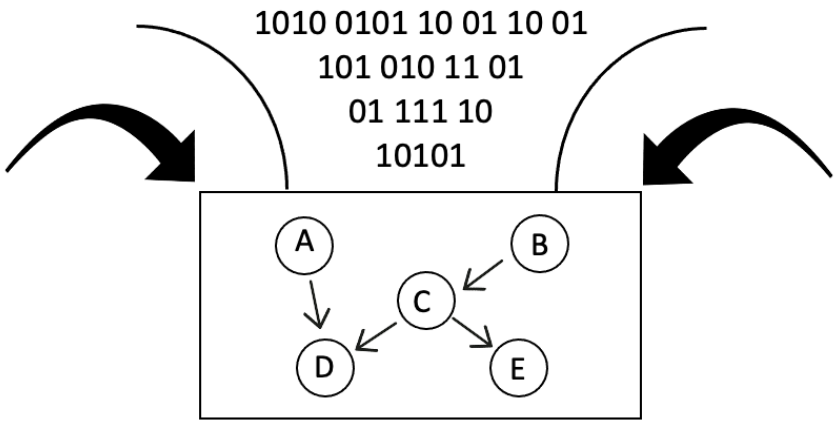
TÉCNICO LISBOA



Dynamic Bayesian networks for clinical decision support on ankylosing spondylitis patients under biological therapies

Ana Rita Lopes Borges Martins

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisor(s): Prof. Susana de Almeida Mendes Vinga Martins
Prof. Alexandra Sofia Martins de Carvalho

Examination Committee

Chairperson: Prof. João Miguel Raposo Sanches
Supervisor: Prof. Susana de Almeida Mendes Vinga Martins
Members of the Committee: Prof. Maria da Conceição Esperança Amado
Prof. Ana Maria Ferreira Rodrigues

July 2021

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work in this thesis was performed at INESC-ID and IT, in collaboration with the EpiDoC Unit, CEDOC, during the period of March 2020-June 2021, under the supervision of Prof. Alexandra Carvalho and Prof. Susana Vinga, and within the scope of the Project PREDICT (PTDC/CCI- CIF/29877/2017), funded by Fundo Europeu de Desenvolvimento Regional (FEDER), through Programa Operacional Regional LISBOA (LISBOA2020), and by national funds, through Fundação para a Ciência e Tecnologia (FCT).

Acknowledgments

The completion of this journey was only possible due to the support of many people that accompanied me, and to whom I would truly like to thank.

First of all, I would like to express my deepest gratitude to my supervisors Prof. Susana Vinga and Prof. Alexandra Carvalho for all their guidance and support given throughout this period, for all their useful insights and suggestions, and also for their kindness and friendship.

I would also like to thank all the people involved in the PREDICT project and in the Systems Biomedicine group, for all the insightful discussions.

Moreover, I would like to thank my family, specially my parents, for all their help and support. For always believing in me, for motivating me to do my best and for giving me essential values that made me the person that I am today. Without them, nothing of this would have been possible, and for that, I am forever grateful.

Additionally, I would like to thank my friends, for all the support and amazing moments, and for making this journey so much better.

Last but not least, I would like to thank Nuno for all his love and support. Thank you for always being there for me, for cheering me up in my worst moments and for always encouraging me to keep going. It wouldn't have been the same without you.

Resumo

As doenças reumáticas e musculoesqueléticas (DRMs) encontram-se entre as doenças crónicas mais comuns do mundo, constituindo um enorme fardo para os indivíduos. A espondilite anquilosante (EA) é uma doença inflamatória que afeta principalmente as articulações da coluna, levando os pacientes a experienciar não só limitações físicas, mas também uma redução na qualidade de vida.

Atualmente, não existe cura para a doença, no entanto, existem tratamentos que permitem retardar a progressão e aliviar os sintomas, apresentando-se como promissoras terapias biológicas recentemente desenvolvidas, incluindo antagonistas do fator de necrose tumoral (TNF) alfa. Contudo, nem todos os pacientes respondem a estes tratamentos, podendo até experienciar efeitos adversos severos. Adicionalmente, estas são terapias dispendiosas, sendo essencial identificar quais os pacientes que delas poderão beneficiar.

O objetivo deste trabalho é identificar preditores de resposta à terapia e distinguir entre três diferentes agentes anti-TNF, adalimumab, etanercept e infliximab, de modo a avaliar o que melhor se adequa a cada paciente. Este objetivo foi conseguido através de redes de Bayes dinâmicas construídas a partir de dados do Registo Nacional de Doentes Reumáticos.

Resultados obtidos com este estudo corroboram alguns dos encontrados na literatura, como os homens apresentarem melhores respostas às três terapias estudadas. Por outro lado, novas associações foram identificadas, com pacientes com IMC mais elevado apresentando uma pior resposta quando tratados com adalimumab. Para além disso, pacientes HLA-B27 negativos parecem ter uma pior resposta quando tratados com infliximab. Finalmente, para pacientes mais velhos ou com maior duração da doença, infliximab parece ser uma melhor opção terapêutica.

Palavras-chave: Espondilite anquilosante, redes de Bayes dinâmicas, terapias anti-TNF-alfa, previsão de resultados terapêuticos

Abstract

Rheumatic and musculoskeletal diseases (RMDs) are among the most common chronic diseases worldwide, having an enormous burden on individuals. Ankylosing spondylitis (AS) is a type of inflammatory arthritis that affects primarily the spine and the sacroiliac joints, with patients experiencing not only physical limitations, but also reduced quality of life.

Currently, there is no cure for the disease, but there are treatments that aim at slowing its progression and relieving its symptoms, appearing as promising recently developed biological therapies, as tumor necrosis factor (TNF) antagonists. Nevertheless, not all patients respond to these therapies and some of them even experience severe adverse events. Furthermore, TNF blocking therapy is costly, being essential to identify which patients are likely to benefit from these agents and which are not.

The goal of this work is to find clinical predictors of therapy response and to distinguish between three different anti-TNF agents, adalimumab, etanercept and infliximab, in order to evaluate which therapy best suits each patient. This was done using dynamic Bayesian network (DBN) models built from data from the Rheumatic Diseases Portuguese Register.

This study allowed to corroborate existing literature findings, as male patients responding better to all therapies studied, and also to produce new insights. Firstly, patients with greater BMI seem to response worse when treated with adalimumab. Secondly, HLA-B27 negative patients seem to respond worse to infliximab. Finally, for older patients or patients with higher disease duration, infliximab seems to be a better therapeutic option.

Keywords: Ankylosing spondylitis, dynamic Bayesian networks, anti-TNF-alpha therapies, therapy outcome prediction

Contents

List of Tables

List of Figures

Nomenclature

AIC Akaike Information Criterion

ASAS Assessment of Spondyloarthritis International Society

ASDAS Ankylosing Spondylitis Disease Activity Score

AS Ankylosing Spondylitis

axSpa Axial Spondyloarthritis

BASDAI Bath Ankylosing Spondylitis Disease Activity Index

BASFI Bath Ankylosing Spondylitis Functional Index

BDeu Bayesian-Dirichlet Equivalent Uniform

BIC Bayesian Information Criterion

BMI Body Mass Index

CPT Conditional Probability Table

CRP C-Reactive Protein

CT Computed Tomography

DAG Directed Acyclic Graph

DBN Dynamic Bayesian Network

DMARD Disease-Modifying Anti-Rheumatic Drug

ERAP1 Endoplasmic Reticulum Aminopeptidase 1

ESR Erythrocyte Sedimentation Rate

HC Hill Climbing

HLA-B27 Human Leukocyte Antigen B27

IL Interleukin

JIA Juvenile Idiopathic Arthritis

LOCF Last Observation Carried Forward

MHC Major Histocompatibility Complex

MMHC Max-Min Hill-Climbing

MMPC Max-Min Parent-and-Children

MRI Magnetic Resonance Imaging

NOCB Next Observation Carried Backward

nr-axSpa Non-Radiographic Axial Spondyloarthritis

NSAID Non-steroidal Anti-Inflammatory Drug

PsA Psoriatic Arthritis

RA Rheumatoid Arthritis

RMD Rheumatic and Musculoskeletal Disease

RNDR Rheumatic Diseases Portuguese Regist

SEM Structural Expectation-Maximization

SM Silander-Myllymäki

SpAs Spondyloarthropathies (or Spondyloarthritis)

SPR Portuguese Society of Rheumatology

TNF Tumor Necrosis Factor

VAS Visual Analog Scale

Chapter 1

Introduction

1.1 Motivation

Rheumatic and musculoskeletal diseases (RMDs) are among the most common chronic diseases worldwide, with around one-quarter of the population in Europe suffering from a chronic RMD [1]. RMDs comprise more than 200 different diseases affecting mainly the joints, tendons, ligaments, bones and muscles and affect individuals from all age groups, having an enormous burden on individuals, families and societies. These diseases cause great functional limitations, representing the leading cause of inability to work and early retirement and one of the greatest causes of short-term sick-leave [2]. Besides the impairment of physical functions, RMDs also have a huge impact on people's mental health, being associated with anxiety symptoms and emotional problems [3]. Moreover, these diseases have a very negative impact on the economy, inducing high costs to healthcare systems, due to expenses in medicines, surgeries, hospital stay and physiotherapy, among others [4].

Ankylosing spondylitis (AS) is a type of inflammatory arthritis that affects primarily the spine and the sacroiliac joints [5, 6]. It belongs to a group of rheumatic diseases called spondyloarthropathies, or spondyloarthritis (SpAs), being the most common among these, as well as the one that presents more severe outcomes. AS patients experience not only physical limitations, which can lead to implications for employment and in extreme cases even result in inability of working, but also to a reduced quality of life, possibly experiencing anxiety and depression. It is a complex and unpredictable disease, being its pathogenesis poorly understood [7]. Nevertheless, due to its serious consequences, there is an urgent need of better understanding the disease and improving the knowledge of the causes and mechanisms of AS, in order to improve early diagnosis and prevention and develop innovative therapies [1].

Currently, there is no cure for the disease, but there are treatments that aim at slowing its progression and relieving its symptoms [8]. Recently developed biological therapies, as tumor necrosis factor (TNF) antagonists, or anti-TNF, appear as promising treatments. Nevertheless, not all patients respond to these therapies and some of them even experience severe adverse events. Indeed, TNF antagonists seem to increase the risk of infections, including increase of tuberculosis susceptibility and reactivation, and other opportunistic infections [9, 10]. Furthermore, TNF blocking therapy is costly, being associated with a great economic burden. Therefore, to minimise risks and costs associated with these therapies, it is essential to identify which patients are likely to benefit from anti-TNF agents and which are not.

This way, there is the need of identifying predictors of response to therapy, to help in clinical decisions. Additionally, there is no specific drug selection criteria regards deciding which biologic should be given to each patient, with clinical decisions regarding this choice being in some cases based on logistic issues, rather than on scientific evidence. This leads to patients switching between different therapies until finding one to which they respond positively. Nevertheless, this lack of accuracy in therapeutic choice should not be underestimated due not only to the high costs associated with these therapies, but also to the negative effects that the patients might experience. This way, being able to predict in advance the patient's response for each treatment, according to the patient's characteristics, would be of great benefit.

1.2 Objectives

The main goal of this work is to identify predictors of response to the different biological therapies and to be able to produce insights about which treatment suits each patient best, in the context of a personalized medicine approach. Indeed, several works have studied anti-TNF therapies as a whole, and more precisely predictors of response to these group of therapies, nevertheless, to the extent of our knowledge, there are yet no studies investigating differences between the existing anti-TNF therapies and trying to tailor treatment to each individual patient. Hence, this work is a pioneer in the field, with the main goal of producing first insights on this topic.

One proposes to achieve this objective by developing Bayesian network (BN) models, more concretely dynamic Bayesian networks (DBNs), using data of AS patients on biological therapies taken from the Rheumatic Diseases Portuguese Register (RNDR), the Reuma.pt database. The aim is that the DBN models help understand patient's response to therapy by detecting probabilistic relationships among clinical variables and identifying risk factors related to therapy outcome. The goal is that these models can help doctors in the prognosis task, by generating insights about each patient's response to therapy.

On the other hand, to build these models, data should be processed and treated until the required format is achieved. Indeed, the full potential and wealth of the Reuma.pt dataset has not been thoroughly exploited so far, and so this work is required to extract further profits from it in order to generate clinical hypotheses that might be useful for healthcare providers.

1.3 Claim of contributions

With this work it was possible to identify differences between the anti-TNF therapies studied and several potential predictors of biological therapy inefficiency, which should be used as starting points for future studies on this topic.

Another contribution was the pre-processing of the dataset, with the help of medical professionals, which allowed the identification and correction of inconsistencies in the data. Furthermore, several datasets were created, where data was organized into a structured format ready to be used in future works investigating AS.

1.4 Thesis Outline

This thesis is organized as follows.

First, in Chapter 2 the main theoretic concepts of this work are described. Indeed, a theoretical background on ankylosing spondylitis is provided, addressing topics like etiology and pathogenesis, diagnosis and treatments, assessment and predictors of response to therapy. Additionally, a theoretical background on Bayesian networks is also given, including the definition, related concepts and some examples, and the concept of dynamic Bayesian networks is also introduced.

Next, in Chapter 3 the object of study of this dissertation and the methodology followed are presented, being described all steps since data acquisition and processing, to the implementation of the DBN models.

In Chapter 4, the dataset used is described in detail and the results obtained from the DBN models are presented and discussed.

Finally, in Chapter 5 the conclusions reached with this work are summarized, its strengths and weaknesses are discussed and suggestions for future work are presented.

Chapter 2

Background

2.1 Ankylosing Spondylitis

Ankylosing Spondylitis (AS) is a type of inflammatory arthritis that affects primarily the spine and the sacroiliac joints, causing inflammation and leading to chronic pain and discomfort [5, 6]. This disease can be associated with new bone formation in the spine and fusion of the vertebrae, as schematized in Figure 2.1, resulting, in extreme situations, in the total fusion of the spine, a characteristic called "bamboo spine" [11]. This causes stiffness and pain, and leads to possible loss of physical function and spinal mobility. Nevertheless, AS doesn't affect only the spine and peripheral arthritis and extra-articular features can also be present [12].

AS starts at a relatively young age, usually in the second or third decade of life, with men being two to three times more affected than women. Also, the pattern of the disease is different between the two, with women having milder disease and more extra spinal involvement [7]. The worldwide prevalence of AS is believed to be between 0,1% and 1,4% [13], while in Europe is around 0,25% [14].

2.1.1 Etiology and pathogenesis

The etiology and pathogenesis of AS are still poorly understood. Nevertheless, it is believed that it develops through complex interactions between genetic and environmental factors [7, 16].

AS is strongly associated with the major histocompatibility complex (MHC) class I allele human leukocyte antigen B27 (HLA-B27) [17]. HLA-B27 is a protein located on the surface of white blood cells that is found in 90-95% of AS patients. However, only 1-2% [16] of people that have the antigen HLA-B27, which is around 8% of the total population [18], develop AS, indicating that there are other factors that lead to the appearance of the disease.

There are other genes thought to be involved in the etiology of AS. One of these is *ERAP1*, which codes an enzyme called endoplasmic reticulum aminopeptidase 1, that trims peptides for binding to HLA class 1 molecules, such as HLA-B27 [12, 19]. Indeed, some studies indicate that interactions between HLA-B27 and variants of *ERAP1* have a significant role in the development of AS [16, 19].

The interleukin (IL)-23/IL-17 pathway is also believed to be implicated in the pathogenesis of AS [16]. IL-17 is a proinflammatory cytokine involved in defense against bacteria and fungi, produced by

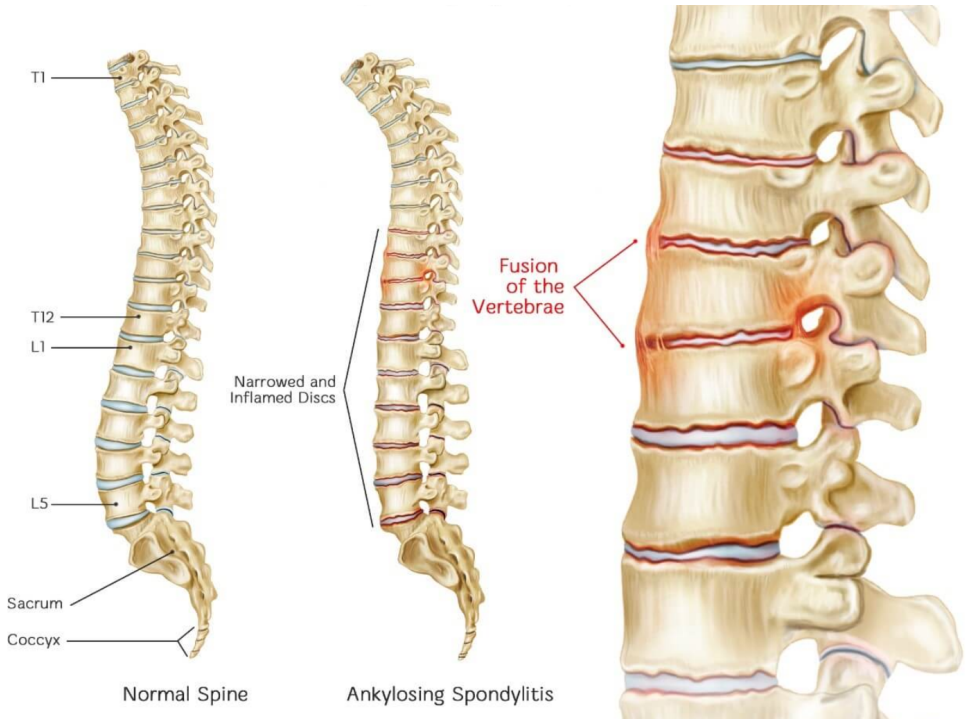


Figure 2.1: Schematic representation of AS effects on the spine. Taken from [15].

Th17 cells. These cells result from the differentiation of CD4 T cells, induced by cytokine IL-23. [20]

Another proinflammatory cytokine believed to be implicated in the pathogenesis of AS is the tumor necrosis factor alpha (TNF-alpha). TNF-alpha is a cytokine involved in bone remodeling, which has been found to be raised in the serum and synovial tissue of AS patients [12, 21]. In fact, for the case of AS, TNF-alpha in combination with IL-17 are thought to induce bone matrix formation, a characteristic of the disease [22].

2.1.2 Diagnosis and treatments

Since unequivocal sacroiliac changes may take many years after the beginning of symptoms to be evident on radiographs, diagnosis can be delayed. To some patients these changes may not appear at all, a condition called non-radiographic axial spondyloarthritis (nr-axSpa), which together with AS constitutes a group of diseases called axial spondyloarthritis (axSpa). To tackle this issue, magnetic resonance imaging (MRI) may be used due to its higher sensitivity in detecting inflammation, being able to detect it even when computed tomography (CT) scans are normal [12].

To help clinicians with the diagnosis task, several criteria sets are available, being one of the most widely used the modified New York Criteria [7], presented in Table 2.1. The Assessment of Spondyloarthritis International Society (ASAS) classification criteria for axial spondyloarthritis, which can be seen in Figure 2.2, is another criteria set that can also help physicians in diagnosing this condition.

Currently, there is no cure for the disease. Nonetheless, there are treatments that aim to relieve symptoms, like pain and stiffness, prevent or delay spinal deformity and reduce complications, in order to promote better quality of life [8, 24]. Besides, independently of any drug treatment, people should

Table 2.1: Modified New York Criteria for Ankylosing Spondylitis.

Clinical criteria
Low back pain and stiffness which improves with activity for more than 3 months.
Limited range of motion of the lumbar spine in both forward and lateral bending.
Limitation of chest expansion relative to normal values correlated for age and sex.
Radiological criteria
Sacroiliitis grade greater or equal to 2 bilaterally.
Sacroiliitis grade 3 to 4 unilaterally.
Note: Diagnosis of AS is made if the patient fulfills at least one radiological and one clinical criteria.

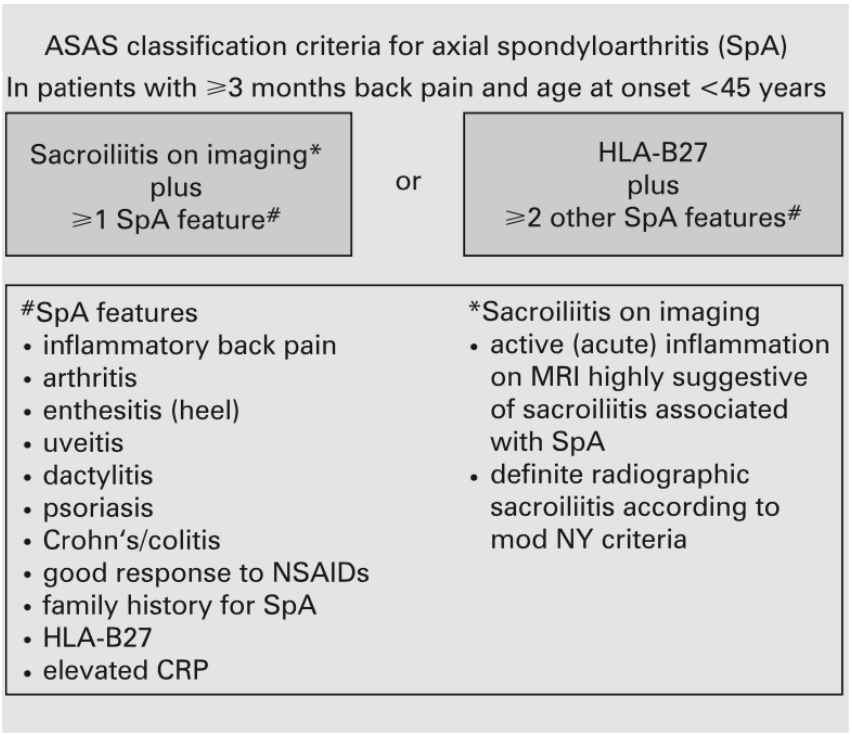


Figure 2.2: ASAS criteria for classification of axial spondyloarthritis (to be applied in patients with chronic back pain and age at onset of back pain less than 45 years). Taken from [23].

consider some changes on their lifestyle, like practicing physical activity, going to physiotherapy and stop smoking [16].

Non-steroidal anti-inflammatory drugs (NSAIDs) are the first-line treatment for AS. These are helpful relieving inflammatory symptoms, however, they do not alter the course of the disease neither prevent structural damage. Also, their long term use can have negative effects on patients, like hypertension, adverse gastrointestinal effects and renal disease, which restricts their use [12].

Disease-modifying antirheumatic drugs (DMARDs) and corticosteroids are sometimes used as second-line treatments in AS.

If with the above therapeutics patients don't show significant improvements, then, the next treatment option is to use recently developed biologic agents. Biological therapies, also known as biologics, are complex proteins manufactured within a living system, which target specific molecules thought to be involved in AS [25]. These agents, besides providing relief of symptoms, are potentially capable of

altering the course of the disease, by targeting inflammatory processes.

One type of biological therapy are TNF-inhibitors (anti-TNF), which target the tumor necrosis factor. Adalimumab, etarnercept and infliximab are three of the most commonly used anti-TNFs in clinical practice, being adalimumab and infliximab both monoclonal antibodies directed against TNF, the first fully humanized and the latter chimeric. On the other hand, etanercept is a fusion protein that blocks the interaction between TNF-alpha and the receptors on the cells' surface [26].

Another type of biological therapy are interleukin (IL)-17 inhibitors, that can be an alternative to anti-TNF therapy, in case of failure of the latter [12].

If after 3 months of application of a first biological therapy patients don't show considerable improvement, or if they are unable to tolerate the drug, they may need to change to a second one [16]. Nevertheless, before initiating therapy with IL-17 inhibitors, doctors should try an alternative TNF-inhibitor [12]. Figure 2.3 presents a graphical representation of the drug treatment strategy followed for AS patients.

2.1.3 Predictors of response to therapy

Although TNF-alpha blocking therapies are effective in a great number of AS patients, some don't show any response, others show less improvement, and some may even suffer from serious adverse events. Identifying patients who are most likely to benefit from anti-TNF therapy is important to minimize the costs and potential side effects of these therapeutic options. To this end, several studies have been conducted over the years in order to identify predictors of response for these biological agents, that might identify responders and help doctors make better therapeutic decisions.

Some of the previously identified predictors of response include younger age [10, 28, 29], male gender [10, 29], higher ASDAS [29] and BASDAI [10, 28] scores at baseline, lower BASFI at baseline [10, 28], raised levels of ESR [28, 30] and CRP [10, 28, 30] at baseline, presence of peripheral arthritis [29], smaller disease duration [28] and higher patient's global assessment of disease activity [29]. Furthermore, HLA-B27-positive patients have been reported to respond better to TNF antagonists [10, 31].

On the other hand, female gender [29, 32], absence of peripheral arthritis [29], higher baseline BASFI score [30] and lower baseline levels of ESR and CRP [29] have been identified as predictors of discontinuation of TNF-alpha blocking therapy. Additionally, higher levels of BMI have been associated with non-response [33, 34].

2.1.4 Disease assessment

The need of measuring outcomes and assessing disease status and response to therapy became even more relevant after the emergence of anti-TNF therapies [35, 36]. To this end, the Assessment of SpondyloArthritis international Society, ASAS, formerly Assessment in Ankylosing Spondylitis, was formed in 1995 [23]. The group's work is focused on supporting and promoting the study of spondyloarthritis, by increasing awareness in the disease, promoting early diagnosis, developing and validating assessment instruments and studying treatment response in clinical trials [23, 36, 37]. Therefore, over time, a lot of

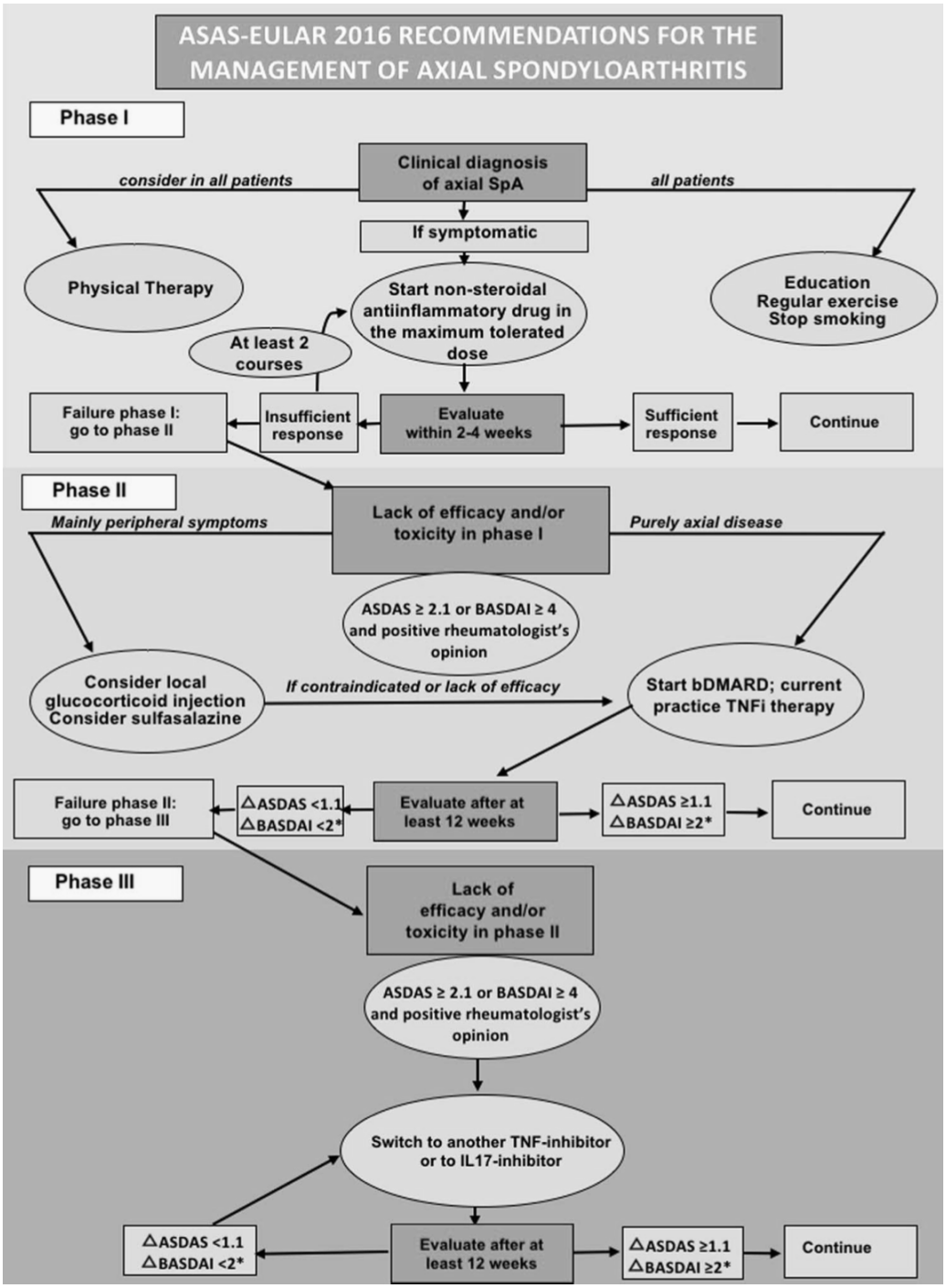


Figure 2.3: Drug treatment strategy for AS patients. Taken from [27].

clinical measures have been developed and used in clinical practice, being the Bath Ankylosing Spondylitis Functional Index (BASFI), the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and the Ankylosing Spondylitis Disease Activity Score (ASDAS) among the most commonly used.

BASDAI is a score used to assess patient-reported disease activity during the week prior to answering the corresponding questions, using visual analogue scales (VAS) [35, 38]. It consists of 6 questions related to fatigue, spinal pain, peripheral arthritis, enthesitis and duration and severity of morning stiffness [39], which are presented in Table 2.2. Each question is scored between 0 and 10, where 0 means no disease

activity and 10 means very active disease, except for question 6, where 0 is zero hours, 5 is one hour, and 10 is two or more hours. The formula to calculate the BASDAI score is presented in Table 2.3.

Next, BASFI is used to assess the degree of functional limitation of patients. This index is based on 10 questions about daily activities, presented in Table 2.2, like dressing, bending and standing, and ability to cope with everyday life, using, again, visual analog scales. BASFI score ranges from 0 to 10, being calculated as the simple average of the answers to the 10 questions, with higher scores meaning greater functional impairment [35, 39].

Table 2.2: BASDAI and BASFI Questions.

BASDAI Questions
Q1. How would you describe the overall level of fatigue/tiredness you have experienced?
Q2. How would you describe the overall level of AS neck, back, or hip pain you have had?
Q3. How would you describe the overall level of pain/swelling in joints other than neck, back, hips you have had?
Q4. How would you describe the level of discomfort you have had from an area tender to touch or pressure?
Q5. How would you describe the level of morning stiffness you have had from the time you wake up?
Q6. How long does your morning stiffness last from the time you wake up?
BASFI Questions
Please indicate your level of ability with each of the following activities in the past week.
Q1. Putting on your socks or tights without help or aids (e.g. sock aids)?
Q2. Bending forward from the waist to pick up a pen from the floor without an aid?
Q3. Reaching up to a high shelf without help or aids (e.g. helping hand)?
Q4. Getting up from an armless chair without using your hands or any other help?
Q5. Getting up off the floor without any help from lying on your back?
Q6. Standing unsupported for 10 minutes without discomfort?
Q7. Climbing 12-15 steps without using a handrail or walking aid (one foot on each step)?
Q8. Looking over your shoulder without turning your body?
Q9. Doing physically demanding activities (e.g. physiotherapy exercises, gardening or sports)?
Q10. Doing a full day activities whether it be at home or work?

More recently, the ASDAS was developed to also assess disease activity in AS, being the first score to use both self-reported items and objective measures [35]. This score includes patient-reported assessments of back pain (BP - BASDAI Question 2), duration of morning stiffness (DMS - BASDAI Question 6), global assessment of disease activity (PG) and peripheral joint pain/swelling (PP/S - BASDAI Question 3), using visual analog scales from 0 to 10, and an acute phase-reactant. Acute phase reactants are inflammation markers that exhibit significant changes in serum concentration during inflammation [40] and the ones used in ASDAS are either the erythrocyte sedimentation rate (ESR), in mm/h, or the C-reactive protein (CRP), in mg/L, depending on the ASDAS score's version used, being ASDAS-CRP the preferred one. The formulas for the calculation of both ASDAS's versions are presented in Table 2.3.

Table 2.3: Formulas for the calculation of ASDAS, BASDAI and BASFI scores.

Score	Formula
ASDAS-CRP	$0.12 \times \text{BP} + 0.06 \times \text{DMS} + 0.11 \times \text{PG} + 0.07 \times \text{PP/S} + 0.58 \times \ln(\text{CRP}+1)$
ASDAS-ESR	$0.08 \times \text{BP} + 0.07 \times \text{DMS} + 0.11 \times \text{PG} + 0.09 \times \text{PP/S} + 0.29 \times \sqrt{ESR}$
BASDAI	((Q1 + Q2 + Q3 + Q4) + ((Q5 + Q6) / 2)) / 5
BASFI	((Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q7 + Q8 + Q9 + Q10) / 10)

Then, based on the ASDAS score, cutoffs for disease activity states and improvement scores were defined. Four disease activity states were created: inactive disease, low disease activity, high disease activity and very high disease activity, and the correspondent cut-offs 1.3, 2.1 and 3.5, as shown in Figure 2.4 (A). Moreover, cut-offs for assessing improvement were also developed, being a change of at least 1.1 units a clinically important improvement and a change of at least 2.0 units a major improvement, as shown in Figure 2.4 (B) [41].

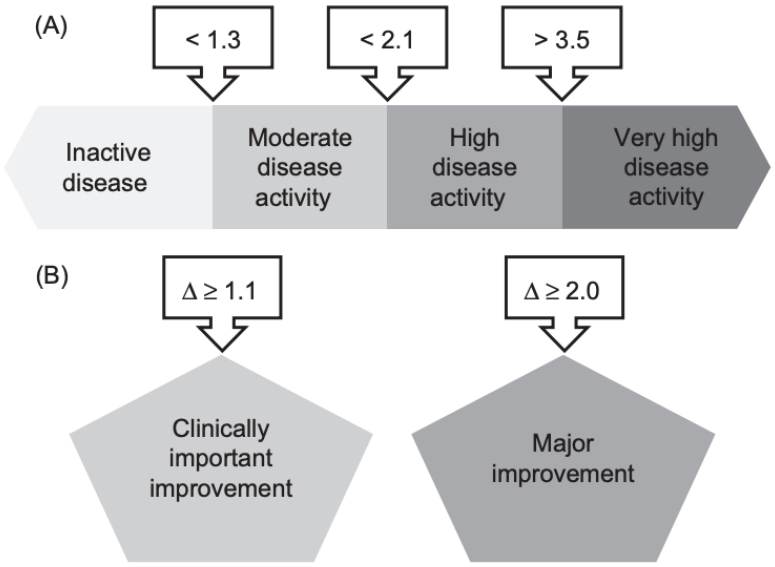


Figure 2.4: Selected cut-offs for (A) disease activity states and (B) improvement scores according to the Ankylosing Spondylitis Disease Activity Score (ASDAS). Taken from [41].

2.2 Bayesian Networks

2.2.1 Definition and basic concepts

A Bayesian Network (BN) is a graphical model for representing relationships among random variables, using conditional probabilities. A BN $\mathbf{B}$ is defined as a triple $\mathbf{B} = (\mathbf{X}, G, \Theta)$ where:

- $\mathbf{X} = \{X_1, ..., X_n\}$ is a n-dimensional finite random vector, where each random variable X_i, with $i \in \{1, ..., n\}$, takes values in the set $\{x_{i1}, ..., x_{\text{ir}_i}\}$, being x_{ik} the k-th value of X_i ;

- $G = (N, E)$ is a directed acyclic graph (DAG), with $\mathbf{N}$ being the set of nodes/vertices in $\mathbf{X}$, representing the random variables and $\mathbf{E}$ is the set of directed edges that represents the dependencies between the variables. If there is an edge from node i to node j, i is said to be a parent of j and j is said to be a child of i;

- $\Theta = \{\theta_{\text{ijk}}\}$ is a set of parameters specifying the conditional probability distribution of each variable according to the structure of the graph G, being

$$\theta_{\text{ijk}} = P_B(X_i = x_{\text{ik}} | \mathbf{pa}(X_i) = w_{\text{ij}}), \tag{2.1}$$

  where $\mathbf{pa}(Xi)$ denotes the set of parents of X_i in G, and w_{ij} is the j-th configuration of the set of parents $\mathbf{pa}(Xi)$ [42, 43].

A BN defines a unique joint probability distribution given by the product, over all the nodes of the graph, of a conditional distribution for each node conditioned on its parents. Therefore, for a BN with n nodes, the joint distribution is given by [17]:

$$P_B(X_1, ..., X_n) = \prod_{i=1}^{n} P_B(X_i | \mathbf{pa}(X_i)), \tag{2.2}$$

meaning that given its parents, each node is conditionally independent of all other non descendant nodes [43, 44]. One argument against the use of probability theory is that specifying completely a probability distribution requires lots of numbers. However, due to built-in independence assumptions present in Bayesian Networks, regarding conditional independence, it is possible to completely specify a probability distribution using less values, and, thus, simplifying the structure of the model. [45, 46]

Definition (Conditional Independence). In probability theory, two random events A and B are conditionally independent given a third event C, denoted by (A $\perp\!\!\!\perp$ B) | C, if and only if, given knowledge that C occurs, knowledge of whether A occurs provides no information on the likelihood of B occurring, and knowledge of whether B occurs provides no information on the likelihood of A occurring.

To ascertain whether a particular conditional independence statement (A $\perp\!\!\!\perp$ B) | C is implied by a given DAG, it is necessary to resort to the notion of d-separation. Considering A, B and C three disjoint subsets of nodes in a DAG, A is said to be d-separated from B by C if all paths from any node in A to any node in B are blocked by C, which happens due to two possible situations:

1. the connection is serial ($A \rightarrow \mathbf{C} \rightarrow B$) or diverging ($A \leftarrow \mathbf{C} \rightarrow B$) and the state of C is known
2. the connection is converging and neither C nor any of its descendants are known ($A \rightarrow C \leftarrow B$)

This is, therefore, the reason why only the parents of a node need to be in the conditioning portion of each term in the factorization of the joint distribution.

Since represented by a DAG, it should also be noted that there are no directed cycles in the structure of a BN, which means that there is no path that starts at a node, moves from node to node following the direction of the arrows and ends up back at the starting node [46].

An example of a BN is presented in Figure 2.5. This example was taken from [47] and describes a situation regarding compensation and assistance given by a airline company for passengers with delayed

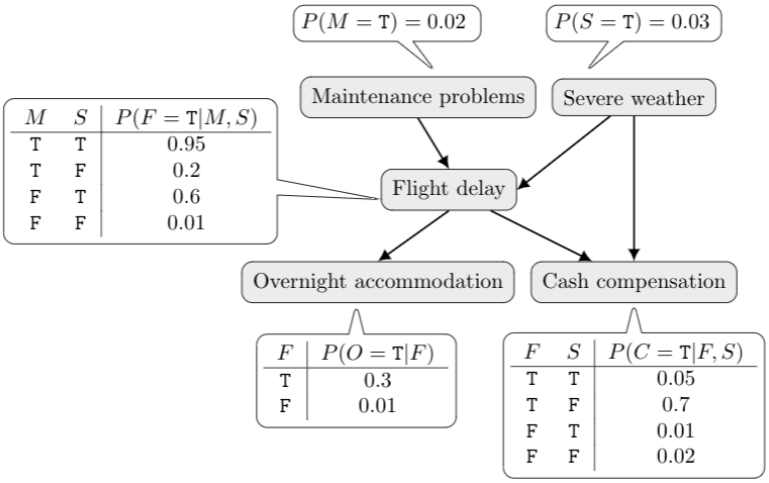


Figure 2.5: A BN example regarding airline regulations with conditional probability tables. Taken from [47].

flights. In this case, there are two reasons for flight delay: maintenance problems or severe weather, like a hurricane or a blizzard. If the delay is not caused by an external event to the company, the passenger may receive a cash compensation. If the delay is long enough the passenger can even receive an overnight accommodation, regardless of the cause of the delay. These relations between variables lead to the dependencies encoded by the graph, being the joint probability distribution of the network given by:

$$P(M, S, F, O, C) = P(M)P(S)P(F|M, S)P(O|F)P(C|F, S), \tag{2.3}$$

where each variable is Boolean and represented by the first letter of its name: M—Maintenance problems, S—Severe weather, F—Flight delay, O—Overnight accommodation and C—Cash compensation.

2.2.2 Learning Bayesian networks

The problem of learning a Bayesian Network given a dataset of observations of a set of variables, consists in finding the network that is most probable to have generated the data, this is, learning its structure and parameters [48].

Learning the structure of the network consists in finding the DAG that may have generated the data. Various algorithms exist for this purpose, nevertheless a complete search is only feasible for networks with a small number of nodes [49]. On the other hand, for larger networks, heuristic methods exist. Most of the methods proposed for learning the structure of a BN are either constraint based or search and score, or even a hybrid algorithm combining the previous two. On the one hand, constraint based methods check conditional independence relations among all variables with statistical tests, linking nodes not found to be independent. On the other hand, algorithms based in the search and score paradigm, have two components, a scoring function and a search procedure, in which the search for a structure is guided by a scoring function that computes a score that reflects how good the structure fits to the data. The goal is, therefore, to search for a structure that maximizes the scoring function. Finally, hybrid

methods combine the two mentioned approaches, narrowing the space of candidate DAGs first, using a constraint based method, and then using a score-based method to find the optimal DAG structure in the restricted space [50].

On the other hand, parameter learning consists in discovering the conditional probabilities that represent the data, given the structure of the network previously learnt [49].

2.2.3 Dynamic Bayesian networks

Dynamic Bayesian networks (DBNs) are an extension of Bayesian networks to represent the evolution in time of a system. In fact, DBNs consist of a series of time slices that represent the state of all variables at a certain time t, being allowed not only intra-slice connections, that are the connections within the time slices, but also inter-slice connections, that are connections between variables from different time slices. These inter-slice connections follow the direction of time, meaning that a variable cannot have descendants in previous time slices [51, 52].

An example of a dynamic Bayesian network, taken from [53], is presented in Figure 2.6. This example represents a simplified situation of monitoring a vehicle with a possibly faulty sensor. Each variable in the model is encoded by its simplified name, where "Location" is the car's current location, "Velocity" is the car's current velocity, "Weather" is the current weather, "Failure" is the failure status of the sensor and, finally, "Obs" is the current observation. In this model, it is possible to note the dependencies between variables in different time points: for example, the car's location depends on the previous location and velocity.

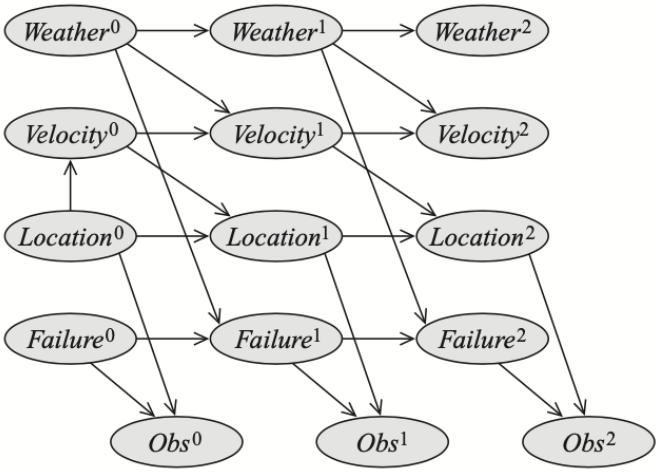


Figure 2.6: Highly simplified DBN for monitoring a vehicle. Taken from [53].

2.2.4 Applications of Bayesian networks

Nowadays, BNs are a well-established tool used in numerous fields for modeling intelligent systems in domains in which is difficult to model all the different conditions that can underlie a finite set of observations. In fact, BNs are highly accepted since they provide an easy-to-understand graphical representation of the dependencies between variables, guaranteeing mathematical consistency and correctness [54].

A common application of Bayesian Networks is as a decision support system in healthcare. One example of this is medical diagnosis, in which they allow the reproduction of the doctor's diagnostic process, making possible to reach a conclusion regarding the pathological situation of the patient, based on the set of symptoms and signals, or other test results. Additionally, computational biology is another area that makes great use of Bayesian networks, through various applications including gene expression analysis and inferring cellular networks and pathway modelling [55].

On the other hand, DBNs also have several applications in modelling processes that evolve over time. One common use in healthcare is to make predictions regarding different variables of interest such as patient survival, disease progression, effect of treatments and development of complications [6]. In fact, DBNs are being used as prognostic models, offering an approach that allows for the incorporation of the causal and temporal nature of medical domain knowledge, thus allowing great prognostic predictions. In this context, some examples of applications are the construction of a DBN for prognosis of patients that suffer from low-grade midgut carcinoids by van Gerven et al. (2008) [56], for diagnosing ventilator-associated pneumonia in ICU patients by Charitos et al. (2009) [57] and for predicting the occurrence of exacerbation events in patients with chronic obstructive pulmonary disease (COPD) by van der Heijden et al. (2013) [58].

Chapter 3

Materials and Methods

3.1 Reuma.pt database

The object of study of this dissertation is the Rheumatic Diseases Portuguese Register (RNDR), i.e. the Reuma.pt database, more precisely the data from patients with AS who undergo or have undergone biologic therapy. The data analysed was retrieved on July 22nd 2019, complying with all ethical and data protection issues. This online platform was developed by the Portuguese Society of Rheumatology (SPR) and it became active in June 2008 due to a need of improving monitoring in patients with rheumatic diseases, including not only AS, but also rheumatoid arthritis (RA), psoriatic arthritis (PsA) and juvenile idiopathic arthritis (JIA). Reuma.pt allows better data production and evaluation, leading to a better insight about these diseases and therefore allowing the development of better therapies. Its main goal is to eventually register all patients with rheumatic diseases in Portugal that are being treated with biological therapies, in order to determine treatment's efficacy and safety. Reuma.pt also includes patients receiving synthetic DMARDs and other treatment strategies such as non-steroidal anti-inflammatory drugs in AS [59–61]. Currently, the database has more than 150 thousand clinical registries, from more than 12 thousand patients and more than 80 clinical centers [62, 63].

The database contains several information for each patient, including identification data, demographic data, work status, life style habits, anthropomorphic data (such as height, weight and body mass index), comorbidities, previous medical history, past and current therapies, disease activity and functional assessment scores, laboratory measurements and adverse events. Among the data, both static and dynamic variables can be found. On the one hand, static variables are either time-independent covariates, like gender, or data gathered only at the patient's first visit, like smoking status. On the other hand, dynamic variables are measurements that are not constant over the whole study, collected over subsequent patient's visits [60]. Patient's dynamic data was presented as a panel structure, more precisely as an unbalanced panel structure, with the number of appointments registered differing between patients, being the smallest follow-up time of a patient a single appointment, while other patients were followed for several years.

3.2 Data processing

The dataset was presented as an Excel file with different worksheets, each one containing different information. For the present dissertation, there were 3 relevant ones: one with general information on the patients, a second one with patients' past and current therapies and, finally, a third one with the registered appointments, and corresponding measurements.

To identify inconsistencies in the data, a detailed analysis of the dataset was performed. Several issues were spotted and corrected, as described below, being all the processing steps implemented in Python.

3.2.1 Therapies' data

The processing of the dataset containing information on the patient's therapies appears in line with the work carried out by Barata (2020) [64] on the same database, which goal was to identify predictors of response for spondyloarthritis. To this end, it was necessary to analyse and process the Reuma.pt database, with particular focus on the data regarding patient's therapies. Indeed, Barata's work allowed the identification of several inconsistencies in the dataset, and their subsequent correction, as described in the following paragraphs.

First, for some patients, the start date of a biologic was reported to be equal to or later than the corresponding end date, as presented in Table 3.1. In these cases, it was not possible to understand what would be the correct start and end dates, so these patients were removed from the database.

Table 3.1: Example of data before processing for different patients where the start date of one biologic was equal to/ later than the end date.

Patient ID	Biologic	Start date	End date
36****	Adalimumab	2013-01-31	2013-01-31
39*****	Adalimumab	2009-08-28	2009-06-30

Second, it was noticed that there were cases for two consecutive biological therapies of the same patient where the end date of the first was reported to be after the start date of the next, as presented in Table 3.2, which constitutes an inconsistency since a patient cannot be taking two biologics at the same time. To deal with this issue, the start date of the second biologic was assumed to be correct and the end date of the first biologic was changed to one day before the start date of the second biologic, if the biologics were the same, and to one month before, if the biologics were different. An example of this situation can be seen in Table 3.3, where the outcome of processing data in Table 3.2 is presented. This difference of one month for different biologics is called the washout period, which is defined as the period between stopping one biologic and initiating a new one in which the patient is not taking any biologic, and was established under the assistance of medical experts.

Next, after performing these corrections on the end dates of the patients' therapies, the interval between the different biologics was evaluated. In cases where consecutive records had the same registered biologic, if the pause between the two therapy sequences was less than a year, these registries were grouped into only one registry, assuming the patient was taking the biologic without any interruption during that

Table 3.2: Example of data for one single patient where the end date of one biologic is later than the start date of next biologic.

Patient ID	Biologic	Start date	End date
39*****	Infliximab	2006-10-24	2010-06-30
39*****	Adalimumab	2010-03-19	2015-09-06

Table 3.3: Outcome achieved after processing data in Table 3.2.

Patient ID	Biologic	Start date	End date
39*****	Infliximab	2006-10-24	2010-02-16
39*****	Adalimumab	2010-03-19	2015-09-06

period. This is shown in Tables 3.4 and 3.5, which represent respectively the original data and the data after processing.

Table 3.4: Example of data where the same biologic is presented as two different therapies, with a pause smaller than a year.

Patient ID	Biologic	Start date	End date
36****	Etanercept	2008-05-26	2009-08-17
36****	Etanercept	2010-03-29	NaN

Table 3.5: Outcome after processing data in Table 3.4.

Patient ID	Biologic	Start date	End date
36****	Etanercept	2008-05-26	NaN

3.2.2 Appointments' data

The second dataset studied was the one containing information on the patients' appointments, where each row in the data consisted in a register for a single appointment, where all the clinical variables measured, together with the corresponding date of the appointment, were recorded. For this dataset, contrarily to the dataset with therapies' information, no processing had yet been done on this data.

First, it was noticed the existence of single appointments with multiple registries. Furthermore, in some cases, the same variable presented different values in the different registries for the same appointment, as shown in Table 3.6. To deal with this issue, registries for the same appointment were grouped into only one. In cases where only one value was filled for a certain variable, or the existing values were all the same, the corresponding value was the one considered for that variable. On the other hand, if different values were introduced for the same variable, it was kept the one in the last registry inserted. An example of this issue is presented in Tables 3.6 and 3.7, where the original data and the resultant outcome after data processing are shown, respectively.

Table 3.6: Example of data where one appointment has more than one registry, with different values for the same variable.

Patient ID	Appointment's date	ASDAS	SF36 1
39*****	2014-01-06	2.7	2
39*****	2014-01-06	nan	3.4
39*****	2014-01-06	nan	nan

Table 3.7: Outcome after processing data in Table 3.6.

Patient ID	Appointment's date	ASDAS	SF36 1
39*****	2014-01-06	2.7	3.4

3.2.3 Patients' data

Finally, regarding patients' general information, the processing steps applied were more targeted to individual situations.

On the one hand, some patients were removed from the data due to inconsistent records, being certain variables fulfilled with invalid information, as exemplified in Table 3.8.

Table 3.8: Example of incorrect data of a patient.

Patient ID	Age at diagnosis	Age at beginning of disease	Age at first biologic
42******	2002-06-03	1994-06-30	Yes

On the other hand, as shown in Table 3.9, for some patients the height was presented in meters instead of in centimeters, as it should, which resulted in an incorrect value for the body mass index (BMI). Both these values were corrected in the database.

Table 3.9: Example of inconsistent data of a patient.

Patient ID	Weight (kg)	Height (cm)	BMI (kg/m2)
66*****	56	1.5	248888.89

Table 3.10: Outcome after processing data in Table3.9.

Patient ID	Weight (kg)	Height (cm)	BMI (kg/m2)
66*****	56	150	24.89

Moreover, regarding the BMI, by crossing information between different tabs it was possible to fill previously empty records. Indeed, it was noted that in the tab with information on the patients, some of them did not have information about this variable, nevertheless in the tab with the patient's appointments some of these patients had the variables "height" and "weight" filled, which are enough to calculate the BMI of the patient. This value was, therefore, calculated for these patients, also contributing to a decrease of the amount of missing values, which was one of the greatest issues with the datase. Indeed, this is

a very common problem when dealing with clinical data, with missing values even outnumbering the available ones for certain variables.

3.3 From data to Bayesian networks

To study AS, and the patient's different response to therapy, different DBNs were learn from data using an R package called `bnstruct`, which uses state-of-the art algorithms to learn the network that may have generated the data, particularly its structure and its parameters [48, 51]. To achieve these models, several steps had to be followed, from variable selection, to datasets preparation and creation, which are described in detail below. Before moving forward, it should be noted that only the first biological agent received by each patient was considered in this study, so, more precisely, this study evaluates the patient's response for the first biological therapy received, which is known to be different from the response of the subsequent biologics.

3.3.1 Variable selection and discretization

First, it was necessary to select the variables that were going to be used in the models. This process was done based on expert knowledge, after discussion with medical professionals who pointed out relevant variables, and also based on the amount of missing data for each variable.

The selected variables are presented in Tables 3.11, 3.12 and 3.13, respectively static, dynamic and outcome variables, along with the corresponding discretization. Indeed, among the chosen variables, there are categorical and quantitative ones, both continuous and discrete, nevertheless, `bnstruct` only accepts discretized variables, so the discretizaton process had to be performed. The package allows the possibility of discretizing in place, by defining the number of levels that the variable can have, nevertheless, it does not allow the definition of thresholds for the different levels. In fact, the in place discretization available is based only on quantiles, which was not desired for all the variables, so the discretization was made prior to inserting data into `bnstruct`, using Python.

Table 3.11: Static variables used in the study, together with the corresponding description and discretization performed.

Variable	Description	Discretization (label: corresponding elements)
Static variables		
Gender	Self-explanatory	0: Male ; 1: Female
Age at biologic onset	Self-explanatory	0:]0,36] ; 1:]36,47] ; 2:]47,+∞[
BMI	Body Mass Index at onset (kg/m2)	0:]0;25[; 1: [25,30[; 2: [30,+∞[
HLA-B27	Whether a patient carries, or not, the HLA-B27 gene	0: Negative ; 1: Positive
Education level	Self-explanatory	0: No education or elementary school; 1: Middle school ; 2: High school ; 3: Higher education
Disease duration	Time in years since the beggining of disease	0: [0,7] ; 1:]7,16] ; 2:]16,+∞[

Table 3.12: Dynamic variables used in the study, together with the corresponding description and discretization performed.

Variable	Description	Discretization (label: corresponding elements)
Dynamic variables		
DMARD	Whether, or not, a patient is taking a DMARD at the time of the appointment	0: No ; 1: Yes
Corticoid	Whether, or not, a patient is taking a corticoid at the time of the appointment	0: No ; 1: Yes
VAS patient	Patient's global self evaluation (on a scale from 0 to 100)	0: [0,20[; 1: [20,50[; 2: [50,100]
VAS doctor	Doctor's evaluation of the patient's global state (on a scale from 0 to 100)	0: [0,10[; 1: [10,30[; 2: [30,100]
BASDAI Q1	Patient's answer to the 1^{st} BASDAI question (on a scale from 1 to 10)	0: [0,4[; 1: [4,7[; 2: [7,10]
BASDAI Q2	Patient's answer to the 2^{nd} BASDAI question (on a scale from 1 to 10)	0: [0,4[; 1: [4,7[; 2: [7,10]
BASDAI Q3	Patient's answer to the 3^{rd} BASDAI question (on a scale from 1 to 10)	0: [0,4[; 1: [4,7[; 2: [7,10]
BASDAI Q4	Patient's answer to the 4^{th} BASDAI question (on a scale from 1 to 10)	0: [0,4[; 1: [4,7[; 2: [7,10]
BASDAI Q5 & Q6	Average of patient's answers to the 5^{th} and 6^{th} BASDAI questions (on a scale from 1 to 10)	0: [0,4[; 1: [4,7[; 2: [7,10]
BASFI	Patient's BASFI index (average of patient's answers to the 10 questions composing it, on a scale from 1 to 10)	0: [0,2] ; 1:]2,5] ; 2:]5,10]
CRP	C-Reactive Protein's level (mg/L)	0: [0,1] ; 1:]1,3] ; 2:]3,8.5] ; 3:]8.5,$+\infty$[

Table 3.13: Outcome variables used in the study, together with the corresponding description and discretization performed.

Variable	Description	Discretization (label: corresponding elements)
Outcome variables		
ASDAS 12M	Composite score to assess disease activity in AS, evaluated at month 12	0: [0,1.3[; 1: [1.3,2.1[; 2: [2.1,3.5] ; 3:]3.5,$+\infty$[
ASDAS improvement 12M	Difference between ASDAS after 12 months of treatment with the biologic and ASDAS at the begging of the treatment (ASDAS 12m - ASDAS 0m)	0: $]-\infty, 1.1[$; 1 : $[1.1, 2[$; 2 : $[2, +\infty[$
Inefficiency	Whether a patient stopped, or not, the therapy due to inefficiency	0: No ; 1: Yes

For categorical variables, like gender and HLA-B27, the discretization process was trivial. For quantitative variables, on the other hand, this process required more effort. On one side, some variables have reference levels based on clinical knowledge, as it is the case for BMI, BASDAI questions, CRP, ASDAS

and ASDAS improvement. However, by discretizing using the existing reference values as thresholds, some categories were very little populated. Therefore, to avoid a great discrepancy in the number of elements of the different categories, some adjustments had to be made in the thresholds considered for discretizing some variables.

Indeed, for BMI, five reference levels exist: below 18.5 is the underweight range, between 18.5 and 24.9 is normal range, between 25 and 29.9 is the overweight range, between 30 and 34.9 is the obese range and finally above 35 is the extremely obese range. Nevertheless, the extreme categories were very little populated, so the underweight category was put together with the normal one, while the extremely obese category was put together with the obese one.

On the other hand, for CRP, according to the theoretical ranges, a value lower than 3 mg/L is considered normal, a value between 3 and 10 mg/L is considered mildly elevated and a value above 10 mg/L represents an active and more significant inflammation [65]. Nevertheless, using this reference ranges originated an assymetric distribution of data between the ranges, so it was opted to use a discretization based on 4 quantiles, which resulted in a distribution with thresholds relatively close to the reference values.

Finally, the remaining variables were discretized based on quantiles, being the selected number of bins equal to 3.

The goal of this work is to understand the impact of different variables in the outcome of interest. For the present study, three outcomes were defined: 1) ASDAS score at 12 months, 2) ASDAS improvement at 12 months and 3) inefficiency. Concerning inefficiency, it should be noted that it is not limited to a certain period of time after the beginning of therapy. Instead, therapy is considered inefficient independently of how much time has passed since therapy initiation, including, therefore, both primary and secondary failure. Indeed, primary failure corresponds to the situation where patients that don't present any positive response to therapy, while secondary failure regards situations where patients initially present some clinical response, but eventually lose it. Regarding ASDAS improvement, it should be noted that this variable is not directly found in the database and therefore needed to be calculated from the existing data.

3.3.2 Data restructuring into bnstruct's required format

To input data into the models, it was required that the different observations, i.e. the dynamic variables acquired in the different appointments, were equally spaced. Nevertheless, in the data, the spacing between appointments differed between patients. Therefore, in order to have records at the same time points for all patients, it was necessary to select those time points and the interval between them. It was opted to consider two different time intervals between appointments, 3 and 6 months. To this end, an algorithm to select the relevant appointments according to the chosen granularity was developed and implemented in Python, being the corresponding steps described below. The time frame considered was restricted to the first 12 months after the beginning of the therapy, due to a trade-off between considering a time frame big enough to notice therapy effects and small enough to avoid the reduced number of records for appointments at more advanced time points.

The first step of the algorithm was to assign for each record a corresponding month. Indeed, when considering an interval of 3 months between appointments, it would be required to have appointments at 0, 3, 6, 9 and 12 months, while for an interval of 6 months, it would be required to have appointments at 0, 6 and 12 months. The first challenge faced was that not all patient's first registered appointment in the appointments' dataset corresponded to the beginning of the therapy: some patients had appointments registered when they were still not taking any biologic, while others only had the first registered appointment only awhile after therapy initiation. Therefore, it was necessary to cross information between the appointments' data and the therapies' data, where the start date of each therapy was registered, and compare the date of each appointment with the start date of the biologic, considered time 0. This was done through the creation of an auxiliary variable named "time since beginning of therapy", which is calculated as the difference in days between the date of the appointment and the start date of the respective biologic. Then, auxiliary variables were created labeling each appointment into the desired month, based on the "time since beginning of therapy", as presented in Table 3.14, by considering an interval with middle value the exact desired time: 91 and 182 days for 3 and 6 months, respectively. Therefore, for assigning appointments into the corresponding month considering the 3-month interval, it was considered a time frame of one and a half months before the 91 days and one and a half months after, while for the 6-month interval 3 months before and after were considered.

Table 3.14: Example of data with auxiliary month variables.

Patient ID	**Appointment's date**	**Time since beginning of therapy**	**Auxiliary month (3 by 3)**	**Auxiliary month (6 by 6)**
36****	2009-07-20	0	0	0
36****	2009-08-03	14	0	0
36****	2009-08-31	42	0	0
36****	2009-10-12	84	3	0
36****	2009-11-23	126	3	6
36****	2010-01-04	168	6	6
36****	2010-02-15	210	6	6

After having these auxiliary variables, it was noted that there were cases where only one appointment was assigned for each month, but there were also cases where more than one appointment was assigned for a particular month. In situations like this, as only one appointment per month should be chosen, the choice was based on the proximity to the desired month, being chosen the closest one. Nevertheless, to aim at reducing the amount of missing values in the data, if a variable in the chosen appointment was not filled, but in another appointment for the same month there was a value for the variable, this value would be considered as a value for that variable for that specific month. The process described consists in the application of two imputation techniques: last observation carried forward (LOCF), where a missing value from a follow-up visit is replaced by the previously observed value for that subject, i.e. the last observation is carried forward, and next observation carried backward (NOCB), which is the reverse of LOCF, where the recovered value is, instead of the previous one, the following one. This situation is

exemplified in Tables 3.15 and 3.16. In fact, in Table 3.15 can be seen two appointments that according to the variable "time since beginning of therapy" correspond to the same month, in this case, month 6. Since 3 months were considered to be 91 days, 6 months correspond to 182 days, which means that the closest appointment is the one on 2010-01-04, registered 168 days after therapy beginning. Nevertheless, it is possible to notice that the ASDAS value in this appointment was not filled, while in the appointment on 2010-02-15 it was. Therefore, this value was recovered (NOCB) and considered as the value of that month, as exemplified in Table 3.16

Table 3.15: Example of data with variables not filled in the desired appointment, but filled in other appointments of the same month.

Patient ID	Appointment's date	Time since beginning of therapy	BASDAI	ASDAS	Auxiliary month (3 by 3)
36****	2010-01-04	168	4.62	NaN	6
36****	2010-02-15	210	3.61	3.3	6

Table 3.16: Outcome of data processing after NOCB.

Patient ID	Appointment's date	Time since beginning of therapy	BASDAI	ASDAS	Auxiliary month (3 by 3)
36****	2010-01-04	168	4.62	3.3	6

On the other hand, if there were no registries for a particular month, an empty row would be added into the data, to meet the need of having the same time points for all patients. This is shown in Tables 3.17 and 3.18, in which a record was added for month 6. This example also summarizes the different processing steps performed.

Table 3.17: Example of data before appointments' selection.

Patient ID	Appointment's date	Time since beginning of therapy	BASDAI	ASDAS	Auxiliary month (3 by 3)
36****	2004-03-22	0	6.96	4.1	0
36****	2004-04-05	14	3.05	2	0
36****	2004-06-14	84	0.67	0.9	3
36****	2004-12-09	262	NaN	NaN	9
36****	2005-01-10	294	0.81	1	9
36****	2005-04-04	378	3.05	2.4	12

Table 3.18: Example of data after appointments' selection.

Patient ID	Appointment's date	Time since beginning of therapy	BASDAI	ASDAS	Auxiliary month (3 by 3)
36****	2004-03-22	0	6.96	4.1	0
36****	2004-06-14	84	0.67	0.9	3
36****	NaN	NaN	NaN	NaN	6
36****	2004-12-09	262	0.81	1	9
36****	2005-04-04	378	3.05	2.4	12

3.3.3 Creation of testing datasets

In order to have more confidence on the results, different datasets representing the same data were created, with the aim of building several different models and being able to compare them. Therefore, for creating these datasets, different parameters were varied: the time interval between appointments, as already mentioned, of 3 or 6 months, the presence or not of imputation on the variables, the maximum percentage of missing data allowed per patient and, finally, the biologic. The steps followed to build the datasets with an interval between appointments of 3 months, along with the number of patients for each dataset, are schematically represented in Figure 3.1. On the other hand, the same schematic representation for the 6-month interval datasets is presented in Figure A.1, in Appendix A.

As mentioned, two different datasets were initially created, one where appointments were selected with a 3-month interval between them, and another one with a 6-month interval. In this step, it was noticed that some patients did not have any records of dynamic variables, either because there were no registered appointments or because the variables were not filled. Besides, some patients presented a considerable amount of missing values for dynamic variables, and so, to set an exclusion criteria, patients with more than 70% of missing dynamic data were removed from the database.

The next step intended to deal with the missing values in the data. In fact, training a model with a dataset that has a considerable amount of data missing can significantly impact the quality of the model and, subsequently, the conclusions that can be drawn. It is, therefore, essential to handle missing values, existing several strategies for this purpose. One of them consists in removing from the data observations with a significant amount of missing values. However, by removing observations, it is possible to be losing valuable information. On the other hand, another method consists in infering missing values from the existing part of the data, i.e. to perform imputation. Nevertheless, when performing imputation, it is important to take into account the percentage of missings for each observation, since performing imputation when there is a lot of missing data may lead to the introduction of errors into the model learnt. This may lead to the need of removing patients with greater amount of missing before performing imputation. It is, therefore, important to find a trade-off between the observations removed and the existing amount of missings.

In this study, this was the strategy followed. Indeed, data imputation was performed and some patients with an increased amount of missing values were removed.

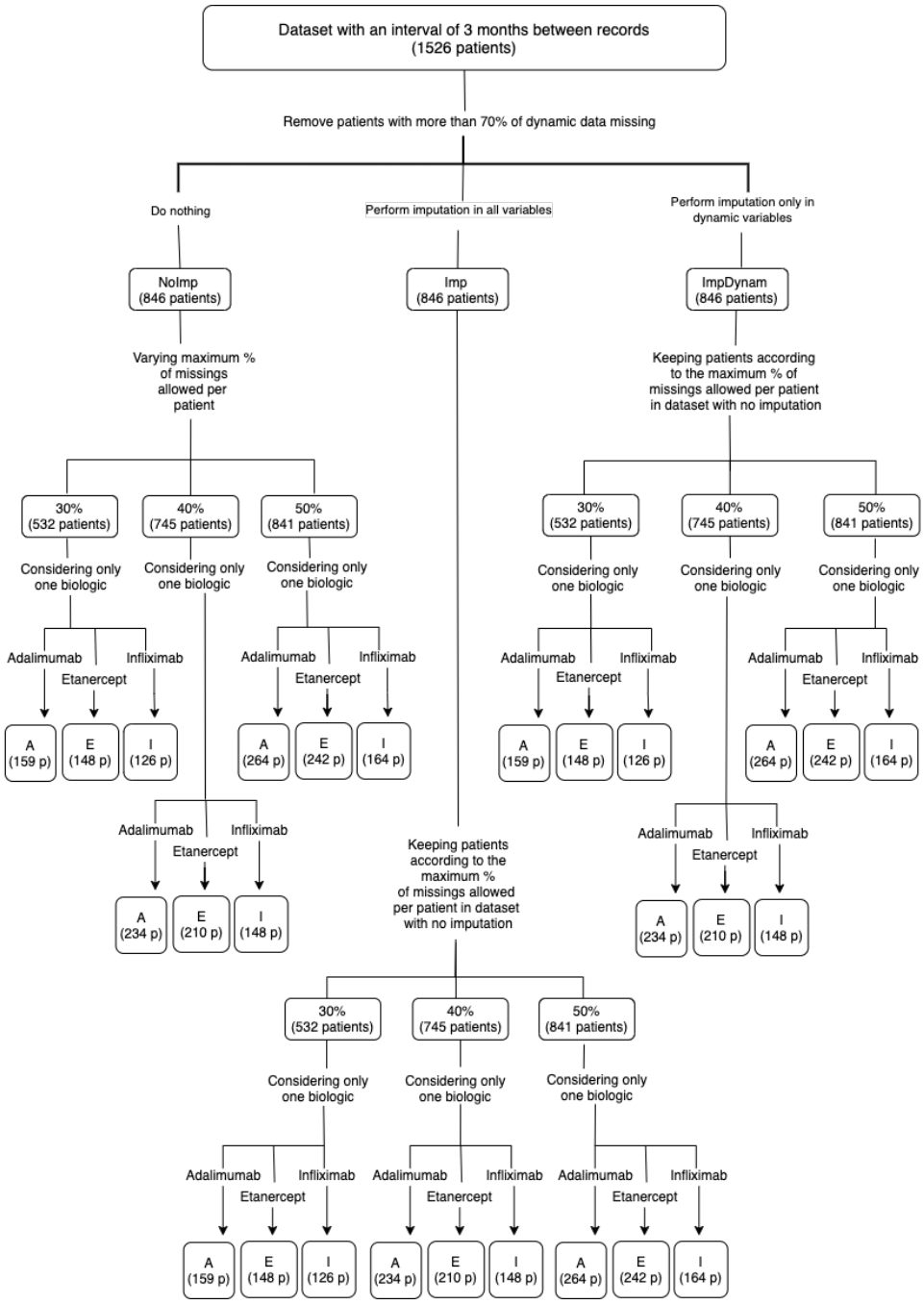


Figure 3.1: Schematic representation of datasets' creation process (interval of 3 months between records).

To perform imputation, it was necessary to have in mind that there were two types of variables in the data, static and dynamic, that needed to be treated differently.

Concerning static data, the main approach used was the k-nearest neighbor (kNN) imputation, which identifies, based on other variables' values, the k points that are closest to a point with a missing value, and estimates the missing value using the known values of the neighboring observations [66, 67]. Indeed, kNN imputation was the prefered technique, with k=5, instead of replacement with the mean, median or mode values, other commonly used techniques, since it allows to keep variability in the data. The only exception for this was education, where mode was prefered. In fact, regarding the patient's education,

there are two variables in the data that translate it, the "academic degree" and the "number of years of education", existing patients with one of them filled and not the other. The information between these two variables was therefore combined into only one variable by first using as reference the variable with less missings, which was the academic degree. Second, the patients with "number of years of education" filled but not "academic degree" were identified. Then, looking into the complete data, it was evaluated which was the most common academic degree for the different years of education, which was used to fill the "academic degree" of patients that only had the "number of years of education" filled. For the remaining values that were still missing after this approach, a replacement with the variables' mode was performed.

On the other hand, regarding dynamic data, the `interpolate` function from Python was used to impute missing data. This function fills NaNs surrounded by valid values by applying linear interpolation, which allows to estimate unknown values between known values, by assuming that there is a straight line between them. Aditionally, the `interpolate` function also fills NaNs outside valid values, with NOCB being used before the first valid value and LOCF after the last.

By performing imputation on the two existing datasets, two more were originated from each one: in one of them imputation was performed on both static and dynamic data and in the other one imputation was performed only on dynamic data.

The next step was to create different datasets from each of the existing ones, varying the maximum amount of missings allowed for each patient, as shown in Figure 3.1.

Finally, further datasets were created keeping data from only one biologic, among the three most common ones: adalimumab, etanercept and infliximab.

3.3.4 Structure and parameter learning

The `bnstruct` R package provides five algorithms to learn the structure of the network: the Silander-Myllymäki (SM), the Max-Min Parent-and-Children (MMPC), the Hill Climbing (HC), the Max-Min Hill-Climbing (MMHC) and the Structural Expectation-Maximization (SEM) [49].

The SM algorithm is an exact search-and-score method, which returns the best network by performing a complete evaluation of the search space. Nevertheless, as expected, it requires a lot of computational effort and it's not feasible for networks with more than 25-30 nodes [2], which is the case of the dataset used for creating the models in this dissertation. Therefore, this algorithm was not used to learn the structure of the network.

The MMPC is a constraint-based method that returns the skeleton of the network, with non-directed edges, therefore not allowing to perform parameter learning, which was one of the goals of this work.

The HC is a greedy search algorithm, which improves networks iteratively through local modifications: edge addition, edge deletion or edge reversal. The search can start either by the complete possibilities' space or by a subset of possible edges.

The MMHC is an heuristic algorithm that combines the MMPC and the HC algorithms. First, the MMPC returns a set of edges connecting the variables and then, the HC algorithm returns the edges' orientation by starting the search in the possibilities' space from the configuration provided by MMPC.

This was the algorithm used to perform the DBN structure learning.

Finally, the SEM algorithm is used to learn a network when there are missing values in the dataset, which are filled using the expectation-maximization algorithm. In fact, the dataset used in this dissertation contains numerous missing values, nevertheless, it was chosen to perform imputation separately, in order to choose the best way to deal with each variable, so this algorithm was also not used to learn the structure of the networks.

Search-and-score methods also require a scoring function to evaluate how good the structure fits the data. Three scoring functions are provided by `bnstruct`: the Bayesian-Dirichlet equivalent uniform (BDeu), the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). Indeed, the use of the MMHC algorithms required the choice of a scoring function in order to find the optimal network structure. For this work, all three available scoring functions were chosen to create the networks, to then perform a comparison of the different structures obtained, which represent the same underlying data.

Additionally, `bnstruct` also allows to define constrains to the network structure, which work as a way of inputing domain knowledge, by defining layers for the variables where a variable can only have parents in lower-numbered layers, in order to avoid relations with no clinical or biological sense. This way, variables were divided into 4 general layers, as presented in Table 3.19, separating static variables from dynamic and outcome ones. Moreover, each temporal layer was also separated in 4 layers, according to layering scheme presented in Table 3.20.

Besides defining the layers, it is also possible to specify the influence of upper lower-numbered layers in lower high-numbered layers, by defining which layers can contain parent nodes for variables in other layers. In the case of the present work, the possible influences between layers were specified as follows. First, static variables were all placed in the same layer, the upper one. These variables were allowed to influence variables in all other layers, but not each other, assuming to be independent. Next, variables at each time point were allowed to influence: 1) each other, based on a layering scheme presented in Table 3.20, 2) variables at the following time point, this is, variables at time t influencing variables at time $t+1$, and 3) the outcome variables. In turn, variables at the last time point were allowed to influence each other and the outcome variables. Finally, outcome variables are in the lowest layer and edges between them are forbidden. Furthermore, regarding layers defined within each temporal layer, the first one is composed of concomitant DMARD and concomitant corticosteroids, which were allowed to influence each other and all other dynamic variables at that time. The second layer is composed of BASDAI questions, BASFI and CRP, which in turn are allowed to influence themselves and the following layers, which are VAS patient and VAS doctor. When layering these two variables, VAS patient was placed above VAS doctor based on causality assumptions, where the doctor's evaluation on the patient was assumed to be influenced by all other variables for the same time point, including patient's self evaluation (VAS patient). All these variables were also allowed to influence variables in the next time point, according to the general layering.

Table 3.19: General layering of variables in the DBN.

Layer	Variables
1	Gender, HLA-B27, BMI, Education level, Age at biologic onset, Disease duration
2	Variables at time $t-1$
3	Variables at time t
4	ASDAS improvement at 12 months, ASDAS at 12 months, Inefficiency

Table 3.20: Layering of variables inside each temporal layer.

Layer	Variables
1	DMARD, Corticoid
2	CRP, BASDAI questions, BASFI
3	VAS patient
4	VAS doctor

Finally, to conclude network learning, parameter learning is done by performing a Maximum-A-Posteriori (MAP) estimate of the parameters.

3.3.5 Representation and visualisation of the networks

The structure of a Bayesian network is represented as an adjacency matrix. An adjacency matrix is a $n \times n$ matrix, with n being the number of nodes, where an entry (i, j) takes the value of 1 if there is an edge from node i to node j, and 0 otherwise.

Several networks were built to represent the same underlying data, allowing to obtain more reliable results. Indeed, by comparing the different networks obtained, it was possible to evaluate the confidence on the edges found, where greater confidence was given to edges that appeared more frequently in the models.

To provide an easier and more intuitive visual summary of data, different adjacency matrices representing the same underlying data were summed, condensing different networks into a single graphical representation. This way, a matrix where each cell's value ranges from 0 up to the number of adjacency matrices summed was created. The resultant matrix was then used to create a heatmap, a two-dimensional data visualization technique where the magnitude of a phenomenon is represented through color, using the `gplots` R package. By varying the interval between appointments, if the dataset was imputated or not and which variables were imputed, the maximum percentage of missings allowed and the biologic, 18 different heatmaps were created, with 9 different networks each. Indeed, for each heatmap, 3 datasets with different maximum percentages of missings allowed per patient were used, being these values of 30%, 40% and 50% for the datasets with a 3-month interval between records and of 20%, 30% and 40% for datasets with an interval of 6-months between records. Furthermore, each of the previous datasets resulted in 3 different networks, by varying the scoring function used: AIC, BIC and BDeu.

Chapter 4

Results and Discussion

This chapter starts with a description of the dataset studied in this work. Next, results achieved with the creation of DBNs are presented and discussed. Finally, the outcomes of the Chi-square tests performed to study further variable's associations are also presented and discussed.

As previously described, three outcome variables were defined, therapy inefficiency, ASDAS score 12 months after therapy initiation and change in ASDAS at the same time, being the main focus of this work to evaluate which variables relate to these outcome ones. Particularly, since the main goal of this study is to identify associations that help predict therapy outcome, the most interesting ones would be between outcome variables and variables that can be measured before therapy initiation, i.e., static and baseline variables. Nevertheless, if these associations are not found in the networks, it can also be of interest to identify correlations with variables shortly after therapy initiation, to avoid extending therapies that will most likely end up being inefficient.

4.1 Description of the dataset studied

This section intends to describe the dataset used in this study. As mentioned, the data is from the Reuma.pt database concerning AS patients, retrieved on July 22nd 2019, and has undergone several processing steps until the desired format was obtained.

First, in the original data, 2293 patients were retrieved from the database. However, since the aim of this work is to study patients under biological therapies, after removing patients with only non biological therapies registered, the number of patients decreased to 1732. After further processing, other patients ended up being deleted, as mentioned in Section 3.2, resulting in a dataset with 1526 patients.

Then, according to the required data structure for input into the models, records of dynamic variables should be equally spaced, and so, two-different datasets were created, one with an interval of 3 months between records and another with an interval of 6 months. When doing this process, patients with more than 70% of dynamic data missing were removed from the database, resulting in two datasets that after further processing will give rise to the ones used in the models. Some descriptive statistics of these two datasets are presented in Tables B.1 to B.4, in Appendix B.

The first difference to be noted between the two datasets is the number of individuals in each one of them, 846 and 1116 for datasets with an interval of 3 and 6 months between records, respectively.

Indeed, by broadening the time period accepted for record selection, which was the case for the 6-month interval dataset, like explained in Section 3.3.2, it is expected that the number of missing values is smaller than when using a stricter interval, as it was the case for the dataset with the 3-month interval between records. Therefore, when removing patients with more than 70% of missing dynamic data, as the 3-month interval dataset has more missing values, it is expected that more patients will be removed.

Next, looking into the number of patients receiving each therapy, which motivated the choice of the biologics studied in this work, it is noticed that the most common biologic used as first option is adalimumab, with 266 patients in the 3-month interval dataset and 354 patients in the 6-month one. Adalimumab is followed by etanercept, infliximab, golimumab, certolizumab and secucinumab, with 243/300, 165/219, 142/203, 22/27 and 8/13 patients in the 3-month/6-month interval dataset, respectively. Indeed, the three most common biologics were the ones used in this study, for being the ones with more available data.

Additionally, it is important to note the amount of missing data presented by each variable. Moreover, all variables were categorized, as required for building the DBN models, being also important to know the distribution of the data into the different categories.

Concerning static variables, and starting with gender, it can be noted that it is one of the few variables with no missing data. In terms of patient's distribution into the different genders, it is noted that more than half of the patients are male (55-60%), which is in line with the literature, reporting men to be more affected than women. On the other hand, regarding HLA-B27, there is no information on this gene for around a quarter of the patients present in the study. However, for the patients with available information, around three quarters have this gene, which is also in line with the literature, reporting HLA-B27 to be found in 90-95% of AS patients, although with a less accentuated predominance of HLA-B27 positive patients. Regarding BMI, only around half of the patients present the respective BMI value, corresponding to a considerable amount of missing data. Nevertheless, it was opted to keep this variable in the study, due to its relevance. The categorization of this variable was done according to reference values in three categories, firstly underweight and normal weight patients, secondly over weight patients and, thirdly, obese patients. Concerning age at biologic onset and disease duration, both these variables were discretized based on quantiles. Regarding missing data, the first one has no missing values, while the latter has around 10%. Finally, regarding education, this variable has around 30% missing values, being the discretization based on existing education levels.

On the other hand, regarding dynamic variables there are two main patterns that should be noted. First, regarding data availability, it is possible to note that the amount of missing values for each variable increases with time. This may be due to not having some variable's values filled for certain appointments, but also to a decrease in the number of registered appointments through time. Indeed, this was the main reason for choosing an evaluation period of 12 months after therapy initiation, to find a trade-off between the available data and obtaining relevant results after sufficient time has passed after initiating therapy. Furthermore, regarding data distribution into the different categories, it is possible to note a general shift of data distribution from higher values to lower values as time passes, i.e., more patients present lower values for variables that reflect disease activity. These findings translate the positive effect that

therapy presents on patients, which can be seen already after 3 months of therapy initiation. Information regarding missing values and distribution of these data into the different categories can also be found in the mentioned tables. Additionally, in Figures 4.1 and 4.2 it is possible to see the evolution of the average values of the time dependent variables used in the study, also allowing to verify the pattern described, being noticed a clear decrease in the average value of each of the variables, after 3 months of therapy.

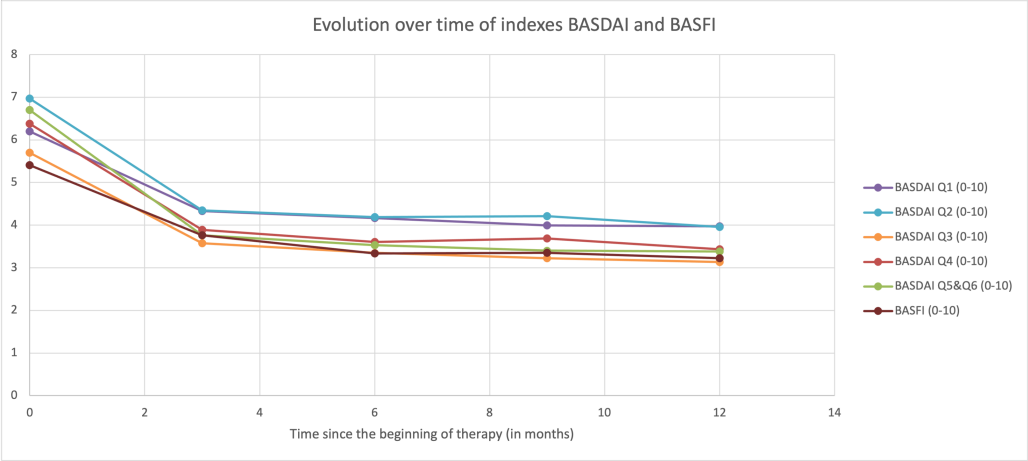


Figure 4.1: Evolution over time of indexes BASDAI and BASFI.

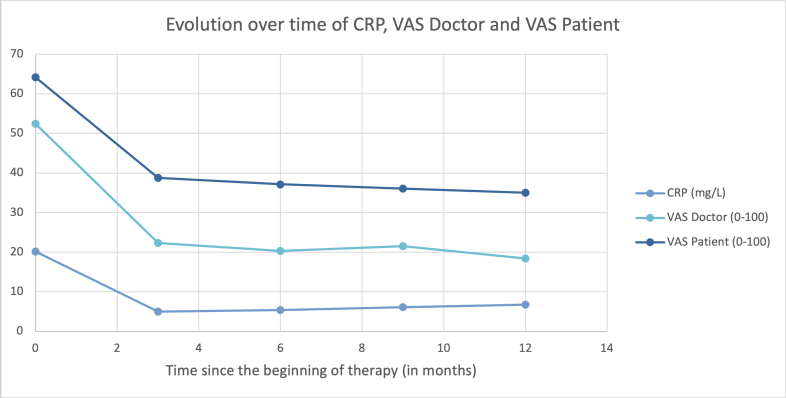


Figure 4.2: Evolution over time of CRP, VAS Doctor and VAS Patient.

Finally, regarding the outcome variables, inefficiency is the only one with no missing data and it has two possible outcomes, *yes* or *no*, depending if the therapy was inefficient, or not, respectively. It is possible to note that, for the great majority of the patients, therapy was not stopped due to inefficiency, with only around a fifth of the patients stopping therapy due to inefficiency. Next, both ASDAS at 12 months and ASDAS change were discretized based on reference values, previously presented. Regarding the activity disease levels translated by each ASDAS category, presented in Figure 2.4 of Section 2.1.4, it is possible to observe that less than 10% of patients with available data have a very high disease activity after 12 months of receiving a biological therapy. For the other 2 levels of disease activity, patients are relatively equally distributed between them. Finally, ASDAS improvement is the outcome variable with the greatest amount of missing values, being the patients' distribution over the different categories relatively equal.

4.2 DBN models

As previously mentioned, heatmaps, which represent the magnitude of a phenomenon through color, were the chosen technique to visualize the obtained DBNs. In the present case, the phenomenon is the appearance of a certain edge in the networks, whose direction should be read from rows to columns, i.e., the row variable is a parent of the column variable.

Several heatmaps were built, with a single heatmap translating the information of several networks. This way, before presenting the obtained heatmaps, it is essential to clarify the differences between each of them, which can be understood with the help of the schematic representation presented in Figure 4.3, where each heatmap is represented by a purple rectangle. Indeed, 18 heatmaps were built varying the interval between appointments of the datasets used, the biologic, and the imputation performed.

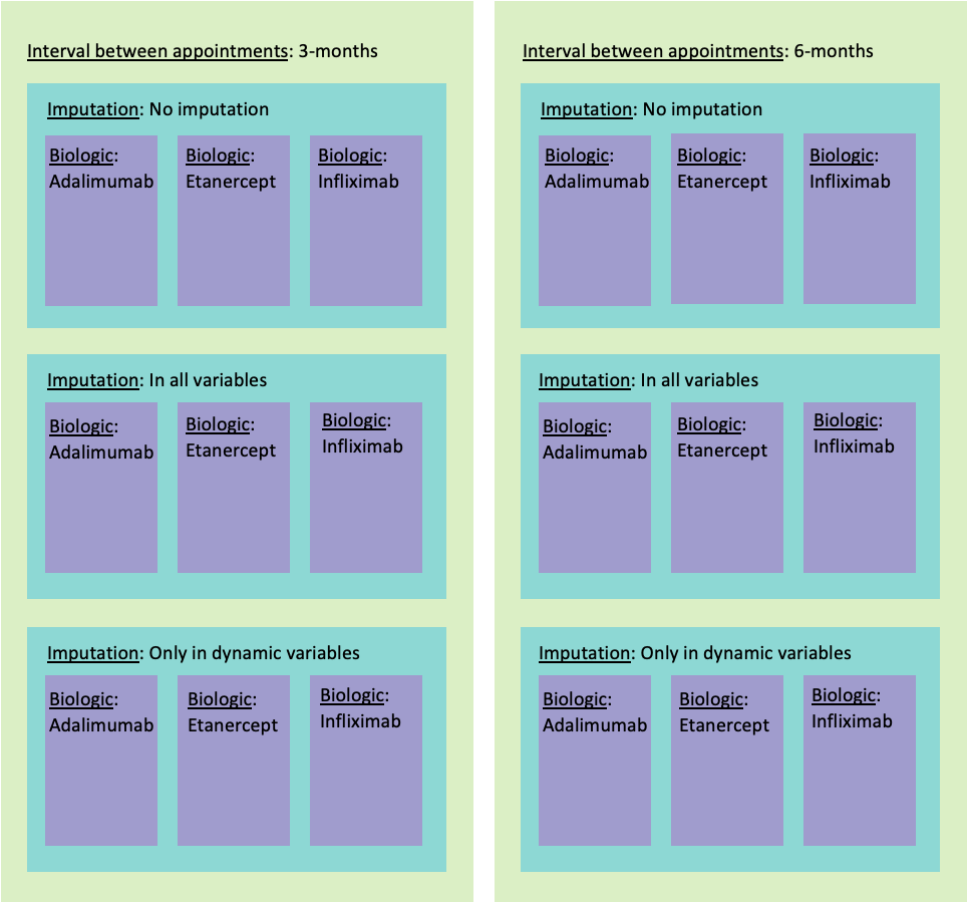


Figure 4.3: Schematic representation of the different heatmaps obtained.

Furthermore, a single heatmap translates the information of several networks, as schematized in Figure 4.4. As can be seen, one heatmaps results from varying in each situation the maximum percentage of missing data allowed per patient(among 3 possible ones) and the scoring function used (also among 3 existing ones), resulting in nine networks per heatmap, as already mentioned. The structure of a single network can be seen in Figure 4.5. Due to the numerous amount of variables, it becomes less practical to present and consequently analyse the results in the format of a network, and so, as previously mentioned, heatmaps were the representation method chosen to overcome this disadvantage.

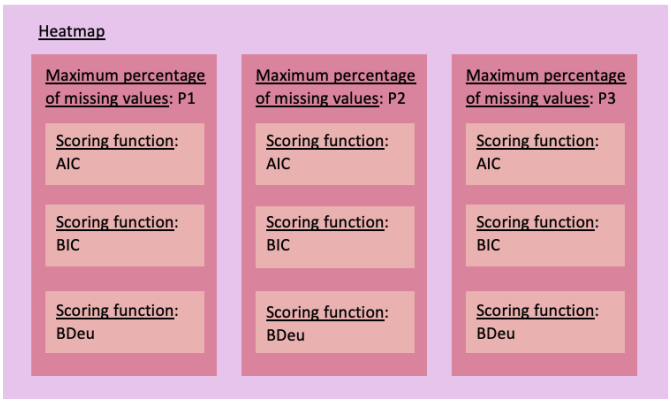


Figure 4.4: Schematic representation of the networks used to build a single heatmap.

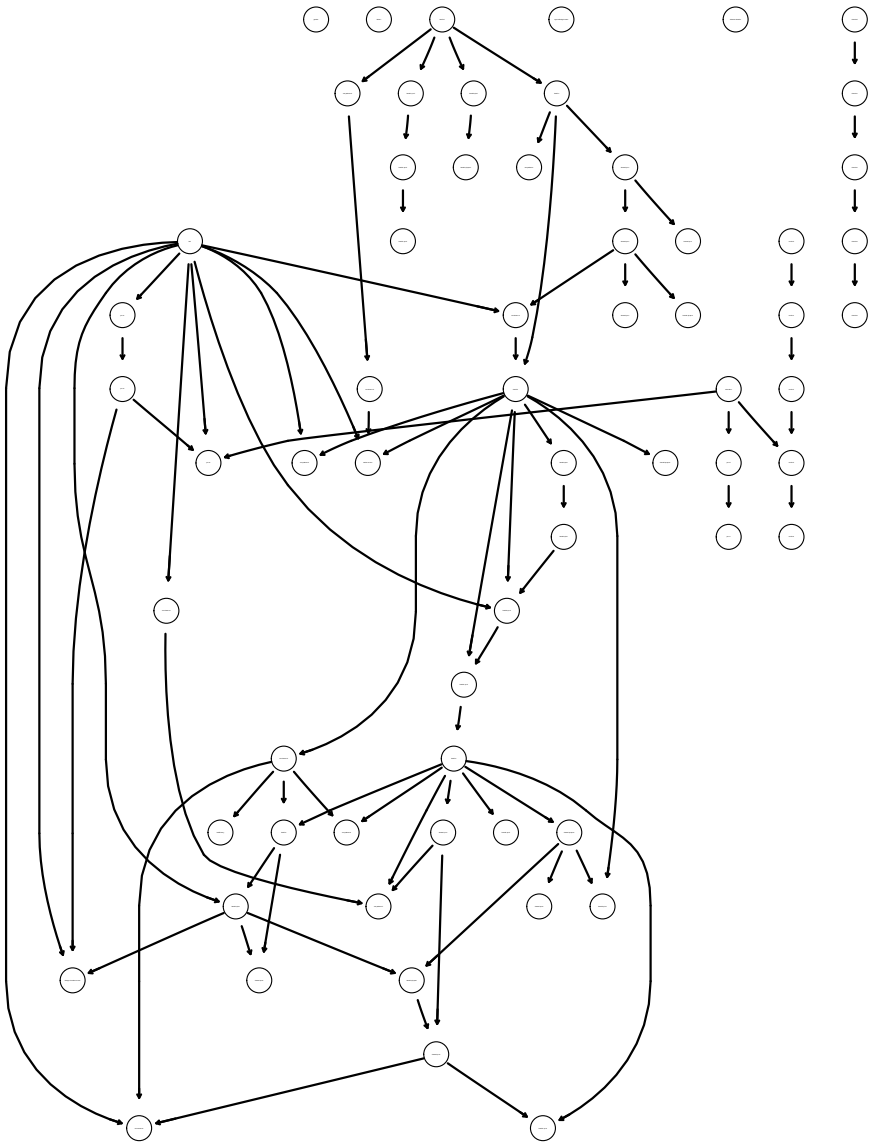


Figure 4.5: DBN obtained for adalimumab using datasets with a 3-month interval between records, no imputation, 30% maximum allowed percentage of missing data and BDeu scoring function.

This way, the different colors in the heatmap translate the number of times that edge appeared in the different networks used to build the heatmap. Therefore, it is essential to know the color scale used in the plots, which was the same for all heatmaps and which is presented in Figure 4.6. On one side of the scale, light yellow represents absence of edges connecting the variables, followed by light green, light blue and dark blue, representing an increasing number of times of appearance of a particular edge in the networks. Finally, on the other end of the scale, black means that a certain edge appeared in all the built networks, i.e., in all the nine networks that constitute one single heatmap, obtained by combining the 3 maximum percentages of missing data allowed with the 3 scoring functions used. This way, the magnitude of appearance in all networks can be seen as a measure of confidence in that edge, with greater times of appearance meaning greater confidence.

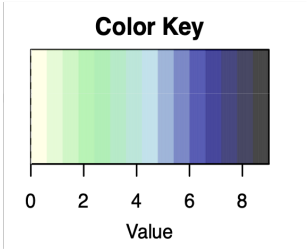


Figure 4.6: Color key used in all the heatmaps.

In the following subsections, each obtained heatmap will be analyzed, and comparisons between different heatmaps will be performed. It should be noted that the present analysis will be conducted via a manual process, by looking into the conditional probability tables (CPTs) of the different networks and trying to understand how each variable that is a parent of an outcome variable influences its value. Additionally, it should be noted that variables can have more than one parent in each network, as expected, since in real life no disease outcome depends only on one parameter, being patient's response a combination of multiple factors. To simplify the present analysis, and the obtained networks, the maximum number of parents was limited to three. As mentioned, in BNs, each variable's value is totally defined by its parents' values and therefore, the greater the number of parents, the more complex it is to evaluate the impact of a single variable in the outcome variable by looking into the CPTs, since the number of different parent combinations increases.

Given the large number of networks built, 162 in total, and given that, for each network, each node has a corresponding CPT, and having in mind that the 3-month and 6-month networks have, respectively, 64 and 42 nodes, it would be both unfeasible and impractical to present all the CPTs. This way, it was opted to present only a few of them, to exemplify the conclusions stated. Moreover, considering the mentioned increase in complexity when evaluating the CPTs given the increase in the number of parents for each node, when more than one CPT allowed to reach the same conclusion, when possible, it was opted to present the tables for nodes with only one parent, where the parent is the relevant variable.

4.2.1 Datasets with no imputation with a 3-month interval between records

First, let's look into the results achieved using datasets for the three biological therapies where no imputation was performed and considering an interval between records of three months.

Starting with adalimumab, whose corresponding heatmap can be seen in Figure 4.7, it is possible to observe that both static and dynamic variables correlate with the outcome variables.

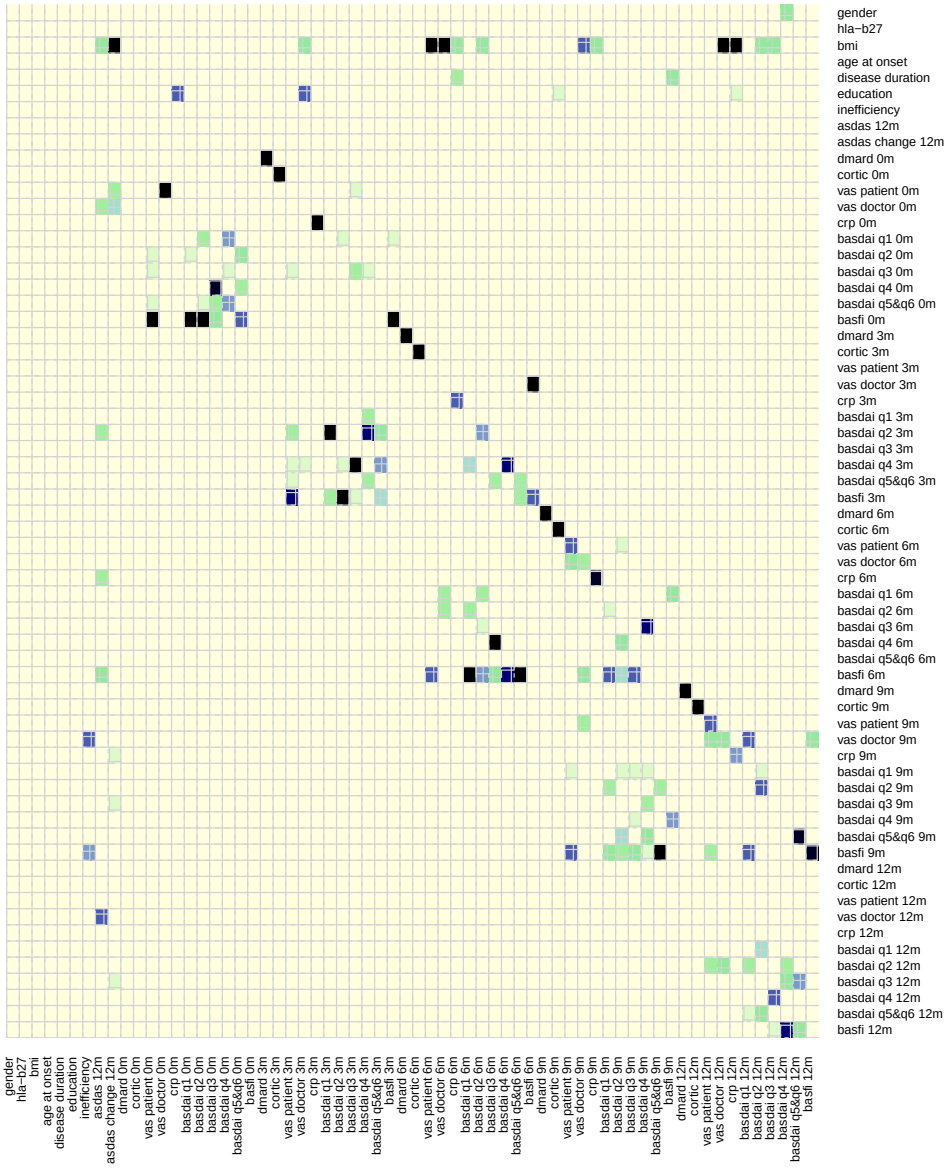


Figure 4.7: Heatmap obtained for adalimumab using datasets with a 3-month interval between records and no imputation.

First, for therapy inefficiency, it is noted that it correlates with two dynamic variables recorded 9 months after therapy beginning, doctor's global evaluation of the patient (VAS doctor) and BASFI score.

Table 4.1: Probability distribution of inefficiency conditioned on VAS doctor at 9 months, in one of adalimumab's networks.

	Inefficiency	
	0	1
VAS doctor at 9 months		
0	0.96	0.04
1	0.95	0.05
2	0.45	0.55

Table 4.2: Probability distribution of inefficiency conditioned on BASFI at 9 months, in one of the networks.

	Inefficiency	
	0	1
BASFI at 9 months		
0	0.96	0.04
1	0.78	0.22
2	0.66	0.34

By looking into the CPTs presented in Tables 4.1 and 4.2, it is possible to conclude that for both these variables, a higher value is associated with a higher probability of therapy failure due to inefficiency, when compared to lower values for these variables. Moreover, if both variables present increased values at the same time, the most likely outcome is that therapy will fail, as can be seen in Table 4.3.

Table 4.3: Probability distribution of inefficiency conditioned on VAS doctor at 9 months and BASFI at 9 months, in one of adalimumab's networks.

	Inefficiency	
	0	1
BASFI at 9 months|VAS Doctor at 9 months		
0 | 0	0.97	0.03
0 | 1	0.99	0.01
0 | 2	0.5	0.5
1 | 0	0.91	0.09
1 | 1	0.99	0.01
1 | 2	0.83	0.17
2 | 0	0.97	0.03
2 | 1	0.99	0.01
2 | 2	0.01	0.99

In this case, it can be hypothesised that, if after 9 months of therapy patients don't show significant response and these variables still present an increased value, it can be an indicator that therapy will be inefficient. Nevertheless, in the context of therapy outcome prediction, these are not the most relevant associations since they are recorded at an advanced time in therapy and it is expected that records at more advanced time points are more associated with the outcomes.

Next, regarding ASDAS score evaluated 12 months after therapy initiation, it appears to be correlated with BMI, VAS doctor at baseline, question number 2 of BASDAI score evaluated 3 months after beginning therapy, CRP level and BASFI score after 6 months and, finally, VAS doctor after 12 months. Concerning correlation with BMI, it was noted that the smaller the BMI, the higher the probability of presenting a smaller ASDAS score at 12 months, and the greater the BMI, the higher the probability of presenting a greater ASDAS score at 12 months. This indicates that people with excess weight seem to respond worse to therapy, which is in fact confirmed by literature about anti-TNF therapy predictors

[33, 34]. An example of a CPT which allowed to achieve this conclusion can be seen in Table 4.4. In the case presented, in the corresponding network presented, the only parent of ASDAS at 12 months was BMI, allowing to easily identify its impact on the outcome variable.

Another possibly interesting correlation for ASDAS at 12 months in the context of therapy prediction is with doctor's evaluation at baseline. Nevertheless, this variable appears as a parent of ASDAS at 12 months only together with other variables, including BMI and BASDAI question number 2 at 3 months, making the visual interpretation of VAS doctor's influence on the outcome variable more difficult. In this case, by looking at the CPTs, being an example presented in Table 4.5, it is possible to notice that with a small BMI and a greater VAS doctor at baseline, the most probable outcome is a smaller ASDAS at 12 months, which doesn't happen with a small BMI and a small VAS doctor at baseline. This seems to indicate that a greater VAS doctor at baseline may be a good predictor of response, which is actually, corroborated by other studies that associate higher disease activity at baseline, with higher therapy response [28].

Table 4.4: Probability distribution of ASDAS at 12 months conditioned on BMI, in one of adalimumab's networks.

	ASDAS at 12 months			
	0	1	2	3
BMI				
0	0.58	0.21	0.17	0.04
1	0.12	0.41	0.47	0.0
2	0.20	0.20	0.49	0.11

Table 4.5: Probability distribution of ASDAS at 12 months conditioned on BMI and VAS doctor at 0 months, in one of adalimumab's networks.

	ASDAS at 12 months			
	0	1	2	3
BMI|VAS Doctor at 0 months				
0 | 0	0.25	0.25	0.25	0.25
0 | 1	0.91	0.03	0.03	0.03
0 | 2	0.73	0.25	0.01	0.01
1 | 0	0.25	0.25	0.25	0.25
1 | 1	0.25	0.25	0.25	0.25
1 | 2	0.03	0.03	0.92	0.04
2 | 0	0.25	0.25	0.25	0.25
2 | 1	0.25	0.25	0.25	0.25
2 | 2	0.33	0.01	0.65	0.01

Regarding BASDAI question 2 after 3 months and CRP and BASFI at 6 months, it was not possible to capture their single effect in ASDAS at 12 months just by looking into the CPTs, since this variables only appear as a parent of this outcome variable together with others. Finally, regarding correlation between VAS doctor at 12 months and ASDAS at the same time, a greater VAS doctor is associated with a greater ASDAS value, and a smaller VAS doctor is associated with a smaller ASDAS value, as can be seen in Table 4.6. Indeed, since ASDAS is a score that evaluates disease activity, and since the doctor makes a global evaluation of the patient, it is expected that the values of these two variables evolve in the same direction.

To conclude, concerning change in ASDAS after 12 months of therapy, it appears to be correlated

Table 4.6: Probability distribution of ASDAS at 12 months conditioned on VAS Doctor at 12 months, in one of adalimumab's networks.

	ASDAS at 12 months			
	0	1	2	3
VAS doctor at 12 months				
0	0.62	0.19	0.17	0.02
1	0.04	0.32	0.60	0.04
2	0.01	0.17	0.49	0.33

with BMI, VAS doctor and patient's self evaluation (VAS patient) at baseline, BASDAI question number 3 and CRP levels evaluated after 9 months and, lastly, question number 3 of BASDAI score at 12 months. Regarding BMI, patients with smaller values for this variable seem to be more likely to experience greater improvements in ASDAS score, while patients with higher BMI seem to be more likely to experience smaller improvements in ASDAS score. These findings are in accordance with previous results for ASDAS at 12 months, where patients with greater BMI seemed more likely to respond worse to therapy, by presenting greater ASDAS values at 12 months. An example of CPT which allows to achieve this conclusion can be seen in Table 4.7. In the case presented, the only parent of ASDAS change at 12 months was BMI, allowing to easily identify its impact on the outcome variable.

Table 4.7: Probability distribution of ASDAS at 12 months conditioned on BMI, in one of adalimumab's networks.

	Change in ASDAS at 12 months		
	0	1	2
BMI			
0	0.23	0.09	0.68
1	0.37	0.37	0.26
2	0.50	0.30	0.20

Table 4.8: Probability distribution of ASDAS at 12 months conditioned on BMI and VAS doctor at 0 months, in one of adalimumab's networks.

	Change in ASDAS at 12 months		
	0	1	2
BMI | VAS doctor at 0 months			
0 | 0	0.33	0.33	0.33
0 | 1	0.02	0.49	0.49
0 | 2	0.25	0.01	0.74
1 | 0	0.33	0.33	0.33
1 | 1	0.33	0.33	0.33
1 | 2	0.66	0.01	0.33
2 | 0	0.33	0.33	0.33
2 | 1	0.33	0.33	0.33
2 | 2	0.33	0.66	0.01

Additionally, change in ASDAS at 12 months appears to be correlated with VAS doctor and VAS patient at baseline. Once more, by looking at the CPTs, it is possible to notice that with a small BMI and a great VAS doctor at baseline, the most probable outcome is a greater ASDAS improvement, as can be seen in Table 4.8. Furthermore, with a small BMI and both VAS doctor and VAS patient increased at

baseline, the most likely outcome is, again, a greater improvement, suggesting once more that a greater disease activity at baseline may be a good predictor of response.

Regarding the remaining variables correlated with ASDAS improvement, BASDAI question 3 and CRP after 9 months and BASDAI question 3 at 12 months, these are not the most interesting ones in the context of therapy outcome prediction, since recorded at advanced times after therapy initiation. Moreover, these correlations only appeared once, which might indicate that these associations may not greatly represent the data.

The second biologic to be evaluated was etanercept, being the corresponding heatmap presented in Figure 4.8.

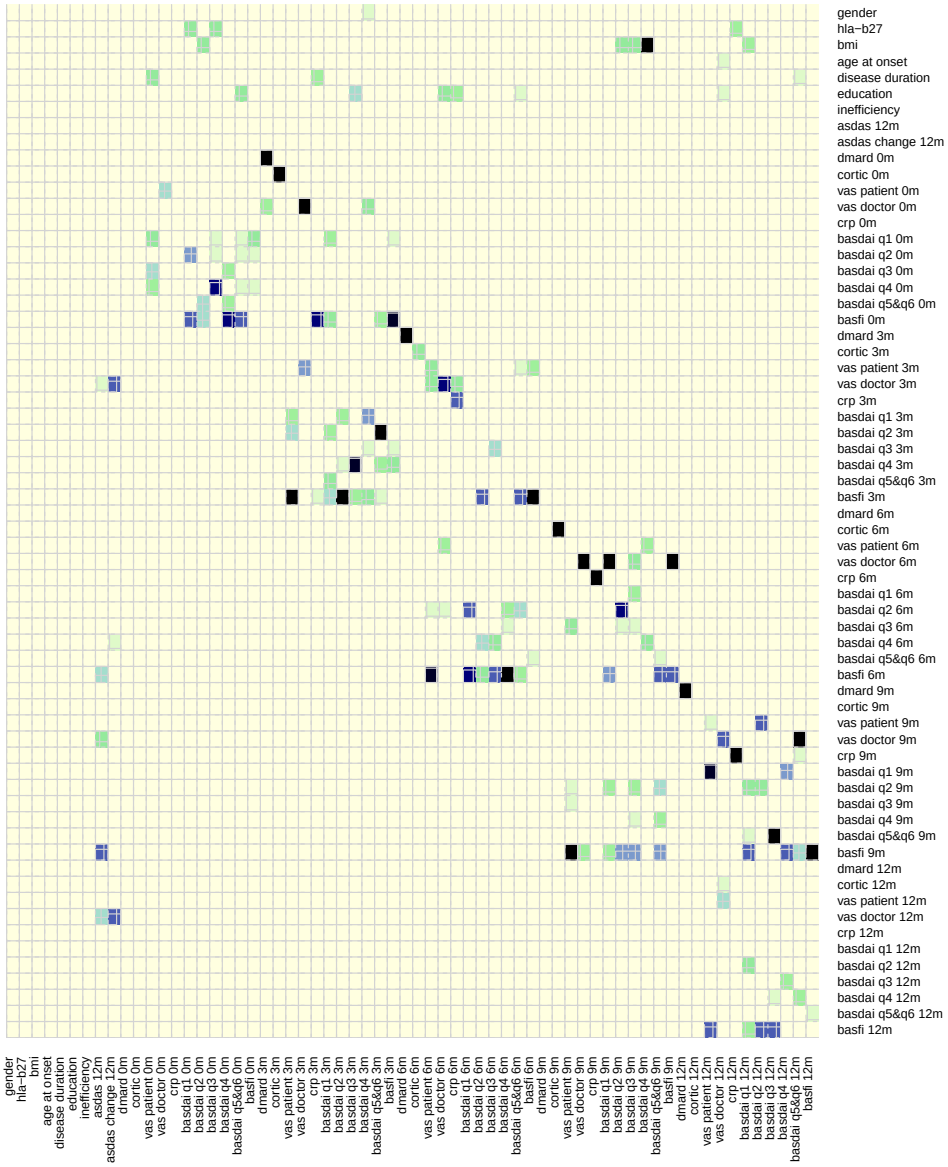


Figure 4.8: Heatmap obtained for etanercept using datasets with a 3-month interval between records and no imputation.

Starting with the outcome variable inefficiency, it is noted that the models don't show any edge

connecting with this variable, which seems to indicate that possible relations with this variable are not the ones that better represent the data.

Then, looking into ASDAS at 12 months, it is possible to note that it correlates with VAS doctor at 3 months, nevertheless, this edge is only present in one of the nine networks obtained, so it is possible to conclude that this edge may not represent a significant association. Moreover, by looking into the network's conditional probabilities table, it is difficult to manually understand the influence of this variable on ASDAS at 12 months, since two other variables are ASDAS's parents in the same network. Indeed, additionally, ASDAS at 12 months correlates with some other dynamic variables in more advanced time points, namely BASFI at 6 months, VAS doctor and BASFI at 9 months and VAS doctor at 12 months, being therefore not so relevant for the purpose of the present research.

Lastly, change in ASDAS appears to be correlated with VAS doctor after 3 months of therapy, BASDAI question 4 at 6 months and VAS doctor at 12 months. By looking into the CPTs, a smaller value for VAS doctor at 3 months seems to be more associated with greater ASDAS improvements, as can be seen in Table 4.9, which may be an indicator of therapy response. Possibly, if VAS doctor already decreases/presents a reduced value after 3 months of therapy, it might be an indicator that the patient will respond to the therapy and show great improvements.

Table 4.9: Probability distribution of ASDAS change at 12 months conditioned on VAS Doctor at 3 months, in one of etanercept's networks.

	Change in ASDAS at 12 months		
	0	1	2
VAS Doctor at 3 months			
0	0.24	0.16	0.60
1	0.44	0.44	0.12
2	0.25	0.50	0.25

On the other hand, as expected, a higher VAS doctor at 12 months seems to be associated with a smaller ASDAS improvement and, contrarily, a smaller VAS doctor at 12 months was more associated with a higher improvement. Finally, correlation with BASDAI question 4 at 6 months only appears in one of the nine networks, again leading to the conclusion that this might not be a significant association.

To conclude, the third therapy evaluated was infliximab, being the respective heatmap shown in Figure 4.9.

Regarding inefficiency, as it was also the case for etanercept, no edge was found in the networks between this outcome variable and any other variable, which, again, may be due to the fact that there is no specific variable whose correlation with inefficiency significantly explains the data.

Secondly, regarding ASDAS after 12 months, it seems to be correlated with BASDAI questions number 5 and 6 after 3 months of therapy, VAS doctor at 9 months and, finally, VAS patient at 12 months. These last two variables are evaluated at advanced time points, and therefore do not pose particular interest in initial prediction of therapy outcome. On the other hand, regarding questions 5 and 6 of BASDAI, the CPTs do not allow one to infer a relation between this variable and ASDAS 12 alone, since in all

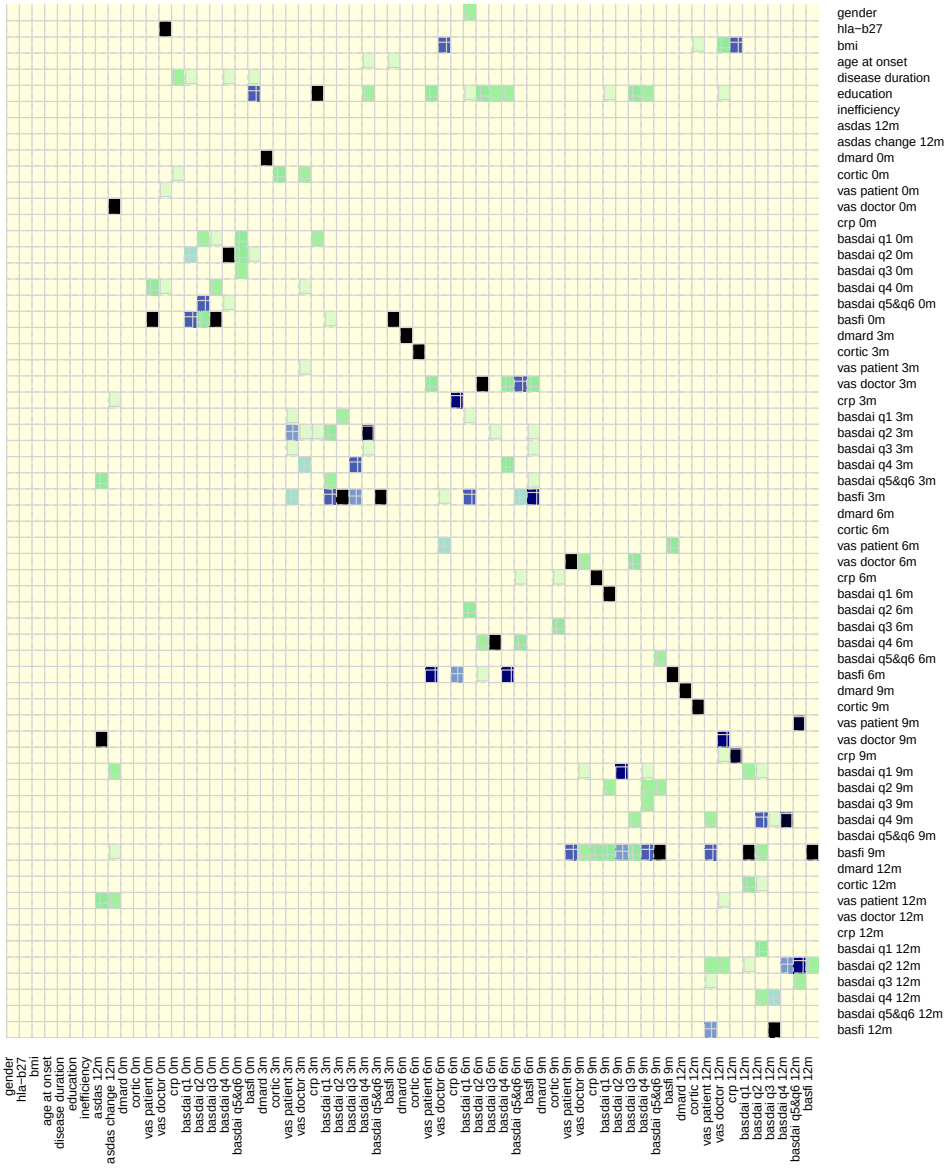


Figure 4.9: Heatmap obtained for infliximab using datasets with a 3-month interval between records and no imputation.

the networks where this edge appears, it appears together with edges from VAS doctor at 9 months and VAS patient at 12 months. Therefore, data is explained by a combination of these variables, and the individual influence of each one becomes more difficult to interpret.

Finally, regarding ASDAS improvement after 12 months of therapy, results show correlation with VAS doctor at baseline, CRP after 3 months, BASDAI question 1 and BASFI after 9 months and VAS patient at 12 months. Looking into the heatmap, it is possible to identify a strong correlation between VAS doctor at baseline and ASDAS improvement, through the appearence of the corresponding edge in all the networks. Next, evaluating the CPTs, it is shown that a smaller VAS doctor at baseline is associated with smaller improvements, as confirmed in Table 4.10. Correlation with CRP levels after 3 months of therapy could also be an interesting association, nevertheless besides only appearing in one of the

networks, it does not appear as a single parent to the variable, which makes the manual interpretation of the variable's effect more difficult. Finally, the remaining variables do not present interesting associations in the context of therapy outcome prediction.

Table 4.10: Probability distribution of ASDAS change at 12 months conditioned on VAS Doctor at 0 months, in one of infliximab's networks.

	Change in ASDAS at 12 months		
	0	1	2
VAS Doctor at 0 months			
0	0.96	0.02	0.02
1	0.40	0.40	0.20
2	0.39	0.39	0.22

Comparing the three heatmaps for the three different biologics, it is possible to identify some common aspects. First, there is a clear diagonal line on the three heatmaps, which corresponds to the correlation between variables with themselves in the next time point, which is an expected relation. Moreover, there are also, for all the heatmaps, correlations between variables evaluated at the same time, which can be seen as 'squared shapes' on the left of the diagonal line. Indeed, it is also expected that variables at the same time point evolve in the same direction and therefore present correlation between them.

Now, focusing on the differences, one of the most relevant ones consists on the influence of BMI in the different networks. In fact, for patients receiving adalimumab, BMI was strongly correlated not only with the outcome variables ASDAS at 12 months and ASDAS improvement, but also with dynamic variables evaluated at 12 months, as can be seen in the corresponding heatmap. This association of BMI with outcome variables is not identified in any other network from other biologic, and even regarding identified correlations with variables evaluated at 12 months, the association is not as common as in networks for adalimumab's patients. These findings seem to indicate that BMI has a higher impact in patients being treated with adalimumab than in patients being treated with etanercept or infliximab, which leads to the conclusion that when dealing with obese patients, adalimumab may not be the best option.

4.2.2 Datasets with imputation in all variables with a 3-month interval between records

The second step consisted in building networks using datasets with an interval of 3 months between records and where imputation was performed in all variables, both static and dynamic, according to the explanation provided in Section 3.3.3. The corresponding heatmaps for the three biologics can be seen in Figures 4.10 to 4.12.

First, looking into adalimumab, whose heatmap is presented in Figure 4.10, it is possible to notice that only dynamic variables were identified to correlate with the outcome variables, contrarily to the heatmap obtained using datasets where no imputation was performed.

Starting with inefficiency, it correlates with VAS doctor at 3 months, VAS patient at 9 months and BASFI at 12 months. Regarding VAS doctor at 3 months, looking at the CPTs it is possible to conclude

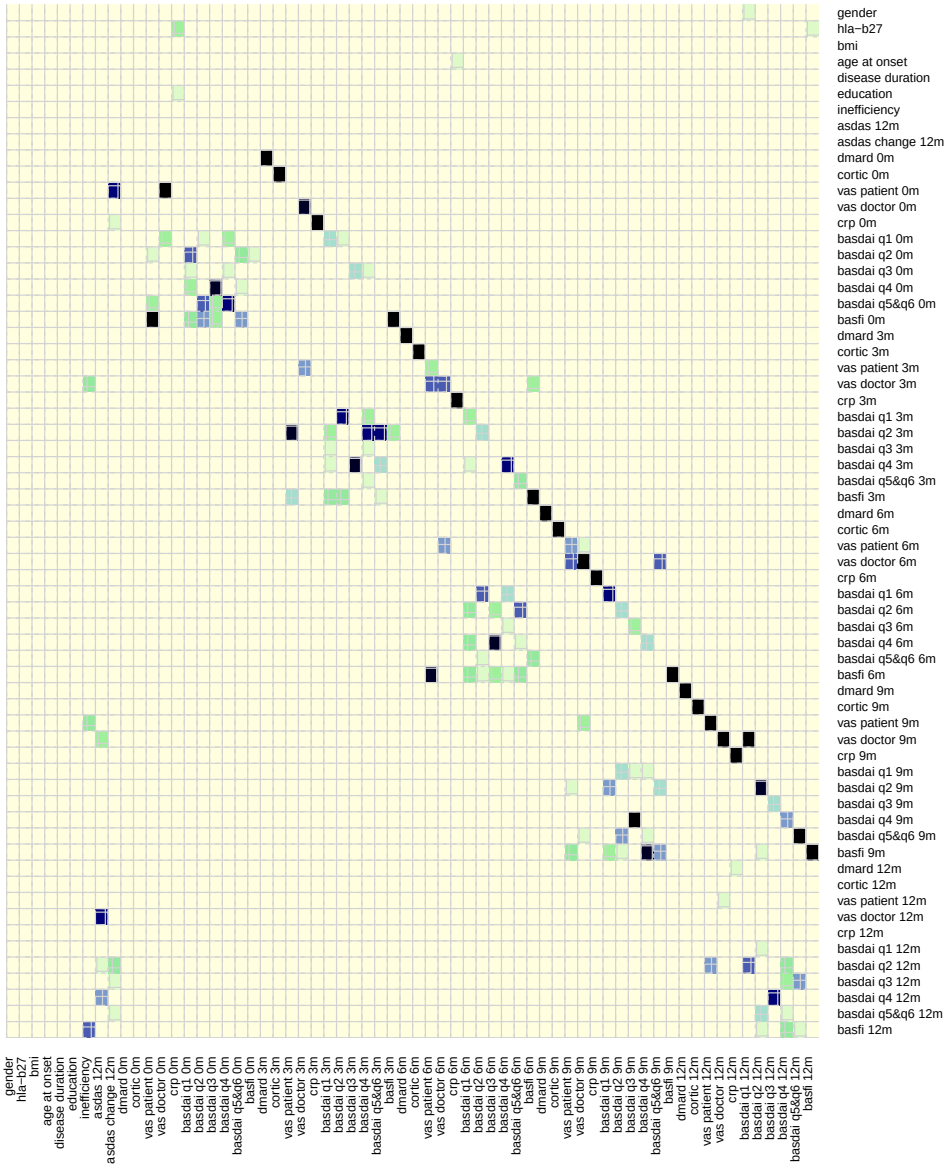


Figure 4.10: Heatmap obtained for adalimumab using datasets with a 3-month interval between records and imputation in all variables.

that patients where doctor's evaluation after 3 months of therapy was higher, have greater probability of therapy inefficiency, compared to patient's with lower VAS doctor, as exemplified in Table 4.11. This may lead to the conclusion that this variable is an indicator of therapy response, in the sense that, if after 3 months of treatment initiation doctor's evaluation is still high, it is more likely that the therapy will fail in the future. On the other hand, regarding the other correlations found, it is expected that inefficiency correlates with variables at more advanced time points, where it is already possible to see if patient is responding or not to therapy. Indeed, for BASFI at 12 months, it is possible to conclude that, as expected, a higher value of this score is more associated with therapy inefficiency. Nevertheless, these relations are not interesting for therapy outcome prediction, which is the focus of this dissertation.

For ASDAS at 12 months, the only correlations identified in networks are with dynamic variables

Table 4.11: Probability distribution of inefficiency conditioned on VAS Doctor at 3 months, in one of adalimumab's networks.

		Inefficiency	
		0	1
VAS Doctor at 3 months			
	0	0.91	0.09
	1	0.94	0.06
	2	0.66	0.34

in advanced time points, VAS doctor at 9 months and BASDAI questions 2 and 4 and VAS doctor at 12 months. Again, it is expected that variables that globally evaluate the condition of the patient, as VAS doctor, at advanced time points may express disease activity, also expressed by ASDAS. Moreover, patient's answer to BASDAI question number 2 is used in the calculation of ASDAS, so it is also expected that these two variables are correlated.

Finally, regarding ASDAS improvement at 12 months, it appears correlated with VAS patient and CRP level at baseline and BASDAI questions 2, 3, 5 and 6 after 12 months of therapy. Regarding VAS patient at the beginning of therapy, looking at the CPTs it is possible to conclude that a smaller VAS patient is more associated with a smaller improvement, while a greater VAS patient is more associated with a greater improvement, as shown in Table 4.12.

Table 4.12: Probability distribution of ASDAS change at 12 months conditioned on VAS Patient at 0 months, in one of adalimumab's networks.

		Change in ASDAS at 12 months		
		0	1	2
VAS Patient at 0 months				
	0	0.75	0.21	0.03
	1	0.52	0.34	0.13
	2	0.33	0.23	0.44

Next, concerning CRP at baseline, it appears as a parent together with VAS patient at the same time, with small values of both variables being highly associated with a small improvement, and greater values being mostly associated with a greater improvement. Finally, the remaining variables are evaluated at an advanced stage of therapy, not posing particular interest for the present study.

Secondly, networks were built using etanercept's data, whose corresponding heatmap can be seen in Figure 4.11.

Concerning inefficiency, it was not found any edge connecting this variables with any other variable in any of the networks.

Regarding ASDAS at 12 months, besides being correlated with variables at advanced time points, particularly VAS doctor at 6 months and VAS patient, VAS doctor and BASDAI question 3 at 12 months of therapy, one of the networks also presented an edge from gender to ASDAS at 12 months, which appears together with an edge from VAS doctor at 6 months. Looking into the CPTs, presented

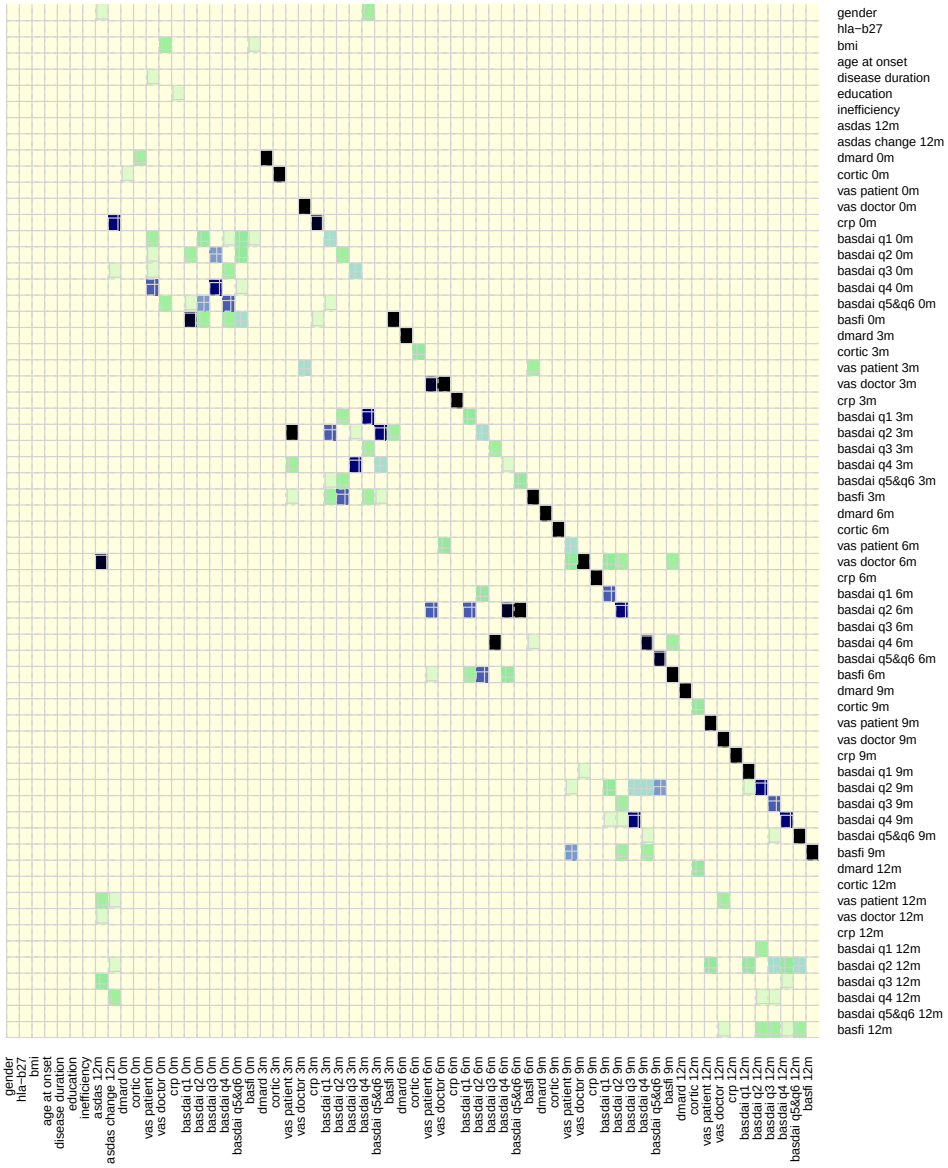


Figure 4.11: Heatmap obtained for etanercept using datasets with a 3-month interval between records and imputation in all variables.

in Table 4.13, when VAS doctor at 6 months is small and the patient is male, the ASDAS value is more likely to be lower after 12 months. On the contrary, for a female patient with small VAS doctor at 6 months, the most likely outcome is that ASDAS will be higher. This way, it seems that male gender would be a good predictor of therapeutic response, which is in fact confirmed in literature for anti-TNF medication [10, 29]. This finding could lead to the assumption that gender poses a more significant impact in etanercept's patients, nevertheless, this conclusion should be evaluated carefully. First, this relation did not appear when imputation was not performed, and gender has a considerable amount of missing data in the initial dataset. In fact, imputation is known to be a complex process, which not always produces good results, being particularly difficult to impute time independent variables. Moreover, this correlation was only found in one of the nine networks built and so, it should be further studied.

Table 4.13: Probability distribution of ASDAS at 12 months conditioned on Gender and VAS Doctor at 6 months, in one of etanercept's networks.

	ASDAS at 12 months			
	0	1	2	3
Gender|VAS Doctor at 6 months				
0 | 0	0.53	0.41	0.03	0.03
0 | 1	0.08	0.61	0.31	0.00
0 | 2	0.17	0.01	0.49	0.33
1 | 0	0.26	0.22	0.52	0.00
1 | 1	0.20	0.32	0.48	0.00
1 | 2	0.00	0.11	0.66	0.22

Finally, regarding ASDAS improvement, it was correlated with CRP and BASDAI question 3 at baseline and VAS patient and BASDAI questions 2 and 4 at 12 months. Concerning variables at baseline, which are the ones that allow to predict therapy response before its initiation, it is possible to conclude, looking into CPTs, that a smaller CRP at baseline is more associated with a smaller ASDAS improvement, while a greater CRP at baseline is associated with a higher change in the ASDAS score, as can be seen in 4.14. In the same way, patients with higher CRP and BASDAI question 3 are more likely associated with a greater ASDAS improvement, while patients with lower CRP and BASDAI question 3 are more likely associated with a smaller ASDAS improvement.

Table 4.14: Probability distribution of ASDAS change at 12 months conditioned on CRP at 0 months, in one of etanercept's networks.

	Change in ASDAS at 12 months		
	0	1	2
CRP at 0 months			
0	0.70	0.19	0.11
1	0.63	0.23	0.13
2	0.58	0.37	0.05
3	0.24	0.24	0.52

The last biologic evaluated was infliximab, whose heatmap is presented in Figure 4.12.

First, regarding inefficiency, it was only correlated with BASDAI questions 5 and 6 at 12 months. Nevertheless, besides not being relevant in the context of outcome prediction, this relation does not seem to be of particular interest since it only appeared in one of the nine networks learnt.

Next, concerning ASDAS at 12 months, it only correlates with variables evaluated at that time, namely VAS doctor and BASDAI questions 2 and 4, which is expected.

Finally, for ASDAS improvement, it correlates with CRP and BASFI at baseline, VAS doctor after 3 months of therapy and VAS patient and BASFI after 12 months. Evaluating the impact of the baseline variables in ASDAS improvement, both appear in the networks as a parent of ASDAS change together with VAS patient after 12 months. In both cases, it is noted that a great value for these variables at baseline and a small value of VAS patient after 12 months is more associated with greater improvements,

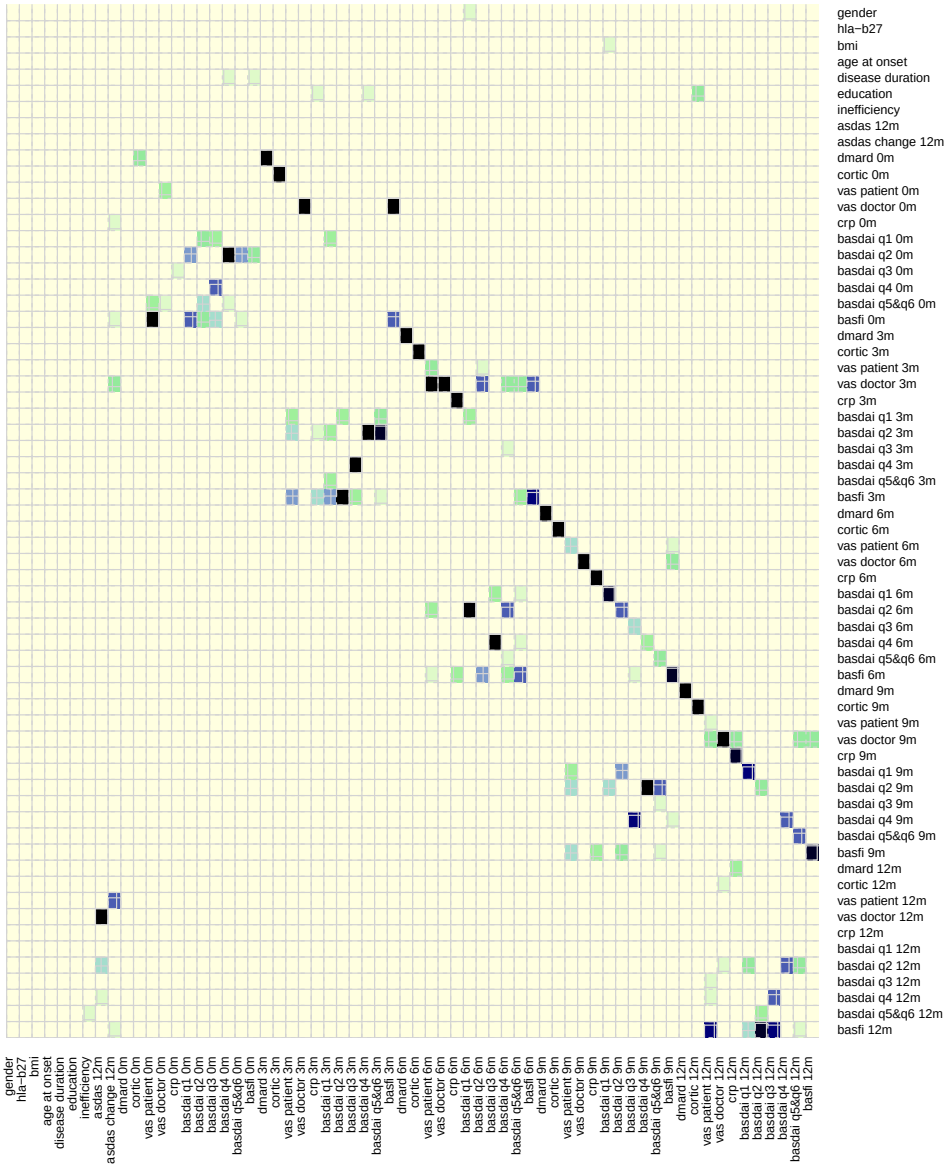


Figure 4.12: Heatmap obtained for infliximab using datasets with a 3-month interval between records and imputation in all variables.

while smaller values at baseline and greater values of VAS doctor at 12 months are more associated with smaller improvements, as expected due to the calculation formula, which constitutes the difference in the ASDAS value at the start of the therapy and the ASDAS value after 12 months of therapy. Nevertheless, it becomes more difficult to assess each variable's individual influence on the outcome variable. Regarding VAS doctor after 3 months of therapy, by looking into the CPTs, as presented in Table 4.15, it is noted that greater values of this variable are more likely to be associated with smaller ASDAS improvements, while smaller values are associated with greater improvements.

This finding may lead to the assumption that this variable can be seen an indicator of therapy response, in the sense that if after 3 months of therapy patients already present a decreased VAS doctor, then it is more likely that the improvement in ASDAS after 12 months will be greater. Finally, a greater value

Table 4.15: Probability distribution of ASDAS change at 12 months conditioned on VAS Doctor at 3 months, in one of infliximab's networks.

	Change in ASDAS at 12 months		
	0	1	2
VAS Doctor at 3 months			
0	0.25	0.43	0.32
1	0.61	0.19	0.19
2	0.64	0.24	0.12

of VAS patient after 12 months is associated with smaller improvements, while smaller values of this variable are associated with greater improvements. Nevertheless, correlations with variables evaluated at this time are not the most interesting ones for the purpose of this work.

Comparing the heatmaps obtained considering imputation with the ones where variables were not imputed, it is possible to notice that the ones considering imputation have a more clear diagonal line. Indeed, as already mentioned, this diagonal line represents interactions of variables with themselves in the next evaluated time point, so, when linear interpolation is performed in dynamic variables, time dependencies between them become more clear, which is noted in the networks.

Moreover, it is also possible to notice that the number of correlations between static and dynamic variables is significantly lower when imputation is performed, for all therapies. A possible motive for this is that imputation of static variables does not provide reliable results, possibly canceling the existing relationships between these variables and outcome ones. Another possible explanation is that the strength of associations between dynamic variables was increased and suppressed the smaller strength of correlations between static and outcome variables.

4.2.3 Datasets with imputation only in dynamic variables with a 3-month interval between records

Finally, due to the known small reliability of imputation of time independent variables, datasets where only dynamic variables were imputed were also created, and the corresponding networks were built, originating the heatmaps presented in the following figures.

Starting with adalimumab, the corresponding heatmap can be seen in Figure 4.13.

The first outcome variable to be evaluated was inefficiency, which appeared correlated with VAS doctor at 3 months, VAS patient at 9 months and BASFI at 12 months. Regarding VAS doctor at 3 months, as already mentioned, it may be an indicator of possible therapy failure, since a higher VAS doctor at 3 months is more associated with biologic failure. In fact, if after 3 months of therapy the patient still presents a higher VAS doctor, it may be an early indicator that therapy might be more likely to fail. Again, the remaining variables found to be correlated with inefficiency do not pose significant interest in the context of therapy outcome prediction.

Secondly, regarding ASDAS score after 12 months of therapy, it appears to be correlated with VAS doctor at 9 months and VAS doctor and BASDAI questions 2 and 4 after 12 months of treatment. As

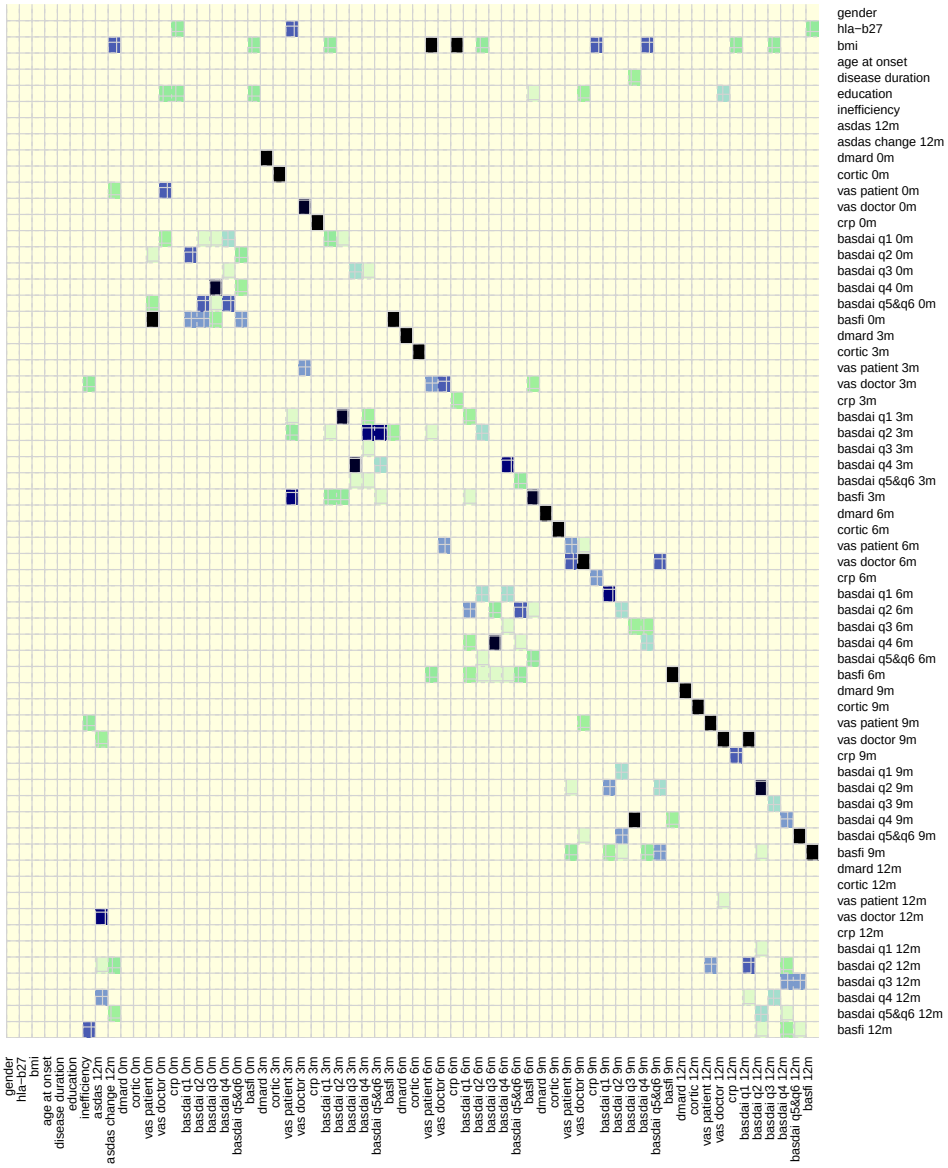


Figure 4.13: Heatmap obtained for adalimumab using datasets with a 3-month interval between records and imputation in dynamic variables only.

mentioned, it is expected that the value of this score correlates with variables evaluated at that time. Indeed, ASDAS score evaluates disease activity, which is also intended to be translated by VAS doctor. Moreover, BASDAI question 2 is also part of the calculation of the ASDAS score.

Finally, regarding ASDAS change after 12 months, it correlates with some dynamic variables, VAS patient at baseline and BASDAI questions 2, 5 and 6 at 12 months, but also with BMI. This correlation had already been found when studying the dataset where no imputation was performed, and the conclusion taken is the same. Indeed, a smaller BMI is more often associated with a higher ASDAS improvement after 12 months, while a greater BMI is more likely to be associated with a smaller change in ASDAS score. Regarding VAS patient at baseline, it does not appear as a single parent of ASDAS change, appearing together with BMI. Indeed, for patients with smaller BMI and smaller VAS patient at

baseline, the percentage of people that presents greater improvements is smaller than the percentage of people that present greater improvements with smaller BMI but greater VAS patient at baseline. Again, this may lead to the assumption that a greater disease activity at baseline is a good indicator of therapy response, which is in fact stated in literature.

Moving into the second biologic studied, etanercept's heatmap can be found in Figure 4.14.

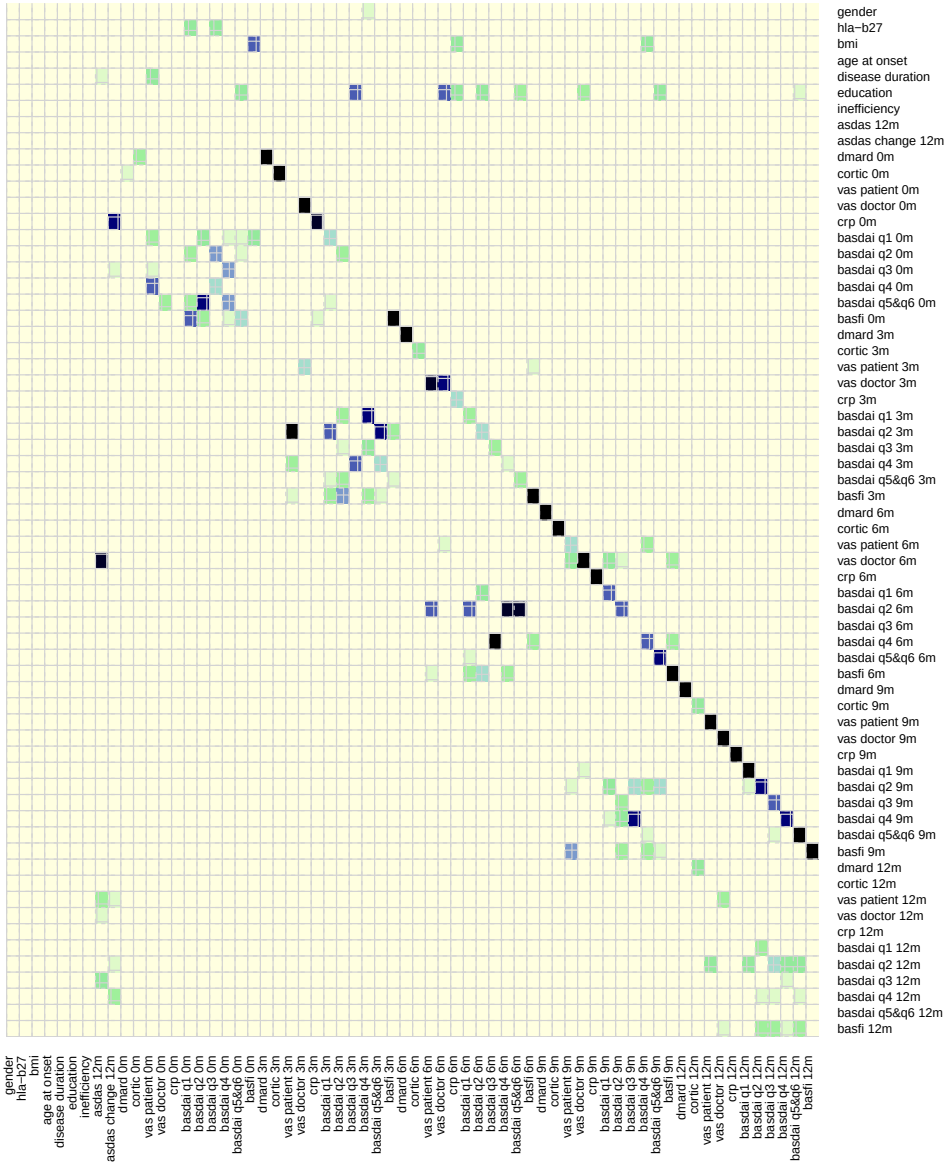


Figure 4.14: Heatmap obtained for etanercept using datasets with a 3-month interval between records and imputation in dynamic variables only.

First, looking into outcome variable inefficiency, it does not appear to correlate with any other variable.

Next, regarding ASDAS at 12 months, besides appearing correlated with dynamic variables VAS doctor at 6 months and VAS patient, VAS doctor and BASDAI question 3 after 12 months of therapy, one of the networks also correlates ASDAS with disease duration at the beginning of treatment. In this network, ASDAS at 12 months has two parents, disease duration and VAS doctor at 6 months.

When looking into the CPTs, presented in Table 4.16, it is possible to note that when VAS doctor has an intermediate value, if disease duration is smaller, the most probable outcome is a moderate disease activity at 12 months, while with a higher disease duration the patient is most likely to present a high disease activity after 12 months of treatment. Moreover, considering patients with smaller VAS doctor at 6 months, in patients where also disease duration is smaller there is a higher percentage of people to be in a state of inactive disease after 12 months, than when disease duration is higher.

Table 4.16: Probability distribution of ASDAS at 12 months conditioned on Disease Duration and VAS Doctor at 6 months, in one of etanercept's networks.

	ASDAS at 12 months			
	0	1	2	3
Disease Duration|VAS Doctor at 6 months				
0 | 0	0.48	0.19	0.33	0.00
0 | 1	0.15	0.50	0.35	0.00
0 | 2	0.01	0.01	0.65	0.33
1 | 0	0.29	0.53	0.12	0.06
1 | 1	0.08	0.54	0.38	0.00
1 | 2	0.25	0.01	0.49	0.25
2 | 0	0.33	0.33	0.33	0.01
2 | 1	0.10	0.30	0.60	0.00
2 | 2	0.01	0.01	0.66	0.32

This may seem to indicate that higher disease duration may have a negative impact on treatment outcome, which is in fact confirmed by literature [28]. Nevertheless, this correlation only appeared in one of the nine networks, so it should be further studied.

Lastly, regarding ASDAS change, it appears correlated with CRP and BASDAI question 3 at baseline and with VAS patient and BASDAI questions 2 and 4 after 12 months of treatment. Correlation with CRP at baseline appears to be relatively strong, since it appears in the majority of the networks. Looking into the CPTs, the conclusion taken is that smaller CRP at baseline is associated with a smaller improvement, while higher CRP at baseline is more associated with higher improvements. On the other hand, correlation with BASDAI question 3 at baseline only appears in one of the nine networks, together with the edge from CRP at baseline to ASDAS improvement. The same idea is verified, in which greater values of both these variables are likely to be associated with higher improvements, while lower values of both this variables are more likely to be associated with smaller improvements.

To conclude, the heatmap obtained for infliximab is presented in Figure 4.15.

First, regarding inefficiency, only one correlation was found involving this variable, with BASDAI questions 5 and 6 after 12 months of therapy, being therefore a correlation with little interest in the context of this work.

Next, regarding ASDAS value at 12 months, it appears correlated with dynamic variables VAS doctor and BASDAI questions 2 and 4 at that time. Additionally, it also appears correlated with education, although in only one of the nine networks. Nevertheless, in this heatmap, it is possible to notice that, besides being correlated with this outcome variable, education is also correlated with several other dy-

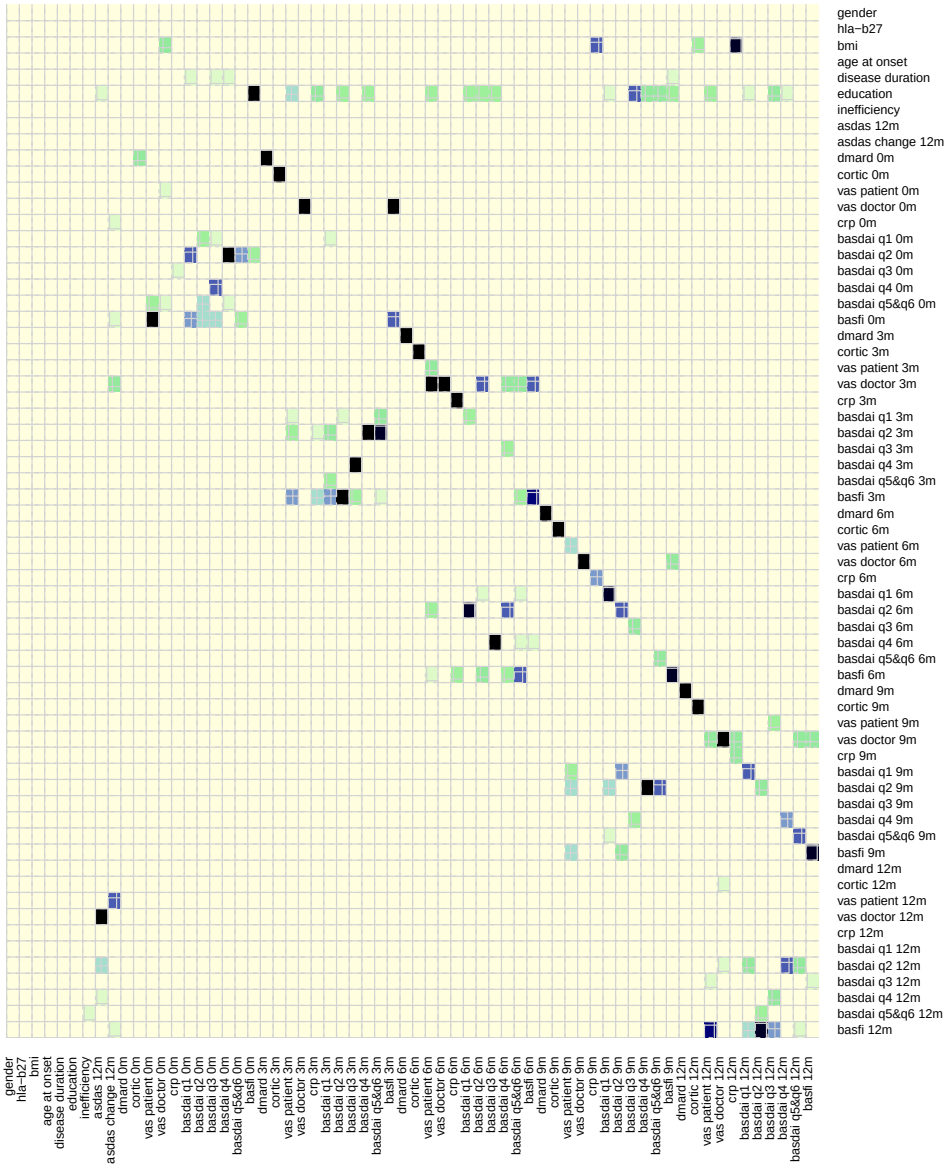


Figure 4.15: Heatmap obtained for infliximab using datasets with a 3-month interval between records and imputation in dynamic variables only.

namic variables, specially in advanced time points, including after 12 months of therapy, which may indicate that education has a greater influence on therapy response for this biologic. In the network where education is a parent of ASDAS at 12 months, this edge is accompanied by edges from VAS doctor and BASDAI question 2, and so it is not so intuitive to look at the CPTs and understand education's impact. However, further studies should be conducted to further study this possible association.

Finally, regarding ASDAS improvement, all correlations are with dynamic variables, namely CRP and BASFI at baseline, both appearing only once in the different networks, VAS doctor at 3 months and finally, VAS patient and BASFI recorded after 12 months of therapy. Concerning the two mentioned variables correlated at baseline, they do not appear as single parents of ASDAS improvement, both appearing together with VAS patient after 12 months, being therefore more difficult to understand their

single influence. On the other hand, regarding VAS doctor at 3 months, it is noticed that smaller values are more correlated with greater improvements, while greater values are more correlated with smaller improvements, indicating again that this variable might be an early indicator of therapy failure.

To conclude, comparing these last three heatmaps with the ones where no imputation was performed, it is noticeable that that diagonal line became more clear, as it was the case for heatmaps built with datasets where all variables were imputed. On the other hand, the correlations of static variables with dynamic variables did not disappear, as happened for datasets where all variables were imputed. Indeed, by imputing only dynamic variables, correlations of static variables were not lost, and by performing imputation on dynamic variables, which is more reliable, it was even possible to unravel previously non significant correlations of static variables with dynamic ones.

4.2.4 Datasets with no imputation with a 6-month interval between records

Besides evaluating datasets with an interval of 3 months between records, datasets with an interval of 6 months were also created. This way, it would be possible to reduce the amount of missing data by increasing the allowed time range for record selection, as previously described. As before, heatmaps were built using datasets with no imputation, with imputation on both static and dynamic variables and with imputation only on dynamic variables.

Starting with datasets where no imputation was performed, the corresponding heatmaps can be seen in Figures 4.16 to 4.18.

Regarding adalimumab, whose heatmap is presented in Figure 4.16, inefficiency appears to be correlated with VAS doctor at baseline and VAS doctor and BASFI after 6 months of treatment. While associations with variables recorded after 6 months do not pose particular interest for the purpose of this study, since recorded at an advanced time, VAS doctor at baseline constitutes an interesting association. For this variable, it seems that, for patients with higher VAS doctor at therapy beginning, there is a higher proportion of people in which therapy was inefficient, than in patients where VAS doctor was smaller, as can be seen in Table 4.17.

Table 4.17: Probability distribution of inefficiency conditioned on VAS Doctor at 0 months, in one of adalimumab's networks.

	Inefficiency	
	0	1
VAS Doctor at 0 months		
0	0.99	0.01
1	0.99	0.01
2	0.85	0.15

This finding might seem contradictory comparing to associations previously found, where a greater VAS doctor at baseline was associated with a smaller ASDAS after 12 months. Nevertheless, it can be hypothesized that having higher VAS doctor at baseline can in some cases lead to therapy inefficiency, while in other cases it can have a positive impact on therapy response. An hypothesis can be, for example,

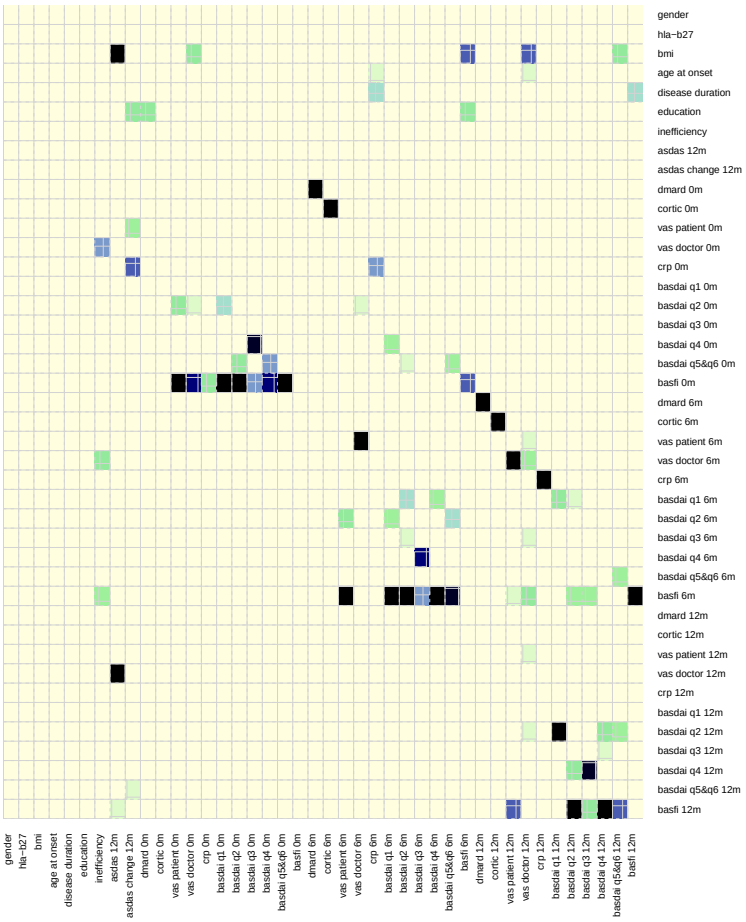


Figure 4.16: Heatmap obtained for adalimumab using datasets with a 6-month interval between records and no imputation.

that for extremely elevated values it is no longer an advantage for therapy response. It should be noted, however, that the findings reached in this dissertation are preliminary results on this topic, that should therefore be further studied, in order to be confirmed or rejected.

Next, regarding ASDAS at 12 months, it appears correlated with two dynamic variables, VAS doctor and BASFI after 12 months, but also with a static variable, BMI. Indeed, this relation was also found in the corresponding heatmap for datasets with a 3 month interval between appointments, however, this time, it does not appear as the only parent of this outcome variable, which makes manual interpretation less intuitive, when looking into the CPTs.

Finally, regarding ASDAS change, it appears correlated with education, VAS patient and CRP at baseline and BASDAI questions 5 and 6 after 12 months of treatment. Looking into the CPTs, education impact on ASDAS change does not seem to have a defined pattern influencing smaller or greater improvement, as can be seen in Table 4.18. On the other hand, regarding CRP at baseline, as previously, a smaller CRP at the beginning of therapy is more associated with smaller improvements, while a higher CRP is more associated with greater improvements. Regarding VAS patient at baseline, when CRP at the same time is higher, if VAS patient is higher, the percentage of people among these that experience

a greater improvement is greater than the same percentage among patients with higher CRP but intermediate VAS patient, which in turn is higher than the same percentage for people with higher CRP but lower VAS patient at baseline. This findings seem to be in line with previous ones, where higher disease activity at baseline was associated with greater improvements.

Table 4.18: Probability distribution of ASDAS change at 12 months conditioned on VAS Doctor at 3 months, in one of infliximab's networks.

	Change in ASDAS at 12 months		
	0	1	2
Education			
0	0.42	0.16	0.42
1	0.39	0.39	0.22
2	0.36	0.09	0.54
3	0.33	0.29	0.39

Moving into the second biologic studied, etanercept, the corresponding heatmap can be observed in Figure 4.17.

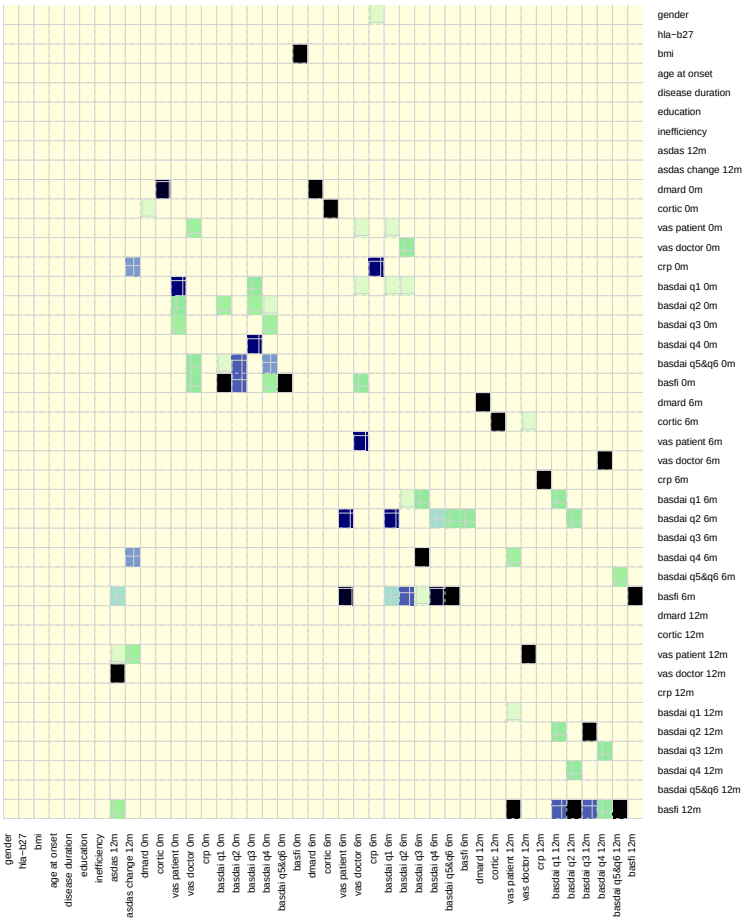


Figure 4.17: Heatmap obtained for etanercept using datasets with a 6-month interval between records and no imputation.

Starting with inefficiency, it is possible to note that it does not appear correlated with any other variable.

Next, regarding ASDAS after 12 months, it only appears correlated with dynamic variables at advanced time points: BASFI after 6 and 12 months of therapy initiation and VAS patient and VAS doctor after 12 months.

Finally, regarding ASDAS improvement, it appears correlated with CRP level at baseline, BASDAI question 4 at 6 months and VAS patient at 12 months. Again, the correlation with CRP at baseline is in accordance with previous findings, where smaller CRP levels at the beginning of therapy have as most probable outcome a smaller change in ASDAS and higher CRP levels are, on the opposite side, most likely associated with greater improvements.

Concluding, for infliximab, the corresponding heatmap is presented in Figure 4.18.

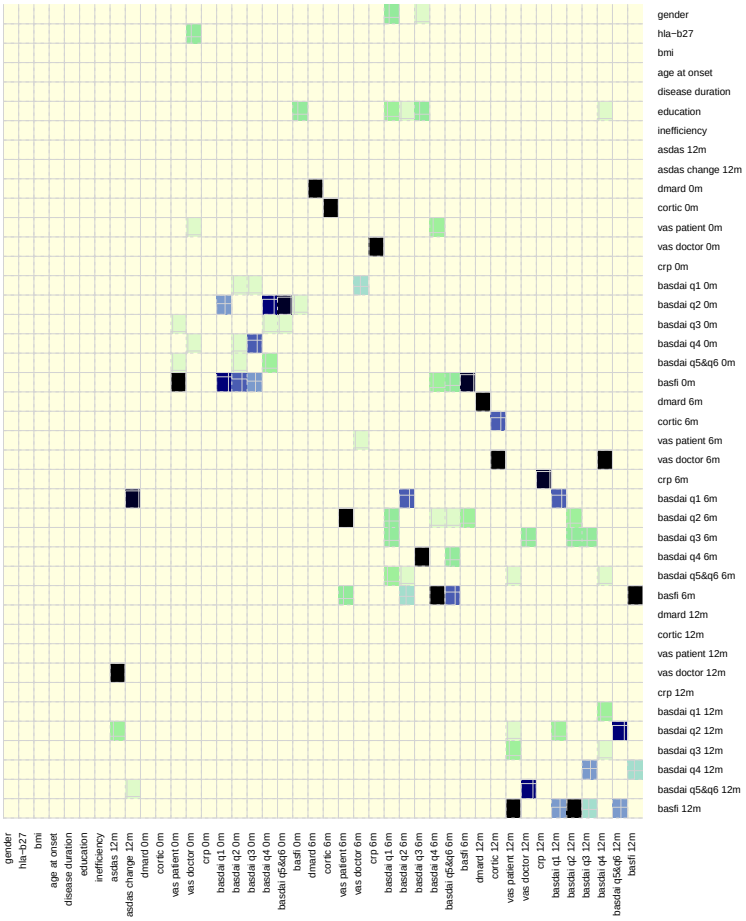


Figure 4.18: Heatmap obtained for infliximab using datasets with a 6-month interval between records and no imputation.

Looking into the first outcome studied, inefficiency, it does not appear correlated with any variable, either static or dynamic, at any time.

Next, looking into ASDAS at 12 months, it only appears correlated with dynamic variables evaluated at the same time, which is, as already mentioned, an expected correlation.

Finally, regarding ASDAS change after 12 months of therapy, it also only appears correlated with dynamic variables at advanced time points, namely BASDAI question 1 at 6 months and BASDAI questions 5 and 6 at 12 months, which are not the most interesting correlations in the context of therapy outcome prediction.

Comparing the three heatmaps for the three different biologics, for datasets with an interval of 6 months between records, it is possible to identify some common characteristics. As previously, a diagonal line is present in the three heatmaps, which corresponds to the correlation between variables with themselves in the next time point, which is an expected relation. Nevertheless, this line is less clear in these last heatmaps, built using datasets with a 6-month interval between records, than in the previously obtained heatmaps considering a 3-month interval. Indeed, since variables are recorded at more distant time points, the values presented may be more different and, thus, correlations may become less clear. Additionally, for all the heatmaps, correlations between variables evaluated at the same time are also present.

On the other hand, some differences can also be identified, being again one of the most relevant ones the influence of BMI in the different networks. Indeed, for adalimumab, BMI was strongly correlated not only with the outcome variable ASDAS at 12 months, but also with dynamic variables evaluated at 12 months, which was not identified in any other network from any other biologic.

4.2.5 Datasets with imputation in all variables with a 6-month interval between records

Next, as previously done for the dataset with a 3 month interval between records, all variables were imputed and the resulting networks analysed.

Starting with adalimumab, the corresponding heatmap is presented in Figure 4.19.

Concerning inefficiency, looking into the heatmap, it is possible to note that it is correlated only with dynamic variables at advanced time points, namely BASFI after 6 months and VAS doctor at 12 months.

Secondly, regarding ASDAS at 12 months, it only appeared correlated with dynamic variables at the same time, specifically VAS doctor, BASDAI questions 2 and 4 and, lastly, BASFI.

Finally, regarding ASDAS improvement, besides appearing correlated with other variables after 12 months of therapy, it also appears correlated with VAS patient and CRP at baseline, in which, again, patients with higher CRP and VAS patient at the beginning of therapy are more likely to be associated with higher improvements.

Moving into the second biologic studied, etanercept, the corresponding heatmap can be seen in Figure 4.20.

Concerning the study of the outcome variables, starting with inefficiency, it does not appear correlated with any other variable.

Next, regarding ASDAS after 12 months of therapy, as for adalimumab, it only appears correlated with variables evaluated at that time: VAS patient, VAS doctor, BASDAI questions 2 and 3 and, lastly, BASFI.

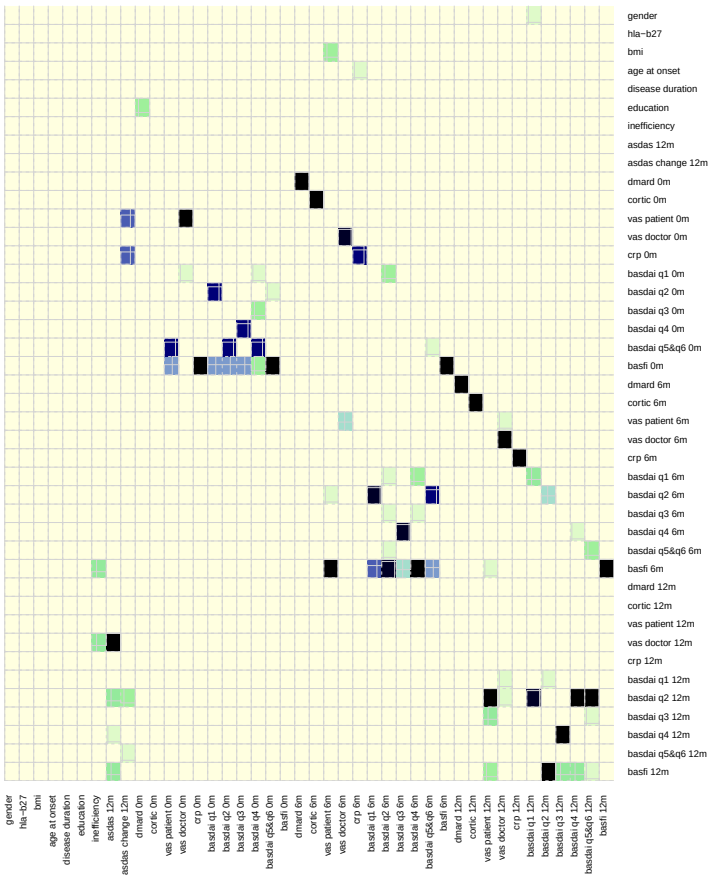


Figure 4.19: Heatmap obtained for adalimumab using datasets with a 6-month interval between records and imputation in all variables.

Finally, regarding ASDAS change, besides correlated with VAS patient and BASDAI questions 1 to 3 after 12 months, it also appears correlated with CRP at baseline, again with smaller CRP more associated with smaller changes in the ASDAS score, and higher CRP more associated with higher changes.

Finally, the heatmap obtained for infliximab is presented in Figure 4.21.

Regarding inefficiency, as for etanercept, it does not appear correlated with any other variable.

In turn, ASDAS at 12 months only appears correlated with VAS doctor and BASDAI question 2 at that time, which are expected associations.

To conclude, regarding ASDAS improvement, it appears to be correlated with several dynamic variables, although some correlations only appear in one of the nine built networks, which is the case of BASDAI questions 5 and 6 at baseline and VAS patient after 6 months. Nevertheless, regarding BASDAI questions 5 and 6 at baseline, by looking into the CPTs, it seems that higher values of this variable are associated with greater improvements, while smaller values are associated with smaller improvements. Additionally, it appears to be correlated with BASDAI question 1 at 6 months and BASDAI questions 3, 5 and 6 after 12 months.

As previously, comparing heatmaps from datasets with imputation performed in all variables with heatmaps from datasets with no imputation, it is possible to notice that there were less correlations be-

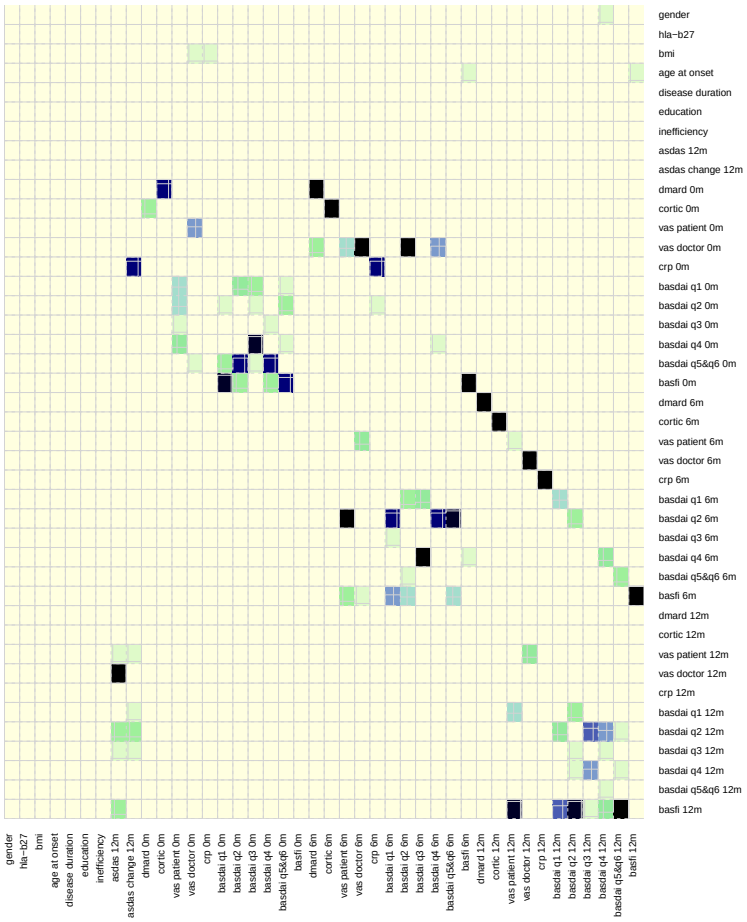


Figure 4.20: Heatmap obtained for etanercept using datasets with a 6-month interval between records and imputation in all variables.

tween static variables and dynamic ones, which may indicate the poorness of the imputation performed on time independent variables, that resulted in the removal of existing associations involving these variables.

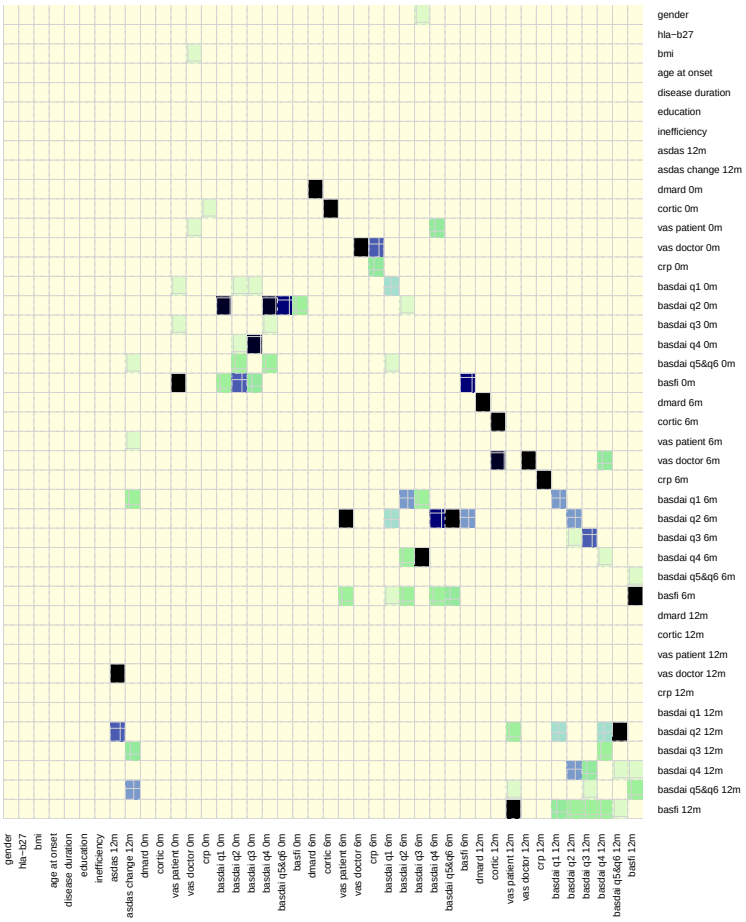


Figure 4.21: Heatmap obtained for infliximab using datasets with a 6-month interval between records and imputation in all variables.

4.2.6 Datasets with imputation only in dynamic variables with a 6-month interval between records

Finally, as before, datasets where only dynamic variables were imputed were also created and studied.

First, the heatmap regarding adalimumab is presented in Figure 4.22.

Starting with the inefficiency outcome, it appears correlated only with variables at advanced time points, BASFI at 6 months and VAS doctor at 12 months.

Next, concerning ASDAS at 12 months, besides appearing correlated with dynamic variables at that time, it also appears correlated with BMI, as previously found. Again, looking into the CPTs, an increased BMI seems to be associated with higher ASDAS scores after 12 months of therapy, meaning higher disease activity, while lower BMI is more likely to be associated with smaller ASDAS score after 12 months.

Finally, regarding ASDAS improvement, it also appears correlated with BMI, with, again, patients with higher BMI being more associated with smaller improvements and patients with smaller BMI being more associated with higher improvements. Additionally, it appears correlated with VAS patient and CRP at baseline, where higher values of this variables are more likely associated with higher improvements.

On the other hand, etanercept's heatmap is presented in Figure 4.23.

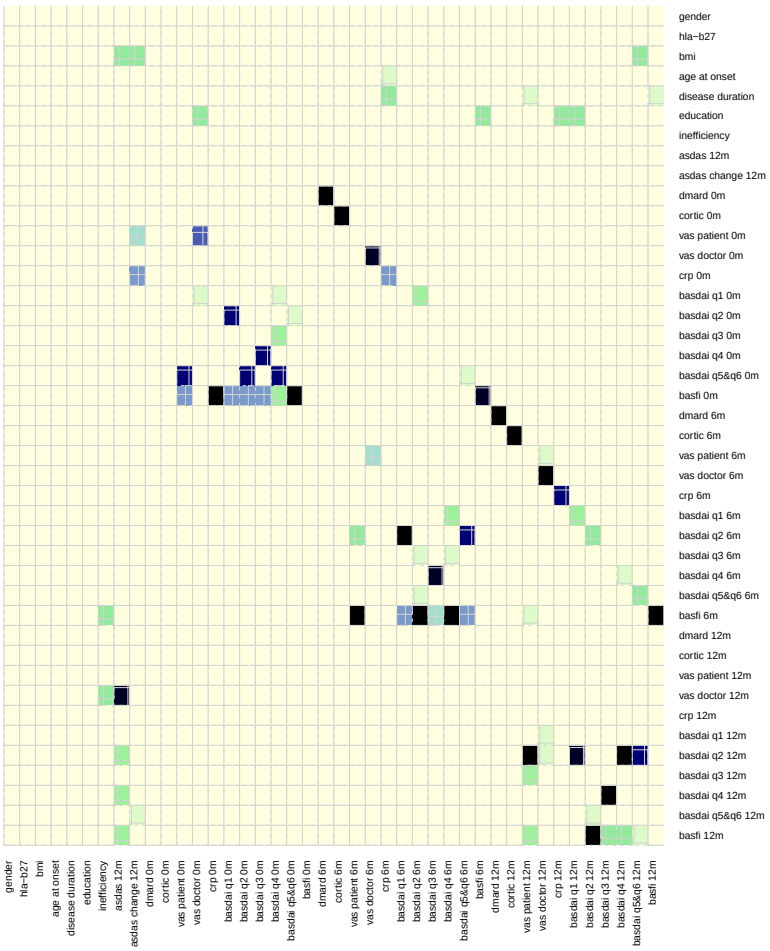


Figure 4.22: Heatmap obtained for adalimumab using datasets with a 6-month interval between records and imputation in dynamic variables only.

Starting with inefficiency, again, it does not appear correlated with any variable.

Regarding ASDAS at 12 months, it only appears to be correlated with variables evaluated at that time, VAS patient, VAS doctor, BASDAI questions 2 and 3 and BASFI.

Finally, regarding ASDAS improvement, it appears correlated with VAS patient and BASDAI questions 1 to 3 after 12 months and also with CRP at baseline, where as before, greater values appear correlated with greater changes.

To conclude, the last biologic studied was infliximab, whose heatmap is presented in Figure 4.24.

Regarding inefficiency, it does not appear correlated with any variable.

Secondly, regarding ASDAS at 12 months, it appears correlated with two static variables, BMI and HLA-B27. For both these variables, they appear as a parent of ASDAS together with other variables, which leads to an increased difficulty in understanding the individual influence by just looking into the CPTs, being the one concerning the association with HLA-B27 presented in Table 4.19.

According to existing studies, the presence of HLA-B27 is associated with better response to TNF antagonists. Looking into the CPT presented, it can be noticed that for patients with small VAS doctor at 6 months, the proportion of people that present smaller ASDAS at 12 months is greater for HLA-B27

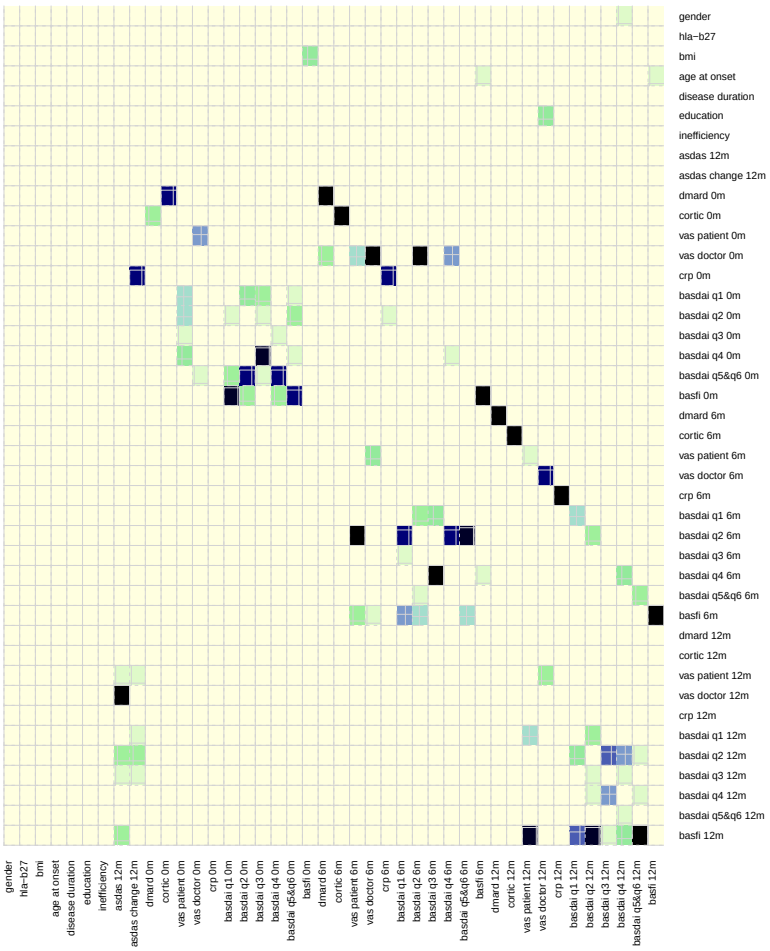


Figure 4.23: Heatmap obtained for etanercept using datasets with a 6-month interval between records and imputation in dynamic variables only.

Table 4.19: Probability distribution of ASDAS at 12 months conditioned on HLA-B27 and VAS Doctor at 6 months, in one of infliximab's networks.

	ASDAS at 12 months			
	0	1	2	3
HLA-B27|VAS Doctor at 6 months				
0 | 0	0.01	0.81	0.17	0.01
0 | 1	0.01	0.20	0.59	0.20
0 | 2	0.01	0.49	0.33	0.17
1 | 0	0.35	0.32	0.32	0.02
1 | 1	0.00	0.39	0.61	0.00
1 | 2	0.20	0.07	0.53	0.20

positive patients, which could corroborate the hypothesis presented, however, it is not possible to clearly confirm this association. The fact that this association was only present for infliximab's data may indicate that this relation is more notorious for this biologic. Nevertheless, further studies should be conducted to confirm or reject this hypothesis.

Finally, regarding ASDAS improvement, it appears correlated with BASDAI questions 5 and 6 at

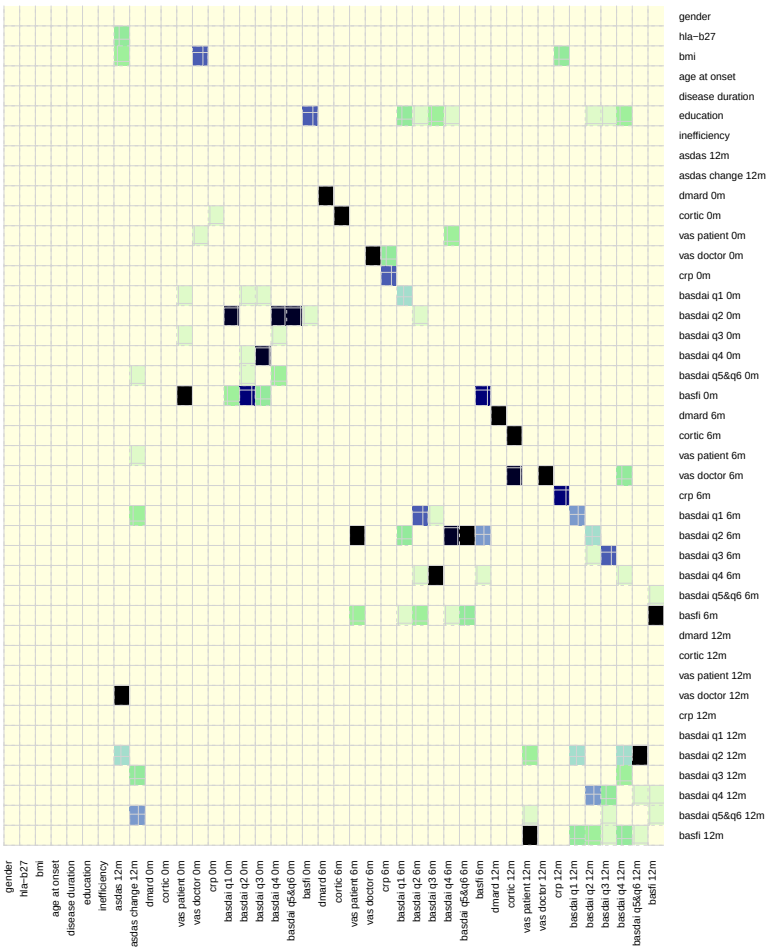


Figure 4.24: Heatmap obtained for infliximab using datasets with a 6-month interval between records and imputation in dynamic variables only.

baseline, where greater values for this variable appear to be associated with greater improvements.

To conclude, comparing these last three heatmaps to the ones built from datasets with a 3-month interval between records where also only dynamic variables were imputed, the same pattern is noted, with associations involving static variables being kept and associations between dynamic data becoming more clear, when comparing with the heatmaps created for datasets where no imputation was performed.

4.2.7 Summary of results for the different datasets

In this section, all correlations identified in the networks with each outcome variable, for each heatmap, together with the number of times that the respective association appeared in different networks are presented for adalimumab, etanercept and infliximab in Tables 4.20, 4.21 and 4.22, respectively.

Starting with inefficiency, it is possible to note that only adalimumab's networks present significant correlations between inefficiency and other variables. Indeed, etanercept's networks do not show any variable correlating with this outcome, while networks for infliximab present a single correlation with BASDAI questions 5 and 6 after 12 months of therapy initiation.

Table 4.20: Summary of variables correlating with outcome variables for adalimumab. The different outcome variables are presented in the columns, while the different heatmaps (varying interval between appointments and type of imputation) can be seen in the rows.

	Adalimumab		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
3MNoImp	VAS doctor 9m (6) BASFI 9m (5)	BMI (3) VAS doctor 0m (2) BASDAI q2 3m (2) BASFI 6m (2) CRP 6m (3) VAS doctor 12m (6)	BMI (9) VAS patient 0m (2) VAS doctor 0m (4) CRP 9m (1) BASDAI q3 9m (1) BASDAI q3 12m (1)
3MImp	VAS doctor 3m (3) VAS patient 9m (3) BASFI 12m (6)	VAS doctor 9m (2) VAS doctor 12m (7) BASDAI q2 12m (1) BASDAI q4 12m (5)	VAS patient 0m (7) CRP 0m (1) BASDAI q2 12m (3) BASDAI q3 12m (1) BASDAI q5&q6 12m (1)
3MImpDynam	VAS doctor 3m (3) VAS patient 9m (3) BASFI 12m (6)	VAS doctor 9m (2) VAS doctor 12m (7) BASDAI q2 12m (1) BASDAI q4 12m (5)	BMI (6) VAS patient 0m (2) BASDAI q2 12m (3) BASDAI q5&q6 12m (2)
6MNoImp	VAS doctor 0m (5) VAS doctor 6m (3) BASFI 6m (2)	BMI (9) VAS doctor 12m (9) BASFI 12m (1)	Education (3) VAS patient 0m (2) CRP 0m (6) BASDAI q5&q6 12m (1)
6MImp	BASFI 6m (3) VAS doctor 12m (3)	VAS doctor 12m (9) BASDAI q2 12m (3) BASDAI q4 12m (1) BASFI 12m (3)	VAS patient 0m (6) CRP 0m (6) BASDAI q2 12m (2) BASDAI q5&q6 12m (1)
6MImpDynam	BASFI 6m (3) VAS doctor 12m (3)	BMI (3) VAS doctor 12m (8) BASDAI q2 12m (2) BASDAI q4 12m (2) BASFI 12m (2)	BMI (3) VAS patient 0m (4) CRP 0m (5) BASDAI q5&q6 12m (1)

Among associations with inefficiency observed for adalimumab, it is noted that a significant amount of these are with variables recorded at more advanced time points. Indeed, it can be expected that the greater the time the patient is subject to a therapy, the more clearly its impact on the patient can be seen, and therefore these associations become more likely to arise. On the other hand, correlation between inefficiency and VAS doctor at 3 months was identified. In this case, it seems that patients with worse doctor's evaluation after 3 months of therapy have greater probability of therapy inefficiency, compared to patients with lower VAS doctor at that time, leading to the conclusion that this variable can act as an early indicator of therapy response. Indeed, it can be assumed that if after 3 months of treatment initiation the doctor's evaluation on the patient's condition is still high, there are more chances that the therapy will fail in the future, while contrarily, if the patient shows a positive response in the first 3 months, it can be an indicator that therapy will not fail due to inefficiency. Finally, for VAS doctor at baseline, it seems that, for patients with higher VAS doctor at therapy beginning, there is a higher proportion of people in which therapy is inefficient, than in patients where VAS doctor was smaller,

Table 4.21: Summary of variables correlating with outcome variables for etanercept. The different outcome variables are presented in the columns, while the different heatmaps (varying interval between appointments and type of imputation) can be seen in the rows.

	Etanercept		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
3MNoImp	—	VAS doctor 3m (1) BASFI 6m (4) VAS doctor 9m (3) BASFI 9m (6) VAS doctor 12m (4)	VAS doctor 3m (6) BASDAI q4 6m (1) VAS doctor 12m (6)
3MImp	—	Gender (1) VAS doctor 6m (8) VAS patient 12m (2) VAS doctor 12m (1) BASDAI q3 12m (3)	CRP 0m (7) BASDAI q3 0m (1) VAS patient 12m (1) BASDAI q2 12m (1) BASDAI q4 12m (2)
3MImpDynam	—	Disease duration (1) VAS doctor 6m (8) VAS patient 12m (2) VAS doctor 12m (1) BASDAI q3 12m (3)	CRP 0m (7) BASDAI q3 0m (1) VAS patient 12m (1) BASDAI q2 12m (1) BASDAI q4 12m (2)
6MNoImp	—	BASFI 6m (4) VAS patient 12m (1) VAS doctor 12m (9) BASFI 12m (2)	CRP 0m (5) BASDAI q4 6m (5) VAS patient 12m (2)
6MImp	—	VAS patient 12m (1) VAS doctor 12m (9) BASDAI q2 12m (2) BASDAI q3 12m (1) BASFI 12m (2)	CRP 0m (7) VAS patient 12m (1) BASDAI q1 12m (1) BASDAI q2 12m (2) BASDAI q3 12m (1)
6MImpDynam	—	VAS patient 12m (1) VAS doctor 12m (9) BASDAI q2 12m (2) BASDAI q3 12m (1) BASFI 12m (2)	CRP 0m (7) VAS patient 12m (1) BASDAI q1 12m (1) BASDAI q2 12m (2) BASDAI q3 12m (1)

being hypothesized that having higher VAS doctor at baseline can in some situations lead to therapy inefficiency.

Regarding ASDAS at 12 months, looking into the different correlations obtained for the 3 biologics, it is possible to note that this outcome correlates mainly with variables recorded 12 months after therapy initiation. Indeed, since ASDAS is a score that measures disease activity, it is expected that other dynamic variables reflecting any aspect of the patient's condition evaluated at this time may evolve in the same direction than this outcome variable, leading to the existence of correlation between them. As an example, correlation with VAS doctor after 12 months is present in most of the networks. As mentioned, ASDAS is a score that evaluates disease activity, and since VAS doctor reflects the doctor's global evaluation on the patient's condition, it is expected that these two values are correlated. Moreover, CRP and BASDAI questions 2, 3 and 6 are part of the ASDAS calculation formula, being the correlation between them expected.

Table 4.22: Summary of variables correlating with outcome variables for infliximab. The different outcome variables are presented in the columns, while the different heatmaps (varying interval between appointments and type of imputation) can be seen in the rows.

	Infliximab		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
3MNoImp	—	BASDAI q5&q6 3m (3) VAS doctor 9m (9) VAS patient 12m (3)	VAS doctor 0m (9) CRP 3m (1) BASDAI q1 9m (2) BASFI 9m (1) VAS patient 12m (2)
3MImp	BASDAI q5&q6 12m (1)	VAS doctor 12m (9) BASDAI q2 12m (4) BASDAI q4 12m (1)	CRP 0m (1) BASFI 0m (1) VAS doctor 3m (3) VAS patient 12m (6) BASFI 12m (1)
3MImpDynam	BASDAI q5&q6 12m (1)	Education (1) VAS doctor 12m (9) BASDAI q2 12m (4) BASDAI q4 12m (1)	CRP 0m (1) BASFI 0m (1) VAS doctor 3m (3) VAS patient 12m (6) BASFI 12m (1)
6MNoImp	—	VAS doctor 12m (9) BASDAI q2 12m (2)	BASDAI q1 6m (8) BASDAI q5&q6 12m (1)
6MImp	—	VAS doctor 12m (9) BASDAI q2 12m (6)	BASDAI q5&q6 0m (1) VAS patient 6m (1) BASDAI q1 6m (2) BASDAI q3 12m (3) BASDAI q5&q6 12m (5)
6MImpDynam	—	HLA-B27 (3) BMI (2) VAS doctor 12m (9) BASDAI q2 12m (4)	BASDAI q5&q6 0m (1) VAS patient 6m (1) BASDAI q1 6m (2) BASDAI q3 12m (3) BASDAI q5&q6 12m (5)

Concerning static variables associated with ASDAS at 12 months, a strong correlation between this variable and BMI was found for biologic adalimumab. Indeed, it is known that patients with greater BMI respond worse to therapy, while patients with smaller BMI show better response, which was confirmed in the obtained networks by having smaller BMI more associated with smaller ASDAS at 12 months and greater BMI more associated with greater ASDAS at 12 months. Nevertheless, this correlation appeared mainly in patients being treated with adalimumab, which may indicate that the BMI impact is stronger for patients being treated with this therapy, comparing to patients being treated with etanercept or infliximab. In literature, there was no evidence about this fact, however, when discussing this finding with medical professionals, they mentioned that although no studies have yet been done on this field, it was something that they had already noticed in their clinical practice. The proposed reason for this was the fact that for both etanercept and infliximab the amount given to the patients is based on their body weight, which doesn't happen with adalimumab, where instead a fixed amount is given. Therefore, this may lead to obese patients being treated with adalimumab not receiving enough amount of medication needed to produce good therapy response, and therefore, this may lead to the conclusion that this therapy

should be avoided for obese patients. On the other hand, for etanercept, two correlations with static variables were found, being those variables gender and disease duration. Nevertheless, each correlation was only found in a single heatmap, in a single network, indicating that this association may not be very strong. Finally, for infliximab, three static variables were found to correlate with ASDAS at 12 months, education, HLA-B27 and BMI. Regarding association with HLA-B27, it seems through the CPT that HLA-B27 positive patients may respond better to therapy, however, this is not clear. For associations with education and BMI, no clear pattern was identified.

Finally, concerning ASDAS change at 12 months, several correlations were obtained between this outcome variable and other variables, for the different biologics. ASDAS improvement is obtained by subtracting the ASDAS value at 12 months to the ASDAS value at baseline, therefore depending on ASDAS in these two time points. When looking into the identified correlations, a considerable amount was indeed with variables at baseline and 12 months after therapy. It could be expected that the identified correlations were with variables at these two time points simultaneously, nevertheless it was not always the case. Indeed, several single correlations were found.

Regarding variables at baseline, some that presented a greater amount of correlation were CRP, VAS patient and VAS doctor. These variables can be seen as reflecting the level of disease activity of the patient, like ASDAS does. In fact, while CRP correlates positively with ASDAS value, according to the calculation formula, VAS patient and VAS doctor intend to express the global state of the patient, either from a self or a doctor's perspective. When correlating individually with ASDAS improvement, a greater value of these variables at baseline was associated with a greater improvement. Nevertheless, these findings should be evaluated carefully, since patients with higher disease activity at baseline, and subsequently a greater ASDAS, have more room for improvement, i.e., for decreasing their ASDAS score, than patients who already have smaller ASDAS at baseline. Thus, it is expected that the ones that suffered greater improvements were the ones with higher disease activity at therapy initiation, but it does not mean that patients with a smaller disease activity at beginning will not respond positively to these therapies. Indeed, this hypothesis gains strength when noticing that these variables did not appear correlated with ASDAS at 12 months, possibly meaning that the association is more with ASDAS improvement than with the disease activity level itself.

The same logic can be applied for dynamic variables after 12 months of therapy, where greater values of VAS doctor/ patient, which reflect a greater disease activity, are expected to be associated with smaller improvements, since the activity is still elevated after 12 months of therapy, while smaller values have a greater probability of being associated with greater improvements.

On the other hand, correlation of ASDAS change with static variables can be of great interest to understand which patient's characteristics makes them more likely of responding well to the therapy. For adalimumab, a great correlation between this outcome variable and BMI was identified, with greater BMI being associated with a smaller ASDAS improvement, confirming the hypothesis that adalimumab might not be the best therapeutic option for obese patients. Correlation between education and ASDAS change was also identified for adalimumab in one of the networks, however, when looking into the CPTs it was not possible to identify any pattern in the influence of education on ASDAS change. On the other

hand, for etanercept and infliximab, no correlation with static variables was found.

To conclude, associations between VAS doctor at 3 months and ASDAS change at 12 months were also found for biologics etanercept and infliximab, with smaller values for the doctor's evaluation being more associated with greater ASDAS improvements, which can indicate that VAS doctor might be an early indicator of patient's response to therapy.

4.3 Chi-square tests of independence

In order to verify some previously obtained associations between variables, Chi-square (χ^2) tests of independence were performed. A Chi-square test determines whether categorical variables are independent or related. This is a statistical hypothesis test in which there is a null hypothesis that assumes that no relationship exists between the two categorical variables, i.e., that they are independent, with the purpose of evaluating how likely the observed frequencies would be if the null hypothesis is true. This test uses a contingency table to analyze the data, in which categories of one variable appear in the rows and categories of the other appear in the columns, with each cell presenting the observed count for a pair of categories [68–70]. Then, the χ^2 value is calculated using the formula:

$$X^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}, \tag{4.1}$$

with O_i being the observed values and E_i the expected ones. In turn, this value is used to calculate a p-value, which represents the probability of obtaining results 'as extreme' or 'more extreme' than the ones obtained, assuming the null hypothesis is true, i.e, given that there is no correlation between the variables. Choosing a significance level, usually α=0.05 (5%), which was the significance level also chosen in the present study, if the p-value is smaller than this value, then the null hypothesis should be rejected, meaning that the variables are not independent of each other and that there is a statistically significant association between them. On the contrary, if the p-value is greater than the significance level, we fail to reject the null hypothesis [71].

Since the main goal of this study is to help predict therapy outcome, and the ideal is to obtain associations between outcomes and variables that can be measured before therapy initiation, these tests were performed only for static variables and for VAS patient, VAS doctor and CRP at baseline, which were the three dynamic variables at baseline that more commonly appeared correlated with outcome variables in the networks, and for which the identification of a pattern was possible. The tests were performed for all different biologic therapies and the corresponding results are presented in Tables 4.23, 4.24 and 4.25, for adalimumab, etanercept and infliximab, respectively. Moreover, the contingency tables for the pairs of variables where the p-value was smaller than 0.05, i.e., the pairs of variables that present a statistically significant relationship between them, are presented in Appendix C.

First, starting with inefficiency, previously learnt networks found few associations of other variables with this outcome variable, being most of the ones found with dynamic variables at advanced time points. Looking into the p-values obtained for tests with this variable, the null hypothesis can only be rejected

Table 4.23: P-values obtained from Chi-square tests for adalimumab.

	Adalimumab		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
Gender	0.104	0.000	0.002
HLA-B27	0.112	0.399	0.442
BMI	0.922	0.004	0.008
Age at biologic onset	0.686	0.003	0.102
Disease duration	0.821	0.001	0.855
Education	0.552	0.277	0.064
VAS patient at 0M	0.064	0.090	0.000
VAS doctor at 0M	0.093	0.531	0.060
CRP at 0M	0.805	0.825	0.000

Table 4.24: P-values obtained from Chi-square tests for etanercept.

	Etanercept		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
Gender	0.338	0.001	0.028
HLA-B27	0.315	0.471	0.403
BMI	0.126	0.809	0.748
Age at biologic onset	0.881	0.002	0.966
Disease duration	0.370	0.000	0.211
Education	0.661	0.180	0.085
VAS patient at 0M	0.140	0.002	0.001
VAS doctor at 0M	0.814	0.129	0.043
CRP at 0M	0.758	0.733	0.000

Table 4.25: P-values obtained from Chi-square tests for infliximab.

	Infliximab		
	Inefficiency	**ASDAS at 12 months**	**ASDAS change at 12 months**
Gender	0.741	0.033	0.042
HLA-B27	0.014	0.038	0.068
BMI	0.904	0.131	0.105
Age at biologic onset	0.231	0.522	0.118
Disease duration	0.859	0.185	0.794
Education	0.424	0.018	0.259
VAS patient at 0M	0.351	0.130	0.707
VAS doctor at 0M	0.213	0.570	0.601
CRP at 0M	0.860	0.769	0.002

in one situation. Indeed, for infliximab, inefficiency and HLA-B27 seem to be related, although this association was not found in the networks. When looking into the contingency table obtained for the pair inefficiency & HLA-B27, presented in Table C.1, it is noticed that for patients that don't have the HLA-B27 gene, the number of patients that experienced therapy inefficiency was higher than expected if there was no statistically significant relationship between the two variables, while the number of patients that didn't experience therapy inefficiency was smaller. Contrarily, for HLA-B27 positive patients, the number of patients experiencing therapy failure due to inefficiency was smaller than expected, while the number of patients not experiencing therapy inefficiency was greater than expected. These findings

seem to indicate that infliximab might not be the best therapeutic option for HLA-B27 negative patients, nevertheless, this association should be further studied in order to be confirmed or rejected. Additionally, it should be noted that despite association between inefficiency and HLA-B27 haven't been found in the networks, association between HLA-B27 and ASDAS at 12 months was indeed found in the networks for infliximab.

On the other hand, for ASDAS and ASDAS change at 12 months, several associations with other variables were found to be statistically significant. Therefore, these associations will be evaluated by looking into the static/baseline variables and evaluating to what outcomes they correlate, instead of looking into the outcome variable itself and seeing which variables correlate with it.

Following the previous idea, looking into HLA-B27, this variable also appears to be correlated with ASDAS at 12 months for patients taking infliximab. Indeed, infliximab was the only biologic for which HLA-B27 presented any statistically significant association with any outcome variable. As mentioned, this finding is in line with the association previously found in the networks, where HLA-B27 appeared as a parent of ASDAS at 12 months. In this case, as ASDAS can take a greater number of values, it becomes more difficult to find a pattern in the data just by looking at the contingency table, presented in Table C.2. Nevertheless, looking at it, it is possible to note that for HLA-B27 negative patients, the number of patients in the lowest ASDAS disease activity level is smaller than expected, while for HLA-B27 positive patients the number of patients in a disease inactive state is higher than expected, which might again indicate that HLA-B27 positive patients respond better to infliximab's therapy than HLA-B27 negative patients. This association should be further studied in order to be confirmed or rejected.

Moving into gender, the networks revealed a single correlation between this variable and an outcome variable, in this case ASDAS at 12 months, for etanercept. On the other hand, looking into the p-values resultant from the Chi-square tests, it is possible to note that the values obtained between gender and both ASDAS and ASDAS improvement at 12 months are smaller than the significance level 0.05 for all three biologics studied. Therefore, the null hypothesis should be rejected for all therapies, meaning that there is a statistically significant relationship between gender and these two variables, for every considered therapy. According to the literature, male patients respond better to anti-TNF therapies, nevertheless, Chi-square tests only reveal if there is a dependency between the two variables, but not the direction of the correlation. However, by looking into the contingency tables obtained, presented in Tables C.3 to C.8, it is noted that the number of male patients with smaller ASDAS is greater than expected, while with greater ASDAS is smaller than expected. On the other hand, for female patients, the number of the ones with smaller ASDAS is smaller than expected, while the number of female patients with higher ASDAS is higher than expected, for all three biologics. In turn, for ASDAS change, male patients also seem to experience better improvements, by presenting a number of greater improvements higher than expected and a number of smaller improvements smaller than expected, for all three biologics. Following the same logic, for female patients the exact opposite outcomes are verified, being more associated with smaller improvements. Indeed, this pattern was identified in patients treated with all biologic therapies, which is in accordance to the literature.

Next, regarding BMI, correlation between this variable and ASDAS/ASDAS improvement in patients

treated with adalimumab was one of the strongest found in the previously learnt networks, leading to the hypothesis that this correlation may be stronger for this biologic. To confirm this results Chi-square tests were performed for all 3 biologics. When looking into the obtained p-values for the different biologics, and considering a significance level of 0.05, it is possible to conclude that only for adalimumab is possible to reject the null hypothesis both for ASDAS at 12 months and ASDAS change at 12 months, i.e., only for patients being treated with adalimumab it's possible to conclude that the pairs BMI & ASDAS at 12 months and BMI & ASDAS change at 12 months are related. Looking into the contingency table for BMI & ASDAS at 12 months, presented in Table C.9, it is noticed that, for smaller BMI values the number of patients with smaller ASDAS after 12 months is greater than expected. On the other hand, for higher BMI, the number of patients with smaller ASDAS is smaller than expected and the number of patients with higher ASDAS is higher than expected, corroborating the results found through the DBNs. Furthermore, for the pair BMI & ASDAS change at 12 months, looking into the contingency table shown in Table C.10, it is noticed that the number of patients with smaller BMI experiencing greater improvements is higher than expected, while the number of patients experiencing smaller improvements is lower than expected. On the other hand, for patients with higher BMI, the opposite pattern is found, again confirming the previously obtained results. On the contrary, for etanercept and infliximab, an association between these variables was not identified.

On the other hand, concerning age at biologic onset, this variable did not appear correlated with any outcome variable in the learnt networks. However, looking into the p-values obtained for Chi-square tests between this variable and outcome variables, the corresponding value is smaller than 0.05 for correlation with ASDAS at 12 months for adalimumab and etanercept. According to the literature, younger age is a predictor of good clinical response to anti-TNF therapy. When looking into the contingency table obtained for both adalimumab and etanercept, presented in Tables C.11 and C.12, respectively, this pattern seems to be confirmed, with the number of younger patients presenting smaller ASDAS after 12 months being higher than expected and presenting higher ASDAS at 12 months being smaller than expected. On the contrary, the number of older patients presenting smaller ASDAS is lower than expected, while the number of older patients presenting higher ASDAS is higher than expected. This association, however, is not noticed in the infliximab's dataset, which might indicate that the age factor is not so relevant in patients taking this biologic, and therefore it might be a good alternative for older patients, while adalimumab or etanercept could be preferred for younger patients. Nevertheless, these associations should be further studied.

The next static variable studied was disease duration. In the networks, this variable appeared correlated with ASDAS at 12 months for etanercept, being the only identified correlation with outcome variables. Looking into the results of the Chi-square tests, the p-value was smaller than 0.05 for the pair disease duration and ASDAS at 12 months for the biologics adalimumab and etanercept, meaning that for these two therapies it is possible to reject the null hypothesis and assume that the two variables are related. When looking into the contingency tables obtained, presented in Tables C.13 and C.14, respectively for adalimumab and etanercept, it seems that having higher disease duration has a negative impact on the ASDAS value after 12 months, with patients experiencing higher ASDAS more frequently

than expected if variables were independent. On the contrary, having a smaller disease duration has the opposite impact.

These findings however, are not observed in the infliximab's dataset, as it was the case for age at biologic onset, which might indicate that disease duration does not influence patients treated with this therapy as much as it influences in patients treated with either adalimumab or etanercept. Since patients don't seem to respond worse with age, this may lead to the assumption that probably infliximab is a good therapeutic option for patients with higher disease duration. It should be noted that disease duration and age at biologic onset might be associated, in the sense that patients in which disease starts earlier may have higher disease duration.

Finally, the last static variable evaluated was education, which appeared correlated with two other variables in the built networks: on the one hand, for adalimumab, it appeared correlated with ASDAS improvement and, on the other hand, it appeared correlated with ASDAS at 12 months for infliximab. For both cases, it was not possible to identify any pattern in the influence of education by looking into the CPTs. Examining the obtained p-values, a value smaller than 0.05 was obtained for the pair ASDAS at 12 months and education for infliximab's data, nevertheless, looking into the contingency table obtained, presented in Table C.15, it is also not possible to identify any pattern in the data. Nevertheless, as the association between education and ASDAS at 12 months was identified both through Bayesian networks and Chi-square tests, the possibility that education influences ASDAS at 12 months in a certain way in patient's being treated with infliximab cannot be excluded and therefore further studies should be conducted.

Regarding VAS patient at baseline, the results from the Chi-square tests present a statistically significant relationship with ASDAS improvement, for patients being treated with adalimumab. Indeed, this association had already appeared several times in the previously learnt networks, where patients with higher VAS patient were associated with higher improvements and patients with smaller VAS patient were associated with smaller improvements, which was indeed the pattern found when looking into the contingency table for this pair of variables, presented in Table C.16. The fact that this variable appears associated only with ASDAS change and not with ASDAS itself seems to indicate that this variable influences the improvement for each patient, but not the final state of disease activity after therapy, with a possible explanation being that patients with higher VAS patient at baseline are expected to have a higher ASDAS at baseline, and therefore they have more room for decreasing it than patients that start the treatment already with a smaller ASDAS. Nevertheless, it does not mean that patients that start already with a smaller ASDAS do not have a small ASDAS after 12 months of therapy. On the other hand, for etanercept, in the obtained networks this variable did not appear to be correlated with any outcome variable, however, looking into the p-values that resulted from the Chi-square tests, VAS patient at baseline and both ASDAS at 12 months and ASDAS improvement appear to be related. For ASDAS change, the same pattern as for adalimumab seems to be identified by looking into the contingency table presented in Table C.18, with smaller values of VAS patient being associated with smaller improvements and higher values being associated with higher improvements. On the other hand, for ASDAS at 12 months, looking into the obtained contingency table in Table C.17, it seems that a smaller VAS patient

at baseline is more associated with smaller ASDAS after 12 months and less with higher ASDAS, while greater VAS patient at baseline seems to be more associated with higher ASDAS values after 12 months and less with smaller ASDAS values. Again, this finding seems to strengthen the conclusion that a higher VAS patient at baseline can be translate into more room for improvement, but not as a predictor of smaller disease activity after 12 months of therapy. Finally, for infliximab, the Chi-square tests did not reveal any statistically significant association between VAS patient at baseline and any other variable, which is in line with the findings from the networks, where no association was found.

Next, concerning VAS doctor, the results from the Chi-square tests only found a statistically significant relationship between this variable and ASDAS improvement 12 months after therapy initiation for patients being treated with etanercept, contrarily to what was found in the networks, where this correlation was obtained for both adalimumab and infliximab, but not etanercept. By looking into the corresponding contingency table, presented in Table C.19, it seems that, as before, the number of patients with a smaller VAS doctor at baseline that experience smaller improvements is greater than expected and that the number of patients with a greater VAS doctor at baseline that experience greater improvements is greater than expected.

Finally, regarding CRP at baseline, looking into the obtained p-values, it presented a statistically significant relationship with ASDAS improvement, but not with ASDAS, for all three biologics studied, in line with the correlations obtained in the networks. Again, as can be seen in the contingency tables presented in Tables C.20 to C.22, for all three biologics the pattern is the same: for smaller levels of CRP at baseline, the number of patients experiencing smaller improvements is higher than expected and the number of patients experiencing higher improvements is smaller than expected. Following the same logic, for higher levels of CRP at baseline, the number of patients experiencing smaller improvements is smaller than expected while the number of patients experiencing higher improvements is higher than expected. In fact, again, this can be seen as an indicator of the hypothesis presented regarding the room for improvement, with patients with higher CRP at baseline, and thus higher ASDAS, having higher possibility of greatly improving their ASDAS value than patients whose ASDAS score is already smaller at the beginning, and not of the possibility of CRP being a predictor of smaller or higher disease activity state after therapy.

To conclude, these tests allowed us to confirm some previously found associations, and to unravel new ones. However, it should be noted that Chi-square tests only measure independence between variables, not being able to provide any inferences about causation, or identify any tendencies in the data. For cases where both variables take only a few values, it is, however, possible to look at the obtained contingency tables and try to identify a pattern, which was done in this study. This way, these tests were done to strengthen previously found associations, even if the direction and pattern of the association was unknown, in order to provide further starting points for future studies.

Chapter 5

Conclusions

Due to the need of improving decision-making in ankylosing spondylitis, it is essential to improve the prognosis process, by being able to predict the most adequate treatment option according to the patient's characteristics. In this work, the approach used allowed to identify potential risk factors associated with the different AS therapies, working as a preliminary study in trying to distinguish between different anti-TNF agents.

DBNs are graphical representations of probabilistic dependencies among random variables over time. A major strength of this approach is, therefore, the explicit representation of the relations between the different variables and the pathways along which each variable influences the clinical outcomes.

Additionally, the R package chosen to estimate the DBNs allows the presence of missing data, which was also a great advantage due to the considerable amount of unobserved values in the dataset. Nevertheless, despite the possibility of having missing values in the data, several processes were conducted in order to try to reduce them, since drawing conclusions from datasets with a great amount of missing values might also not lead to the best conclusions. Therefore, some incomplete observations were deleted and some missing values were replaced with an estimated value based on the information available. However, both these methods can as well significantly impact the conclusions drawn from the data. Indeed, imputation inevitably introduces errors in the dataset, mainly if performed for time independent variables, which was the case in this study, so conclusions should be drawn carefully and having this aspect in mind.

On the other hand, another limitation of this work is that the DBN models required variables to be discretized. For variables where theoretical reference ranges exist, the preferred approach was to use those values to discretize the variables. Nevertheless, for some variables there were no reference values defined, while for others these ranges were not suited to the data, since some categories ended up with very few observations, while others had too many. This way, in cases like this, it was opted to discretize based on quantiles, although the obtained discretization may not be the one that better translates what are considered, or not, normal values for the variables, in clinical context.

Despite the challenges faced during the development of this work, the proposed objectives were achieved. Indeed, it was possible to create different datasets to be used in further studies regarding AS, as well as to identify possible predictors of response for the different biologics.

To achieve more reliable conclusions, Chi-square tests were performed as a complement of Bayesian networks. Indeed, if an association was found both in the networks and through the Chi-square tests, it

was possible to have a better confidence on its validity.

Next, it is noteworthy the different results obtained for each therapy studied, adalimumab, etanercept and infliximab.

First, regarding adalimumab's patients, one of the most relevant associations unravelled through both methods was with BMI, where greater values for this variable are associated with a poorer response to therapy and smaller values are associated with a better response. Indeed, BMI seems to have a greater impact on patients treated with adalimumab than on patients treated with either etanercept or infliximab, indicating that this might not be the best therapeutic option for obese patients.

Next, regarding adalimumab's and etarnecept's patients, associations with age at therapy onset and disease duration were identified, with higher values for both these variables being associated with higher disease activity levels after 12 months of treatment. The fact that this association was not identified for infliximab's patients seems to indicate that for older patients or patients with greater disease duration, infliximab might be a good therapeutic option.

Finally, regarding infliximab's patients, it was hypothesized that the impact of HLA-B27 could be stronger in these patients than in patients treated either with adalimumab or etanercept, with HLA-B27 positive patients responding better to therapy than HLA-B27 negative patients. Therefore, infliximab might not be the best treatment choice for HLA-B27 negative patients.

Moreover, literature findings were also corroborated with this study, as it was the case of gender, where male patients were found to respond better to all three therapies.

Additionally, it was also noticed that the therapy's effect is already visible after 3 months, being hypothesized that an absence of response, or a smaller response after this period might be an indicator of therapy inefficiency.

5.1 Achievements

Firstly, it is important to note that this work was developed using data of AS patients from Reuma.pt, whose potential had not been fully exploited yet. Indeed, the dynamic component of the data, where time-dependent records were selected according to a specific granularity had never been handled, making the whole processing work more time consuming. Nevertheless, this work enabled the creation of several datasets that can be easily used in future studies, allowing to save time in the processing steps, being this a great achievement of this work.

Furthermore, this work allowed to identify differences in the response process of the patients, regarding the therapy to which they have been subjected to. Again, it should be noted that this work was a pioneer in the field, being the first to try to distinguish between different anti-TNF therapies, generating hypothesis and insights that should be further analysed by clinicians.

5.2 Future Work

During the development of this work, several aspects were identified as possible suggestions for future work.

First, for the present analysis, only data regarding patient's first anti-TNF therapy was considered. Since patients are known to respond differently to therapy if it's not the first anti-TNF treatment they are receiving, data for posterior therapies should also be studied, in order to evaluate differences in response.

Secondly, concerning the outcomes studied in this work, the ones selected were therapy inefficiency, ASDAS score value 12 months after therapy initiation and ASDAS improvement at the same time. Nevertheless, adverse events are also one of the greatest reasons for withdrawing therapy, and could also be studied and evaluated, in order to identify predictors that may lead to an adverse event, which are probably different than the ones for the outcome variables used in this study.

Furthermore, the drug survival time, i.e., the time during which the patient keeps taking the same drug, was not taken into account in this study, being only considered if that biologic ended up being inefficient or not. This way, it could also be interesting to study this variable.

On the other hand, in the present dissertation, the `bnstruct` R package was used to learn the BN. Nevertheless, different algorithms can be used the different structures obtained can be compared. Additionally, DBNs can be used to predict the patient's temporal evolution, by predicting how the dynamic variables evolve over time and how they influence each other, and thus perform simulation analyses about the potential effectiveness of different therapies. In this work, this predictive capabilities of DBN were not used, however, in future works, it would be a great functionality to explore.

Bibliography

[1] EULAR. Horizon 2020 framework programme eular's position and recommendations horizon 2020 framework programme eular's position and recommendations. 2011.

[2] American College of Rheumatology. Rheumatic Diseases in America : The Problem, The Impact, The Answers. pages 1–16, 2013.

[3] J. C. Branco, A. M. Rodrigues, N. Gouveia, M. Eusébio, S. Ramiro, P. M. Machado, L. P. Da Costa, A. F. Mourão, I. Silva, P. Laires, A. Sepriano, F. Araújo, S. Gonçalves, P. S. Coelho, V. Tavares, J. Cerol, J. M. Mendes, L. Carmona, and H. Canhão. Prevalence of rheumatic and musculoskeletal diseases and their impact on health-related quality of life, physical function and mental health in portugal: results from epireumapt- a national health survey. *RMD Open*, 2(1), 2016.

[4] B. Fautrel and F. Guillemin. Cost of illness studies in rheumatic diseases. *Current Opinion in Rheumatology*, 14(2):121–126, 2002.

[5] Arthritis Society. Ankylosing Spondylitis. `https://arthritis.ca/about-arthritis/arthritis-types-(a-z)/types/ankylosing-spondylitis`. Accessed: 2020-04-24.

[6] Spondylitis Association of America. Overview of Ankylosing Spondylitis. `https://spondylitis.org/about-spondylitis/types-of-spondylitis/ankylosing-spondylitis/`. Accessed: 2020-04-24.

[7] J. Sieper, J. Braun, M. Rudwaleit, A. Boonen, A. Zink. Ankylosing spondylitis: an overview. *Annals of the Rheumatic Diseases*, 61(Suppl 3), 2002.

[8] National Health Service England. Treatment ankylosing spondylitis. `https://www.nhs.uk/conditions/ankylosing-spondylitis/treatment/`. Accessed: 2020-04-25.

[9] T. Ali, S. Kaitha, S. Mahmood, A. Ftesi, J. Stone, and M. S. Bronze. Clinical use of anti-TNF therapy and increased risk of infections. *Drug, Healthcare and Patient Safety*, 5:79–99, 2013.

[10] J. R. Maneiro, A. Souto, E. Salgado, A. Mera, J. Gomez-Reino. Predictors of response to TNF antagonists in patients with ankylosing spondylitis and psoriatic arthritis: systematic review and meta-analysis. *RMD Open*, 1(1), 2014.

[11] Radiopaedia. Ankylosing spondylitis. `https://radiopaedia.org/articles/ankylosing-spondylitis-1`. Accessed: 2020-04-26.

[12] F. Mahmood and P. Helliwell. Ankylosing spondylitis: A review. *European Medical Journal*, 2(4): 134–139, 2017.

[13] L. E. Dean, G. T. Jones, A. G. Macdonald, C. Downham, R. D. Sturrock, and G. J. Macfarlane. Global prevalence of ankylosing spondylitis. *Rheumatology (United Kingdom)*, 53(4):650–657, 2013.

[14] C. Stolwijk, M. van Onna, A. Boonen, and A. van Tubergen. The Global Prevalence of Spondyloarthritis: A Systematic Review and Meta-Regression Analysis. *Arthritis Care Research*, 68(9): 1320–1331, 2015.

[15] REHAB My Patient. Ankylosing spondylitis. `https://www.rehabmypatient.com/thoracic-spine/ankylosing-spondylitis`. Accessed: 2020-04-27.

[16] W. Zhu, X. He, K. Cheng, L. Zhang, D. Chen, X. Wang, G. Qiu, X. Cao, and X. Weng. Ankylosing spondylitis: etiology, pathogenesis, and treatments . *Bone Research*, 7, 2019.

[17] G. Layh-Schmitt and R. A. Colbert. The interleukin-23/interleukin-17 axis in spondyloarthritis. *Current Opinion in Rheumatology*, 20:392–397, 2008.

[18] M. H. Abdelrahman, S. Mahdy, I. A. Khanjar, A. M. Siam, H. A. Malallah, S. A. Al-Emadi, H. A. Sarakbi, and M. Hammoudeh. Prevalence of HLA-B27 in patients with ankylosing spondylitis in Qatar. *International Journal of Rheumatology*, 2012.

[19] A. R. Roberts, L. H. Appleton, A. Cortes, M. Vecellio, J. Lau, L. Watts, M. A. Brown, and P. Wordsworth. ERAP1 association with ankylosing spondylitis is attributable to common genotypes rather than rare haplotype combinations. *Proceedings of the National Academy of Sciences of the United States of America*, 114(3):558–561, 2017.

[20] V. R. Moulton. *Cytokines*. Elsevier, 2016.

[21] B. Osta, G. Benedetti, and P. Miossec. Classical and Paradoxical Effects of TNF- on Bone Homeostasis. *Frontiers in Immunology*, 5, 2014.

[22] P. Miossec. Update on interleukin-17: A role in the pathogenesis of inflammatory arthritis and implication for clinical practice. *RMD Open*, 3(1), 2017.

[23] J. Sieper, M. Rudwaleit, X. Baraliakos, J. Brandt, J. Braun, R. Burgos-Vargas, M. Dougados, K. G. Hermann, R. Landewe, W. Maksymowych, and D. Van Der Heijde. The Assessment of SpondyloArthritis international Society (ASAS) handbook: A guide to assess spondyloarthritis. *Annals of the Rheumatic Diseases*, 68(Suppl 2):ii1–ii44, 2009.

[24] Mayo Clinic. Ankylosing spondylitis. `https://www.mayoclinic.org/diseases-conditions/ankylosing-spondylitis/diagnosis-treatment/drc-20354813`. Accessed: 2020-04-28.

[25] National Axial Spondyloarthritis Society. Biologic therapys. `https://nass.co.uk/managing-my-as/medication/biologic-therapy/`. Accessed: 2020-04-27.

[26] F. Cantini, L. Niccoli, and D. Goletti. Adalimumab, Etanercept, Infliximab, and the Risk of Tuberculosis: Data from Clinical Trials, National Registries, and Postmarketing Surveillance. *The Journal of Rheumatology*, 91:47–55, 2014.

[27] D. Van Der Heijde, S. Ramiro, R. Landewé, X. Baraliakos, F. Van Den Bosch, A. Sepriano, A. Regel, A. Ciurea, H. Dagfinrud, M. Dougados, F. Van Gaalen, P. Géher, I. Van Der Horst-Bruinsma, R. D. Inman, M. Jongkees, U. Kiltz, T. K. Kvien, P. M. Machado, H. Marzo-Ortega, A. Molto, V. Navarro-Compàn, S. Ozgocmen, F. M. Pimentel-Santos, J. Reveille, M. Rudwaleit, J. Sieper, P. Sampaio-Barros, D. Wiek, and J. Braun. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Annals of the Rheumatic Diseases*, 76(6):978–991, 2017.

[28] A. R. Cravo, V. Tavares, and J. Canas Da Silva. Terapêutica anti-TNF alfa na espondilite anquilosante. *Acta Medica Portuguesa*, 19(2):141–150, 2006.

[29] S. Arends, E. Brouwer, E. van der Veer, H. Groen, M. K. Leijsma, P. M. Houtman, T. L. Th A Jansen, C. G. Kallenberg, and A. Spoorenberg. Baseline predictors of response and discontinuation of tumor necrosis factor-alpha blocking therapy in ankylosing spondylitis: a prospective longitudinal observational cohort study . *RMD Open*, 13(3), 2011.

[30] P. A. Lord, T. M. Farragher, M. Lunt, K. D. Watson, D. P. Symmons, and K. L. Hyrich. Predictors of response to anti-TNF therapy in ankylosing spondylitis: Results from the British Society for Rheumatology Biologics Register. *Rheumatology*, 49(3):563–570, 2009.

[31] N. Vastesaeger, D. Van Der Heijde, R. D. Inman, Y. Wang, A. Deodhar, B. Hsu, M. U. Rahman, B. Dijkmans, P. Geusens, B. V. Cruyssen, E. Collantes, J. Sieper, and J. Braun. Predicting the outcome of ankylosing spondylitis therapy. *Annals of the Rheumatic Diseases*, 70(6):973–981, 2011.

[32] M. Lorenzin, A. Ortolan, P. Frallonardo, F. Oliviero, L. Punzi, and R. Ramonda. Predictors of response and drug survival in ankylosing spondylitis patients treated with infliximab. *BMC Musculoskeletal Disorders*, 16(1):10–17, 2015. ISSN 14712474.

[33] G. J. MacFarlane, E. Pathan, G. T. Jones, and L. E. Dean. Predicting response to anti-TNFα therapy among patients with axial spondyloarthritis (axSpA): Results from BSRBR-AS. *Rheumatology (United Kingdom)*, 59(9):2481–2490, 2020. ISSN 14620332. doi: 10.1093/rheumatology/kez657.

[34] E. Gremese, S. Bernardi, S. Bonazza, M. Nowik, G. Peluso, A. Massara, B. Tolusso, L. Messuti, M. C. Miceli, A. Zoli, F. Trotta, M. Govoni, and G. Ferraccioli. Body weight, gender and response to TNF- blockers in axial spondyloarthritis. *Rheumatology*, 53:875–881, 2014.

[35] J. Zochling. Measures of symptoms and disease status in ankylosing spondylitis. *Arthritis Care and Research*, 63(Suppl 4), 2011.

[36] J. Zochling, J. Braun, and D. van der Heijde. Assessments in ankylosing spondylitis. *Best Practice and Research: Clinical Rheumatology*, 20(3):521–537.

[37] Assessment of SpondyloArthritis. Mission statement. `https://www.asas-group.org/about-asas/mission-statement/`. Accessed: 2020-04-29.

[38] QxMD. Basdai. `https://qxmd.com/calculate/calculator_299/basdai`. Accessed: 2020-04-30.

[39] K. Nas, K. Yildirim, R. Cevik, S. Karatay, A. Erdal, O. Baysal, Z. Altay, A. Kamanli, Y. Ersoy, A. Kaya, B. Durmus, O. Ardicoglu, I. Tekeoglu, M. Ugur, A. J. Sarac, K. Senel, A. Gur, and S. Ozgocmen. Discrimination ability of ASDAS estimating disease activity status in patients with ankylosing spondylitis. *International Journal of Rheumatic Diseases*, 13:240–245, 2010.

[40] Assessment of SpondyloArthritis International Society. Physiology, acute phase reactants. `https://www.ncbi.nlm.nih.gov/books/NBK519570/`. Accessed: 2020-04-30.

[41] NCBI. Asdas calculator. `https://www.asas-group.org/instruments/asdas-calculator/`. Accessed: 2020-04-30.

[42] M. N. de Almeida Rodrigues de Sousa. Advances in probabilistic graphic models, 2017.

[43] A. M. Carvalho. Scoring functions for learning bayesian networks. *Inesc-id Tec. Rep*, 2009.

[44] A. Meyer-Baese and V. Schmid. *Pattern Recognition and Signal Analysis in Medical Imaging*. Elsevier, 2014.

[45] E. Charniak. Bayesian networks without tears. *AI Magazine*, 12(4), 1991.

[46] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[47] M. Sousa and A. M. Carvalho. Polynomial-time algorithm for learning optimal bfs-consistent dynamic bayesian networks. *Entropy*, 20(4):274, 2018.

[48] A. Franzin, F. Sambo, and B. Di Camillo. Bnstruct: An R package for Bayesian Network structure learning in the presence of missing data. *Bioinformatics*, 33(8):1250–1252, 2016.

[49] F. Sambo and A. Franzin. bnstruct : an R package for Bayesian Network Structure Learning with missing data. 2016.

[50] J. Gu, F. Fu, and Q. Zhou. Penalized estimation of directed acyclic graphs from discrete data. *Statistics and Computing*, 29:161–176, 2019.

[51] A. Zandonà, R. Vasta, A. Chiò, and B. D. Camillo. A Dynamic Bayesian Network model for simulation of disease progression in Amyotrophic Lateral Sclerosis patients. *BMC Bioinformatics*, 5 (Suppl 4):1–11, 2017.

[52] L. E. Sucar. Dynamic and temporal bayesian networks. `https://ccc.inaoep.mx/~esucar/Clases-mgp/Notes/c9-dbn.pdf`. Accessed: 2020-05-02.

[53] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques*. Adaptive computation and machine learning. The MIT Press, 2009.

[54] W. Wiegerinck, W. Burgers, and B. Kappen. Bayesian Networks, Introduction and Practical Applications. *Handbook on Neural Information Processing. Intelligent Systems Reference Library*, 49: 401–431, 2013.

[55] C. J. Needham, J. R. Bradford, A. J. Bulpitt, and D. R. Westhead. Inference in Bayesian networks. *Nature Biotechnology*, 24(1):51–53, 2006.

[56] M. A. van Gerven, B. G. Taal, and P. J. Lucas. Dynamic Bayesian networks as prognostic models for clinical patient management. *Journal of Biomedical Informatics*, 41:515–529, 2008.

[57] T. Charitos, L. C. van der Gaag, S. Visscher, K. A. Schurink, and P. J. Lucas. A dynamic Bayesian network for diagnosing ventilator-associated pneumonia in ICU patients. *Expert Systems with Applications*, 36:1249–1258, 2009.

[58] M. Van der Heijden, M. Velikova, and P. J. Lucas. Learning Bayesian networks for clinical time series analysis. *Journal of Biomedical Informatics*, 48:94–105, 2014.

[59] Sociedade Portuguesa de Reumatologia. Reuma.pt. `http://reuma.pt/pt_PT/Default.aspx`. Accessed: 2020-05-11.

[60] H. Canhão, A. Faustino, F. Martins, and J. E. Fonseca. Reuma. PT - the rheumatic diseases portuguese register. *Acta Reumatologica Portuguesa*, 36:45–56, 2011.

[61] H. Canhão, A. Faustino, and J. E. Fonseca. Registo nacional de doentes reumáticos, Reuma.pt. pages 192 – 213, 2012.

[62] News Farma. Mais de 12 mil doentes integram registo nacional reuma.pt. `https://www.newsfarma.`

pt/noticias/2335-mais-de-12-mil-doentes-integram-registo-nacional-reuma-pt.html. Accessed: 2020-05-11.

[63] Saúde Online. Plataforma reuma.pt comemora 10 anos de existência. https://saudeonline.pt/plataforma-reuma-pt-comemora-10-anos-de-existencia/. Accessed: 2020-05-11.

[64] C. de Seixas Serra Domingos Barata. Longitudinal and survival data analysis for rheumatic diseases' prognosis, 2020.

[65] M. N. Today. What does it mean if you have a high c-reactive protein? https://www.medicalnewstoday.com/articles/322138#ranges. Accessed: 2020-06-08.

[66] Towards Data Science. The use of KNN for missing values. https://towardsdatascience.com/the-use-of-knn-for-missing-values-cf33d935c637. Accessed: 2020-09-24.

[67] Analytics Vidhya. KNNImputer: A robust way to impute missing values (using Scikit-Learn). https://www.analyticsvidhya.com/blog/2020/07/knnimputer-a-robust-way-to-impute-missing-values-using-scikit-learn/. Accessed: 2020-09-24.

[68] Kent State University. Spss tutorials: Chi-square test of independence. https://libguides.library.kent.edu/spss/chisquare. Accessed: 2021-01-27.

[69] Chi-square statistic: How to calculate it / distribution. https://www.statisticshowto.com/probability-and-statistics/chi-square/, . Accessed: 2021-01-27.

[70] A gentle introduction to the chi-squared test for machine learning. https://machinelearningmastery.com/chi-squared-test-for-machine-learning/, . Accessed: 2021-01-27.

[71] The data school by Chartio. Correlation and p value. https://dataschool.com/fundamentals-of-analysis/correlation-and-p-value/. Accessed: 2021-01-28.

Appendix A

Datasets' creation process - Flowchart

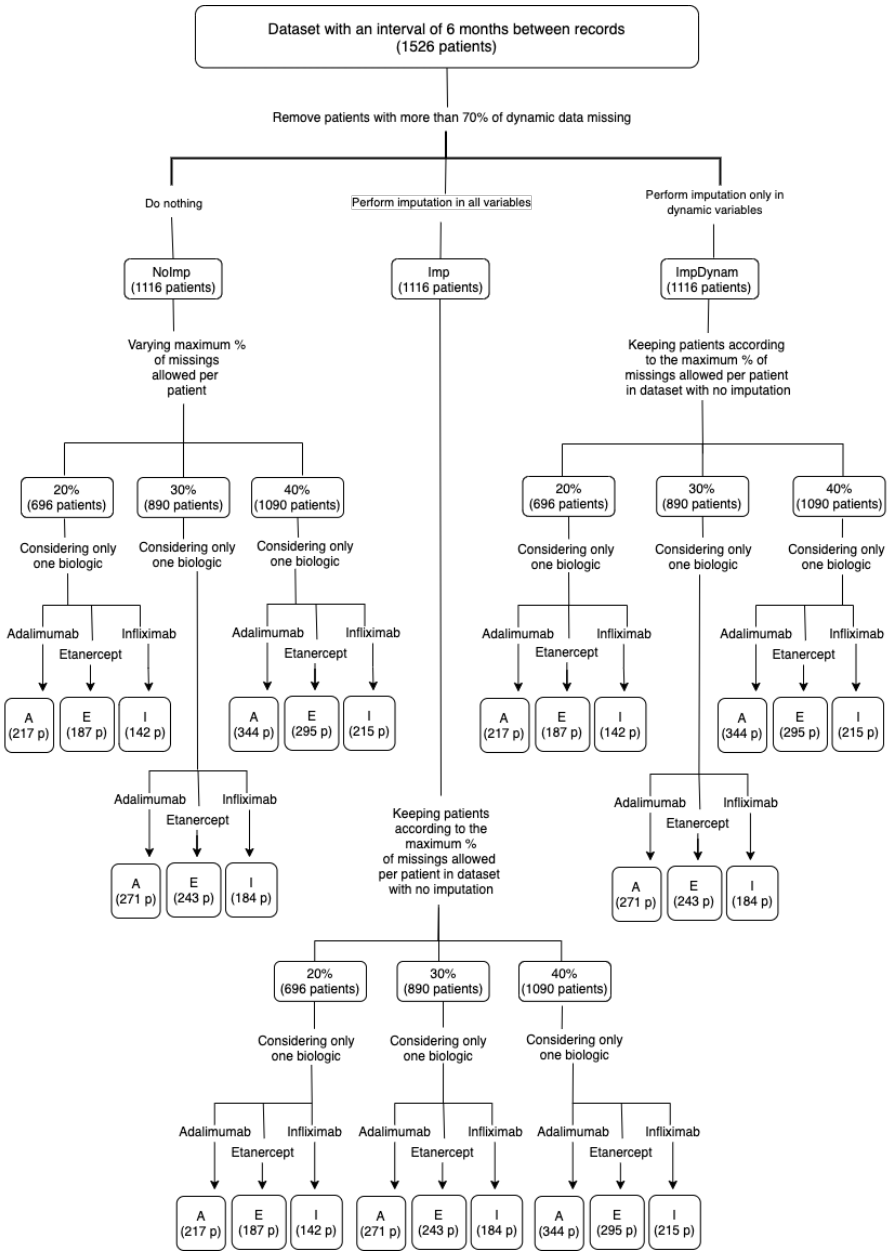


Figure A.1: Schematic representation of datasets' creation process (interval of 6 months between records).

Appendix B

Datasets' descriptive statistics

Table B.1: Descriptive statistics of static and outcome variables in 3-month interval dataset.

Variable	Number of patients with available data, n (percentage)	Observed values, n (percentage)
Gender	846 (100%)	
Male		482 (57%)
Female		364 (43%)
HLA-B27	659 (77.9%)	
Negative		162 (24.6%)
Positive		497 (75.4%)
BMI	408 (48.2%)	
]0,25[		184 (45.1%)
[25,30[		152 (37.3%)
[30,+∞[		72 (17.6%)
Age at biologic onset	846 (100%)	
]0,36]		312 (36.9%)
]36,47]		258 (30.5%)
]47,+∞[		276 (32.6%)
Disease duration	744 (87.9%)	
[0,7]		276 (37.1%)
]7,16]		244 (32.8%)
]16,+∞[		224 (30.1%)
Education	609 (72%)	
No education/ Elementary school		100 (16.4%)
Middle school		178 (29.2%)
High school		164 (26.9%)
High education		167 (27.4%)
Biologic	846 (100%)	
Adalimumab		266 (31.4%)
Certolizumab		22 (2.6%)
Etanercept		243 (28.7%)
Golimumab		142 (16.8%)
Infliximab		165 (19.5%)
Secucinumab		8 (0.9%)
Inefficiency	846 (100%)	
No		693 (81.9%)
Yes		153 (18.1%)
ASDAS at 12 months	520 (61.5%)	
[0, 1.3[		151 (29%)
[1.3, 2.1[		143 (27.5%)
[2.1, 3.5]		194 (37.3%)
]3.5, +∞[		32 (6.2%)
Change in ASDAS at 12 months	462 (54.6%)	
]-∞, 1.1[		153 (33.1%)
[1.1, 2[		149 (32.3%)
[2, +∞[		160 (34.6%)

Table B.2: Descriptive statistics of dynamic variables in 3-month interval dataset.

Variable	Observed values, n (percentage) t=0 months	Observed values, n (percentage) t=3 months	Observed values, n (percentage) t=6 months	Observed values, n (percentage) t=9 months	Observed values, n (percentage) t=12 months
Concomitant DMARD	**805 (95.2%)**	**742 (87.7%)**	**739 (87.4%)**	**629 (74.3%)**	**623 (73.6%)**
No	452 (56.1%)	456 (61.2%)	462 (62.5%)	404 (64.2%)	399 (64.0%)
Yes	353 (43.9%)	288 (38.8%)	277 (37.5%)	225 (35.8%)	224 (36.0%)
Concomitant corticoid	**805 (95.2%)**	**742 (87.7%)**	**739 (87.4%)**	**629 (74.3%)**	**623 (73.6%)**
No	633 (78.6%)	617 (83.2%)	629 (85.1%)	547 (87.0%)	545 (87.5%)
Yes	172 (21.4%)	125 (16.8%)	110 (14.9%)	82 (13.0%)	78 (12.5%)
VAS patient	**763 (90.2%)**	**688 (81.3%)**	**681 (80.5%)**	**562 (66.4%)**	**567 (67.0%)**
[0,20[	63 (8.3%)	289 (42.0%)	298 (43.8%)	259 (46.1%)	270 (47.6%)
[20,50[	176 (23.1%)	234 (34.0%)	229 (33.6%)	195 (34.7%)	182 (32.1%)
[50,100	524 (68.7%)	165 (24.0%)	154 (22.6%)	108 (19.2%)	115 (20.3%)
VAS doctor	**457 (54%)**	**373 (44.1%)**	**410 (48.5%)**	**327 (38.7%)**	**355 (42%)**
[0,10[	27 (5.9%)	170 (45.6%)	222 (54.1%)	174 (53.2%)	215 (60.6%)
[10,30[	72 (15.8%)	131 (35.1%)	129 (31.5%)	106 (32.4%)	107 (30.1%)
[30,100]	358 (78.3%)	72 (19.3%)	59 (14.4%)	47 (14.4%)	33 (9.3%)
CRP	**766 (90.5%)**	**672 (79.4%)**	**679 (80.3%)**	**571 (67.5%)**	**568 (67.1%)**
[0,1]	80 (10.4%)	214 (31.8%)	208 (30.6%)	183 (32.0%)	173 (30.5%)
]1,3]	108 (14.1%)	208 (31.0%)	222 (32.7%)	151 (26.4%)	160 (28.2%)
]3, 8.5]	169 (22.1%)	152 (22.6%)	148 (21.8%)	139 (24.3%)	127 (22.4%)
]8.5,+∞[	409 (53.4%)	98 (14.6%)	101 (14.9%)	98 (17.2%)	108 (19%)
BASDAI question 1	**779 (92.1%)**	**697 (82.4%)**	**693 (81.9%)**	**579 (68.4%)**	**572 (67.6%)**
[0,4[	135 (17.3%)	336 (48.2%)	362 (52.2%)	320 (55.3%)	319 (55.8%)
[4,7[	294 (37.7%)	226 (32.4%)	214 (30.9%)	174 (30.1%)	165 (28.8%)
[7,10]	350 (44.9%)	135 (19.4%)	117 (16.9%)	85 (14.7%)	88 (15.4%)
BASDAI question 2	**780 (92.2%)**	**697 (82.4%)**	**694 (82.0%)**	**578 (68.3%)**	**576 (68.1%)**
[0,4[	98 (12.6%)	351 (50.4%)	370 (53.3%)	321 (55.5%)	334 (58.0%)
[4,7[	205 (26.3%)	204 (29.3%)	204 (29.4%)	159 (27.5%)	157 (27.3%)
[7,10]	477 (61.2%)	142 (20.4%)	120 (17.3%)	98(17.0%)	85 (14.8%)
BASDAI question 3	**780 (92.2%)**	**697 (82.4%)**	**693 (81.9%)**	**575 (68.0%)**	**575 (68.0%)**
[0,4[	284 (36.4%)	474 (68.0%)	501 (72.3%)	422 (73.4%)	426 (74.1%)
[4,7[	228 (29.2%)	142 (20.4%)	120 (17.3%)	109 (19.0%)	105 (18.3%)
[7,10]	268 (34.4%)	81 (11.6%)	72 (10.4%)	44 (7.7%)	44 (7.7%)
BASDAI question 4	**779 (92.1%)**	**697 (82.4%)**	**693 (81.9%)**	**576 (68.1%)**	**572 (67.6%)**
[0,4[	176 (22.6%)	429 (61.6%)	453 (65.4%)	381 (66.1%)	399 (69.8%)
[4,7[	245 (31.4%)	158 (22.7%)	157 (22.7%)	122 (21.2%)	116 (20.2%)
[7,10]	358 (46.0%)	109 (15.7%)	83 (12%)	73 (12.7%)	57 (10.0%
BASDAI questions 5 and 6	**778 (92.0%)**	**698 (82.5%)**	**693 (81.9%)**	**578 (68.3%)**	**576 (68.1%)**
[0,4[	129 (16.6%)	426 (61.0%)	458 (66.1%)	386 (66.8%)	400 (69.4%)
[4,7[	207 (26.6%)	173 (24.8%)	155 (22.4%)	127 (22.0%)	120 (20.8%)
[7,10]	442 (56.8%)	99 (14.2%)	80 (11.5%)	65 (11.2%)	56 (9.7%)
BASFI	**701 (82.9%)**	**627 (74.1%)**	**607 (71.7%)**	**484 (57.2%)**	**499 (59.0%)**
[0,2]	81 (11.6%)	214 (34.1%)	242 (39.9%)	209 (43.2%)	228 (45.7%)
]2,5]	207 (29.5%)	226 (36%)	220 (36.2%)	161 (33.3%)	157 (31.5%)
]5,10]	413 (58.9%)	187 (29.8%)	145 (23.9%)	114 (23.6%)	114 (22.8%)

Table B.3: Descriptive statistics of static and outcome variables in 6-month interval dataset.

Variable	Number of patients with available data, n (percentage)	Observed values, n (percentage)
Gender	1116 (100%)	
Male		624 (55.9%)
Female		492 (44.1%)
HLA-B27	851 (76.3%)	
Negative		206 (24.2%)
Positive		645 (75.8%)
BMI	522 (46.8%)	
]0,25[		232 (44.4%)
[25,30[		192 (36.8%)
[30,+∞[		98 (18.8%)
Age at biologic onset	1116 (100%)	
]0,36]		406 (36.4%)
]36,47]		349 (31.3%)
]47,+∞[		361 (32.3%)
Disease duration	968 (86.7%)	
[0,7]		352 (36.4%)
]7,16]		313 (32.3%)
]16,+∞[		303 (31.3%)
Education	794 (71.1%)	
No education/ Elementary school		127 (16%)
Middle school		244 (30.7%)
High school		209 (26.3%)
High education		214 (27%)
Biologic	1116 (100%)	
Adalimumab		354 (31.7%)
Certolizumab		27 (2.4%)
Etanercept		300 (26.9%)
Golimumab		203 (18.2%)
Infliximab		219 (19.6%)
Secucinumab		13 (1.2%)
Inefficiency	1116 (100%)	
No		890 (79.7%)
Yes		226 (20.3%)
ASDAS at 12 months	793 (71.1%)	
[0, 1.3[		205 (25.9%)
[1.3, 2.1[		235 (29.6%)
[2.1, 3.5]		296 (37.3%)
]3.5, +∞[		57 (7.2%)
Change in ASDAS at 12 months	678 (60.8%)	
]0, 1.1[		248 (36.6%)
[1.1, 2[		199 (29.4%)
[2, +∞[		231 (34.1%)

Table B.4: Descriptive statistics of dynamic variables in 6-month interval dataset.

Variable	Observed values, n (percentage) t=0 months	Observed values, n (percentage) t=6 months	Observed values, n (percentage) t=12 months
Concomitant DMARD	**1048 (93.9%)**	**1074 (96.2%)**	**914 (81.9%)**
No	597 (57.0%)	686 (63.9%)	594 (65.0%)
Yes	451 (43.0%)	388 (36.1%)	320 (35.0%)
Concomitant corticoid	**1048 (93.9%)**	**1074 (96.2%)**	**914 (81.9%)**
No	830 (78.6%)	902 (84.0%)	795 (87.0%)
Yes	218 (21.4%)	172 (16.0%)	119 (13.0%)
VAS patient	**992 (88.9%)**	**1014 (90.9%)**	**844 (75.6%)**
[0,20[	110 (11.1%)	439 (43.3%)	402 (47.6%)
[20,50[	313 (31.6%)	420 (41.4%)	328 (38.9%)
[50,100]	569 (57.4%)	155 (15.3%)	114 (13.5%)
VAS doctor	**639 (57.3%)**	**626 (56.1%)**	**549 (49.2%)**
[0,10[	56 (8.8%)	326 (52.1%)	302 (55.0%)
[10,30[	114 (17.8%)	227 (36.3%)	186 (33.9%)
[30,100]	469 (73.4%)	73 (11.7%)	61 (11.1%)
CRP	**988 (88.5%)**	**993 (89.0%)**	**850 (76.2%)**
[0,1]	112 (11.3%)	329 (33.1%)	269 (31.6%)
]1,3]	157 (15.9%)	302 (30.4%)	248 (29.2%)
]3,8.5]	271 (27.4%)	236 (23.8%)	209 (24.6%)
]8.5,+∞[	448 (45.3%)	126 (12.7%)	124 (14.6%)
BASDAI question 1	**1012 (90.7%)**	**1032 (92.5%)**	**858 (76.9%)**
[0,4[	188 (18.6%)	506 (49.0%)	460 (53.6%)
[4,7[	365 (36.1%)	332 (32.2%)	263 (30.7%)
[7,10]	459 (45.4%)	194 (18.8%)	135 (15.7%)
BASDAI question 2	**1013 (90.8%)**	**1033 (92.6%)**	**861 (77.2%)**
[0,4[	146 (14.4%)	513 (49.7%)	474 (55.1%)
[4,7[	247 (24.4%)	322 (31.2%)	240 (27.9%)
[7,10]	620 (61.2%)	198 (19.2%)	147 (17.1%)
BASDAI question 3	**1013 (90.8%)**	**1031 (92.4%)**	**861 (77.2%)**
[0,4[	382 (37.7%)	724 (70.2%)	634 (73.6%)
[4,7[	295 (29.1%)	193 (18.7%)	161 (18.7%)
[7,10]	336 (33.2%)	114 (11.1%)	66 (7.7%)
BASDAI question 4	**1011 (90.6%)**	**1031 (92.4%)**	**857 (76.8%)**
[0,4[	244 (24.1%)	641 (62.2%)	590 (68.8%)
[4,7[	303 (30.0%)	239 (23.2%)	171 (20.0%)
[7,10]	464 (45.9%)	151 (14.6%)	96 (11.2%)
BASDAI questions 5 and 6	**1014 (90.9%)**	**1033 (92.6%)**	**861 (77.2%)**
[0,4[	188 (18.5%)	651 (63.0%)	580 (67.4%)
[4,7[	264 (26.0%)	241 (23.3%)	189 (22.0%)
[7,10]	562 (55.4%)	141 (13.6%)	92 (10.7%)
BASFI	**896 (80.3%)**	**915 (82.0%)**	**763 (68.4%)**
[0,2]	128 (14.3%)	381 (41.6%)	351 (46.0%)
]2,5]	291 (32.5%)	314 (34.3%)	252 (33.0%)
]5,10]	477 (53.2%)	220 (24.0%)	160 (21.0%)

Appendix C

Chi-square tests of independence: contingency tables

Table C.1: Contingency table HLA-B27 * Inefficiency for infliximab's dataset.

		Inefficiency		**Total**
		0	1	
HLA-B27 = 0				
	Count	24	18	42
	Expected Count	30.3	11.7	42.0
HLA-B27 = 1				
	Count	105	32	137
	Expected Count	98.7	38.3	137.0
Total				
	Count	129	50	179
	Expected Count	129.0	50.0	179.0

Table C.2: Contingency table HLA-B27 * ASDAS at 12 months for infliximab's dataset.

		ASDAS at 12 months				**Total**
		0	1	2	3	
HLA-B27 = 0						
	Count	1	11	12	2	26
	Expected Count	4.5	6.1	13.1	2.3	26.0
HLA-B27 = 1						
	Count	21	19	52	9	101
	Expected Count	17.5	23.9	50.9	8.7	101.0
Total						
	Count	22	30	64	11	127
	Expected Count	22.0	30.0	64.0	11.0	127.0

Table C.3: Contingency table Gender * ASDAS at 12 months for adalimumab's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
Gender = 0					
Count	60	40	36	6	142
Expected Count	45.7	40.7	46.2	9.5	142.0
Gender = 1					
Count	22	33	47	11	113
Expected Count	36.3	32.3	36.8	7.5	113.0
Total					
Count	82	73	83	17	255
Expected Count	82.0	73.0	83.0	17.0	255.0

Table C.4: Contingency table Gender * ASDAS change at 12 months for adalimumab's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
Gender = 0				
Count	32	28	52	112
Expected Count	43.1	28.0	40.9	112.0
Gender = 1				
Count	48	24	24	96
Expected Count	36.9	24.0	35.1	96.0
Total				
Count	80	52	76	208
Expected Count	80.0	52.0	76.0	208.0

Table C.5: Contingency table Gender * ASDAS at 12 months for etanercept's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
Gender = 0					
Count	35	48	33	6	122
Expected Count	29.6	38.8	45.0	8.6	122.0
Gender = 1					
Count	17	20	46	9	92
Expected Count	22.4	29.2	34.0	6.4	92.0
Total					
Count	52	68	79	15	214
Expected Count	52.0	68.0	79.0	15.0	214.0

Table C.6: Contingency table Gender * ASDAS change at 12 months for etanercept's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
Gender = 0				
Count	30	34	44	108
Expected Count	38.5	31.6	37.9	108.0
Gender = 1				
Count	37	21	22	80
Expected Count	28.5	23.4	28.1	80.0
Total				
Count	67	55	66	188
Expected Count	67.0	55.0	66.0	188.0

Table C.7: Contingency table Gender * ASDAS at 12 months for infliximab's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
Gender = 0					
Count	21	24	46	10	101
Expected Count	14.8	27.0	48.2	10.9	101.0
Gender = 1					
Count	2	18	29	7	56
Expected Count	8.2	15.0	26.8	6.1	56.0
Total					
Count	23	42	75	17	157
Expected Count	23.0	42.0	75.0	17.0	157.0

Table C.8: Contingency table Gender * ASDAS change at 12 months for infliximab's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
Gender = 0				
Count	28	29	28	85
Expected Count	32.7	30.2	22.0	85.0
Gender = 1				
Count	24	19	7	50
Expected Count	19.3	17.8	13.0	50.0
Total				
Count	52	48	35	135
Expected Count	52.0	48.0	35.0	135.0

Table C.9: Contingency table BMI * ASDAS at 12 months for adalimumab's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
BMI = 0					
Count	23	17	12	3	55
Expected Count	16.5	16.5	19.0	3.0	55.0
BMI = 1					
Count	8	13	16	0	37
Expected Count	11.1	11.1	12.8	2.0	37.0
BMI = 2					
Count	2	3	10	3	18
Expected Count	5.4	5.4	6.2	1.0	18.0
Total					
Count	33	33	38	6	110
Expected Count	33.0	33.0	38.0	6.0	110.0

Table C.10: Contingency table BMI * ASDAS change at 12 months for adalimumab's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
BMI = 0				
Count	16	5	28	49
Expected Count	17.0	9.3	22.7	49.0
BMI = 1				
Count	7	10	13	30
Expected Count	10.4	5.7	13.9	30.0
BMI = 2				
Count	10	3	3	16
Expected Count	5.6	3.0	7.4	16.0
Total				
Count	33	18	44	95
Expected Count	33.0	18.0	44.0	95.0

Table C.11: Contingency table Age at biologic onset * ASDAS at 12 months for adalimumab's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
Age at biologic onset = 0					
Count	42	27	26	1	96
Expected Count	30.9	27.5	31.2	6.4	96.0
Age at biologic onset = 1					
Count	25	19	28	10	82
Expected Count	26.4	23.5	26.7	5.5	82.0
Age at biologic onset = 2					
Count	15	27	29	6	77
Expected Count	24.8	22.0	25.1	5.1	77.0
Total					
Count	82	73	83	17	255
Expected Count	82.0	73.0	83.0	17.0	255.0

Table C.12: Contingency table Age at biologic onset * ASDAS at 12 months for etanercept's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
Age at biologic onset = 0					
Count	29	22	18	3	72
Expected Count	17.5	22.9	26.6	5.0	72.0
Age at biologic onset = 1					
Count	15	20	25	7	67
Expected Count	16.3	21.3	24.7	4.7	67.0
Age at biologic onset = 2					
Count	8	26	36	5	75
Expected Count	18.2	23.8	27.7	5.3	75.0
Total					
Count	52	68	79	15	214
Expected Count	52.0	68.0	79.0	15.0	214.0

Table C.13: Contingency table Disease duration * ASDAS at 12 months for adalimumab's dataset.

	ASDAS at 12 months				Total
	0	1	2	3	
Disease duration = 0					
Count	25	17	19	1	62
Expected Count	17.4	17.7	22.3	4.6	62.0
Disease duration = 1					
Count	25	14	28	5	72
Expected Count	20.2	20.5	25.9	5.4	72.0
Disease duration = 2					
Count	10	30	30	10	80
Expected Count	22.4	22.8	28.8	6.0	80.0
Total					
Count	60	61	77	16	214
Expected Count	60.0	61.0	77.0	16.0	214.0

Table C.14: Contingency table Disease duration * ASDAS at 12 months for etanercept's dataset.

	ASDAS at 12 months				Total
	0	1	2	3	
Disease duration = 0					
Count	25	16	27	3	71
Expected Count	15.9	21.8	27.7	5.5	71.0
Disease duration = 1					
Count	12	29	16	4	61
Expected Count	13.7	18.7	23.8	4.8	61.0
Disease duration = 2					
Count	6	14	32	8	60
Expected Count	13.4	18.4	23.4	4.7	60.0
Total					
Count	43	59	75	15	192
Expected Count	43.0	59.0	75.0	15.0	192.0

Table C.15: Contingency table Education * ASDAS at 12 months for infliximab's dataset.

	ASDAS at 12 months				Total
	0	1	2	3	
Education = 0					
Count	2	3	18	0	23
Expected Count	3.3	6.8	10.8	2.1	23.0
Education = 1					
Count	2	17	14	6	39
Expected Count	5.6	11.5	18.3	3.5	39.0
Education = 2					
Count	8	10	15	3	36
Expected Count	5.2	10.6	16.9	3.3	36.0
Education = 3					
Count	7	9	15	3	34
Expected Count	4.9	10.0	16.0	3.1	34.0
Total					
Count	19	39	62	12	132
Expected Count	19.0	39.0	62.0	12.0	132.0

Table C.16: Contingency table VAS patient 0M * ASDAS change at 12 months for adalimumab's dataset.

		ASDAS change at 12 months		**Total**
	0	1	2	
VAS patient 0M = 0				
Count	13	5	1	19
Expected Count	7.3	4.8	6.9	19.0
VAS patient 0M = 1				
Count	31	23	17	71
Expected Count	27.3	17.8	25.9	71.0
VAS patient 0M = 2				
Count	36	24	58	118
Expected Count	45.4	29.5	43.1	118.0
Total				
Count	80	52	76	208
Expected Count	80.0	52.0	76.0	208.0

Table C.17: Contingency table VAS patient 0M * ASDAS at 12 months for etanercept's dataset.

	ASDAS at 12 months				**Total**
	0	1	2	3	
VAS patient 0M = 0					
Count	14	7	3	0	24
Expected Count	6.2	7.6	8.6	1.6	24.0
VAS patient 0M = 1					
Count	11	25	20	4	60
Expected Count	15.4	19.1	21.5	4.0	60.0
VAS patient 0M = 2					
Count	25	30	47	9	111
Expected Count	28.5	35.3	39.8	7.4	111.0
Total					
Count	50	62	70	13	195
Expected Count	50.0	62.0	70.0	13.0	195.0

Table C.18: Contingency table VAS patient 0M * ASDAS change at 12 months for etanercept's dataset.

		ASDAS change at 12 months		**Total**
	0	1	2	
VAS patient 0M = 0				
Count	12	5	5	22
Expected Count	7.8	6.4	7.7	22.0
VAS patient 0M = 1				
Count	30	15	12	57
Expected Count	20.3	16.7	20.0	57.0
VAS patient 0M = 2				
Count	25	35	49	109
Expected Count	38.8	31.9	38.3	109.0
Total				
Count	67	55	66	188
Expected Count	67.0	55.0	66.0	188.0

Table C.19: Contingency table VAS doctor 0M * ASDAS change at 12 months for etanercept's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
VAS doctor 0M = 0				
Count	12	5	5	22
Expected Count	7.8	6.4	7.7	22.0
VAS doctor 0M = 1				
Count	30	15	12	57
Expected Count	20.3	16.7	20.0	57.0
VAS doctor 0M = 2				
Count	25	35	49	109
Expected Count	38.8	31.9	38.3	109.0
Total				
Count	67	55	66	188
Expected Count	67.0	55.0	66.0	188.0

Table C.20: Contingency table CRP * ASDAS change at 12 months for adalimumab's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
CRP = 0				
Count	14	5	0	19
Expected Count	7.3	4.8	7.0	19.0
CRP = 1				
Count	22	10	7	39
Expected Count	14.9	9.8	14.3	39.0
CRP = 2				
Count	23	21	13	57
Expected Count	21.8	14.3	20.9	57.0
CRP = 3				
Count	20	16	56	92
Expected Count	35.1	23.1	33.8	92.0
Total				
Count	79	52	76	207
Expected Count	79.0	52.0	76.0	207.0

Table C.21: Contingency table CRP * ASDAS change at 12 months for etanercept's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
CRP = 0				
Count	9	7	3	19
Expected Count	6.8	5.6	6.7	19.0
CRP = 1				
Count	20	9	3	32
Expected Count	11.4	9.4	11.2	32.0
CRP = 2				
Count	19	24	6	49
Expected Count	17.5	14.3	17.2	49.0
CRP = 3				
Count	19	15	54	88
Expected Count	31.4	25.7	30.9	88.0
Total				
Count	67	55	66	188
Expected Count	67.0	55.0	66.0	188.0

Table C.22: Contingency table CRP * ASDAS change at 12 months for infliximab's dataset.

	ASDAS change at 12 months			**Total**
	0	1	2	
CRP = 0				
Count	5	6	1	12
Expected Count	4.6	4.3	3.1	12.0
CRP = 1				
Count	12	5	1	18
Expected Count	6.9	6.4	4.7	18.0
CRP = 2				
Count	19	12	7	38
Expected Count	14.5	13.6	9.9	38.0
CRP = 3				
Count	15	25	26	66
Expected Count	25.1	23.6	17.2	66.0
Total				
Count	51	48	35	134
Expected Count	51.0	48.0	35.0	134.0