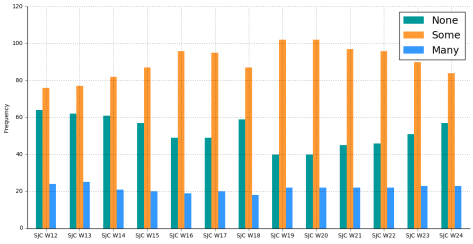


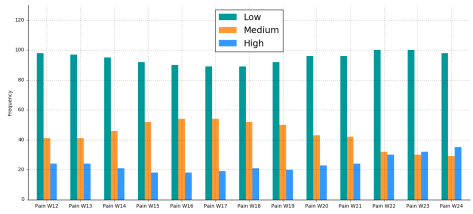
Fig. A.4 Pair plot of scatter plots of TJC, SJC, GH, ESR, and CRP (DAS28 components) in the first follow-up visit, joint with kernel density estimation of histogram of each variable, and colour encoding of diagnosis records of RA, UA, MonoA, PsA, and Others.



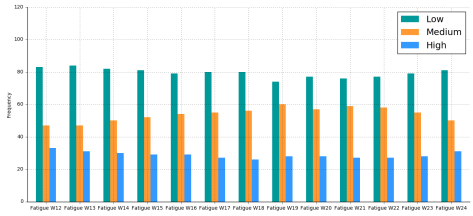
(a) TJC.



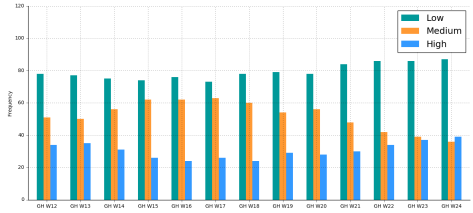
(b) SJC.



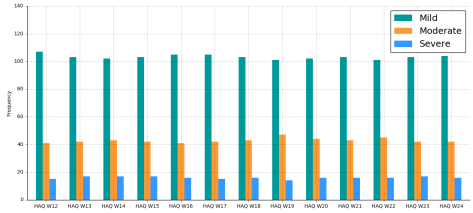
(c) JP.



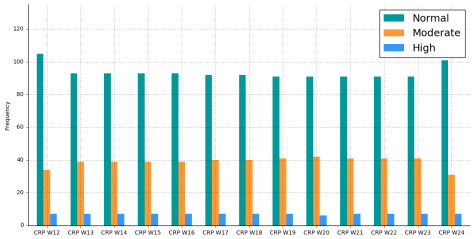
(d) Fatigue.



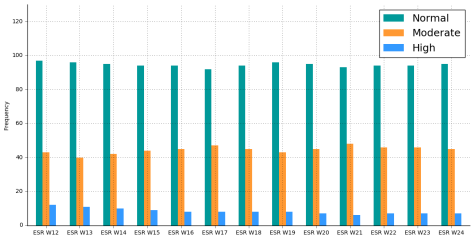
(e) GH.



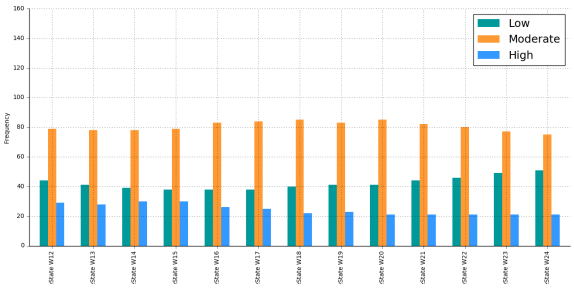
(f) HAQ.



(g) CRP.

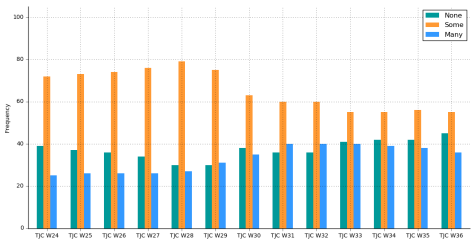


(h) ESR.

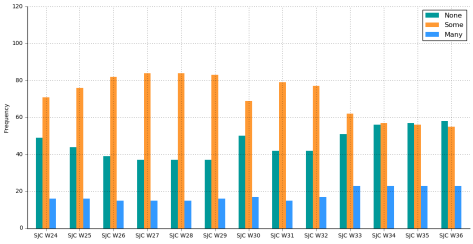


(i) Disease State.

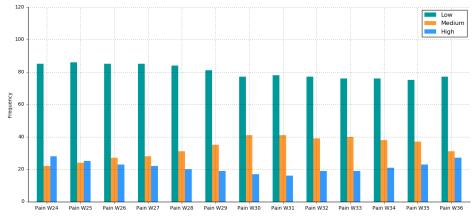
Fig. A.5 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the first follow-up visit (3 months) and second follow-up visit (6 months) and interpolation of 11 weeks between the first and second follow-ups.



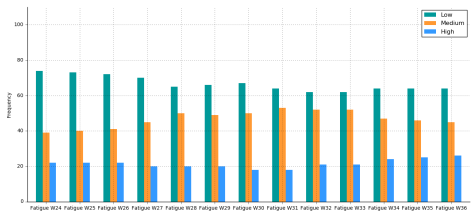
(a) TJC.



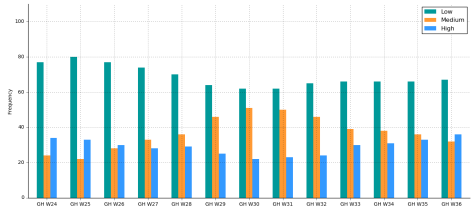
(b) SJC.



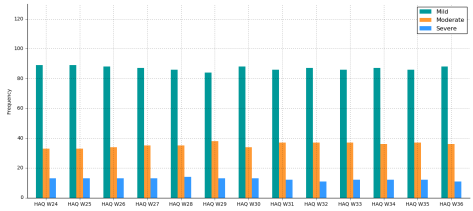
(c) JP.



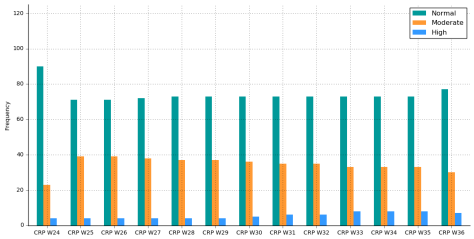
(d) Fatigue.



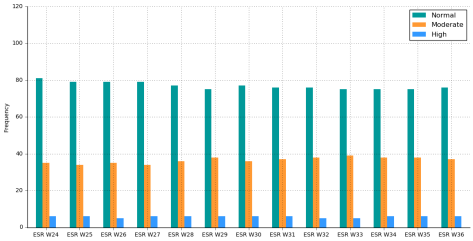
(e) GH.



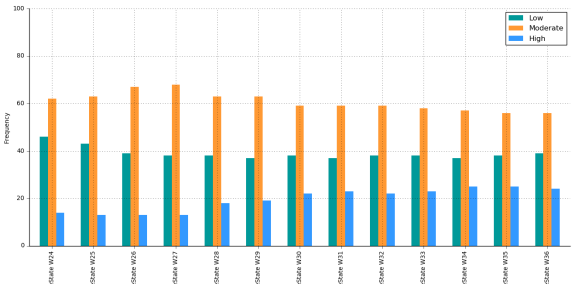
(f) HAQ.



(g) CRP.

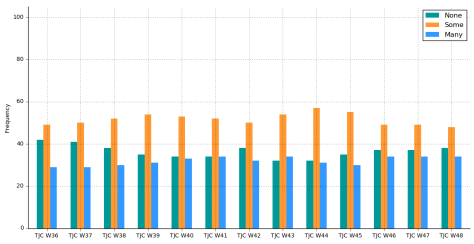


(h) ESR.

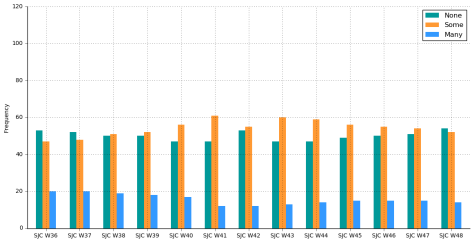


(i) Disease State.

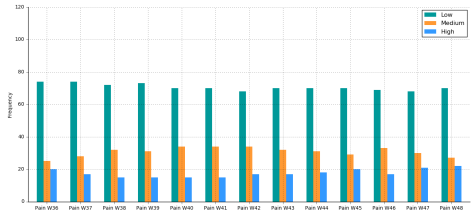
Fig. A.6 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the second follow-up visit (6 months) and third follow-up visit (9 months) and interpolation of 11 weeks between the second and third follow-ups.



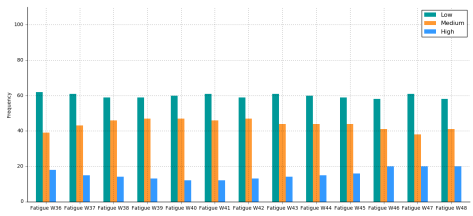
(a) TJC.



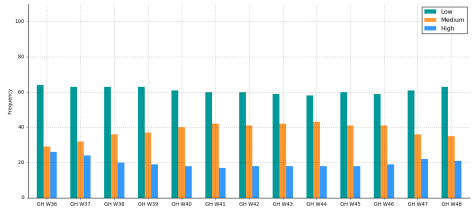
(b) SJC.



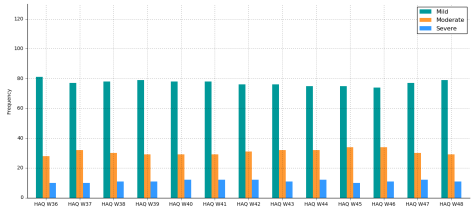
(c) JP.



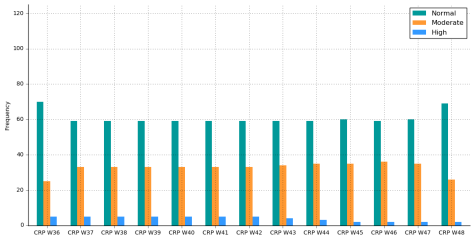
(d) Fatigue.



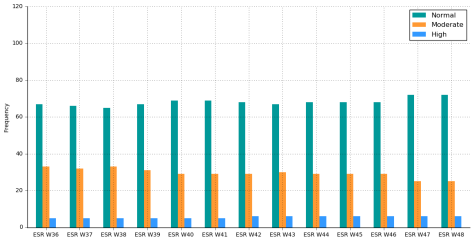
(e) GH.



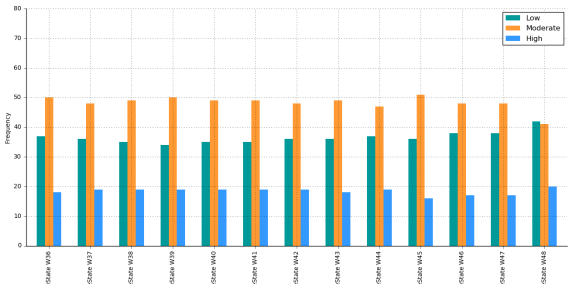
(f) HAQ.



(g) CRP.

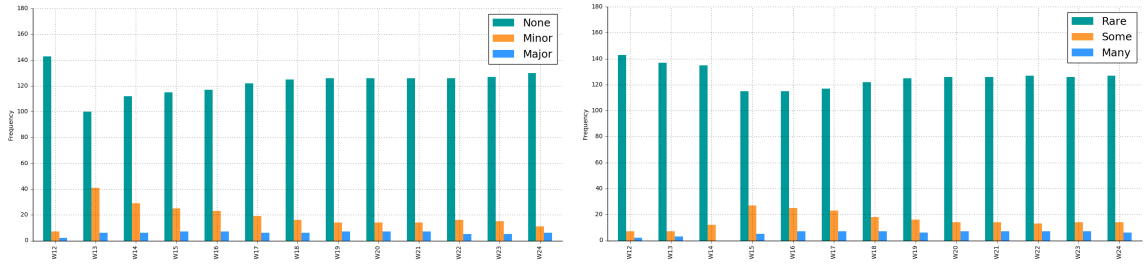


(h) ESR.



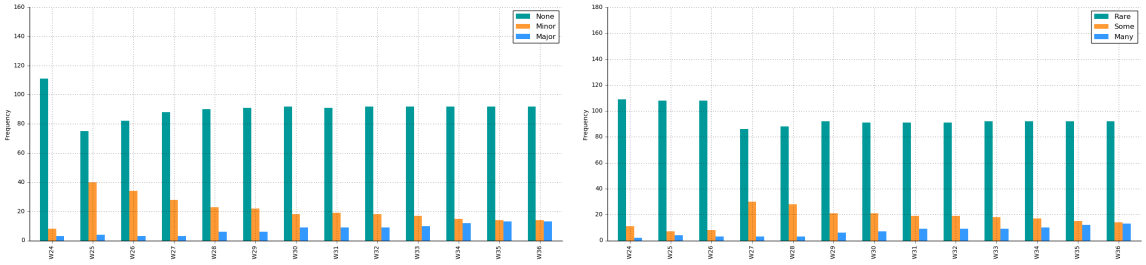
(i) Disease State.

Fig. A.7 Bar plot of discretised signs, symptoms, serology results, and disease state collected in the third follow-up visit (9 months) and fourth follow-up visit (12 months) and interpolation of 11 weeks between the third and fourth follow-ups.



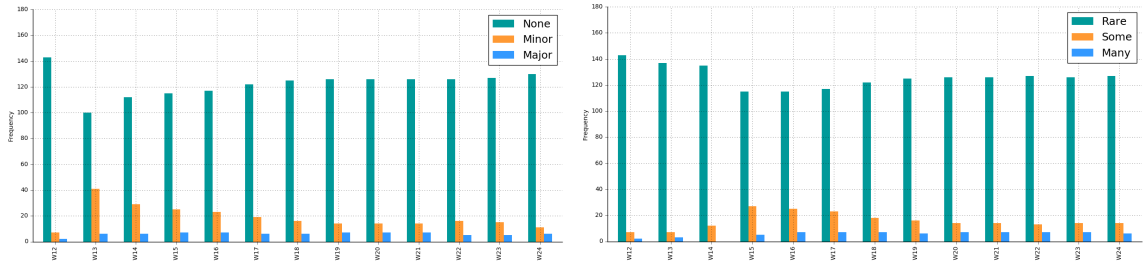
(a) Flare. (b) Flare Frequency.

Fig. A.8 Flare variables computed using disease state values of the first follow-up visit (3 months), the second follow-up visit (6 months), and interpolation of 11 weeks between the first and second follow-ups.



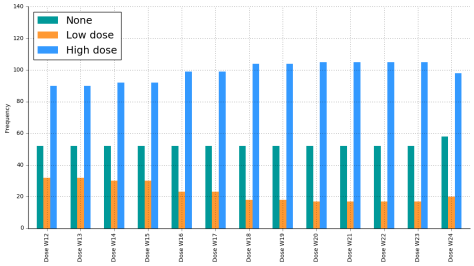
(a) Flare. (b) Flare Frequency.

Fig. A.9 Flare variables computed using disease state values of the second follow-up visit (6 months), the third follow-up visit (9 months), and interpolation of 11 weeks between the second and third follow-ups.

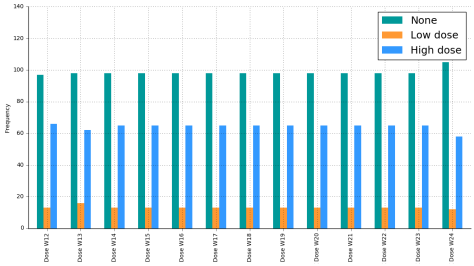


(a) Flare. (b) Flare Frequency.

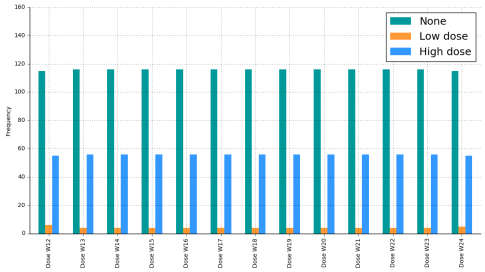
Fig. A.10 Flare variables computed using disease state values of the third follow-up visit (9 months), the fourth follow-up visit (12 months), and interpolation of 11 weeks between the third and fourth follow-ups.



(a) MTX.

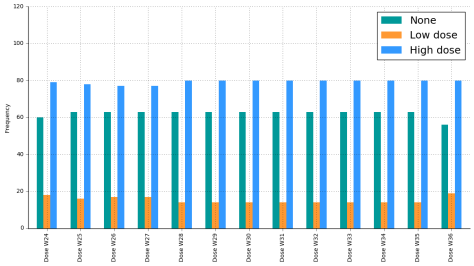


(b) SSZ.

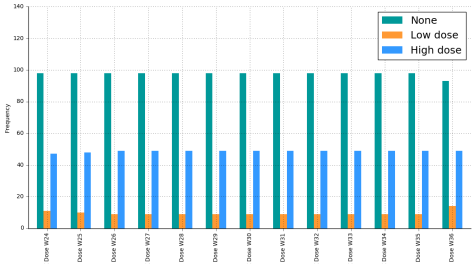


(c) HCQ.

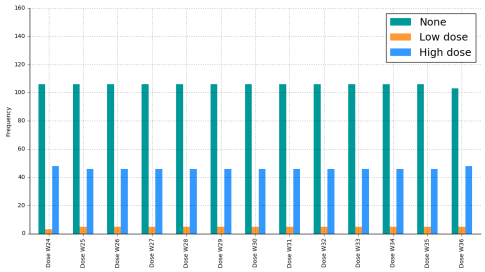
Fig. A.11 Discretised expanded csDMARD values between the first and second follow-up visits with 11 weeks in between.



(a) MTX.

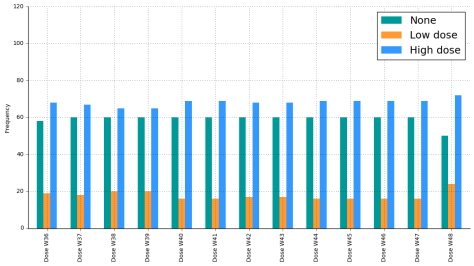


(b) SSZ.

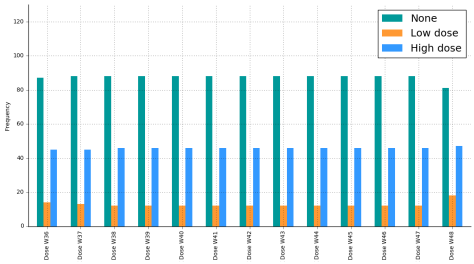


(c) HCQ.

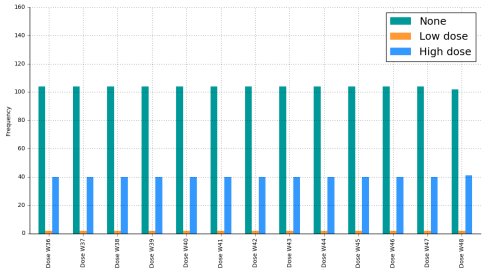
Fig. A.12 Discretised expanded csDMARD values between the second and third follow-up visits with 11 weeks in between.



(a) MTX.



(b) SSZ.



(c) HCQ.

Fig. A.13 Discretised expanded csDMARD values between the third and fourth follow-up visits with 11 weeks in between.

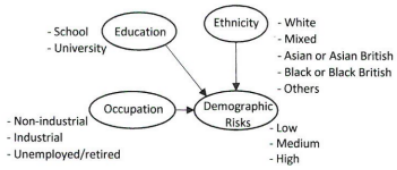
Appendix B

Bayesian Networks for Diagnosis of Rheumatoid Arthritis

B.1 Elicitation of Parameters from Experts

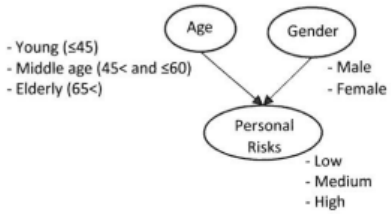
We elicit experts' knowledge to parameterise of six synthetic variables of the Bayesian network model for diagnosis of Rheumatoid Arthritis (RA), namely 'Demographic Risks', 'Personal Risks', 'Medical Background Risks', 'Lifestyle Risks', 'FSH Effect on RA', and Serostatus.

B.1.1 Initial Rules



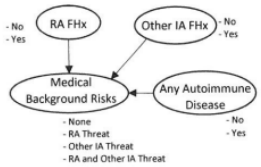
Rule	Accepted	Rejected	Modification
If *Education* is *University*, *Occupation* is *Non-industrial*, and *Ethnicity is Black or Black British*, then *Demographic Risks* is *Low*	✓		
If *Occupation is Non-industrial* and *Ethnicity* is *White*, then *Demographic Risks* is *Medium*	✓		Medium/low.
If *Education* is *School*, *Occupation* is *Industrial*, *Ethnicity* is *White*, then *Demographic Risks* is *High*	✓		

Fig. B.1 Initial rules to label demographic factors and parameterise 'Demographic Risks' variable provided by Dr. Amy MacBrayne.



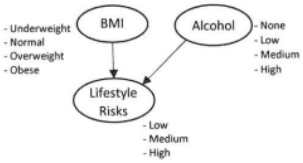
Rule	Accepted	Rejected	Modification
If *Age* is *Young* and *Gender* is *Male*, then *Personal Risks* is *Low*	✓		
If *Age* is *Middle Age* and *Gender* is *Female*, then *Personal Risks* is *Medium*		✓	High.
If *Age* is *Elderly* and *Gender* is *Male*, then *Personal Risks* is *Medium*	✓		
If *Age* is *Elderly* and *Gender* is *Female*, then *Personal Risks* is *High*		✓	? High / Medium.

Fig. B.2 Initial rules to label personal factors and parameterise 'Personal Risks' variable provided by Dr. Amy MacBrayne.



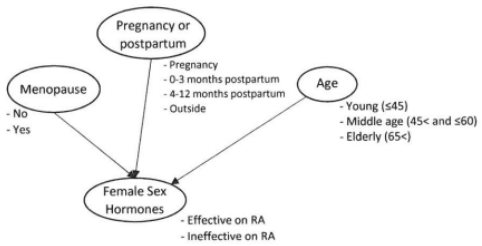
Rule	Accepted	Rejected	Modification
If *RA FHx* is *No*, *Other IA FHx* is *No*, and *Any Autoimmune Disease* is *No*, then *Medical Background Risks* is *None*	✓		-Maybe not 'none' but minimal.
If *RA FHx* is *Yes*, and *Other IA FHx* is *No*, then *Medical Background Risks* is *RA Threat*	✓		
If *Other IA FHx* is *Yes*, *RA FHx* is *No*, and *Other IA FHx* is *No*, then *Medical Background Risks* is *Other IA Threat*	✓		
If *RA FHx* is *Yes*, *Other IA FHx* is *Yes*, and *Any Autoimmune Disease* is *Yes*, then *Medical Background Risks* is *RA and Other IA Threat*	✓		
If *RA FHx* is *No*, *Other IA FHx* is *Yes*, and *Any Autoimmune Disease* is *Yes*, then *Medical Background Risks* is *RA and Other IA Threat*	✓		

Fig. B.3 Initial rules to label medical background factors and parameterise 'Medical Background Risks' variable provided by Dr. Amy MacBrayne.



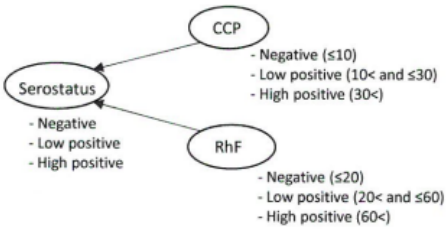
Rule	Accepted	Rejected	Modification
If *BMI* is *Normal* or *Underweight*, and *Alcohol* is *Medium*, then *Lifestyle Risks* is *Low*	✓		
If *BMI* is *Normal*, and *Alcohol* is *None* or *Low*, then *Lifestyle Risks* is *Medium*		✓	still low.
If *BMI* is *Overweight*, and *Alcohol* is *None* or *Low* or *High*, then *Lifestyle Risks* is *Medium*	✓		
If *BMI* is *Obese*, and *Alcohol* is *None* or *High*, then *Lifestyle Risks* is *High*	✓		

Fig. B.4 Initial rules to label lifestyle factors and parameterise 'Lifestyle Risks' variable provided by Dr. Amy MacBrayne.



Rule	Accepted	Rejected	Modification
If *Age* is *Young*, and *Menopause* is *No*, or *Pregnancy or Postpartum* is *Outside*, then *Female Sex Hormones* is *Ineffective on RA*	✓		
If *Age* is *Young*, and *Menopause* is *Yes*, then *Female Sex Hormones* is *Effective on RA*	✓		
If *Age* is *Young*, and *Pregnancy or Postpartum* is *0-3 months postpartum or 4-12 months postpartum*, then *Female Sex Hormones* is *Effective on RA*	✓		
If *Age* is *Middle Age*, and *Pregnancy or Postpartum* is *Outside*, then *Female Sex Hormones* is *Ineffective on RA*			Unable to answer as don't know if menopausal.
If *Age* is *Middle Age*, and *Pregnancy or Postpartum* is *0-3 months postpartum*, then *Female Sex Hormones* is *Effective on RA*	✓		
If *Age* is *Elderly*, then *Female Sex Hormones* is *Ineffective on RA*	✓		

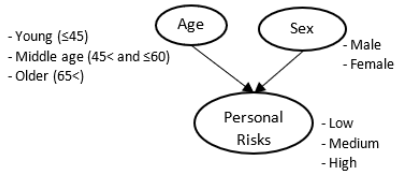
Fig. B.5 Initial rules to label hormonal pathogenesis factors and parameterise 'FSH Effect on RA' variable provided by Dr. Amy MacBrayne.



Rule	Accepted	Rejected	Modification
If *CCP* is *Negative* and *RhF* is *Negative*, then *Serostatus* is *Negative*	✓		
If *CCP* is *Low positive* and/or *RhF* is *Low positive*, then *Serostatus* is *Low positive*	✓		
If *CCP* is *High positive* and/or *RhF* is *High positive*, then *Serostatus* is *High positive*	✓		

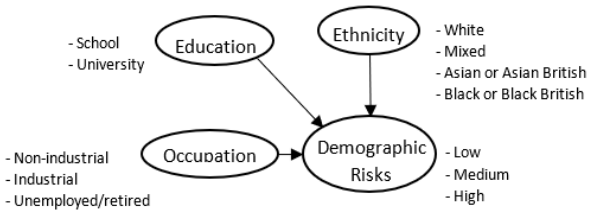
Fig. B.6 Initial rules to label two antibodies and parameterise Serostatus variable provided by Dr. Amy MacBrayne.

B.1.2 Secondary Rules



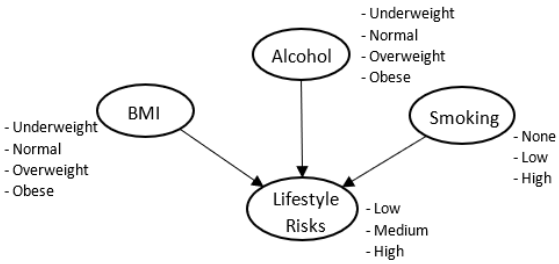
Status	Rule	Accepted	Rejected	Modification
Old rules defined by Fran, Amy, and Gloria	If *Age* is *Young* and *Gender* is *Male*, then *Personal Risks* is *Low*			
	If *Age* is *Middle Age* and *Gender* is *Female*, then *Personal Risks* is *Medium*			
	If *Age* is *Elderly* and *Gender* is *Male*, then *Personal Risks* is *Medium*			
	If *Age* is *Elderly* and *Gender* is *Female*, then *Personal Risks* is *High*			
New rules suggested by Ali (based on Amy's previous views)	If *Age* is *Young* or *Middle Age*, then *Personal Risks* is *Low*		-	If *Age* is *Young* or *Middle Age*, and *Sex* is *Male*, then *Personal Risks* is *Low*
	If *Age* is *Middle Age* and *Sex* is *Female*, then *Personal Risks* is *Medium*			If *Age* is *Young* or *Middle Age* and *Sex* is *Female*, then *Personal Risks* is *Medium*
	If *Age* is *Older* and *Sex* is *Male*, then *Personal Risks* is *Medium*	Yes		
	If *Age* is *Older* and *Sex* is *Female*, then *Personal Risks* is *High*	Yes		

Fig. B.7 Secondary rules to label personal factors and parameterise 'Personal Risks' variable provided by Dr. Amy MacBrayne.



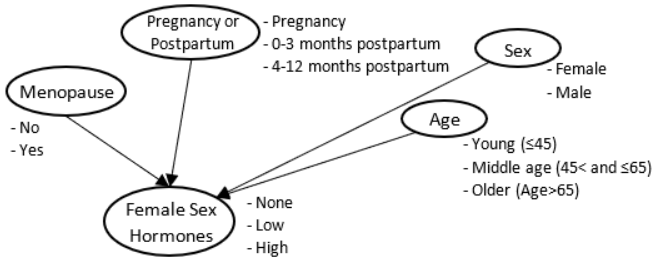
Status	Rule	Accepted	Rejected	Modification
Old rules defined by Fran, Amy, and Gloria	If *Education* is *University*, *Occupation* is *Non-industrial*, and *Ethnicity is Black or Black British*, then *Demographic Risks* is *Low*			
	If *Occupation is Non-industrial* and *Ethnicity* is *White*, then *Demographic Risks* is *Medium*			
	If *Education* is *School*, *Occupation* is *Industrial*, *Ethnicity* is *White*, then *Demographic Risks* is *High*			
New rules suggested by Ali	If *Occupation* is *Industrial, Ethnicity is Black or Black British*, then *Demographic Risks* is *Medium*		No	-
	If *Occupation* is *Industrial*, then *Demographic Risks* is *High*	Yes		
	If *Occupation* is *Non-Industrial, Ethnicity* is *White*, then *Demographic Risks* is *Medium*	-	-	No information
	If *Education* is *School*, *Occupation* is *Non-Industrial*, then *Demographic Risks* is *Medium*	-	-	No information
	If *Occupation* is *Non-Industrial*, then *Demographic Risks* is *Low*	Yes		

Fig. B.8 Secondary rules to label demographic factors and parameterise 'Demographic Risks' variable provided by Dr. Amy MacBrayne.



Status	Rule	Accepted	Rejected	Modification
Old rules defined by Fran, Amy, and Gloria	If *BMI* is *Normal* or *Underweight*, and *Alcohol* is *Medium*, then *Lifestyle Risks* is *Low*			
	If *BMI* is *Normal*, and *Alcohol* is *None* or *Low*, then *Lifestyle Risks* is *Medium*			
	If *BMI* is *Overweight*, and *Alcohol* is *None* or *Low* or *High*, then *Lifestyle Risks* is *Medium*			
	If *BMI* is *Obese*, and *Alcohol* is *None* or *High*, then *Lifestyle Risks* is *High*			
New rules suggested by Ali	If *Smoking* is *None*, then *Lifestyle Risks* is *Low*	Yes		
	If *Smoking* is *Low*, then *Lifestyle Risks* is *Medium*			If *Smoking* is *Low*, then *Lifestyle Risks* is *High*
	If *Smoking* is *High*, then *Lifestyle Risks* is *High*	Yes		

Fig. B.9 Secondary rules to label lifestyle factors and parameterise 'Lifestyle Risks' variable provided by Dr. Amy MacBrayne.



Status	Rule	Accepted	Rejected	Modification
Old rules defined by Fran, Amy, and Gloria	If *Age* is *Young* (or *Middle Age*?), and *Menopause* is *No*, or *Pregnancy or Postpartum* is *Outside*, then *Female Sex Hormones* is *Low* risk on RA			
	If *Age* is *Young* (or *Middle Age*?), and *Menopause* is *Yes*, then *Female Sex Hormones* is *High* risk on RA			
	If *Age* is *Young* (or *Middle Age*?), and *Pregnancy or Postpartum* is *0-3 months postpartum or 4-12 months postpartum*, then *Female Sex Hormones* is *High* risk on RA			
Additional new rules suggested by Ali	If *Age* is *Older*, and *Menopause* is *Yes*, or *Pregnancy or Postpartum* is *Outside*, then *Female Sex Hormones* is *None* risk on RA	Yes		
	If *Age* is *Older*, and *Pregnancy or Postpartum* is *Outside*, then *Female Sex Hormones* is *None* risk on RA	Yes		

Fig. B.10 Secondary rules to label hormonal pathogenesis factors and parameterise 'FSH Effect on RA' variable provided by Dr. Amy MacBrayne.

B.2 Observations of Dummy Patient Scenarios

Table B.1 Scenario of dummy patient 1 - observations of personal factors, demographic factors, lifestyle factors, medical background, comorbidities, signs, symptoms, and serology results, and expected outcome of diagnosis.

Type	Category	Variable	Evidence
Observations	Personal Factors	Age	44
		Sex	Female
		BMI	25
	Demographic Factors	Ethnicity	Asian or Asian British
		Occupation	Housewife
		Education	School
	Lifestyle Factors	Alcohol	None
		Smoking	None
	Medical Background	RA FHx	No
		Other IA FHx	No
	Comorbidity	OA	Yes
		TAD	No
		Skin Psoriasis	No
		Other IA	No
		CTD	No
		CA	No
	Signs	TJC	26
		SJC	21
		SSJ	Yes
	Symptoms	GH	95
		Fatigue	90
		Pain	100
		Stiffness	Yes
	Serology Results	CCP	108
		RhF	36
		ESR	21
		CRP	2
Outcome	Diagnosis	Diagnosis	RA (100%)

Table B.2 Scenario of dummy patient 1: Observations and expected outcome

Type	Category	Variable	Evidence				
			Baseline	1st follow-up	2nd follow-up	3rd follow-up	4th follow-up
Observation	Signs	TJC	26	16	10	8	5
		SJC	21	20	20	16	8
	Symptoms	GH	95	90	70	70	70
		Fatigue	90	90	80	73	75
		Pain	100	90	70	68	60
		Stiffness	Yes	Yes	Yes	Yes	Yes
	Serology Results	ESR	2	16	8	16	14
		CRP	2	2	2	2	1
	Medication	Biologics	None	None	None	None	None
		MTX	10mg	10mg	10mg	15mg	20mg
		SSZ	None	None	None	500mg (+500mg/1w)	2000mg
		HCQ	None	None	None	None	None
		Steroid	IM depo	None	IM depo	IM depo	None
	AME	Organ Damage	-	None	None	None	None
		Symptomatic Side-Effects	-	None	None	None	None
		Infection	-	None	None	None	None
Outcome	Disease Activity	Disease State	6.82	6.11	5.36	5.04	4.23
	Decision	Medication Review	-	Same Regiment	Same Regiment	Escalation	Escalation

Appendix C

Bayesian Network and Rule-Based System for Personalised Care for Living with Rheumatoid Arthritis

C.1 Data Collection Questions and Information Retrieval

Table C.1 Questions to collect personal information and 'Weather Sensitivity' data from patients.

Variable	Primary Question	Choices	Secondary Question	Choices
Age	What is your age?	(in years)	-	-
Sex	What is your sex?	Male, Female	-	-
BMI	What is your height?	(in cm)	-	-
	What is your weight?	(in kg)	-	-
Disease Duration	For how long have you been living with your chronic disease?	(in years)	-	-
Weather Sensitivity	How satisfied are you with managing your arthritis pain and mood swings due to changing weather conditions over the last six months?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last six months how often were your disease activity (swollen joints, pain) increased due to cold and humid weather conditions?	Always, Very Often, Sometimes, Almost Never, Never
			During the last six months how often were your disease activity (swollen joints, pain)l decreased due to sunny and dry weather conditions?	
			During the last six months how often were your mood and energy levels increased due to cold and humid weather conditions?	
			During the last six months how often were your mood and energy levels decreased due to sunny and dry weather conditions?	

Table C.2 Questions to collect disease manifestation data from patients

Variable	Primary Question	Choices	Secondary Question	Choices
Arthritis Pain	How satisfied are you with managing your arthritis pain over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often did you have severe pain from your arthritis?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often did you have pain in two or more joints at the same time?	
			During the last month how often did your pain limited your regular daily activities?	
			During the last month how often did your pain limited your normal social activities?	
Morning Stiffness	How satisfied are you with managing your morning stiffness over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often did your morning stiffness lasts less than half an hour from the time you woke up?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often did your morning stiffness lasts between 30 minutes to an hour from the time you woke up?	
			During the last month how often did your morning stiffness lasts between one hour to two hours from the time you woke up?	
			During the last month how often did your morning stiffness lasts more than two hours from the time you woke up?	
Fatigue	How satisfied are you with managing your fatigue and energy levels over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often did you feel full of pep?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often did you have a lot of energy?	
			During the last month how often did you feel worn out?	
			During the last month how often did you feel tired?	
Deformation Effect	How satisfied are you with condition and function of your body over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the past month could you do your own housework without help?	Always, Very Often, Sometimes, Almost Never, Never
			During the past month could you easily reach shelves that were above your head?	
			During the past month could you easily put on a pullover sweater?	
			During the past month could you easily open a new jar of food?	

Mobility	How satisfied are you with your physical ability to do what you want to do over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often were you physically able to drive a car or use public transportation?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often did someone have to assist you to get around outside your home?	
			During the last month how often were you in a bed or chair for most or all of the day?	
			During the last month how often were you out of the house for at least part of the day?	

Table C.3 Questions to collect quality of life characteristics data from patients.

Variable	Primary Question	Choices	Secondary Question	Choices
Financial Status	How satisfied are you with your financial status over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied		-
Self-care	How satisfied have you been with self-care of your health over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month did you need help to take a bath or shower?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month did you need help to get dressed?	
			During the last month did you need help to use the toilet?	
			During the last month did you need help to get in or out of bed?	
Mood	How satisfied are you with your mood over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often have you enjoyed the things you do?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often have you been in low or very low spirits?	
			During the last month how often did you feel that nothing turned out the way you wanted it to?	
			During the last month how often did you feel so down in the dumps that nothing would cheer you up?	
Sleep Quality	How satisfied are you with your sleep quality over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	How often did your pain make it difficult for you to sleep?	Always, Very Often, Sometimes, Almost Never, Never
			How often did your mood make it difficult for you to sleep?	
			How often did you use pain medication for sleep?	
			How often did you use sleeping pills?	

Table C.4 Questions to collect lifestyle choice data from patients.

Variable	Primary Question	Choices	Secondary Question	Choices
Work Capacity	How satisfied are you with your capacity for paid work, housework or school work over the past month?	Not Satisfied, Somewhat Satisfied, Satisfied	How often were you unable to do any paid work, housework or school work?	Always, Very Often, Sometimes, Almost Never, Never
			On the days that you did work, how often did you have to work a shorter day?	
			On the days that you did work, how often were you unable to do your work as carefully and accurately as you would like?	
			On the days that you did work, how often did you have to change the way your paid work, housework or school work is usually done?	
Social Activity	How satisfied are you with your social activities over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	How often did you get together with friends or relatives in a physical environment?	Always, Very Often, Sometimes, Almost Never, Never
			How often did you get together with friends or relatives in a virtual/social environment?	
			How often were you on the telephone with close friends or relatives?	
			How often did you go to a meeting of a club, team or other group?	
Physical Activity	How satisfied are you with your physical activities over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month did you have trouble doing vigorous activities such as running, lifting heavy objects, or participating in strenuous sports?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month did you have trouble either walking several blocks zor climbing a few flights of stairs?	
			During the last month did you have trouble bending, lifting or stooping?	
			During the last month were you unable to walk unless assisted by another person or by a cane, crutches, or walker?	
Belonging	How satisfied are you with your friends and family belonging over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often did you feel that your family or friends would be around if you needed assistance?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month to what extent did you have the opportunity for leisure and group activities?	
			During the last month to what extent did you have the opportunity for leisure and group activities?	
			During the last month how often did you feel that your family or friends were sensitive to your personal needs?	

			During the last month how often did you go to a meeting of a church, club, team or other group?	
Level of Tension	How satisfied are you with your level of tension over the last month?	Not Satisfied, Somewhat Satisfied, Satisfied	During the last month how often have you felt tense or high strung?	Always, Very Often, Sometimes, Almost Never, Never
			During the last month how often have you been bothered by nervousness or your nerves?	
			During the last month how often were you able to relax without difficulty?	
			During the last month how often have you felt calm and peaceful?	
Medication Adherence	How satisfied are you with managing your medication over the last 6 months?	Not Satisfied, Somewhat Satisfied, Satisfied	Do you ever forget to take your medicine?	Always, Very Often, Sometimes, Almost Never, Never
			Are you careless at times to take your medicine?	
			If you feel worse when you take the medicine, do you stop taking it?	
			When you feel better, do you stop taking your medicine?	
Disease Knowledge	How satisfied are you with the help and information you have received for your chronic disease over the last 6 months?	Not Satisfied, Somewhat Satisfied, Satisfied	Over the last 6 months how often have you received adequate information about your disease and treatment options?	Always, Very Often, Sometimes, Almost Never, Never
			Over the last 6 months how often have you accessed online resources for adequate information about your disease and treatment options?	
			Over the last 6 months how often have you been in-touch with your medical team for adequate information about your disease and treatment options?	

C.2 Evaluation of Bayesian Network Model for Personalised Care for Living with Rheumatoid Arthritis Using Scenarios

Table C.5 AtTRA 10's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration

Fragment	Variable	Evidence	Quote from AtTRA
Personal Factors	Age	Middle Age	-
	Sex	Male	-
	BMI	Overweight	-
QoL Characteristics	Mood	Depressed	-
	Sleep Quality	Low	-
	Self-care	Low	"his wife needed to help him to the toilet and helped to clean him, helped him stand at the sink to wash."
	Financial Status	Average	-
Disease Manifestation	Morning Stiffness	Less than 30 minutes	-
	Pain	Some	"The pain was then gone too"
	Fatigue	Some	-
	Deformation Effect	None	-
	Disability	Good	"In 2014 then 2015 he had hip replacement operations replacing both his hips. He was given a walker and a stick and he can now slowly walk"
Disease Activity	Disease Duration	More than 5 years	"He hasn't changed or interrupted the medication since starting in 2013"
Lifestyle Choices	Medication Adherence	High	"is still taking them all"
	Level of Tension	Low	"This helped him relax"
	Belonging	Medium	"he would also share his experiences with other patients and giving people suggestions based on his 5-6 years of experience"
	Work Capacity	High	-
	Social Activity	Low	"He only left the house when he had an appointment"
	Physical Activity	Low	-
	Disease Knowledge	High	"The most important thing when unwell is that you have support or moving around the house, being able to go outside and have a social life."

Table C.6 AtTRA 17's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration

Fragment	Variable	Evidence	Quote from AtTRA
Personal Factors	Age	-	-
	Sex	Female	-
	BMI	-	-
QoL Characteristics	Mood	Numb	"She also doesn't feel as clear headed"
	Sleep Quality	Low	"sleeping was difficult"
	Self-care	High	
	Financial Status	-	-
Disease Manifestation	Morning Stiffness	Between 30 minutes and 2 hours	Stiffness improved as the day went on.
	Pain	Severe	-
	Fatigue	Some	-
	Deformation Effect	None	-
	Disability	Good	"medication fairly quickly brought it under control"
Disease Activity	Disease Duration	Between 2 and 5 years	-
Lifestyle Choices	Medication Adherence	High	"Medication has worked well for her"
	Level of Tension	Medium	"She wants more than that for her peace of mind"
	Belonging	-	-
	Work Capacity	Medium	"She was not employed"
	Social Activity	-	-
	Physical Activity	-	-
	Disease Knowledge	High	"She has been well looked after by the medical team and could always get more information when she had issues".

Table C.7 AtTRA 24's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration

Fragment	Variable	Evidence	Quote from AtTRA
Personal Factors	Age	-	-
	Sex	Female	-
	BMI	-	-
QoL Characteristics	Mood	-	-
	Sleep Quality	Medium	"She can achieve a lot more when she has had eight to nine hours sleep than if she is only getting six hours sleep"
	Self-care	High	
	Financial Status	Good	-
Disease Manifestation	Morning Stiffness	More than 2 hours	"What they don't know is it took her three and a half hours to get ready for work"
	Pain	Some	-
	Fatigue	Some	"Fatigue is a lot worse when she is inflamed"
	Deformation Effect	None	-
	Disability	Good	-
Disease Activity	Disease Duration	Less than 1 year	"She was diagnosed only 9-10 months ago"
Lifestyle Choices	Medication Adherence	High	"She takes half of the daily dose in the morning, and half in the evening now"
	Level of Tension	Medium	"Your emotions go through highs and lows"
	Belonging	High	-
	Work Capacity	High	"She kept going to work, because she had just started a new job, to forget she had it, she likes work, and it makes her feel normal. She has never had sick days at school or work."
	Social Activity	Medium	"She wasn't a party animal, but she used to go out, meet friends, etc" and "She didn't do that often now"
	Physical Activity	Medium	-
	Disease Knowledge	Low	"She is starting to understand when her pain is worse and when it's not".

Table C.8 AtTRA 38's evidences of personal factors, environmental factors, quality characteristics, disease manifestation, and disease duration

Fragment	Variable	Evidence	Quote from AtTRA
Personal Factors	Age	Older	-
	Sex	Female	-
	BMI	Obese	-
QoL Characteristics	Mood	Depressed	"She suffers from depression and takes depression tablets."
	Sleep Quality	Medium	"Sometimes she can't sleep when in a lot of pain"
	Self-care	Low	"There was a time where she stopped eating because she couldn't pick up a spoon, couldn't go to the toilet, or shower."
	Financial Status	Bad	-
Disease Manifestation	Morning Stiffness	More than 2 hours	"In the morning because of joint stiffness"
	Pain	Severe	"pain you just can't get up, you can't lift your head up"
	Fatigue	Severe	"The medication has a lot of side-effects making her very drowsy, very tired, and very fatigued."
	Deformation Effect	None	-
	Disability	Average	"She feels disabled, even though she may not look it."
Disease Activity	Disease Duration	More than 5 years	-
Lifestyle Choices	Medication Adherence	High	"She has been taking it over a year now, so is a pro now."
	Level of Tension	Medium	-
	Belonging	Medium	"She would like to be able to use social media (or face to face) to ask questions of other patients. It is nicer to speak to someone who is going through the same thing and talk about the negative and positive side." and "Her husband helps doing most stuff and when he is at work her kids help."
	Work Capacity	Low	(retired)
	Social Activity	Low	-
	Physical Activity	Low	"When unwell the most important thing is just to have someone to give physical help"
	Disease Knowledge	High	-

C.3 Using Tornado Graphs for Sensitivity Analysis of Bayesian Network Model for Personalised Care of Rheumatoid Arthritis

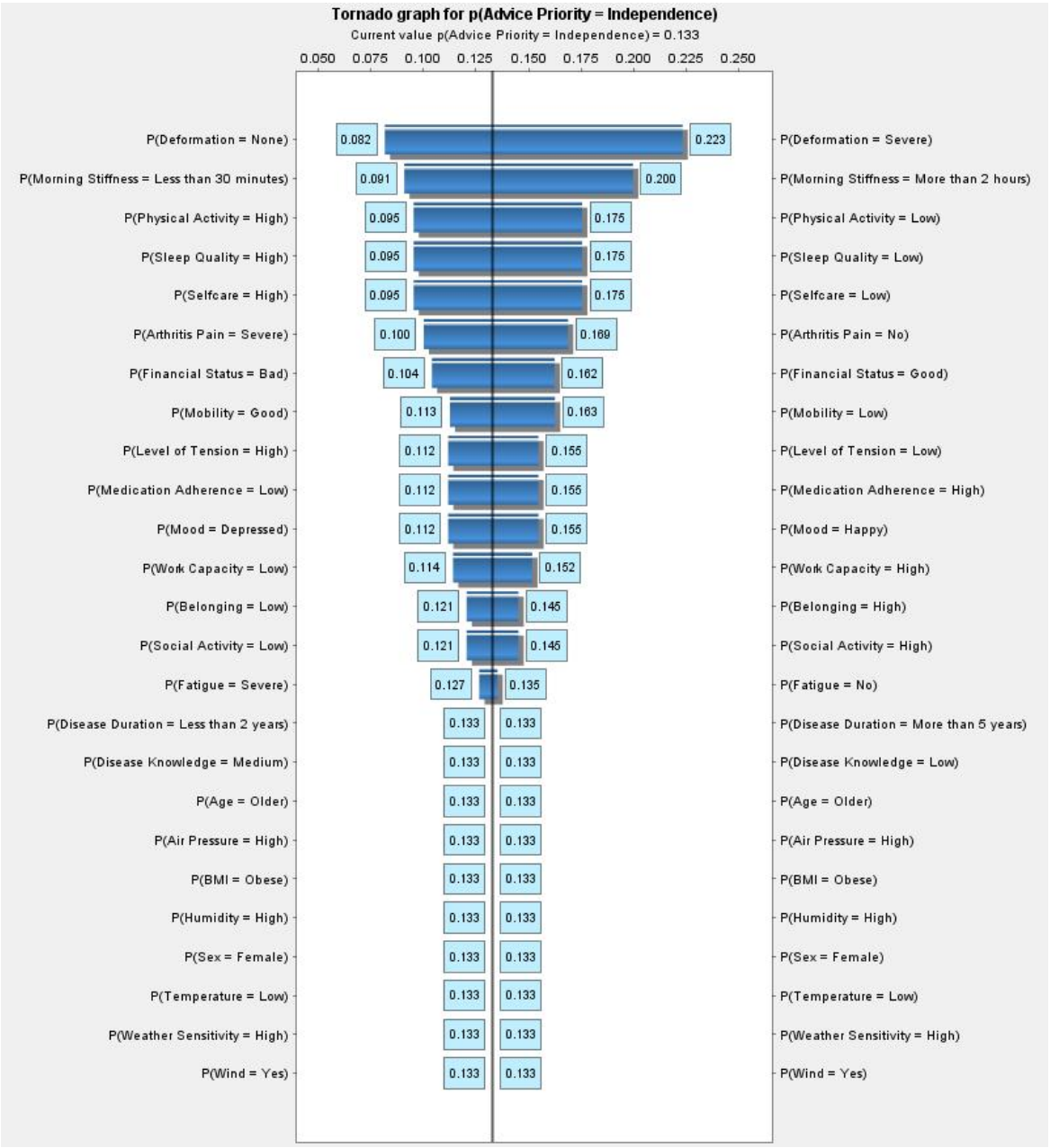


Fig. C.1 Tornado graph of the Independence state of the 'Advice Priority' variable.

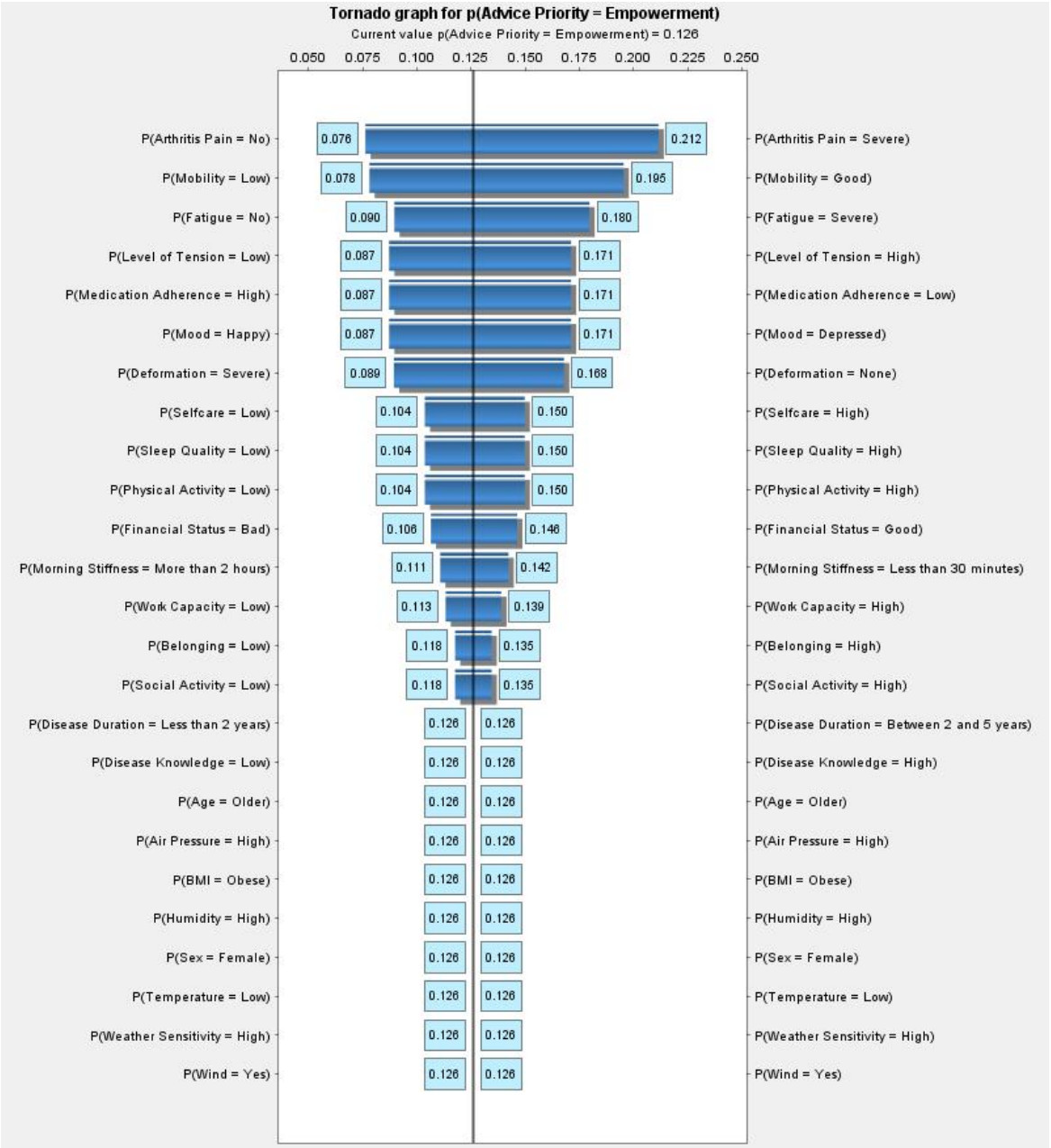


Fig. C.2 Tornado graph of the Empowerment state of the ‘Advice Priority’ variable.

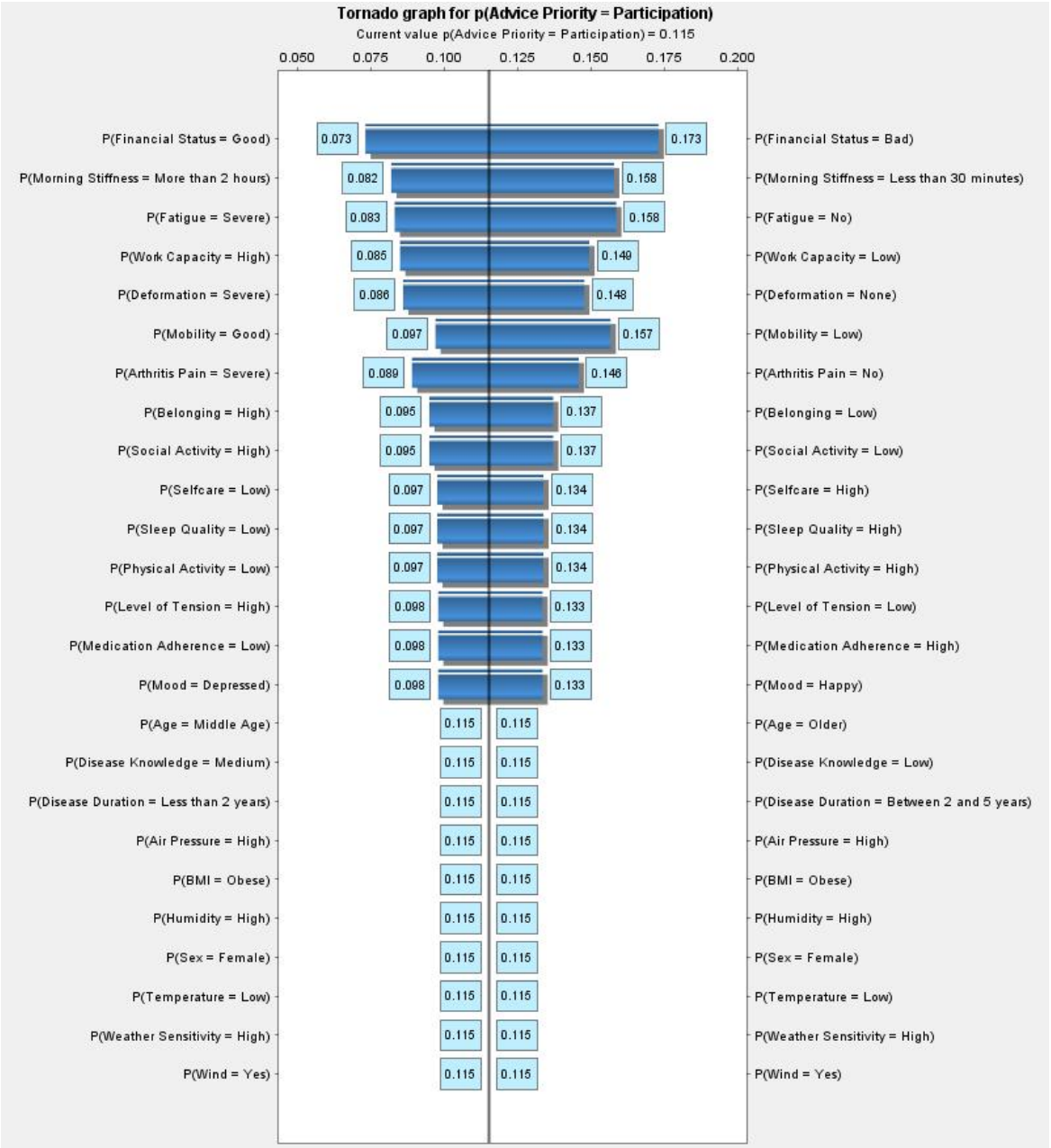


Fig. C.3 Tornado graph of the Participation state of the 'Advice Priority' variable.

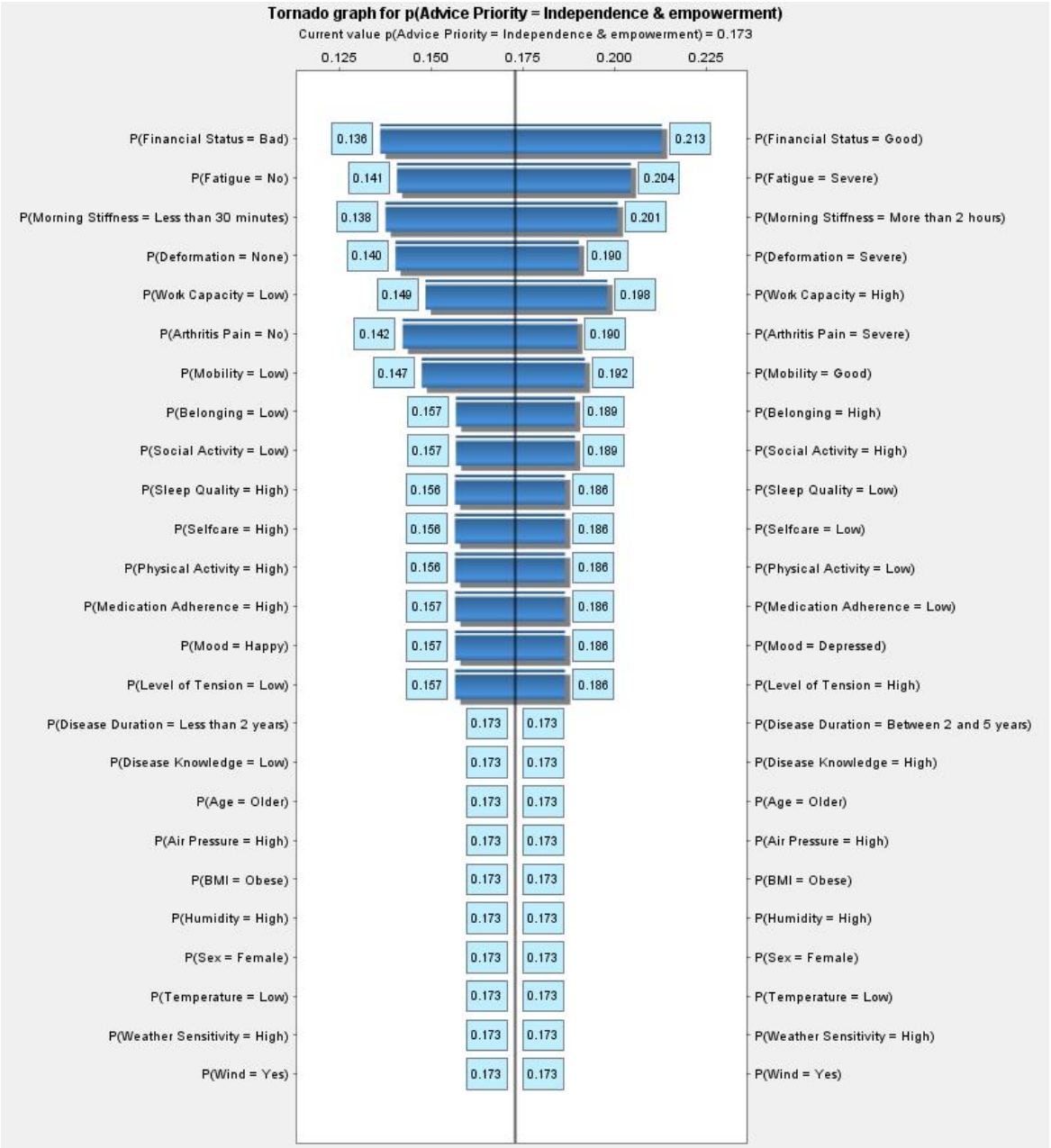


Fig. C.4 Tornado graph of the 'Independence & empowerment' state of the 'Advice Priority' variable.

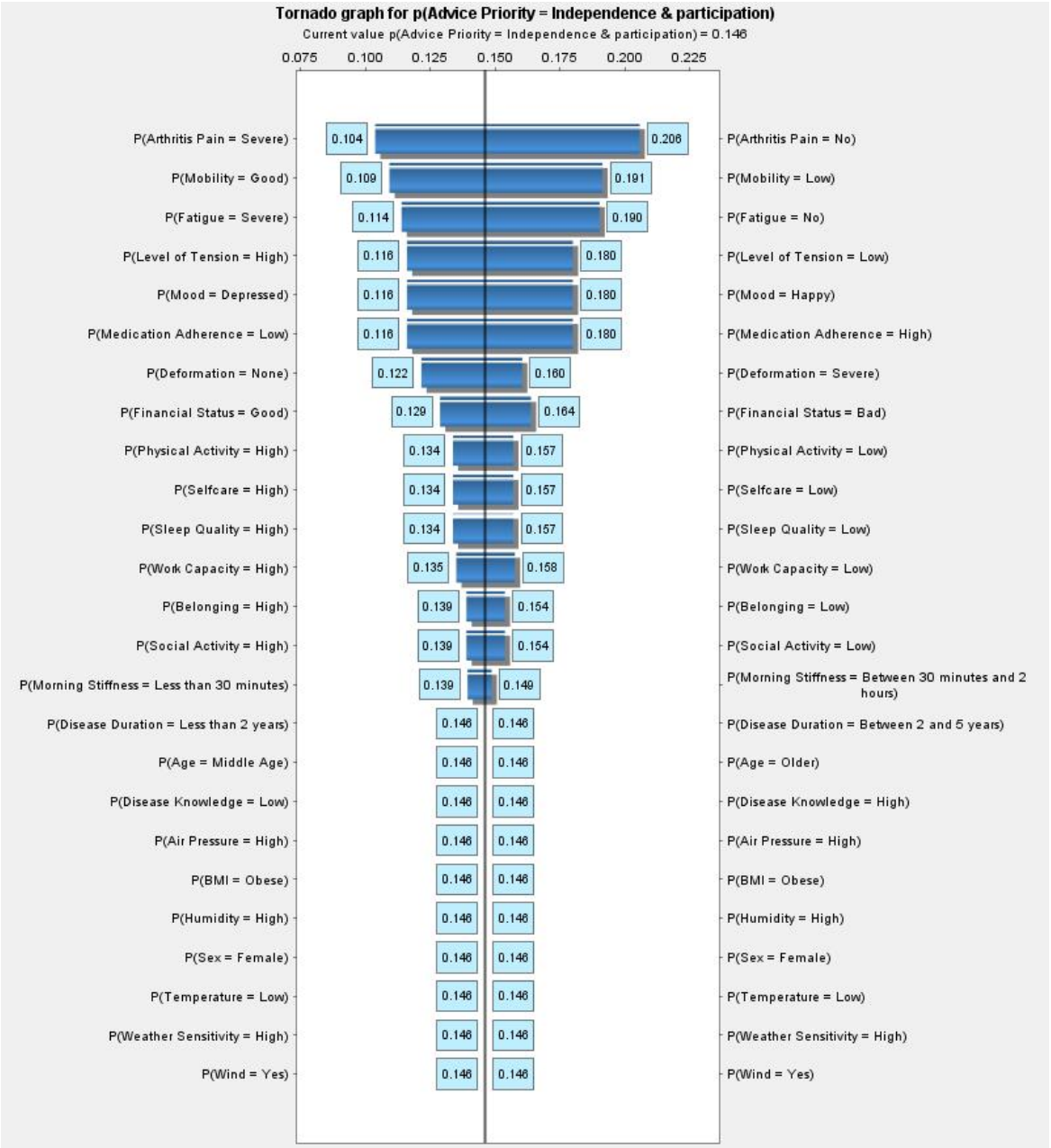


Fig. C.5 Tornado graph of the ‘Independence & participation’ state of the ‘Advice Priority’ variable.

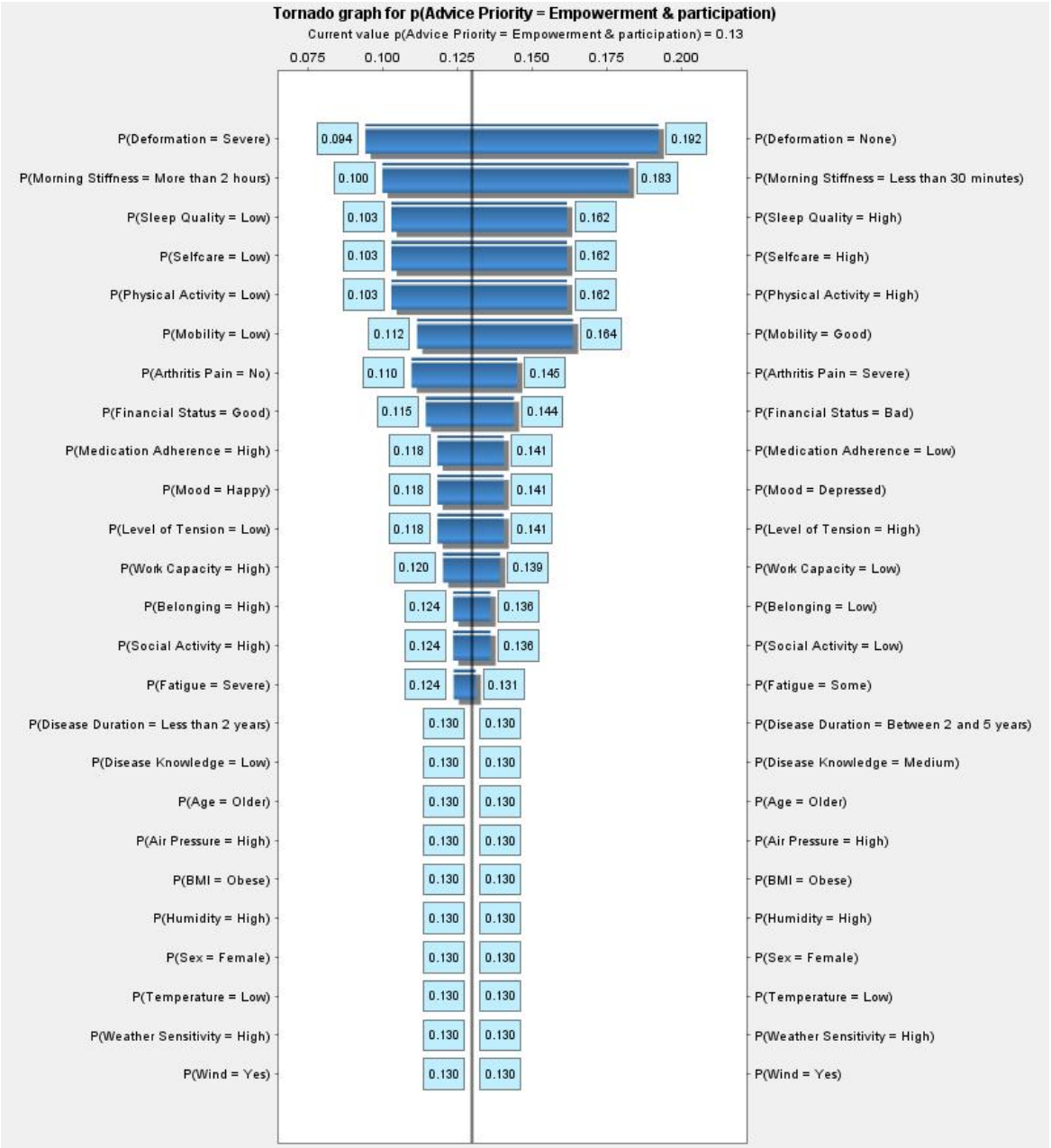


Fig. C.6 Tornado graph of the 'Empowerment & participation' state of the 'Advice Priority' variable.