Study of Predicting Remission in Graves' Disease

Dr. Laura Claire Lane

Thesis submitted for the degree of Doctor of Philosophy

Translational and Clinical Research Institute

Faculty of Medical Sciences

Newcastle University

April 2025

Abstract

Background

Graves' disease (GD) is an autoimmune condition characterised by the presence of thyroid receptor auto-antibodies (TRAbs) that stimulate the thyroid gland to produce excessive thyroid hormones (hyperthyroidism). Only around half of adults will achieve remission of their GD after a course of antithyroid drugs (ATD), and there is no reliable method to accurately predict which patients will relapse their GD. The aim of the prospective observational study described in this thesis was to explore biomarkers of outcome in GD in a cohort at the time of ATD withdrawal.

Method

65 patients with GD who had their ATD stopped after completing at least a 12–18 month course were followed up for 12 months, at which time they were categorised as having relapsed or remitted. Clinical, biochemical, and immunological studies were investigated at baseline and 6-10 weeks later. These factors and $CD19^+$ B cell gene expression were assessed for their ability to predict outcome of GD.

Results

16/65 (25%) relapsed their GD within 12 months after ATD withdrawal. TRAb and sTACI (soluble transmembrane activator and calcium modulator and cyclophilin ligand interactor) both proved to be good baseline biomarkers of GD relapse (OR 3.3; 95% CI 1.5-9.9 (TRAb), OR 2.3; 95% CI 1.3-4.6 (sTACI)). Investigation of the functional and biological characteristics of differentially expressed genes revealed an environment of inflammation and oxidative stress which may precipitate GD relapse. A 5-variable predictive model that included TRAb, sTACI and 3 genetic markers was developed that provided good predictive value for differentiating relapse and remission patients (ROC_{AUC} = 0.99; 95% CI 0.96-1.0).

Conclusions

This study provides insight into potential humoral immunopathological pathways of GD relapse and reveals genetic and immunological B-cell biomarkers of outcome. Validation of these biomarkers in predicting relapse and translation to clinical practice would help guide timing of ATD withdrawal and assist in the formation of personalised therapeutic strategies.

Acknowledgements

I would like to thank the following people who have been invaluable in supporting me to undertake my research.

Firstly, I would like to thank my supervisors, Professor Simon Pearce, and Professor Tim Cheetham whose insight, knowledge and support have been instrumental in helping me complete my research, alongside giving me a great experience of the world of research for which I am very grateful.

Thanks also to Dr Jonathan Coxhead and Raf Hussain (Newcastle University, Core Genomics Facility) for performing the library preparation and RNA sequencing. I would also like to thank the Bioinformatics Support Unit at Newcastle University, particularly Ann Hedley, who performed the initial bioinformatic analysis for the RNAseq data (Chapter 8) and spent time explaining how to interpret huge amounts of data.

A special thanks to Kath Allinson (laboratory technician) who has been so helpful and patient teaching me laboratory skills over many years, especially at the beginning when even using a pipette was a challenge!! I would like to thank Sottini and colleagues, Professor A Gennery, and Dr. Aisling Flinn for providing the triple-insert plasmid used in this thesis (Chapter 5).

I would like to thank the Medical Research Council who awarded me with a Clinical Research Fellowship (MR/S001611/1) which gave me the opportunity to take time out of training to undertake my research. I am also grateful to the Newcastle NIHR BioResource Centre for providing me with control samples.

Importantly, I could not have undertaken this research without all the patients who contributed to my study, so I am immensely grateful to them for their time and copious amounts of blood they donated!

Lastly, I would also like to thank my wonderful and supportive husband Tom, who did many ‘babysitting’ duties of our two girls Rosa and Elsie to allow me to finish my PhD!

Table of Contents

List of Abbreviations

Abbreviation	Definition
ADAM10	A disintegrin and metalloprotease 10
ABCs	Age-associated B Cells
AITD	Autoimmune thyroid disease
APC	Antigen-presenting cell
APRIL	A-proliferating inducing ligand
ATD	Antithyroid drugs
BAFF	B-cell activating factor
BCMA	B-cell maturation antigen
BR	'Block and replace'
Breg	B regulatory cell
cAMP	Cyclic adenosine monophosphate
CBZ	Carbimazole
CD40	Cluster of differentiation 40
CJ	Coding joint
CRP	C-reactive protein
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
CV	Coefficient of variation
CXCL13	Chemokine (C-X-C motif) ligand 13
CXCR5	C-X-C chemokine receptor type 5
DEG	Differentially expressed genes
DGE	Differential gene expression
DNA	Deoxyribonucleic acid
DT	Dose titration
eQTLs	Expression quantitative trait loci
ELISA	Enzyme-linked immunosorbent assay
FACS	Fluorescence activated cell sorting
FBC	Full blood count
FCRL3	Fc receptor-like protein 3
FDR	False discovery rate

FSC-A	Forward-scatter area
FSC-H	Forward-scatter height
FT3	Free triiodothyronine
FT4	Free thyroxine
GD	Graves' disease
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GO	Graves' orbitopathy
'GREAT'	Graves' Recurrent Events After Therapy
HRA	Health Research Authority
HLA	Human leukocyte antigen
IKZF3	Ikaros family zinc finger protein 3
IFN	Interferon
IL	Interleukin
ITP	Immune thrombocytopenic purpura
KRECs	Kappa-deleting recombination excision circles
LLOD	Lower limit of detection
MLR	Monocyte-lymphocyte ratio
MS	Multiple sclerosis
PBMC	Peripheral blood mononuclear cell
PCA	Principal Component Analysis
PD-L1	Programmed death-ligand 1
PTPN22	Protein tyrosine phosphatase, non-receptor type 22
PTU	Propylthiouracil
T3	Triiodothyronine
T4	Thyroxine
TACI	Transmembrane activator calcium modulator and cyclophilin ligand interactor
TG	Thyroglobulin
TNF	Tumour necrosis factor
TREC	T-cell receptor excision circles
RA	Rheumatoid arthritis
RAI	Radioiodine

REC	Research Ethics Committee
RIN	RNA integrity value
RNA	Ribonucleic acid
RNAseq	RNA-sequencing
ROS	Reactive oxidative species
RT-qPCR	Real-time quantitative polymerase chain reaction
sTACI	Soluble transmembrane activator calcium modulator and cyclophilin ligand interactor
sBCMA	Soluble b-cell maturation antigen
SJ	Signal joint
SLE	Systemic lupus erythematous
SNP	Single nucleotide polymorphism
SSC-A	Side-scatter area
TG	Thyroglobulin
TPOAb	Thyroid peroxidase antibody
TRAb	Thyroid receptor auto-antibody
Treg	T regulatory cells
TSH	Thyroid stimulating hormone
TSHR	Thyroid stimulating hormone receptor
UMAP	Uniform Manifold Approximation and Projection
V(D)J	Variable-Diversity-Joining

List of Tables

List of Figures

Chapter 1: Introduction

1.1 General introduction to Graves' disease

Graves' disease (GD) is an autoimmune condition characterised by the presence of thyroid receptor auto-antibodies (TRAbs) that stimulate the thyroid stimulating hormone receptor (TSHR), resulting in uncontrolled excessive secretion of thyroid hormones (hyperthyroidism). Hyperthyroidism can have a substantial impact on the normal physiological activity of many organs including the cardiovascular, gastrointestinal, and central nervous systems (*Chaker L et al. 2024).* In addition, the extrathyroidal manifestations of GD such as Graves' orbitopathy (GO) can threaten sight and result in significant facial disfigurement, with long-lasting implications for an individual's quality of life (*Riguetto CM et al. 2019*)

1.1.1 Epidemiology

GD is the commonest cause of hyperthyroidism with an incidence of 24.8 cases per 100,000 in a UK population, with an adjusted female to male ratio of around 4:1 (*Hussain YS et al. 2017).* The peak incidence is reported to occur at a median (IQR) age of 44 (33–56) years (*Hussain YS et al. 2017)*, however GD is also the commonest cause of hyperthyroidism in young children and adolescents, with a reported incidence of between 0.1-3 per 100,000 (*Williamson S et al. 2010).* Several theories have been proposed to explain the strong female predisposition to developing GD including skewed X chromosome inactivation (*Simmonds MJ et al. 2014*), fetal microchimerism (*Lepez T et al. 2012*) and the presence of circulating sex hormones modulating the immune response (*Ngo ST et al. 2014*)

A global increase in the incidence of autoimmune disease has been described over recent decades and has been observed to be particularly pronounced for GD in a UK population (*Conrad N et al. 2023*). A marked socioeconomic gradient was apparent, with a higher incidence of GD demonstrated in the most deprived socioeconomic groups which notably also included a higher number of smokers (*Conrad N et al. 2023*).

1.1.2 Morbidity and mortality

It is widely recognised that the hyperthyroid state associated with GD can have a significant impact on physical and mental health, particularly with regard to

cardiovascular disease, bone health, and psychological stability (*Lillevang-Johansen M et al. 2019, Brandt F et al. 2012*, Khamisi S et al 2023). Observational studies have shown significantly increased mortality in individuals with untreated or inadequately treated GD (*Lillevang-Johansen M et al. 2019).* Furthermore, many of those who were treated for GD in childhood, adolescence or adulthood report a reduced quality of life (*Lane LC et al. 2021, Abraham-Nordling M et al. 2005*).

1.1.3 Socioeconomic impact

The negative impact of GD on general health results in an increased proportion of those with work disability and a consequent loss of income in the labour market (*Brandt F et al. 2015*). The high risk of relapse after stopping antithyroid drugs (ATD) results in loss of working days and importantly in the paediatric population, negatively affects school education and academic performance with lasting consequences for future career prospects (*Bauer AJ 2011*).

The requirement for regular monitoring of thyroid function and outpatient endocrine clinic review whilst on ATD means there is a significant cost associated with biochemical testing and clinician time (*Donovan PJ at al. 2016*). A cost analysis comparing the different available treatments for GD (ATD, radioiodine (RAI) and thyroidectomy) confirmed previous findings that RAI as a primary treatment is the least expensive option (*Donovan PJ at al. 2016*). Indeed, the recent NICE guidelines that advocate RAI as the first-line treatment for GD in adults (unless ATD likely to achieve remission or RAI unsuitable) has been forecasted to save £1 million per annum largely due to the reduction in use of ATD (*NICE 2019*). However, RAI treatment is not suitable for all as it is contraindicated for some individuals as well as resulting in permanent hypothyroidism. This leaves the patient with a life-long dependence on thyroid hormone replacement and regular monitoring with blood tests.

1.1.4 Clinical presentation

The physiological role of thyroid hormone in multiple systems within the body results in a clinical presentation of GD that can be extremely broad and non-specific (Table 1.1). Triiodothyronine (T3), as opposed to thyroxine (T4), is the more metabolically active of the thyroid hormones and hence is largely responsible for the clinical effects of excess thyroid hormone.

The features of GD at presentation can vary dependant on age and severity of hyperthyroidism (*Nordyke RA et al. 1988*), but cardiovascular symptoms and signs tend to predominate, largely as a result of excess free triiodothyronine (FT3) enhancing beta-adrenergic receptor activity (*Franklyn JA et al. 2012*). Young people with GD may present with poor sleep and reduced academic performance or alternatively, be asymptomatic.

Although the TSHR autoantigen is primarily expressed in thyroid cells, it has also been found at other sites including the upper dermis, orbital fibroblasts, and adipocytes (*Bahn RS et al. 1998*). Therefore, unlike other forms of hyperthyroidism, the presentation of GD can include extrathyroidal manifestations involving the eyes and skin, including GO, thyroid dermopathy (pretibial myxoedema) and thyroid acropachy (finger clubbing) (*Bartalena L et al. 2012*). The patients with higher TRAb concentrations are at greater risk of developing the eye and skin manifestations of GD with a correlation in terms of disease severity and outcome (*Eckstein AK et al. 2006*).

System	Symptoms	Signs
Thyroid	Neck swelling	Goitre
Cardiovascular	Palpitations, shortness of breath	Cardiac failure, atrial arrhythmias, hypertension, tachycardia
Gastrointestinal	Increased appetite	Weight loss
Central nervous system	Fatigue, anxiety, hyperactivity, poor concentration	Hyperactivity, social withdrawal, hyperreflexia
Eyes	Watering, grittiness, discomfort, diplopia	Eyelid retraction and lag, periorbital oedema, conjunctival injection
Muscles	Tremor, weakness	Fine tremor, muscle wasting
Skin	Pruritus, pain	Thyroid dermopathy (skin thickening), finger clubbing
Hair	Hair loss	Hair loss
Reproductive	Oligomenorrhoea, decreased fertility (women); reduced libido (men)	

Table 1.1 – Symptoms and signs of Graves' hyperthyroidism

1.1.5 Aetiology of Graves' disease

The mechanisms that lead to loss of immune tolerance to the TSHR result from a complex interplay of environmental exposure, genetic/epigenetic susceptibility and more recently, the microbiome has been implicated.

Familial and twin studies over the past 50 years have demonstrated that 60-80% of an individual's predisposition to develop GD is attributable to genetic factors (*Skov J et al. 2021, Brix TH et al. 2001).* GD is inherited as a complex genetic trait with more than 80 susceptibility loci identified so far, each contributing to a small increased risk of disease (*Grixti L et al. 2024*). Genomic variants implicated in GD are thyroid-related loci such as *TSHR* (*Płoski R et al. 2010*) and thyroglobulin (*TG) (Stefan M et al. 2011*) *or* immune-related loci including cytotoxic T-lymphocyte-associated protein 4 (*CTLA-4) (Vaidya B et al. 1999*), protein tyrosine phosphatase, non-receptor type 22 (*PTPN22)* (*Velaga MR et al. 2004*), human leukocyte antigen *(HLA) (Chu X et al. 2018*), programmed death-ligand 1 (*PD-L1)* (*Mitchell AL et al. 2009*), cluster of differentiation 40 (*CD40)* (*Chen X et al. 2018*), Fc receptor-like protein 3 (*FCRL3)* (*Owen CJ et al. 2007*) and B-cell activating factor *(BAFF)* (*Lane LC et al. 2019*).

Some of the genetic variants which have been associated with specific phenotypes of GD, such as *CTLA4*, *HLA*, and *PTPN22* have been associated with reoccurrence of GD (Section 1.3). Other genes in the *HLA* region and a polymorphism of *HCP5* have been associated with an earlier age of disease onset (<30 years) (*Kus et al. 2019, Lane LC et al. 2020*), suggesting a greater genetic component for young people with GD compared to their older counterparts for whom exposure to environmental factors may be more important. The functional effects of how these genetic variants influence phenotype remain largely unknown and further next generation whole genome sequencing (WGS) studies are needed to ensure reliable genotype-phenotype correlations that could be clinically useful. A few small studies have suggested that epigenetic dysregulation may contribute to the pathogenesis of GD, however replication of these studies on a larger scale is required (*Limbach M et al. 2016, Cai TT et al. 2015*).

The key environmental risk factors that have been found associated with GD include infections, excessive iodine intake, smoking, stress, and selenium/vitamin D deficiency (*McLachlan SM et al. 2014*). Some of these environmental factors are hypothesised to induce thyroid autoimmunity by 'molecular mimicry' or 'bystander

activation', resulting in the activation of autoreactive lymphocytes (*Rojas M et al. 2018*).

Environmental and endogenous factors such as diet, smoking, stress, and antibiotic use have been hypothesised to cause an imbalance of gut microorganisms that may disrupt immune tolerance in genetically susceptible individuals leading to autoimmune thyroid disease (AITD) (*Masetti G et al. 2020*). The potential importance of the microbiome in GD was highlighted by a small study that demonstrated a reduced risk of hyperthyroid relapse in adults randomised to take a probiotic alongside their ATD (*Salvi M et al. 2019*). Further work on a larger scale would be required to validate these findings and potentially pave the way for future microbiota-targeting therapeutics.

1.1.6 Diagnosis

Biochemistry

The diagnosis of thyrotoxicosis is made by measuring serum thyroid stimulating hormone (TSH), FT3 and free thyroxine (FT4) concentrations. Overt hyperthyroidism will result in a fully suppressed TSH as a result of negative feedback to the pituitary from elevated FT3 and FT4 levels. TSH is a sensitive screening test for thyroid dysfunction, as small changes in circulating FT4 result in relatively large changes in TSH (*Hadlow NC et al. 2013*).

To conclusively diagnose GD as the cause of thyrotoxicosis the TRAb levels are required, which are pathognomonic of GD when present. The main differential diagnosis of thyrotoxicosis includes toxic multinodular goitre, thyroiditis, and solitary toxic adenoma. Incorporating TRAb testing into the GD diagnostic process has been shown to result in a 46% faster time to diagnosis and a 47% reduction in overall cost (*McKee A et al. 2012*). The value of testing TRAb concentrations is also applicable in predicting risk of extrathyroidal manifestations (*Eckstein AK, et al. 2006*) and fetal/neonatal thyrotoxicosis (*Matthews DC et al. 2011*).

The latest third-generation competitive binding assays which are most commonly used in Europe, use human monoclonal stimulating TRAbs that bind to recombinant TSH receptors to achieve a specificity of 99% and sensitivity of 97% in the diagnosis of GD (*Tozzoli R et al. 2012*). Of note, the third-generation assay does not discriminate between the TRAb functional subtypes: stimulating, blocking and

'neutral' (cleavage). The TRAb bioassay can detect levels of stimulating TRAbs by measuring production of cyclic adenosine monophosphate (cAMP) but it is more complex and expensive and as yet there is insufficient evidence to recommend its use in everyday clinical practice (*Barbesino G et al. 2013*). More recently, the automated thyroid-stimulating immunoglobulin bridge immunoassay developed to detect stimulating TRAbs by using a pair of recombinant human TSH receptors in a bridging format has been demonstrated to have good diagnostic performance and the potential to detect earlier relapse, however further clinical validation is required (*van Balkum M et al. 2023, Autilio C et al. 2018*).

All commercially available antibody binding assays rely on competition against M22 (human monoclonal autoantibody to the TSHR) binding (e.g., Roche, Brahms, Siemens). These assays don't distinguish between binding, stimulating, blocking or functionally neutral antibodies. Bioassays may be brought in but will be more time consuming than binding assays and there is less accuracy and reproducibility.

Radiology

If the aetiology of thyrotoxicosis is unclear, isotope imaging using technetium pertechnetate (Tc^{99m}) and observing thyroid uptake can help differentiate between the causes of a thyrotoxic biochemical picture (Table 1.2). Normal uptake in the thyroid gland is described in a UK population as being between 0.2-2.0% (20 minutes post Tc^{99m} administration) (*Macauley M et al. 2018*).

Scan result	Interpretation
Diffuse, increased uptake	Graves' disease Hashitoxicosis (patchy uptake)
Focal, increased uptake	Autonomous nodules
Normal uptake	Hashimoto disease Subacute thyroiditis (recovery phase)
Reduced uptake	Excess iodine intake Subacute thyroiditis Hashimoto disease

Table 1.2 – Interpretation of technetium pertechnetate thyroid uptake scan results

1.1.7 Management

The management of GD involves three main treatment options: ATD, RAI or thyroidectomy. These options have remained largely unchanged since 1951 when carbimazole (CBZ) was first introduced as an ATD to the UK. This is despite each of the three modalities having significant disadvantages with many patients ultimately requiring lifelong thyroid hormone replacement therapy. As such, there remains considerable unmet need in this patient group.

Antithyroid therapy (ATD)

The ATDs (thionamides) that are used to treat GD vary geographically. In the UK, CBZ is the most commonly used ATD and is the first line treatment for both adults and children, with its active metabolite methimazole being primarily used in the USA and Europe. The second line ATD propylthiouracil (PTU), is used in those who develop side effects with CBZ or in the first trimester of pregnancy but is contraindicated in the young because of the risk of hepatotoxicity (*Rivkees SA et al. 2009*). ATDs predominately work by inhibiting the organification of iodine, preventing iodination of tyrosine residues on the thyroglobulin molecule, a vital step in the pathway to synthesise thyroid hormone (*Taurog A 1976*). PTU has an additional role in preventing the peripheral conversion of FT4 to the more metabolically active FT3 (*Kuiper GG et al. 2005).*

It has been proposed that ATDs possess an immunomodulatory function reflected by falling TRAb titres, alongside *in vitro* evidence of methimazole inhibiting thyroid-autoantibody production in cultured lymphocytes and *in vivo* evidence of increased suppressor T cells and decreased intrathyroidal activated T cells (*Weetman AP et al. 1983, Tötterman TH et al. 1987).* This however remains controversial, with others suggesting that the decreasing TRAb titres occur directly as a result of restoring biochemical euthyroidism, and hence reducing exposure to the thyroid antigen or alternatively, is part of the natural history of GD which has the potential to resolve with beta blockers alone (*Codaccioni JL et al. 1988*). The relative contributions of direct or indirect effects of ATD on the immune system to achieve the observed reduction in TRAb titres and subsequent GD remission remain uncertain.

There are two primary methods to administer ATD therapy, either by ‘block and replace’ (B&R) or ‘dose titration’ (DT). These either involve completely blocking endogenous thyroid hormone production with exogenous thyroid hormone

supplementation (B&R), or alternatively titrating ATD to achieve euthyroidism (DT). There is some evidence that using higher doses of ATD (associated with the 'B&R' method) results in a greater incidence of side effects (*Reinwein D et al. 1993, Grebe SK et al. 1998*), with contradicting evidence as to which method results in more stable thyroid function (*Vigone MC et al. 2020, Wood CL et al. 2020).* Ultimately, there is no robust evidence to support higher rates of remission with either method (*Abraham P et al. 2010*).

<u>There are three main issues with the current ATD therapy:</u>

1. The potential serious side effects of ATD are not insignificant, with agranulocytosis and fulminant hepatic failure both potentially fatal adverse effects (*Nakamura H et al. 2013, Wang MT et al. 2014*).
2. ATD are not appropriate for all patients, with an increased risk of congenital anomalies in the pregnant population (CBZ>PTU) (*Andersen SL et al. 2013*) and poor compliance in adolescents (*Léger J et al. 2017*).
3. The high rate of relapse on stopping ATD (discussed in Section 1.3)

It has been demonstrated that thyrotoxicosis itself can cause neutropenia (*Aggarwal N et al. 2016*) and deranged liver function tests (*Huang MJ et al. 1995)* independent of ATD therapy. This can further add to the complexities of management with some patients potentially stopping ATD unnecessarily.

Radioiodine therapy

Ablative RAI therapy (^{131}I) works by radioactive iodine uptake into the thyroid gland, exposing the gland to beta radiation and destroying the thyroid follicles over a period of 6 weeks to 18 months (*Nygaard B et al. 1995*). After therapy, the most likely long-term outcome for the patient is hypothyroidism (5-50% after the first year; followed by a yearly rate of 3-5%) (*Nygaard B et al. 1995),* requiring lifelong thyroid hormone replacement.

RAI is not suitable for all individuals, being contraindicated whilst breastfeeding, during pregnancy or in those planning a pregnancy within 6 months (*Kahaly GJ et al. 2018*). Clinicians are wary about using RAI in the young because of the relatively high radiation exposure of the small body where organs are closer together, and furthermore it is not available in all paediatric units. For those with young children the requirements for limiting radiation exposure post-therapy may not be practical.

Additionally, it is well recognised that RAI can trigger the development or exacerbation of pre-existing eye disease (*Bartalena L et al. 1998*).

Thyroidectomy

Prior to the 1940s surgery was the only treatment for GD, but it is now the least commonly selected treatment for newly diagnosed GD (*Villagelin D et al. 2024*). Alongside the anaesthetic risk, surgery can result in laryngeal nerve palsy causing voice disturbance and any damage to the parathyroid glands can result in either transient or permanent hypocalcaemia. Similar to RAI, post thyroidectomy patients require lifelong thyroid hormone replacement alongside biochemical monitoring.

Novel therapies in Graves' disease

Our increasing understanding of the thyroid autoantigen/antibody interactions has led to a number of potential novel treatment strategies being used in ongoing research studies (*Lane LC et al. 2020)*. Taking advantage of this expanding immunological knowledge, an antigen-specific epitope has been developed (Apitope) as well as thyroid monoclonal stimulating (M22) and blocking (K1-70) antibodies (*Pearce SHS et al. 2019, Furmaniak J et al. 2022*). These targeted TSHR-specific treatment modalities have the potential to revolutionise the treatment of GD.

1.1.8 Summary

GD is a common autoimmune disease with the potential to hugely impact quality of life in affected patients. The treatment strategy of choice is dependent on the individual patient, considering timing of future pregnancies, the presence of comorbidities and extrathyroidal manifestations, as well as factors indicating likelihood of remission after a course of ATD.

However, current treatment options with their potential serious adverse effects, contraindications, and probable need for lifelong thyroxine therapy are far from ideal. Although ATD are effective at treating hyperthyroidism, in the majority of cases they do not cure the disease. Therefore, further research is essential to understand more about the immunopathogenesis of GD which may help in the clinical management of patients and ultimately identify novel therapeutic targets.

1.2 Immunopathogenesis of Graves' disease

Although the specific immune mechanisms driving GD remain largely unknown, there are an increasing number of studies detailing the role and interactions of B and T cells and their respective cytokines and chemokines in affected individuals. Our understanding of the TRAb autoantibodies and their functional complexities has expanded over the past few years, partly explaining the clinical and biochemical spectrum of GD and possibly the variable response to treatment.

1.2.1 Thyroid stimulating hormone receptor (TSHR)

The autoantigen unequivocally involved in GD pathogenesis and responsible for driving the autoimmune process is the TSHR. GD occurs as a consequence of a loss of immunological tolerance to the TSHR antigen. The TSHR located on the surface of thyroid follicular cells, is a G-protein-coupled, transmembrane receptor with 7 transmembrane domains and constitutes of an extracellular alpha (A) subunit and a transmembrane beta (B) subunit. TSH binding to the A-subunit results in the receptor cleaving into its respective subunits (with consequent shedding of the A-subunit), leading to stimulation of the cAMP and phospholipase pathways resulting in thyroid hormone synthesis and secretion (*Rapoport B et al. 2016, Carvalho DP et al. 2017*).

GD represents a model for which the major manifestations of the disease are directly caused by the impact of TRAb on thyroid hormone production. Stimulating TRAbs bind to the extracellular N-terminus of the TSHR which results in the excessive, autonomous production of thyroid hormone (*Hamidi S et al. 2011*). *In vitro* evidence has demonstrated that stimulating TRAbs bind with greater affinity to the shed extracellular A-subunit of the TSHR when compared to the TSH holoreceptor (*Chazenbalk GD et al. 2002*). It has therefore been proposed that the A-subunit contains the primary antigen responsible for driving maturation of the B cells that are producing stimulating TRAbs, initiating or exacerbating GD in genetically susceptible individuals (*Mizutori Y et al. 2009*).

1.2.2 TSH receptor auto-antibodies (TRAbs)

TRAbs are the serological hallmark of GD. As discussed above (Section 1.1), there are three functional subtypes. The relative contributions of these different subtypes

to the clinical phenotype are unclear and evidence is lacking to support their usefulness in predicting the clinical course of GD (*Quadbeck B et al. 2005).*

1.2.3 CD19⁺ B cells and Graves' disease

GD is a B-cell mediated, T cell dependant autoimmune disease. The indisputable role of B cells in GD is evident by their capacity as plasma cells to produce the pathogenic autoantibody, TRAb, which is directly responsible for stimulation of the thyroid gland. However, it is well recognised that B cells also have a vital role as antigen-presenting cells (APC), presenting TSHR epitopes to T cells, as well as modulating the immune response by producing both pro-and anti-inflammatory cytokines helping to drive or ameliorate the autoimmune process, respectively (*Gianoukakis AG et al. 2008*). Studies using animal models of autoimmune disease, including a murine model of GD, have consistently demonstrated that B cells have an important role in the pathogenesis of autoimmunity independent of their ability to secrete autoantibody (*Pichurin, P et al. 2003, O'Neill SK et al. 2005, Takemura S et al. 2001*). Although the autoreactive B cells are fundamental in the pathogenesis of GD, T cells have a crucial role in providing the costimulatory signal to activate the antibody-producing B cells (*Petersone L et al. 2018*).

Evidence for the role of $CD19^+$ B cells in GD is also reflected by the genetic association of B cell related single nucleotide polymorphisms (SNPs), including at *FCRL3*, Ikaros family zinc finger protein 3 (*IKZF3)*, BTB Domain And CNC Homolog 2 (*BACH2*) (*Liu W et al. 2014*), *BAFF* and *CD40 (Li L et al. 2018*). Further evidence for the role of B cells in GD is presented by transcriptomic studies of thyroid cells that have revealed active B cells genes with mRNA overexpression of the antigen-presenting pathway (*Yin X et al. 2014*), which was confirmed in a recent single cell transcriptomic study of thyroid cells from GD patients that reported an increase in antigen processing and presenting pathways and an overexpression of CD40 (*Álvarez-Sierra D et al. 2023*). Indeed, the potential efficacious use of B cell depleting therapies, such as the $CD20^+$ monoclonal antibody Rituximab in B cell mediated autoimmune diseases, including GD (*Cheetham TD et al. 2022, El Fassi D et al. 2007*), demonstrates the importance of B cells in autoimmune pathology.

The role of CD19⁺ B cell subsets in Graves' disease

Specific CD19⁺ B cell subsets in the peripheral blood compartment have been studied in GD. In particular the CD5⁺ transitional and pre-naive mature B lymphocytes, which are recent emigrants from the bone marrow, have been found elevated in individuals with both treated and untreated GD, with correlation to thyroid hormone and TRAb concentrations (*Iwatani Y et al. 1989, Van der Weerd K et al. 2013, Liu Y et al. 2022*). It has been proposed that these increased transitional B cells may reflect aberrant negative selection in the bone marrow resulting in an increased output of immature B cells, or as a direct result of hyperthyroidism, which has been demonstrated to increase bone marrow activity (*Foster MP et al. 1999*).

B cell antibody responses are not only antigen-specific but also result in the generation of immunological memory. Memory B cell subpopulations in the peripheral compartment have been found to be reduced in those with GD compared to healthy controls, however a large proportion of intrathyroidal B cells in GD consist of memory cells (*Segundo C et al. 2001*), suggesting a possible redistribution of this B cell subset rather than an absolute reduction in memory cells.

B cell cytokines and chemokines

B cells produce various cytokines that form an essential part of the pro- and anti-inflammatory response (Figure 1.1). These B cell-derived cytokines play an important role in immunomodulation, either driving or suppressing the activity of B cells themselves, and also other circulating immune cells. B cells can produce interleukin 6 (IL-6) which has been demonstrated to be elevated in GD and associated with disease activity and relapse (*Pedro AB et al. 2011, Salvi M et al. 1996, Meling Stokland AE et al. 2024*). IL-6 is also produced by the thyroid follicular cells and implicated in the generation of ectopic germinal centres (*Arkatkar T et al. 2017*).

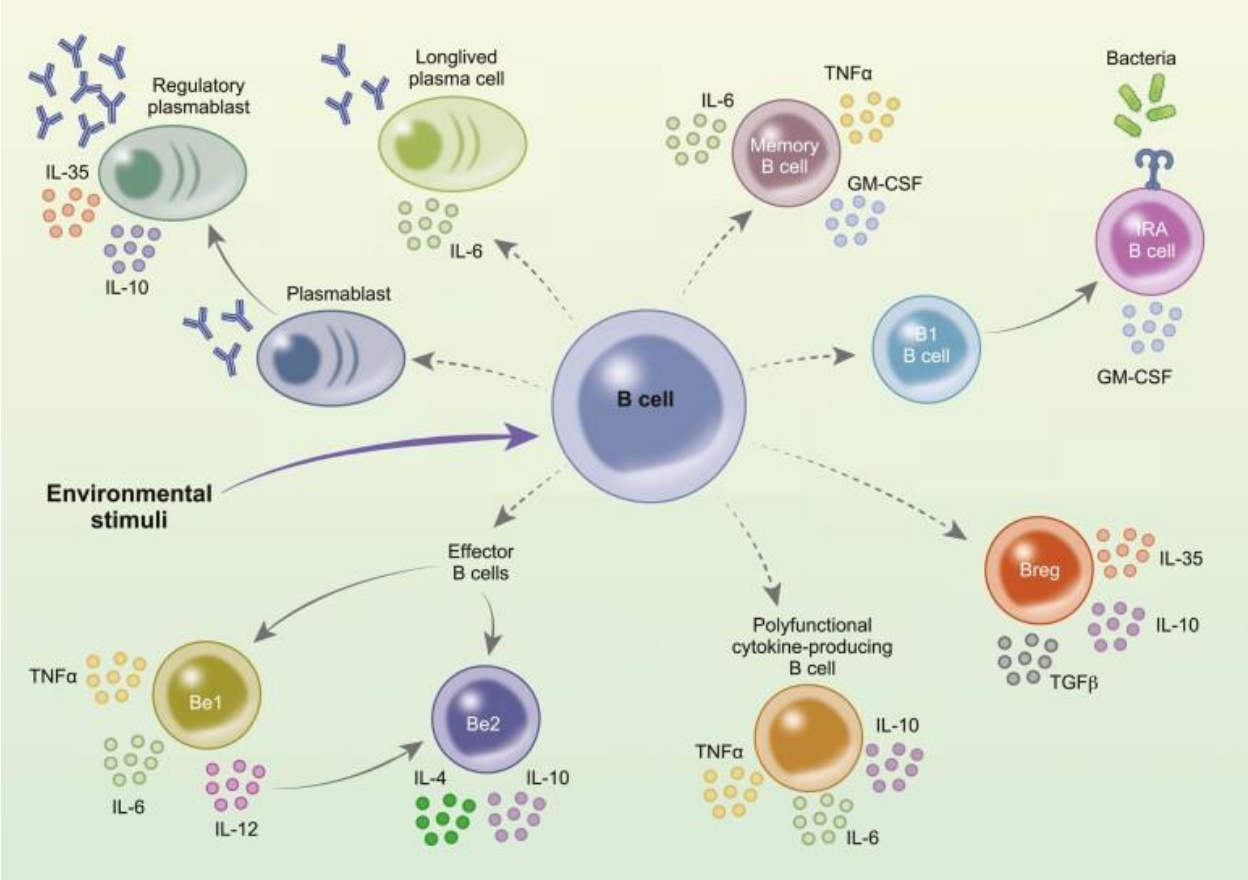


Figure 1.1 – Functional subsets of cytokine-producing B cells (*de Gruijter NM et al. 2022*).

IL: interleukin; TNF: tumour necrosis factor; IFN: interferon; LT: lymphotoxin; GM-CSF: granulocyte-macrophage colony-stimulating factor; Breg: regulatory B cell

The two main cytokines with the capacity to promote the survival and proliferation of B cells, B-cell activating factor (BAFF) and ‘a-proliferating inducing ligand’ (APRIL) have been strongly implicated in autoimmune disease (*Steri M et al. 2017*), with elevated serum BAFF levels demonstrated in individuals with GD (*Vannucchi G et al. 2012, Lin JD et al. 2016*). Indeed, an anti-BAFF human monoclonal antibody (belimumab) is currently approved to treat SLE and has demonstrated efficacy in the treatment of GO and reducing TRAb titres (*Favero V et al. 2021).* Furthermore, blocking BAFF and APRIL signalling pathways with soluble decoy receptors has been shown to ameliorate murine induced hyperthyroidism (*Gilbert JA et al. 2006*). BAFF alongside its receptor BAFFR is expressed in the thyrocytes and infiltrating immune cells in GD (*Campi I et al. 2015*). In addition, mRNA expression of BAFF in thyrocytes stimulated with CD40 has been found to be significantly upregulated in

individuals with GD when compared to healthy controls (*Lee HJ et al. 2017*). Functionally, BAFF has not only been demonstrated to promote the survival of potentially autoreactive B cells, but conversely to induce B cells with a regulatory phenotype that produce IL-10 (*Yang M et al. 2010*).

Both BAFF and APRIL regulate B cell homeostasis by binding to their receptors BAFFR, transmembrane activator calcium modulator and cyclophilin ligand interactor (TACI) and B-cell maturation antigen (BCMA) (*Meinl E et al. 2021*). The soluble forms of the latter two receptors, sTACI and sBCMA, are cleaved and shed from the membrane of activated late-stage B cells (memory/plasma cells) by a disintegrin and metalloprotease 10 (ADAM10) and γ-secretase, respectively (*Meinl E et al. 2021)* (Figure 1.2). Although not previously studied in GD, both sTACI and sBCMA have been found elevated in other autoimmune conditions including systemic lupus erythematous (SLE), rheumatoid arthritis (RA) and multiple sclerosis (MS) and suggested as potential biomarkers of disease activity (*Salazar-Camarena DC et al. 2020, Hoffmann FS et al. 2015, Rodríguez-Carrio J et al. 2018, Laurent, S et al. 2015, Vincent FB 2019*).

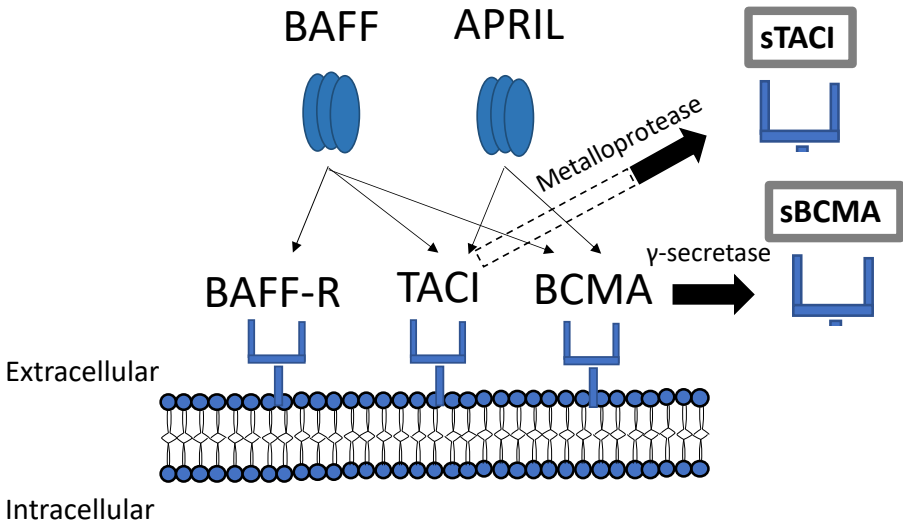


Figure 1.2 – B cell plasma membrane with BAFF and APRIL receptors and their associated cleaved soluble receptors.

Levels of the chemokine (C-X-C motif) ligand 13 (CXCL13) and ligand C-X-C chemokine receptor type 5 (CXCR5) – expressed on mature B cells and follicular B helper T cells, have been found elevated in autoimmune diseases such as SLE and

RA (*Zeng Y et al. 2021, Bugatti S et al. 2014).* The CXCL13/CXCR5 chemokine axis enables the migration of activated B cells to secondary lymphoid tissue where functional germinal centres are formed. There is a demonstrated correlation between CXCL13 and CXCR5 mRNA levels and the number of focal lymphocytic infiltrates and germinal centres in thyroid tissue (*Aust G et al. 2004*).

Kappa-deleting recombination excision circles

Kappa-deleting recombination excision circles (KRECs) are small, circular DNA products created by Variable-Diversity-Joining (V(D)J) gene recombination during B cell maturation in the bone marrow, forming a diverse repertoire of B cell receptors (*van Zelm MC et al. 2007*). The coding joint (CJ) remains stable in the genome, whereas the excised circle of DNA, otherwise known as the signal joint (SJ) or KREC, is diluted two-fold in the periphery with each cell division (Figure 1.3). As a consequence, evaluating KRECs provides not only a quantitative measure of B cell emigration from the bone marrow, but by calculating the CJ/SJ ratio we can understand the replicative history of an individual's B cell population (*van Zelm MC et al. 2007*). Despite their potential relevance, KRECs are relatively understudied in autoimmune disease but have been shown to be elevated in immune thrombocytopenic purpura (ITP) when compared to healthy controls (*Levy-Mendelovich S et al. 2017*). In relapsing GD patients examining KRECs may provide mechanistic insight into the immunopathogenesis of GD or allow earlier identification of the relapse process.

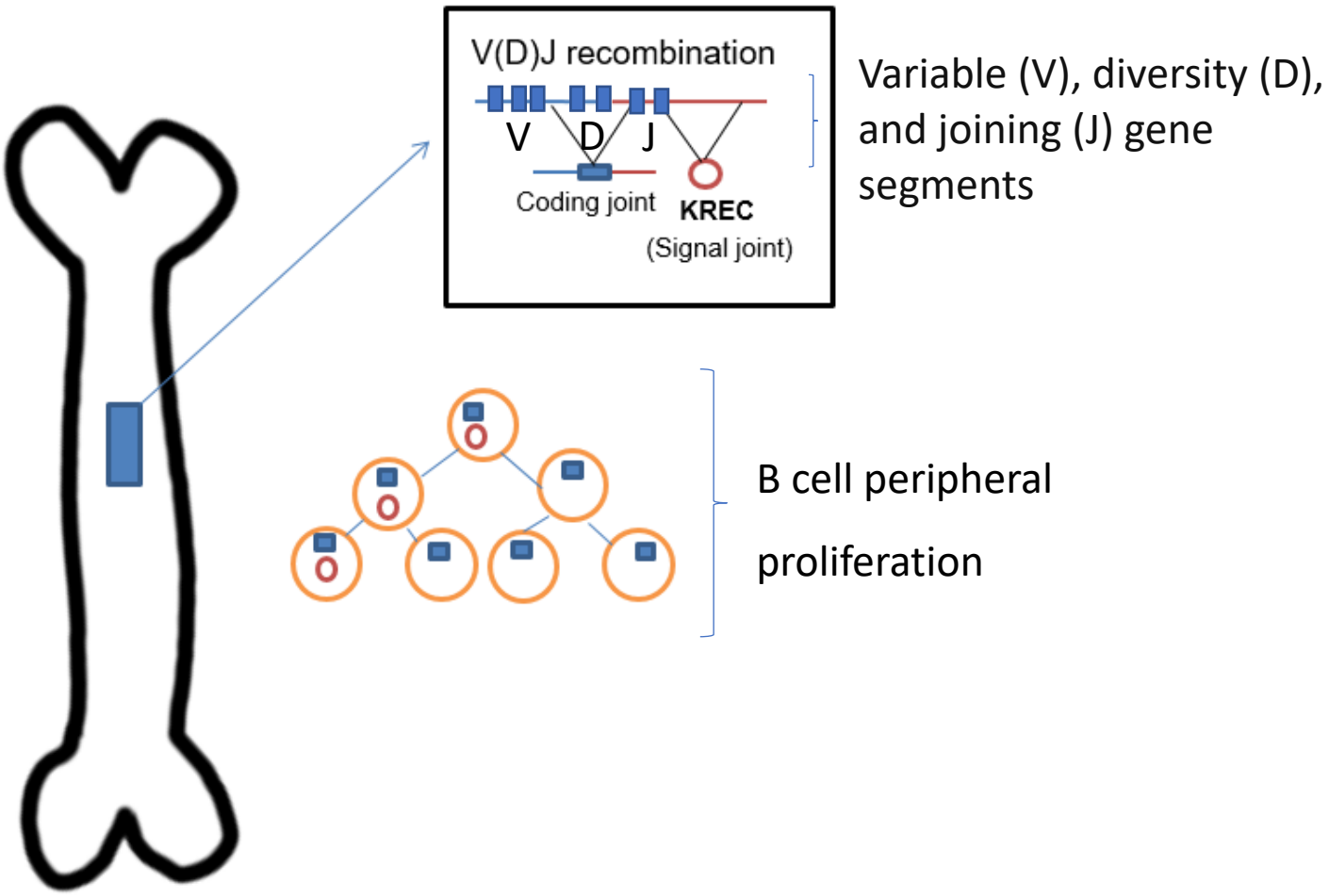


Figure 1.3 – Production of KRECs from the bone marrow during V(D)J recombination, followed by B lymphocyte peripheral proliferation

1.2.4 Summary

It is clear that B cells have a central role in the immunopathogenesis of GD, orchestrating and driving the autoimmune response both directly, by producing the pathogenic autoantibody and indirectly, as APCs with the additional capacity to produce pro-inflammatory cytokines and chemokines.

GD is one of the few autoimmune diseases where full and durable tolerance to endogenous antigens is restored in a large proportion of patients following routine therapy, and unlike other autoimmune diseases, there is a well-recognised single, specific antigen. Both these factors make GD an excellent model of autoimmune disease to study, with a clear target for novel therapies and an opportunity to understand the valuable and relatively unknown mechanisms of immune tolerance restoration.

1.3 Relapse and remission in Graves' disease

One of the main issues with ATD treatment is the high rate of relapse following cessation of therapy. After completing a 12-18 month course of ATD, relapse will occur in around 50-60% of adults (*Abraham P et al. 2010*) and 75% of children (*Rabon S et al. 2016, Wood CL et al. 2020*). In addition, 13% of the GD patients that don't relapse will develop spontaneous hypothyroidism (*Meling Stokland AE et al. 2024*). This has huge health, socioeconomic, educational, and psychological implications for these patients.

1.3.1 Defining relapse and remission

GD is defined as 'relapsed' if there is a return to biochemical hyperthyroidism (suppressed TSH with elevated serum T3/T4 levels) after cessation of ATD treatment, or after RAI where a second higher dose is required. The definition of remission differs between studies, with most patients being considered in 'remission' if they maintain euthyroidism 0.5-2 years after stopping ATD (*Abraham P et al. 2010*). Although 75% of relapses occur within the first 6 months of stopping ATD, up to 10% will relapse 18 months or more after stopping treatment (*Mohlin E et al. 2014*) and relapse of GD has been reported over 20 years after definitive treatment with RAI (*Hegele RA et al. 1985*). A quarter of patients with GD have reported not feeling 'fully recovered' after 6-10 years despite treatment (*Sjölin G et al. 2019*).

1.3.2 Factors associated with relapse

There are many factors that have been associated with an increased likelihood of relapse on stopping ATD (Table 1.3), including younger age (< 40 years), male gender, smoking, larger goitre size, biochemical severity of thyrotoxicosis at presentation and higher TRAb levels at diagnosis and on cessation of ATD.

Studies examining whether a longer duration of ATD increase rates of remission have produced contraindicating outcomes in both the paediatric and adult populations (*van Lieshout JM et al. 2021, Abraham P et al. 2010*). In a 10 year follow up study of GD in the paediatric population, Leger et al. documented a reduced overall relapse rate of 50% (compared to ~68% with a 24-month treatment course), with remission rates after withdrawal of ATD increasing with time (*Léger J et al. 2012*). This is in contrast to Wang et al. who found that a longer duration of treatment was associated with earlier relapse in adults (*Wang PW et al. 2013*), the

findings of which have been replicated (*Masiello E et al. 2018*). Remission rates have been observed to be higher in those who have completed a second course of ATD for relapsed GD (*Liu X et al. 2015*), with continuous ATD and therefore long-term restoration of euthyroidism said to be more effective than repeated cycles of therapy (*Mazza E et al. 2008*).

In addition to the clinical, immunological, biochemical, and genetic factors associated with relapse, there is some evidence that stress-related disorders are associated with an increased risk of autoimmune disease and therefore stressful life events could precipitate a GD relapse (*Song H et al. 2018, Vita R et al. 2015).*

Parameter	Factor	Studies
Clinical	Younger age	Allahabadia A et al. 2000, Kaguelidou F et al. 2008, Vitti P et al. 1997, Glaser NS et al. 2008, Vos XG et al. 2016, Meling Stokland AE et al. 2024
	Male gender	Magri F et al. 2016, Allahabadia A et al. 2000
	Non-Caucasian	Kaguelidou F et al. 2008
	Smoking	Anagnostis P et al. 2013, Kimball LE et al. 2002, Wang PW et al. 2013, Meling Stokland AE et al. 2024
	Increased goitre size/volume at diagnosis and cessation of ATD	Mohlin E et al. 2014, Magri F et al. 2016, Vitti P et al. 1997, Wang PW et al. 2013, Masiello E et al. 2018, Schleusener H et al 1989, Vos XG et al. 2016, Liu L et al. 2016
	Presence of GO at diagnosis	Meling Stokland AE et al. 2024
Biochemical	Elevated TRAb level at diagnosis	Kaguelidou F et al. 2008, Vitti P et al. 1997, Masiello E et al. 2018, Vos XG et al. 2016, Tun

		NN et al. 2016, Cappelli C et al. 2007
	Elevated TRAb level at cessation of ATD	Kaguelidou F et al. 2008, Schleusener H et al 1989, Kashiwai T et al. 2003, Tun NN et al. 2016, Cappelli C et al. 2007
	Severity of thyrotoxicosis at diagnosis (FT4)	Masiello E et al. 2018, Weetman AP et al. 1986, Vos XG et al. 2016
	Severity of thyrotoxicosis at diagnosis (FT3)	Glaser NS et al. 2008
Immunological	IL-6 TNFRSF9 CD40 at cessation of ATD	Meling Stokland AE et al. 2024
Genetic polymorphisms	*CTLA-4 (rs231775)*	Wang PW et al. 2013
	CD40 (rs745307, rs11569309, rs3765457)	Wang PW et al. 2013
	HLA subtypes *(DB1*02, DQA1*05, and DRB1*03)*	Vos XG et al. 2016, Vejrazkova D et al. 2021
	PD-L1 (A8923C)	Hayashi M et al. 2008
	PTPN22 (rs2476601)	Vos XG et al. 2016
	GNAS1 (T393C)	Glowacka D et al. 2009

Table 1.3 – Risk factors associated with relapse in Graves' disease.

1.3.3 Predicting relapse in Graves' disease

The 'gold standard' and best biochemical predictor we currently have to determine the likelihood of an individual relapsing after stopping ATD is the circulating TRAb level (*Barbesino G et al. 2013, Tun NN et al. 2016*). An observational study demonstrated that by four years, a TRAb level of <0.9 IU/L at cessation is associated with a 58% risk of recurrence compared with 82% when the TRAb level >1.5 IU/L (*Tun NN et al. 2016*). In particular, it has been suggested that the younger the age, the greater the predictive value of TRAb antibodies (*Bano A et al. 2019*). However, the TRAb level does not provide the full picture as even those with a negative TRAb level on cessation of ATD can relapse (22% by one year, rising to 60% four years after ATD withdrawal) (*Tun NN et al. 2016*).

In accordance with studies demonstrating that a suppressed TSH at the end of ATD is a poor prognostic factor for relapse (*Schleusener H et al. 1989*), it has been shown that increased TSH levels during ATD treatment provide a favourable indicator of long-term remission (*Choo YK et al. 2010*). It has been proposed that suppressed TSH levels combined with normalised thyroid hormones during treatment with ATD is either due to delayed pituitary-thyroid axis recovery or as a consequence of TRAb antibodies binding to a functioning pituitary TSH receptor, directly suppressing TSH production independent of thyroid hormone levels (*Brokken LJ et al. 2001*).

With the objective of improving the clinician's ability to anticipate relapse, predictive models have been generated using clinical and genetic parameters (*Vos XG et al. 2016*). Applying an individual's genotype for predictability of relapse using HLA subtypes and SNPs of co-stimulatory genes (*CD40* and *CTLA-4*) has provided contradictory results (*Schleusener H et al. 1989, Badenhoop K. et al. 1996, Wang F et al. 2012, Kim KW et al. 2007*). Differences in ethnicities between these studies may explain the variable influence of genetic polymorphisms.

Relapse prediction models

Relapse prediction models such as the 'GREAT' (Graves' Recurrent Events After Therapy) and 'GREAT+' cumulative scores (*Vos XG et al. 2016*), have been developed in an attempt to provide the clinician with the ability to predict relapse at the start of ATD therapy and therefore offer individualised treatment from the outset. The 'GREAT' score categorises patients by the chance of them relapsing their GD into three groups: 16% (I), 44% (II) and 68% (III), and includes the clinical parameters of patient age, FT4 and TRAb concentration, and goitre size at diagnosis. The 'GREAT +' score includes additional genetic markers, PTPN22 (rs2476601) and HLA subtypes DQB1*02, DQA1*05 and DRB*03. Although these prediction models are based solely on a Caucasian population and excluded those with severe hyperthyroidism, large independent retrospective studies including over 1000 patients have reported good external validity for predicting relapse using the 'GREAT' score (*Struja T et al. 2017, Masiello E et al. 2018*, *Zuhur SS et al. 2019*).

Another model intended to help predict relapse at baseline, the Clinical Severity Score (CSS), which includes severity of thyrotoxicosis, thyroid volume and GO has been studied in a clinical setting with 387 GD patients and compared to the 'GREAT' score, with both models demonstrating similar predictive values ('GREAT': AUC 0.63;

'CSS': AUC 0.60) (*Masiello E et al. 2018*). The clinical relevance of these prediction models to include all populations remains unknown, however both patients and clinicians have reported that the 'GREAT +' score would be a valuable addition in the management of GD (*Jansen HI et al. 2024*).

1.3.4 Summary

Although TRAb antibodies are currently the 'gold standard' for predicting GD outcome when stopping ATD, TRAb titres represent only a superficial marker of GD activity and hence individuals with negative TRAb will still go on to relapse their disease. Until we understand more about the relapse process, we are not able to adequately, and with any great certainty, predict when or if a patient will relapse on ATD withdrawal.

If we can predict which individuals are likely to relapse their GD, we can appropriately counsel patients, prevent unnecessary relapses with the subsequent health and socioeconomic consequences, reduce monitoring of those patients we are confident will remit and ultimately form personalised, therapeutic strategies.

Chapter 2: Objectives

Primary objective

The primary objective of this study is to determine whether baseline biomarkers at the time of stopping ATD are predictive of GD relapse, including clinical, biochemical, and immunological markers as well as the differential gene expression of $CD19^+$ B cells.

Secondary objective

The secondary objective of this study is to determine whether changes in clinical, biochemical, and immunological markers 6-10 weeks after stopping ATD are predictive of GD relapse.

Chapter 3: Methods

In this chapter, I describe the methods used in this PhD including the study design and patient recruitment, followed by a detailed account of all laboratory procedures, sample processing, methods of statistical analysis and details of the ethical approval and governance for this study.

3.1 Inclusion and exclusion criteria

The inclusion and exclusion criteria for the study are summarised in Table 3.1.

Inclusion Criteria	1. Aged 18 years or older 2. Diagnosis of Graves' disease (TRAb positive thyrotoxicosis) 3. Have received at least 12 months of antithyroid treatment for Graves' disease and are due to stop treatment. 4. Ability to give informed consent
Exclusion Criteria	1. Less than 18 years old 2. Due to continue with antithyroid treatment for Graves' disease 3. Diagnosed with other autoimmune disease(s) 4. Pregnant, breastfeeding, or plan for pregnancy 5. Serious co-existing cardiorespiratory, renal or liver disease 6. Diagnosed with cancer and undergoing active therapy 7. Diagnosed with an immunological disease 8. Concomitant use of immune-modifying medication other than antithyroid drugs 9. Intercurrent acute infective illness 10. Inability to provide informed consent

Table 3.1 – Inclusion and exclusion criteria for the study.

Patients who met the inclusion and exclusion criteria and were willing to stop ATD were enrolled.

3.1.1 Justification of inclusion criteria

The ethical and Health Research Authority (HRA) approvals for this study are based on an adult population and therefore only those aged 18 years and older were recruited. Additionally, the natural history of GD in the younger population as well as their underlying immunology differs from their older counterparts and therefore studying only those over 17 years of age ensures a more homogenous group of study participants. A formal diagnosis of GD was made by endocrinology clinicians which I then confirmed after review of the patient's biochemistry and medical notes prior to study enrolment. Biochemically, a diagnosis of GD was reflected by a suppressed TSH (<0.05mU/L) with elevated FT3 (> 6.8pmol/L) and /or FT4 (> 22pmol/L) with positive TRAb levels (>1.8IU/L). The specific assays used for the thyroid hormones and TRAb are documented below in Section 3.2.

The usual treatment for adult GD is for patients to receive at least a 12-month course of ATD medication before stopping treatment, therefore only patients that were at this stage of treatment and were willing to stop ATD medication were recruited to the study. It was very important from an ethical perspective that study participants could give informed consent to join the study and therefore they were only recruited if this was felt to be possible.

3.1.2 Justification of exclusion criteria

Given that the purpose of the study was to observe the patient's outcome following ATD cessation, patients were not recruited if they were due to continue with their ATD medication for any reason. Patients were also excluded from the study if they had certain comorbidities, including an additional coexisting autoimmune disease, a serious illness (including organ failure and malignancy) or were taking an immune-modifying medication, all of which have the potential to affect the $CD19^{+}$ B cell transcriptome and therefore the biomarker signature of GD. Additionally, patients who were pregnant/planning pregnancy or had an intercurrent acute infective illness were excluded due to the physiological and immunological changes associated in both conditions that could affect the transcriptome and immunological parameters, including the dynamics of cytokine and chemokine profiling.

3.2 Study definition of Graves' disease relapse and remission

In accordance with most other studies in GD as well as normal clinical practice, for the purposes of this study GD patients were defined as in 'remission' if they remained biochemically euthyroid (FT3 and FT4 in the normal range) 12 months following ATD cessation without either RAI or surgery. Alternatively, relapse was defined as biochemical hyperthyroidism (suppressed TSH (<0.05mU/L) with elevated FT3 (> 6.8pmol/L) and/or FT4 (> 22pmol/L)) within 12 months of stopping ATD.

Serum FT4, FT3 and thyroid peroxidase antibody (TPOAb) levels were determined using the commercial Roche Elecsys immunoassay (Roche Diagnostics) and serum TRAb levels were determined using the commercial BRAHMS TRAK immunoassay (Thermoscientific).

3.3 Study design and patient recruitment

The design of the SPRING (Study of Predicting Remission in Graves' Disease) study is summarised in Figure 3.1. Briefly, patients who were ready to stop ATD had blood samples taken at the timepoint of ATD withdrawal and again 6-10 weeks later. This second timepoint was chosen in order to try and capture the humoral immune response prior to relapse occurring. The participants were then followed up for 12 months at which point they were either categorised as in remission or relapsed as per the definitions discussed previously. During the period of monitoring after stopping ATD, patients had a clinical review and thyroid function checked every 3 months.

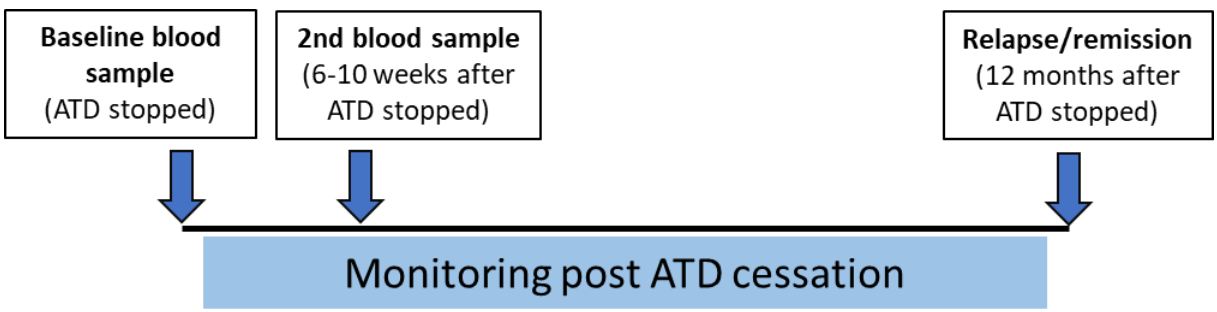


Figure 3.1 – SPRING study design

Patients who satisfied the inclusion criteria were recruited from endocrine NHS clinics at the Royal Victoria Hospital, Newcastle-upon-Tyne NHS Foundation trust and The Queen Elizabeth Hospital, Gateshead Health NHS Foundation Trust. All patients were recruited by me from both sites, but the samples were all processed at the same laboratory of the Royal Victoria Hospital, Newcastle-upon-Tyne to ensure consistency. The patient recruitment pathway is detailed in Figure 3.2.

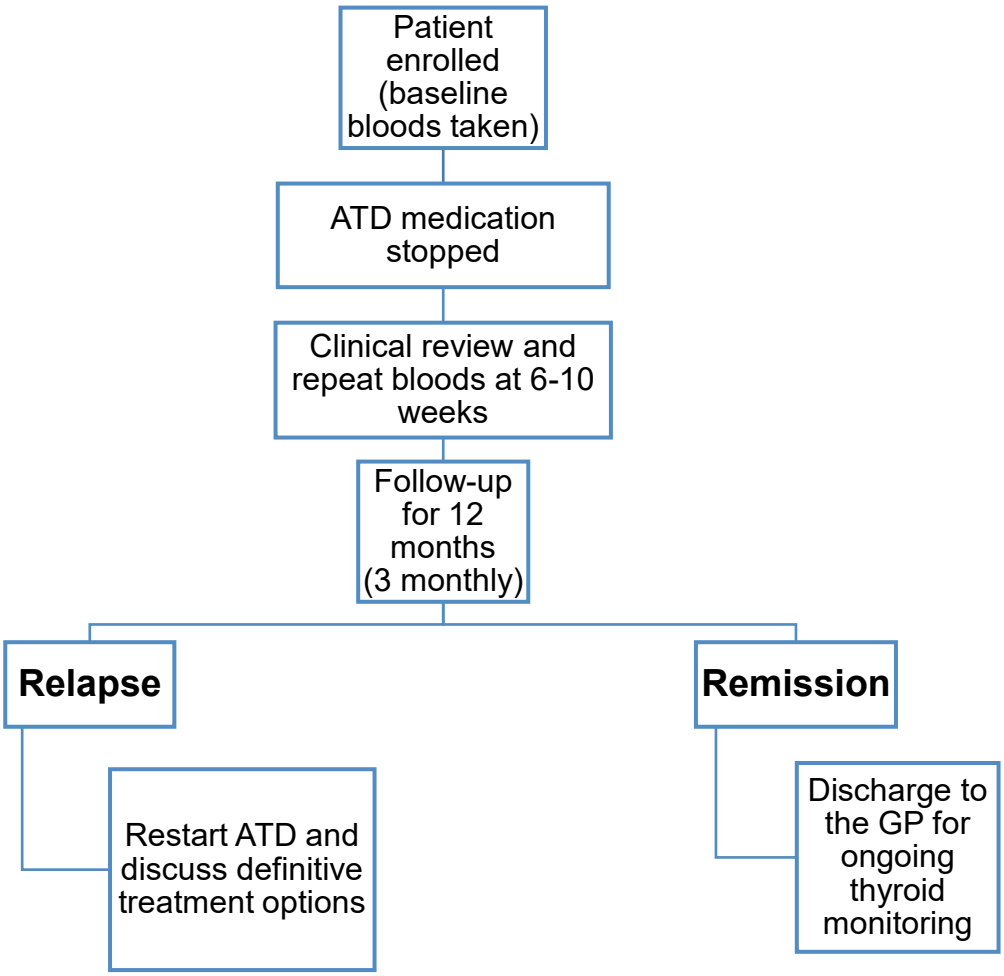


Figure 3.2 – Patient recruitment pathway. ATD: antithyroid drugs

The baseline clinical variables recorded at the time of patient recruitment to the study are summarised in Table 3.2. Retrospective clinical data was collected from the electronic patient record as shown in Table 3.3, including reasons for data collection.

Clinical variable	Data source
Age	Medical notes
Gender	Medical notes
Ethnicity	Patient consultation
Vaccine history	Patient consultation
Intercurrent illness	Patient consultation
Smoking status	Patient consultation
Family history of thyroid disease	Patient consultation
Previous history of GD	Medical notes/patient consultation
Co-morbidities	Medical notes/patient consultation
Current medication	Patient consultation
ATD regimen (BR or DT)	Medical notes/patient consultation
Duration of ATD medication	Medical notes/patient consultation
Goitre size (WHO 1994)	Patient examination
Presence and severity of GO	Patient examination

Table 3.2 – Baseline clinical variables recorded in the SPRING study.

Clinical variable	Reason
Date ATD commenced	To confirm length of ATD treatment
FT4 at diagnosis	To analyse association with outcome
FT3 at diagnosis	To analyse association with outcome
TRAb at diagnosis	To analyse association with outcome
Goitre size at diagnosis	To analyse association with outcome

Table 3.3 – Retrospective clinical variables recorded in the study.

Research blood tests taken included serum (for C-reactive protein (CRP) and cytokine/chemokine profiling) and plasma (for full blood count (FBC), DNA extraction, or for further processing into peripheral blood mononuclear cells (PBMCs) followed by $CD19^+$ B cell isolation and RNA extraction or analysis of lymphocyte subpopulations). Routine blood tests taken as part of usual clinical practice included thyroid function tests (TSH, FT3, FT4) and thyroid autoantibodies (TPOAb/TRAb). These blood tests were all taken between 0900am and 1230pm to minimise the risk of circadian variation between samples.

Patients continued to be reviewed in clinic every 3 months for a year following ATD cessation, where they were monitored clinically and biochemically for relapse which is in accordance with usual NHS clinical practice. Patients were able to get in contact in between these set appointments for testing of their thyroid status if they had symptoms consistent with relapse. If a patient relapsed ATD was re-commenced, and they exited the study to continue to be followed up in the endocrine clinic. A comprehensive schedule of events including blood tests taken is presented in Table 3.4.

Activity	Baseline	Week 6-10	3 monthly	Week 52
Informed consent	X			
Medical history	X			
Patient examination for goitre size/GO	X			
ATD cessation	X			
TSH/FT4/FT3*	X	X	X	X
TPOAb/TRAb[1]	X	X	X	X
CRP	X	X		
FBC	X	X		
Research bloods[2]	X	X		

[1] TRAb level was checked when the patient was close to having had 12 months of ATD to enable a decision regarding stopping treatment.
[2] Research bloods taken included serum and plasma as described above

Table 3.4 – Schedule of events for the SPRING study. **3.4 Sample size estimation and power calculations**

The patient outcome of remission or relapse one year after stopping ATD was used to categorise each patient for analysis.

3.4.1 Power calculation for the RNAseq experiment

Although the SPRING study was conducted with an exploratory approach due to the absence of existing gene expression data in GD relapse, a formal power calculation to estimate the required sample size was performed using a model developed to estimate the statistical power needed to identify differentially expressed genes from RNAseq experiments (*Hart et al. 2013*). This power calculation is presented in Table

3.5. This calculation is based on a sequencing depth of 50 million reads, equating to a coverage of at least 50 reads (μ) for more than 90% of transcripts.

The coefficient of variation (CV) gives an estimate of biological variation. Data obtained from the RNAseq analysis of human B-cells (*Toung et al. 2011*) demonstrated the CV of the FPKM (fragments per kilobase of exon model per million mapped reads) values for the B cell genes were 0.10 ± 0.16 and 0.49 ± 0.53 (mean ± standard deviation). Therefore, an estimate of CV 0.4 was used for this power calculation, along with a standard α of 0.05.

The RNAseq experiment was designed with the aim of having 23 patients in each outcome group, however this would have meant getting 100 patients recruited (~25% relapse rate by one year) which was not possible in the allotted timeframe of the project. Additionally, patient recruitment was hindered by the COVID-19 pandemic suspending research and affecting patient attendance to hospital for at least 6 months. However, it was estimated that having at least 16 patients in each group would be well powered to detect effect sizes of around 1.9 and adequately powered to detect those around 1.6, depending on transcript abundance.

Effect size (fold change; Δ)	Power (β)	Sample size (per group)
1.25	0.65	40
1.3	0.80	40
1.5	0.90	23
1.62	**0.90**	**16**
1.9	**0.99**	**16**
2	0.99	14

Table 3.5 – Power calculation for the RNAseq experiment

3.4.2 Power calculation for BAFF concentrations

The original power calculation for detecting a significant difference in BAFF concentrations was based on a sample size of 50 patients who relapsed and 50 who remained in remission which would have provided >80% power to detect a 10% difference in BAFF concentrations between the groups (given published estimates of BAFF concentration and SD; α=0.05) (*Lin et al. 2016*).

3.5 Laboratory procedures

Details of the reagents and equipment used in the laboratory procedures including catalogue numbers and manufacturer are listed in Appendix A.

An overview of the handling and processing of the blood samples collected are summarised in Figure 3.3. Most of these processes were completed on the freshly collected blood sample and all were performed using an aseptic technique.

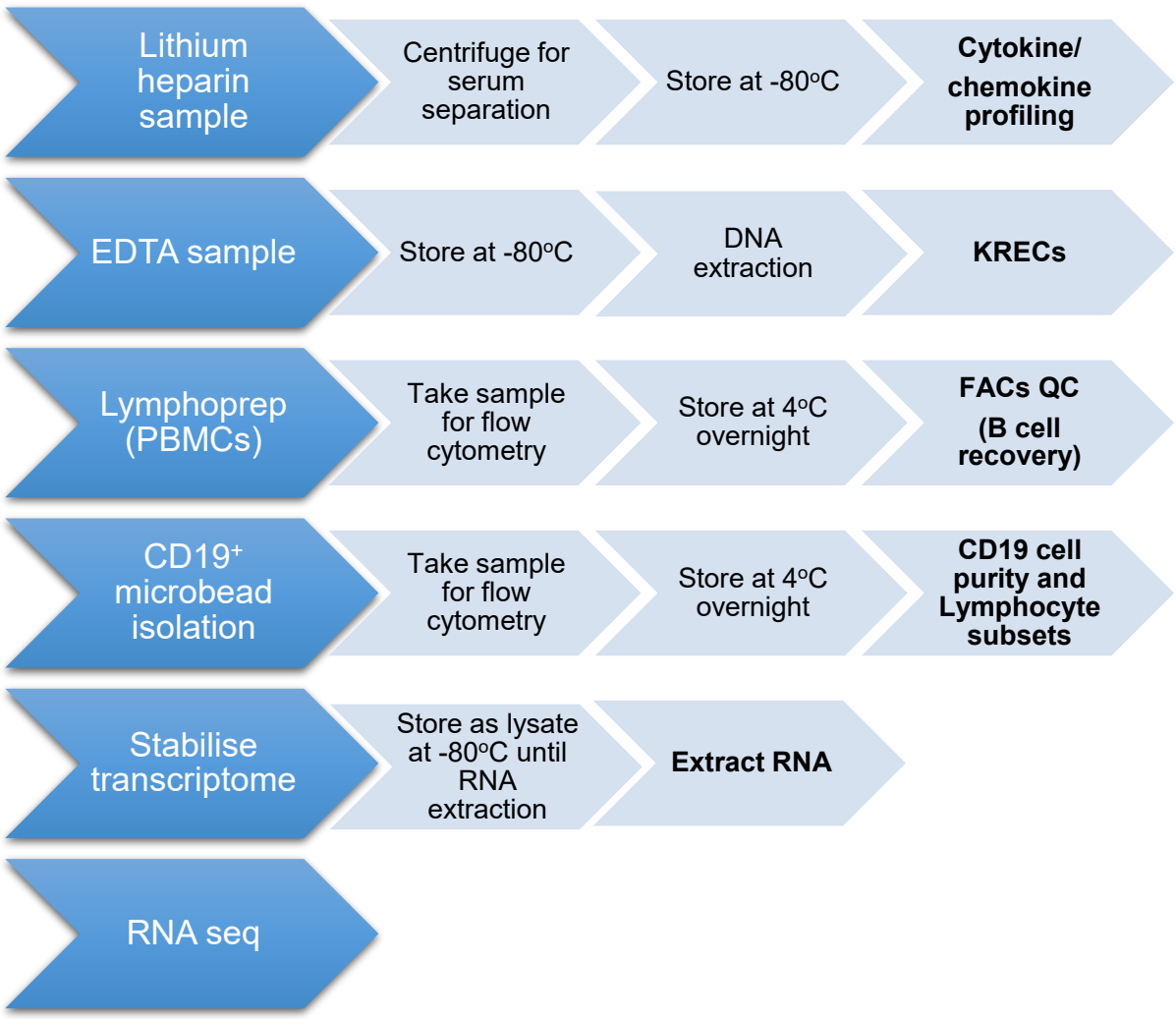


Figure 3.3 – Workflow detailing sample handling, processing, and storage conditions.

3.5.1 Laboratory assays

The TSH, FT4, FT3, TRAb and TPOAb assays used are documented above (Section 3.2). CRP was determined using the commercial particle-enhanced immunoturbidimetric assay (Roche). These laboratory tests were undertaken by the Biochemistry laboratory at the Royal Victoria Infirmary, Newcastle-upon-Tyne.

3.5.2 Serum separation

The lithium heparin tube was centrifuged at 1800xg at room temperature for 10 minutes. The serum was then removed by pipetting and separated into 1ml aliquots and stored at $^{-}80^{O}C$ for future analysis.

3.5.3 DNA extraction from whole blood (non-clotted)

The EDTA samples were stored at -80ºC before being batch processed for DNA extraction. The sample was thawed at room temperature before the addition of buffer (containing Tris-HCL, sucrose, magnesium chloride, distilled water and Triton X-100). The sample was shaken for four minutes before being centrifuged at 1500xg. The cell pellet was resuspended in the buffer and the process repeated.

The cells then underwent lysis with the addition of buffer containing 1% sodium dodecyl sulfate (SDS), before sodium perchlorate was added followed by incubation at 65ºC for 30 minutes. The DNA was extracted by the addition of 2mls of chloroform followed by centrifuge at 2900xg for 10 minutes. Cold ethanol was added to precipitate the DNA and this precipitate was removed and centrifuged in 70% ethanol at 14,000 rpm for 5 minutes. The ethanol was removed, and the pellet allowed to dry at room temperature before the DNA was dissolved in 300µl of TE buffer. The extracted DNA was then quantified on the Thermofisher NanoDrop™ 8000 Spectrophotometer and stored at 4ºC.

3.5.4 DNA extraction from whole blood (clotted)

The healthy control samples from the NIHR BioResource were stored at -80ºC in lithium heparin tubes in the form of clotted blood. Therefore, the DNA extraction process was followed as above except for the addition of some initial steps which resulted in homogenisation of the clot. After the samples were thawed at room temperature, the lithium heparin tubes were inverted in 50ml Falcon tubes and centrifuged at 2,000xg for 3 minutes to remove the clotted blood from the tube. The clot was broken up in the Falcon tube using a scalpel and 2mls of 0.05% trypsin with EDTA was added, before the sample was placed in a 37ºC water bath for one hour. Following this, the DNA extraction method continued as per Method 3.5.2.

3.5.5 Kappa light chain excision circles (KRECs)

KRECs were quantified by real-time quantitative polymerase chain reaction (RT-qPCR) using a triple-insert plasmid (containing T-cell receptor excision circle (TREC) and KREC signal joint (SJ) fragments and the T Cell Receptor Alpha Constant (TCRAC) reference gene) supplied by Sottini and colleagues (*Sottini et al. 2014*). The primers and probes used in this experiment are detailed in Appendix B.

Serial 1:10 dilutions of the plasmid in TE buffer were performed to generate a 6-point standard curve (Figure 3.4), with quantities of 10^1 to 10^6. The highest dilution of the standard curve was undertaken just before performing the assay due to the instability of the small quantity of plasmid DNA when stored.

The primers and probes were thawed at room temperature before being put on ice. Four separate mixes were made: standard curve, CJ, SJ, TCRAC. The standard curve mix had a total volume of 15 µl consisting of: 10 µl 2× TaqMan Universal PCR master mix, 0.9 µl forward/reverse primers (final concentration 0.6µM) for TCRAC and KRECs SJ, 0.4 µl probe (final concentration 0.2µM) for TCRAC and KRECs SJ and 0.6 µl nuclease free water. The CJ, SJ and TCRAC mixes had a total volume of 9 µl consisting of: 5 µl 2× TaqMan Universal PCR master mix, 0.45 µl forward/reverse primers (final concentration 0.6µM) (SJ/CJ/TCRAC), 0.2 µl probe (final concentration 0.2µM) (SJ/CJ/TCRAC) and 2.9 µl nuclease free water.

Genomic DNA was then added to the sample wells (200ng; 2 µl) and nuclease free water was added to the negative control wells. The plate was then taken to a different location to ensure containment of plasmid DNA carry-over. There, each dilution point of the plasmid DNA standard curve was thawed just before use and the last point of the standard curve was diluted. 5 µl of each point dilution of plasmid DNA was added to the standard curve wells. The plate was covered in an adhesive strip before being centrifuged at 1200xg at 4^{o}C for 1 minute to ensure the absence of bubbles at the bottom of the wells.

The assay was run on the QuantStudio™ 7 Flex Real-Time PCR System in standard mode thermal cycling; 50°C for 2 minutes, 95°C for 10 minutes, 95°C for 15 seconds; 60°C for 1 minute – 45 cycles. Both the standard curve and samples were run in triplicate, and 10% of samples were repeated to check the coefficient of variation. Standard curves were run on each plate and PCR efficiency was monitored using the efficiency percentage and slope value, aiming for an efficiency between 90-100% ($-3.6 \geq$ slope ≥ -3.3). The data was exported to the Thermo Fisher Connect Platform where the standard curve was generated, and further analysis was undertaken.

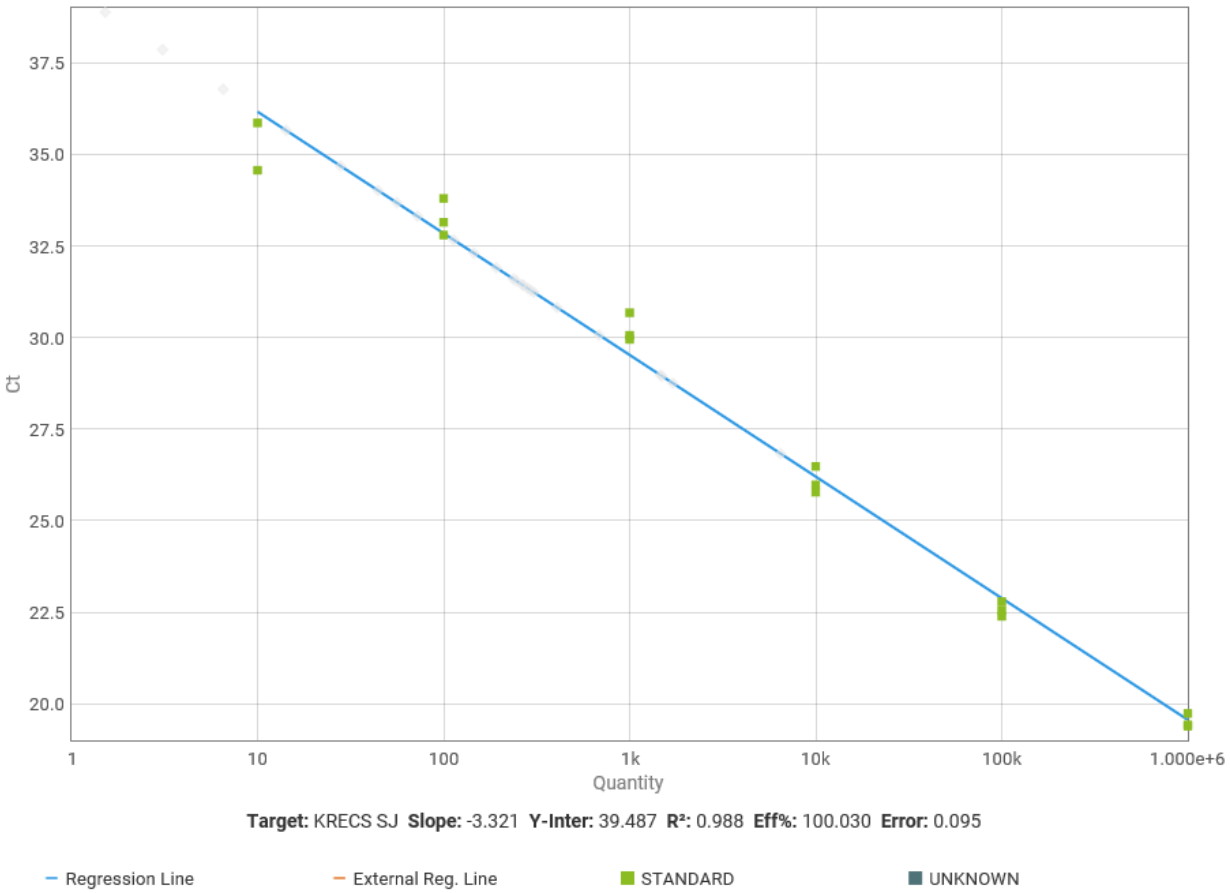


Figure 3.4 – Example of a KREC SJ standard curve generated using serial dilutions of the plasmid

KRECs calculations

The quantity of KRECs is reported as number of KRECs per 10^6 cells. The triple insert plasmid used to quantify the KRECs contained a fragment of TCRAC that served as a reference gene. To determine the number of KRECs per 10^6 cells the following formula was used:

$$\frac{\text{(mean quantity of KRECs)}}{\text{(mean quantity of TCRAC/2) x } 10^6}$$

(The mean quantity of TCRAC is divided by 2 because in each cell there are two TCRAC gene copies)

As described in Chapter 1: Introduction, the CJ remains stable in the genome of the B cell, whereas the excised circle of DNA, otherwise known as the SJ or KREC, is diluted two-fold in the periphery with each cell division. Thus, by using the Ct values to calculate the ΔCT (Ct KREC − Ct CJ)) along with the formula $^2\log\Delta CT$, van Zelm

et al. demonstrated that this is equal to the CJ/SJ ratio (CJ:KREC ratio) and can be used to estimate the replicative history of an individual's B cell population (*van Zelm et al. 2007*) (Figure 3.5).

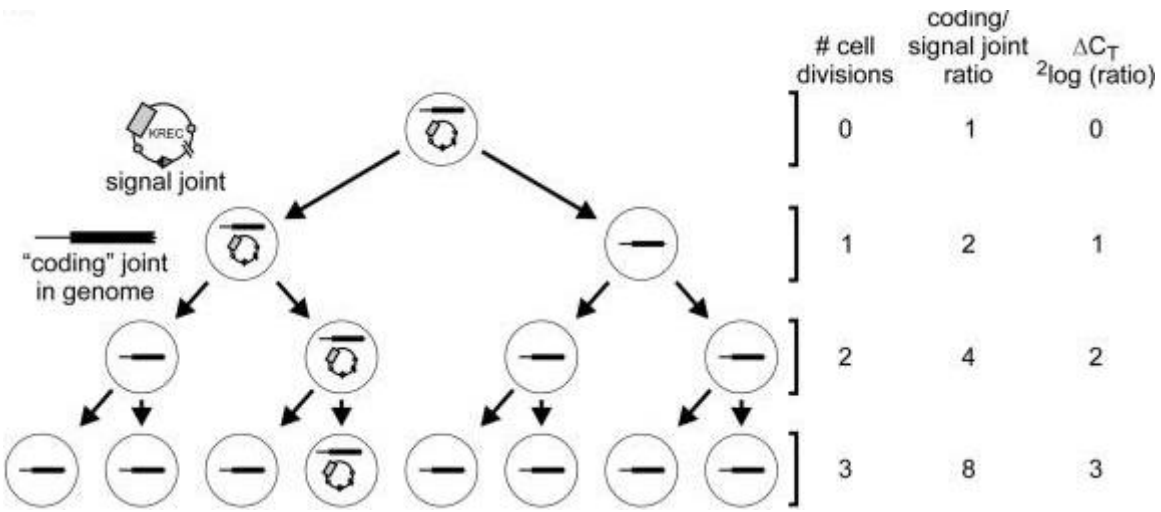


Figure 3.5 – Principle of quantification of the replication history of B cells using KRECs (*van Zelm et al. 2007*).

3.5.6 Peripheral blood mononuclear cell (PBMC) isolation

PBMC isolation was performed at room temperature in a tissue culture hood. PBMCs were isolated from 90-100mls of whole blood drawn into ten 10ml EDTA tubes (sometimes less if the total volume was unable to be drawn from the patient). The blood was diluted 1:1 with Hanks + 2mM EDTA. With 15mls of Lymphoprep in a 50ml Sepmate tube, 25mls of the diluted blood was then layered on top by slow pipetting. These tubes were then centrifuged at 1200xg for 10 minutes at room temperature with the brake on to maintain the separation of the layers. The plasma layer was aspirated without disturbing the buffy coat layer. The tube was then briefly inverted (for no more than two seconds) into a new 50ml tube. The PBMCs were then washed in cold Hanks+1% FCS to a total volume of 50mls, and immediately centrifuged again at 600xg for 7 minutes to remove any contaminating Lymphoprep. The supernatant was then removed, and the cells were resuspended in cold Hanks+1% FCS to a total volume of 50mls, before being centrifuged again at 300xg for 7 minutes to remove any contaminating platelets. The supernatant was removed again, and the cells were resuspended and pooled into one tube with cold Hanks+1%

FCS to a total volume of 50mls. This was then centrifuged again at 300xg for 7 minutes to remove any remaining platelets.

The supernatant was aspirated, and cells resuspended in cold Hanks+1% FCS before being strained through a 70µm nylon filter to remove any debris/clumps. The cells were kept on ice while PBMC cell counting and viability was performed using trypan blue staining and a Burker counting chamber, using the following calculations:

$$\textit{Cell concentration (per ml)} = \textit{[no.of cells in 25 squares]} \times \textit{[dilution factor]} \times 10^4$$

$$\textit{Total no.of cells} = \textit{[no.of cell in 25 squares]} \times \textit{[volume in ml]} \times \textit{[dilution factor]} \times 10^4$$

5 x 10^4 PBMCs were transferred to an Eppendorf tube and stored at 4°C for flow cytometry analysis. The remaining cells were used for the $CD19^+$ B cell isolation.

3.5.7 $CD19^+$ B cell isolation

The $CD19^+$ B cells were isolated by positive selection using Miltenyi CD19 magnetic Microbeads® as shown in Figure 3.6. These Microbeads are 50nm superparamagnetic particles conjugated with monoclonal CD19 antibodies (isotype: mouse IgG1).

The positive selection process involves magnetic labelling of the $CD19^+$ B cells with the Microbeads before they are loaded onto a MACS® LS column which is placed in the magnetic field of a MACS separator. The magnetically labelled $CD19^+$ B cells are retained in the column, while the unlabelled cells are washed out. After removal of the column from the magnetic field, the $CD19^+$ B cells can be eluted.

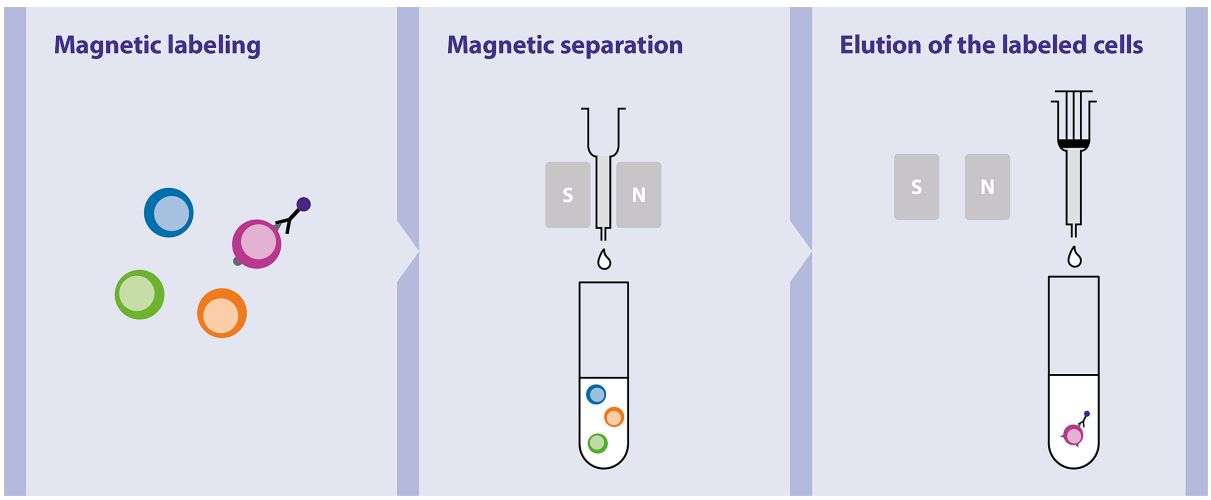


Figure 3.6 – Positive cell selection using Miltenyi® Microbead technology

The $CD19^+$ B cell isolations were carried out using cold reagents with the cells kept on ice and using only calcium and magnesium free media to prevent capping of antibodies on the cell surface and non-specific labelling. The volumes described below are for up to 10^7 total cells, therefore the reagent volumes were scaled up accordingly for larger cell numbers.

The MACS buffer contained 500ml PBS (Ca^{2+}/Mg^{2+} free), 2.5mls (0.5%) FCS and 2ml Endotoxin free 0.5M EDTA. It was filtered through a 0.2µm filter and cooled on ice prior to use.

The tube containing the remaining cells following PBMC isolation was topped up to a total volume of 50mls with cold Hanks+1% FCS before being centrifuged at 400xg at 4°C for 7 minutes. The supernatant was aspirated and the PBMC cell pellet was resuspended in cold MACS buffer (80µl MACS buffer per 10^7 cells) and CD19 Microbeads added (20µl per 10^7 cells). The cells were gently shaken and incubated at 4°C in the refrigerator for 15 minutes, before being washed by adding 1-2mls MACS buffer per 10^7 cells and centrifuged as before (400xg at 4°C for 7 minutes). The supernatant was completely aspirated, and cells resuspended in 500µl MACS buffer (500µl per 10^8 cells).

A positive selection column was placed onto the MACS separator and pre-rinsed with 3mls MACS buffer. The cell suspension was added to the column before being washed 3 times by adding 3mls ice-cold MACS buffer and allowing it to drain through the column. After these washing steps, the column was removed from the separator and placed into a 15ml centrifuge tube where the positive cells were eluted from the column by adding 5mls of MACS buffer and pushing through with the plunger. The tube was filled with MACS buffer and spun at 400xg at 4°C for 7 minutes. The supernatant was aspirated, and cells resuspended in 2mls cold Hanks+1% FCS before being counted as described for the PBMC cells (Method 3.5.6).

5 x 10^4 cells were transferred to an Eppendorf tube and stored at 4°C for purity checking by flow cytometry. A further 2 x 10^5 cells were transferred to an Eppendorf tube and stored at 4°C for lymphocyte subsets by flow cytometry. The remaining cells were then lysed.

3.5.8 CD19⁺ B cell lysis

Throughout the lysis process, RNase and DNase-free filter tips and tubes were used as well as designated RNA pipettes. The laboratory work area was cleaned with RNase*Zap*™ prior to cell processing.

The 15ml centrifuge tube containing the remaining cells was filled with cold Hanks +1% FCS and inverted to mix. It was centrifuged at 400xg at 4°C for 7 minutes, before the supernatant was completely aspirated and cell pellet resuspended. 1ml of TRIzol™ was added to the pellet, after which the lysate was pipetted up and down several times to homogenise. It was left at room temperature for 5 minutes before being transferred to a 1.5ml microcentrifuge tube to be stored immediately at -80°C until RNA extraction.

3.5.9 CD19⁺ B cell RNA extraction

As in Method 3.5.8, RNase and DNase-free filter tips and tubes were used as well as designated RNA pipettes. The laboratory work area was cleaned with RNase*Zap*™ prior to cell processing.

The frozen CD19⁺ B cell lysate (Method 3.5.8) was thawed at room temperature before being mixed by pipetting. To produce phase separation, 200µl of chloroform was added for every 1ml of TRIzol™ before being shaken vigorously for 15 seconds and incubated at room temperature for 3 minutes. The sample was centrifuged at 12,000rcf at 4°C for 15 minutes. Following centrifugation, the mixture separates into a lower red, phenol-chloroform organic phase, an interphase, and a colourless upper aqueous phase (containing the RNA).

The RNeasy® Mini Kits (Qiagen) were used to extract the RNA from the homogenised lysate as shown in Figure 3.7.

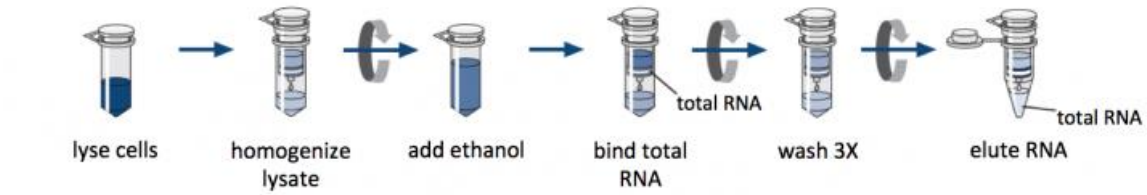


Figure 3.7 – Overview of the RNeasy® process to extract RNA

The aqueous phase was carefully removed by pipetting into a new 2ml tube. One volume of 70% ethanol was added to the homogenised lysate and mixed well by pipetting. Adding ethanol promotes selective binding of RNA to the RNeasy membrane. In the RNeasy spin column, the total RNA binds to the membrane, contaminants are washed away and RNA is eluted. Up to 700µl of the sample was added to a RNeasy spin column and centrifuged for 15 seconds at 13,000rpm. The flow-through was discarded and the process repeated if sample volume exceeded 700µl. 350µl of Buffer RW1 was added to the spin column and centrifuged for 15 seconds at 13,000rpm.

The flow-through was discarded and DNase Digestion was performed to reduce DNA contamination. 10µl of DNase I stock solution was added to 70µl Buffer RDD, before being mixed by inverting and brief centrifuge to collect residual liquid from the sides of the tube. The DNase I incubation mix (80µl) was added directly to the spin column membrane and left at room temperature for 15 minutes. 350µl of Buffer RW1 was added to the spin column and centrifuged for 15 seconds at 13,000rpm. The flow-through was discarded and 500µl of Buffer RPE was added to the spin column and centrifuged for 15 seconds at 13,000rpm. After the flow-through was discarded, this step was repeated but with a centrifuge of 2 minutes at 13,000rpm.

The old collection tube was discarded with the flow-through and the RNeasy spin column was place into a new 2ml collection tube. This was centrifuged for one minute at 13,000rpm. The RNeasy spin column was placed into a new 1.5ml tube and 30µl of RNase-free water was added directly to the spin column membrane. After incubation at room temperature for 10 minutes, the column was centrifuged for one minute at 13,000rpm to elute the RNA. The quantity of RNA was measured using 1µl of each sample and a Qubit™ 4 Fluorometer before being stored at $^{-}$80^{o}C.

3.5.10 PBMC recovery and CD19⁺ purity check by flow cytometry

Flow cytometry was performed for 3 purposes: 1) To determine the starting quantities of T and B cell populations in PBMCs, 2) To determine CD19⁺ B cell purity following B cell isolation, 3) To identify CD19⁺ B cell subpopulations.

Prior to commencing, compensation settings were established using BD™ CompBeads. These anti-mouse immunoglobulin polystyrene microparticles were used to optimise fluorescence compensation settings for the multicolour flow cytometry to correct for spectral overlap. All flow cytometry data was recorded using the BD LSRFortessa™ Cell Analyzer and BD FACSDiva™ Software. Analysis of the flow cytometry data was performed using FCS Express™ (Version 7).

Flow cytometry was carried out within 18-24 hours of the cells being stored (Method 3.5.6). The flow cytometry staining buffer used contained 500mls PBS (Ca2+/Mg2+ free), 2.5g BSA, 1ml Endotoxin free 0.5M EDTA and 500µl 10% sodium azide solution.

Unfixed samples of PBMCs and CD19⁺ B cells stored at 4°C in Eppendorf tubes were resuspended in PBS and centrifuged at 400rcf at 4°C for 3 minutes. The supernatant was removed, and the pellet resuspended by flicking before centrifuge was repeated. The supernatant was again removed, and the pellet was resuspended in flow cytometry staining buffer. Human IgG and antibodies were added to each sample and mixed by pipetting. The quantities of each and flow cytometry wavelengths used are presented in Table 3.6. The flow cytometry antibodies were used to identify the following cell populations: CD3⁺ ⁺/- CD4⁺ T cells, CD14⁺ monocytes, and CD19⁺ B cells. Human IgG (FcX blocker) was used to block nonspecific binding of antibodies to Fc-receptors and prevent background fluorescence.

Cells	FACs buffer	hIgG	405 450/50	488 530/30	561 582/15	635 670/30

			Violet CD3	Blue CD19	Green CD14	Red CD4
PBMC/CD19^{+}	41µl	2µl	1µl	2.5µl	1µl	2.5µl

Table 3.6 – Quantities of FACS buffer, human immunoglobulin G (hIgG), and flow cytometry antibodies for PBMC cell recovery and B cell purity checks.

After the antibody mix was added, the sample was incubated in the dark on ice for 30 minutes. Following this, 100µl of FACs buffer was added before centrifuging at 400rcf at 4^{o}C for 5 minutes. After removal of the supernatant and resuspension by flicking, the stained cells were then resuspended in 150µl of FACS buffer and centrifuged at 400rcf at 4^{o}C for 5 minutes. After further removal of the supernatant and resuspension by flicking, the cells were resuspended in 300µl FACS buffer and transferred to labelled FACS tubes for processing.

Gating strategies for PBMC cell recovery and CD19^{+} B cells for purity are presented in Figure 3.8. The PBMC samples were first gated on side-scatter area (SSC-A) and forward-scatter area (FSC-A) to exclude debris, which were then gated on forward-scatter height (FSC-H) and forward scatter area (FSC-A) to exclude doublets. This population were then gated on SSC-A and FSC-A to identify the lymphocyte population. This lymphocyte population was then gated on CD3 and CD4 to identify CD3^{+}CD4$^{+/-}$ T cells, and SSC-A and CD14 and CD19 to identify CD14^{+} monocytes and CD19^{+} B cells, respectively.

To assess B cell purity, the gating strategy was followed as above, but the total singlet cell population was gated to determine the CD3^{-}CD4^{-} population from which the CD19^{+} B cells and CD14^{+} monocyte populations were gated.

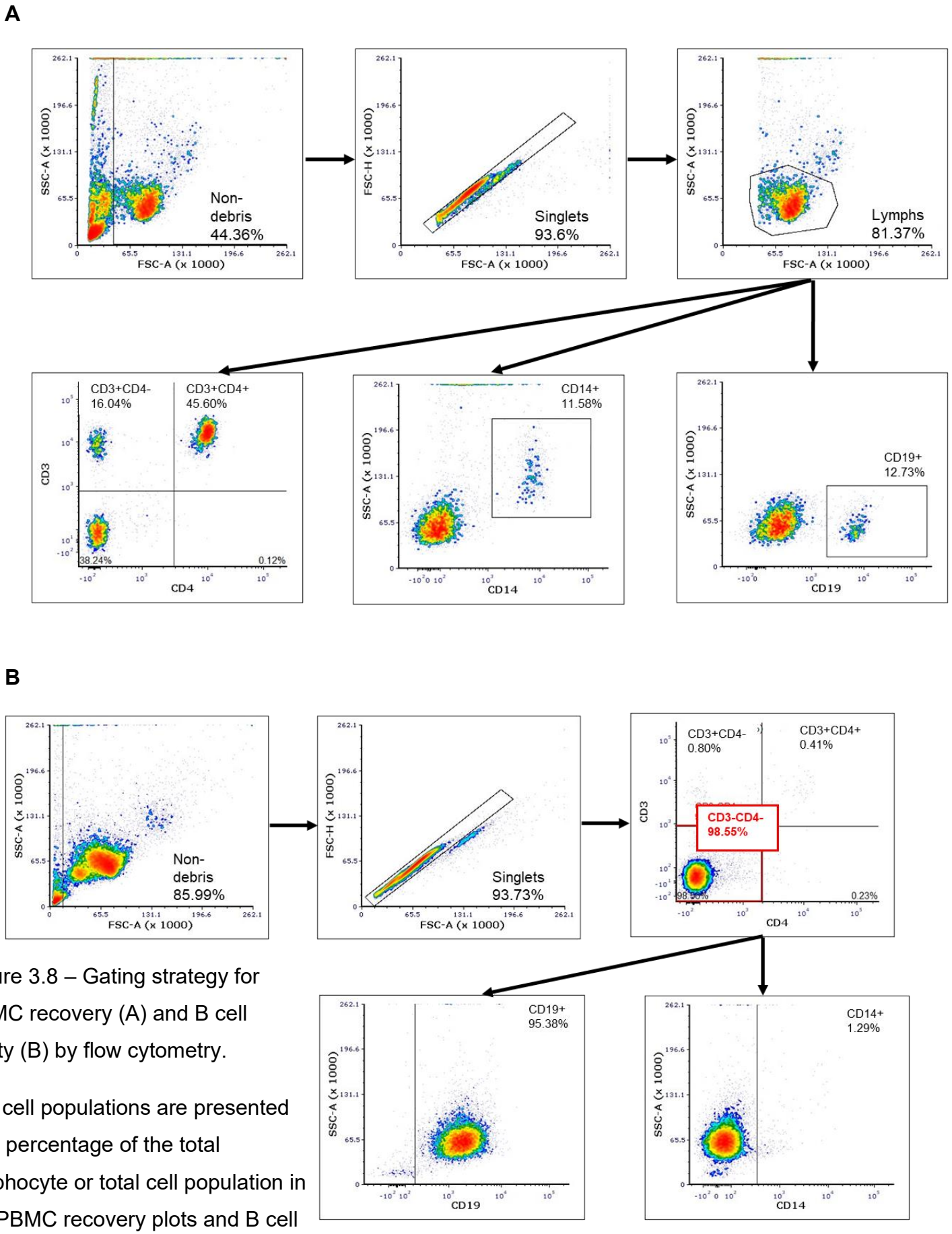


Figure 3.8 – Gating strategy for PBMC recovery (A) and B cell purity (B) by flow cytometry.

The cell populations are presented as a percentage of the total lymphocyte or total cell population in the PBMC recovery plots and B cell purity plots, respectively. SSC-A: side-scatter area; FSC-A: forward scatter area; FSC-H: forward-scatter height.

3.5.11 Immunophenotyping CD19⁺ B cell subpopulations by flow cytometry

The method for staining the B lymphocyte subpopulations was the same as Method 3.5.10, apart from the addition of Zombie dye™ after the first stage of cell resuspension and washing to exclude the dead cell population. Following the addition of 0.5µl of Zombie Aqua™, the cells were incubated in the dark at room temperature for 10 minutes. 100µl of FACS buffer was then added to quench the dye before being further incubated in the dark at room temperature for 5 minutes. The method continues as in 3.5.10, except for the different antibodies added for staining – the quantities of each and flow cytometry wavelengths used are presented in Table 3.7. The phenotypes of the B lymphocyte subpopulations studied are detailed in Chapter 7: Results 4 – FACs data.

Cells	FACs buffer	hIgG	405 450/50 Violet CD38	405 670/30 Violet CD27	488 530/30 Blue CD19	488 710/50 Blue IgD	561 582/15 Green CD1d	561 780/30 Green CD5	635 670/30 Red CD24	635 780/60 Red CD20
CD19⁺	60.5µl	4µl	5µl	4µl	2.5µl	10µl	10µl	2µl	4µl	1µl

Table 3.7 – Quantities of FACS buffer, human immunoglobulin G (hIgG), and flow cytometry antibodies added to each sample for B cell immunophenotyping.

Gating strategies for identifying the B lymphocyte subpopulations are presented in Figure 3.9. The cells were gated to exclude debris and doublets (as described in Method 3.5.8). The singlets were then gated on SSC-A and Zombie dye to exclude the dead cells, before being gated on SSC-A and CD19 to identify the $CD19^+$ B cell population.

The $CD19^+$ cell population was gated on CD1d and CD5 to identify the Breg(2) population. The $CD19^+$ cell population was then gated on CD27 to identify the $CD27^+$ and $CD27^-$ cell populations. The $CD27^+$ population was then gated against IgD to identify the switched and unswitched memory cells, CD24 to identify the Breg(1) cells, and the $CD27^{++}$ cells were used to gate $CD20^-$ and $CD38^{++}$ to identify the plasmablasts. The $CD27^-$ population were gated on IgD to determine the naïve and double negative B cells and they were also further gated on CD24 and CD38 followed by CD5 to determine the transitional cell population.

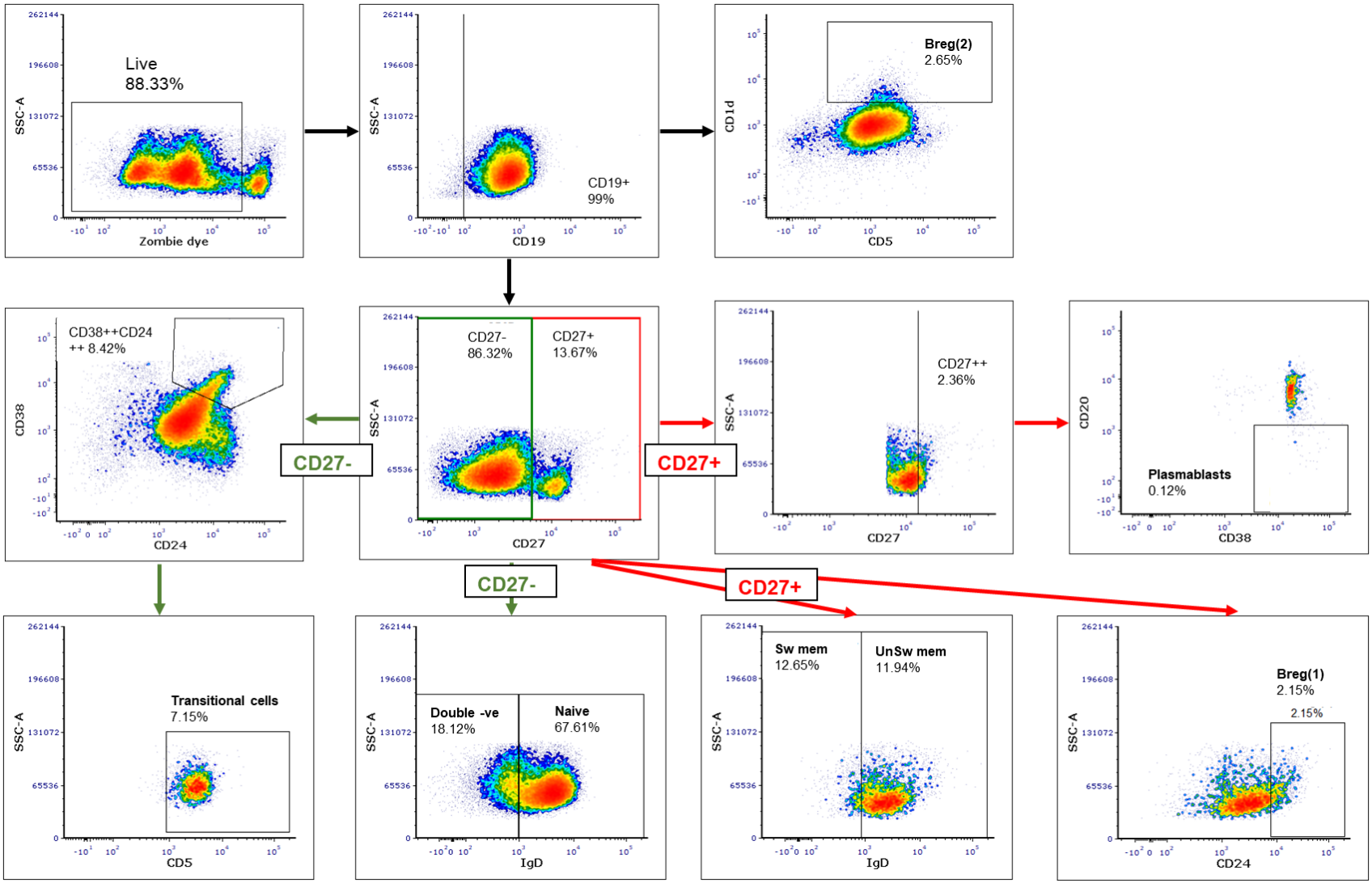


Figure 3.9 – Gating strategy for B lymphocyte immunophenotyping.

To calculate each cell subset per µl of blood, the $CD19^+$ % was extracted from the PBMC flow cytometry data and multiplied by the absolute lymphocyte count (/µl whole blood) to determine the $CD19^+$ B cell quantity (/µl whole blood). The % of $CD19^+$ B cells made up by each cell subset was then used to calculate each cell subset per µl of blood.

3.5.12 Cytokine and chemokine ELISAs

Commercially available sandwich ELISA kits were used to detect five cytokines (BAFF, IL-6, APRIL, sTACI, sBCMA) and one chemokine (CXCL13). ELISAs are a common technique used to measure the presence of soluble proteins in biological samples. Specifically, a sandwich ELISA measures an antigen (in this case a cytokine or chemokine) between two layers of antibodies (capture and detection antibody) (Figure 3.10). All the ELISAs in this study used streptavidin-horseradish peroxidase (HRP) conjugates and Tetramethylbenzidine (TMB) substrate solution.

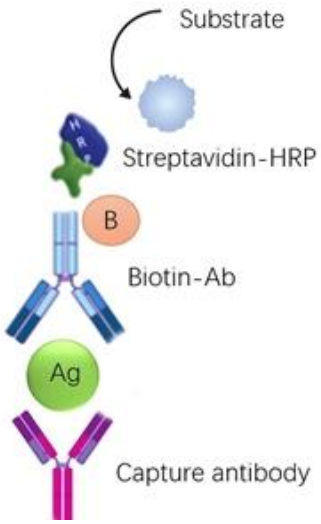


Figure 3.10 – Example of a sandwich ELISA method with Streptavidin-Biotin detection used in this experiment for the measurement of cytokines and chemokines.

In each ELISA plate the serum samples were run in duplicate, and three serum samples were repeated on each plate in order to calculate the intra- and inter-assay coefficients of variation. The assay range (including the upper (ULOD) and lower limits of detection (LLOD)) and sensitivities of each ELISA is shown in Table 3.8.

ELISA	Sensitivity	Assay range

BAFF	6.44 pg/ml	62.5 - 4,000 pg/ml
CXCL13	3.97 pg/ml	7.8 – 500 pg/ml
APRIL	0.4 ng/ml	0.78-50.0 ng/ml
IL-6 HS	0.09 pg/mL	0.2 - 10 pg/mL
sTACI	NA	93.8 - 6,000 pg/mL
sBCMA	NA	31.2 - 2,000 pg/ml

Table 3.8 – Table of the sensitivities and assay ranges of the different ELISA kits used.

BAFF ELISA

The previously separated serum samples stored at 80^{o}C were thawed on ice, mixed by vortexing and centrifuged at 10,000 rpm for 5 mins to remove any precipitate. All reagents were brought to room temperature. The BAFF standard was reconstituted and diluted by 2-fold serial dilutions to produce the standard curve (Figure 3.11). Each serum sample was diluted 2-fold. After adding 100 µL of Assay Diluent RD1-111 to each well, 50 µL of standard, diluted serum sample, or Calibrator Diluent (blank well) were added. The ELISA plate was covered and incubated for 3 hours at room temperature on a horizontal orbital microplate shaker (0.12" orbit) set at 500 rpm.

The ELISA plate was decanted, and then washed four times. The washing process involved filling each well with Wash Buffer (400 µL) using a squirt bottle, leaving the wash buffer in the wells for 30 seconds, and then ensuring complete removal of liquid after at each wash. After the last wash, any remaining Wash Buffer was removed by decanting and blotting the plate against clean paper towels. 200 µL of Human BAFF Conjugate was added to each well, before being incubate for 1 hour at room temperature on the shaker.

The wash procedure was repeated as above, before 200 µL of TMB Substrate Solution was added to each well. The plate was incubated for 30 minutes at room temperature on the benchtop, protected from the light. Following this, 50 µL of Stop Solution was added to each well quickly (in the same order as substrate solution was added).

The optical density of each well was determined within 15 minutes, using a microplate reader (Varioskan™ LUX) set to 450 nm. Wavelength readings at 540nm

were subtracted from the readings at 450 nm. This subtraction corrected for optical imperfections in the plate as readings made directly at 450 nm without correction may be higher or less accurate.

The following calculations were carried out on the microplate reader using SkanIt™ Software:

1. Subtract wavelength readings (450nm-540nm)
2. Average duplicate readings
3. Subtract the reading from the blank cells
4. Create a standard curve (4 PL)
5. Multiple sample readings by dilution factor

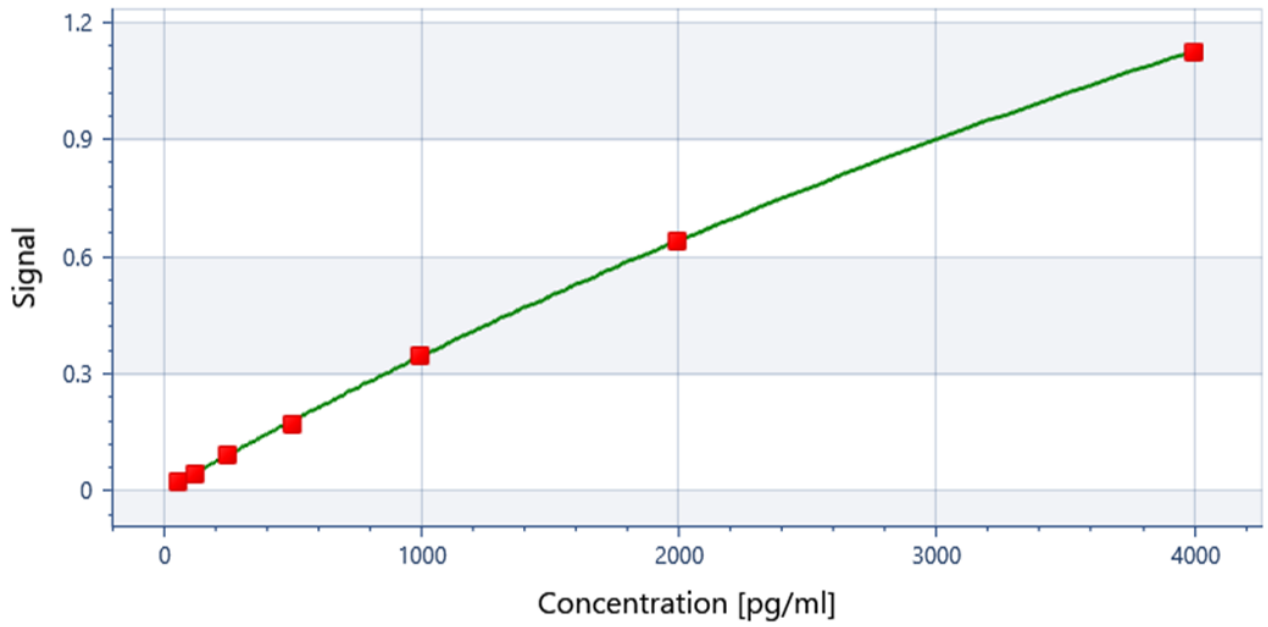


Figure 3.11 – Example of a BAFF standard curve generated using 2-fold serial dilutions of the BAFF standard.

CXCL13 ELISA

The serum samples were processed as in the BAFF ELISA method and the CXCL13 standard was reconstituted and diluted by 2-fold serial dilutions to produce the standard curve (Figure 3.12). Each serum sample was diluted 2-fold. After adding

100 µL of Assay Diluent RD1S to each well, 50 µL of standard, diluted serum sample, or Calibrator Diluent (blank well) were added. The ELISA plate was covered and incubated for 2 hours at room temperature on the benchtop.

The ELISA plate was decanted, and then washed four times as per the BAFF ELISA method. 200 µL of Human CXCL13 Conjugate was added to each well, before being incubated for 2 hours on the benchtop. The wash procedure was repeated, before 200 µL of Substrate Solution was added to each well. The plate was incubated for 30 minutes at room temperature on the benchtop, protected from the light. Following this, 50 µL of Stop Solution was added to each well quickly (in the same order as substrate solution was added).

The optical density of each well was determined within 15 minutes and calculations undertaken as per the BAFF ELISA method; however, the standard curve was a log/log curve fit.

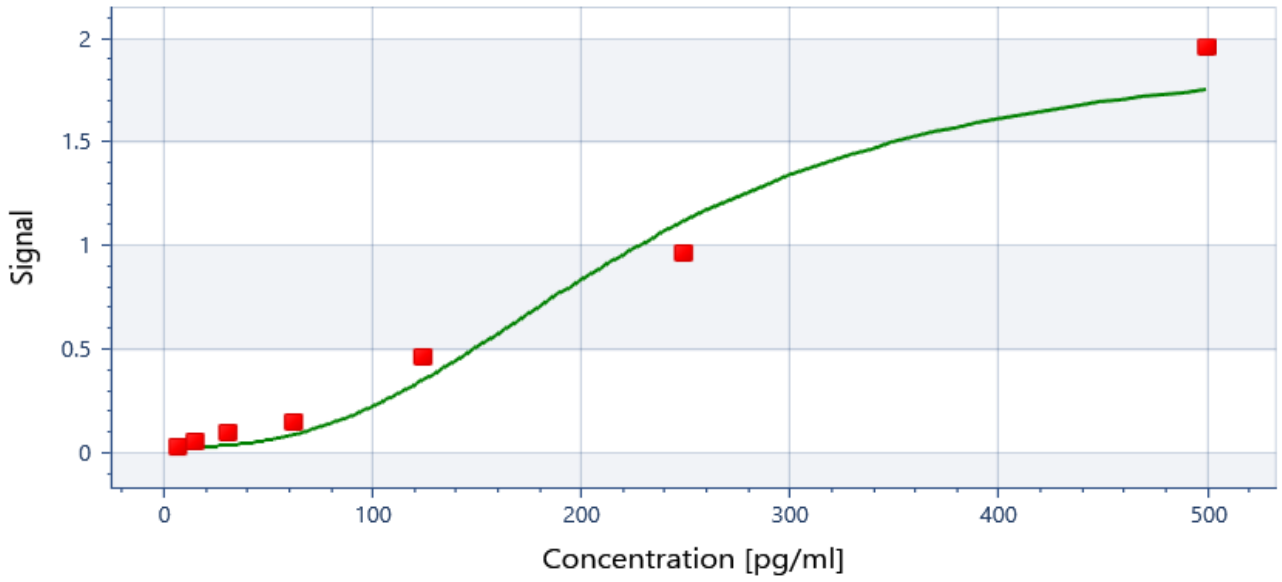


Figure 3.12 – Example of a CXCL13 standard curve generated using 2-fold serial dilutions of the CXCL13 standard.

IL-6 ELISA

The serum samples were processed as in the previous ELISA methods and the IL-6 standard was reconstituted and diluted by 2-fold serial dilutions to produce the standard curve (Figure 3.13). The serum sample was not diluted. After adding 100 µL of Assay Diluent RD1W to each well, 100 µL of standard, serum sample, or

Calibrator Diluent (blank well) were added. The ELISA plate was covered and incubated for 2 hours at room temperature on the shaker set at 500 ± 50 rpm.

The ELISA plate was decanted, and then washed four times as described previously. 200 µL of Human IL-6 Conjugate was added to each well, before being incubated for 1 hour on the shaker. The wash procedure was repeated, before 200 µL of Streptavidin Polymer-HRP was added to each well. The plate was incubated for 30 minutes at room temperature on the shaker. The wash procedure was repeated, before 200 µL of Substrate solution was added to each well after which the plate was incubated for 30 minutes at room temperature on the benchtop, protected from light. Following this, 50 µL of Stop Solution was added to each well quickly (in the same order as substrate solution was added).

The optical density of each well was determined within 15 minutes and calculations undertaken as per the previous ELISA methods (standard curve was a 4PL curve fit).

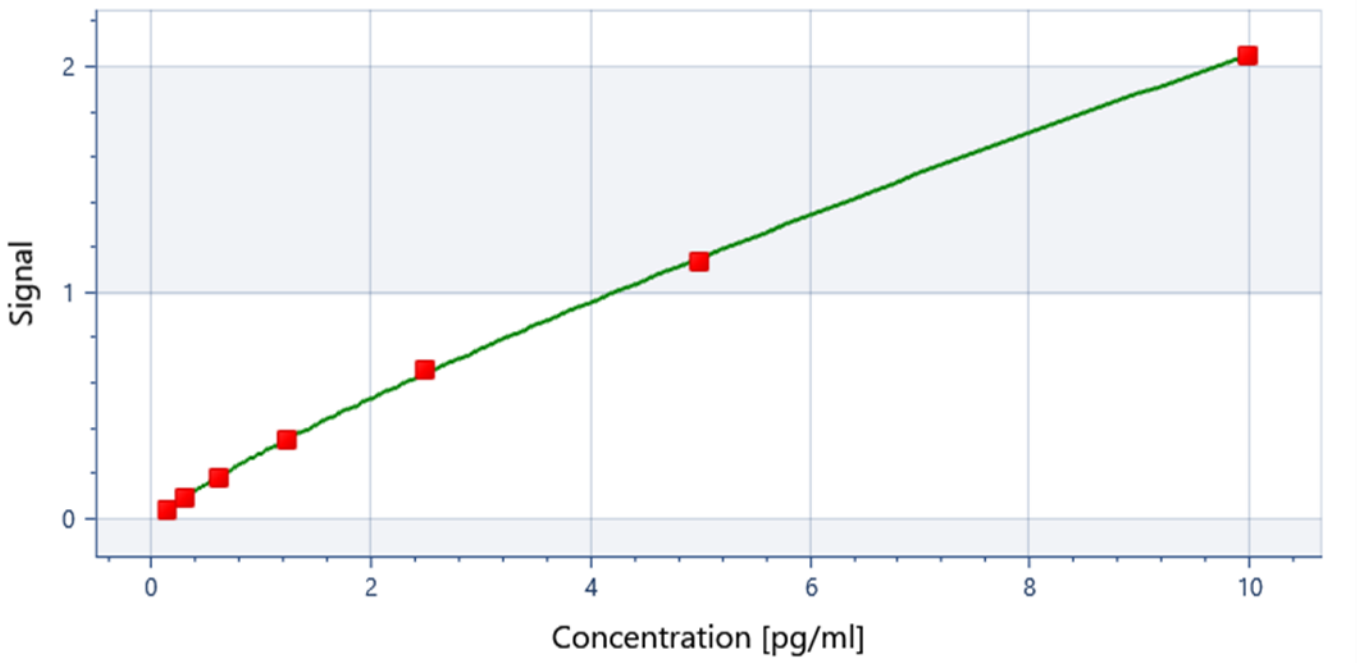


Figure 3.13 – Example of an IL-6 standard curve generated using 2-fold serial dilutions of the IL-6 standard.

sTACI and sBCMA ELISAs

The R&D Systems DuoSet® ELISA Kits were used for both sTACI and sBCMA cytokines and therefore the method undertaken was similar. The serum samples were processed as reported previously, and the sTACI or sBCMA standard was

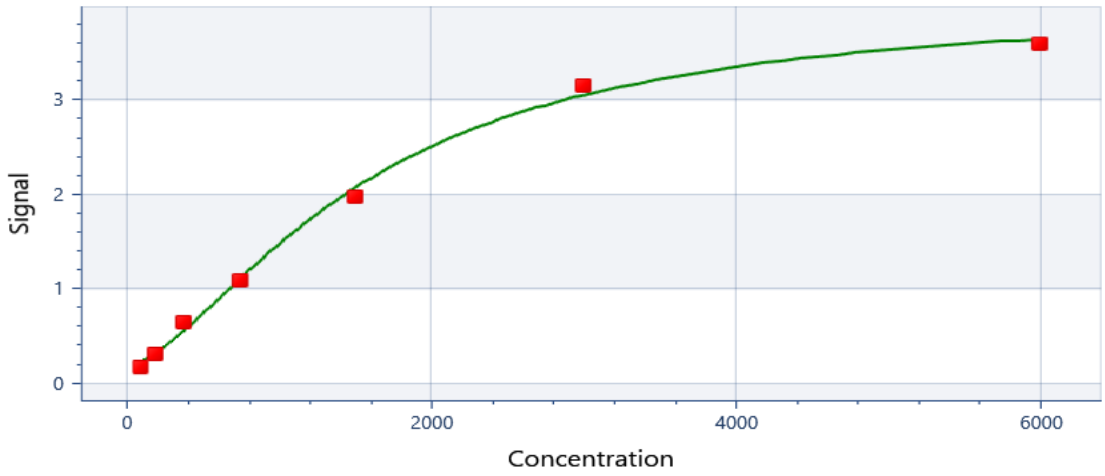
reconstituted and diluted by 2-fold serial dilutions to produce the standard curve (Figure 3.14).

Each serum sample was diluted 2-fold for sTACI, and 50-fold for sBCMA. Following an overnight benchtop incubation of 100 µl per well of capture antibody, the plate underwent three washes using a squirt bottle (as previously described). Each well was blocked by adding 300 µl of Reagent diluent to each well and the plate was incubated for one hour at room temperature, before undergoing a further three washes after which the sample and standards were added. Following a two-hour incubation at room temperature and three washes, 100 µl of detection antibody was added to each well.

After a further incubation of two hours at room temperature and three washes, Streptavidin-HRP was added. The plate was kept out of light and incubated for 20 minutes, before 100 µl of colour substrate solution was added. Following a further 20 minute benchtop incubation (protected from light), 50 µl of stop solution was added to each well.

The optical density of each well was determined within 15 minutes and calculations undertaken as per the BAFF ELISA method.

A



B

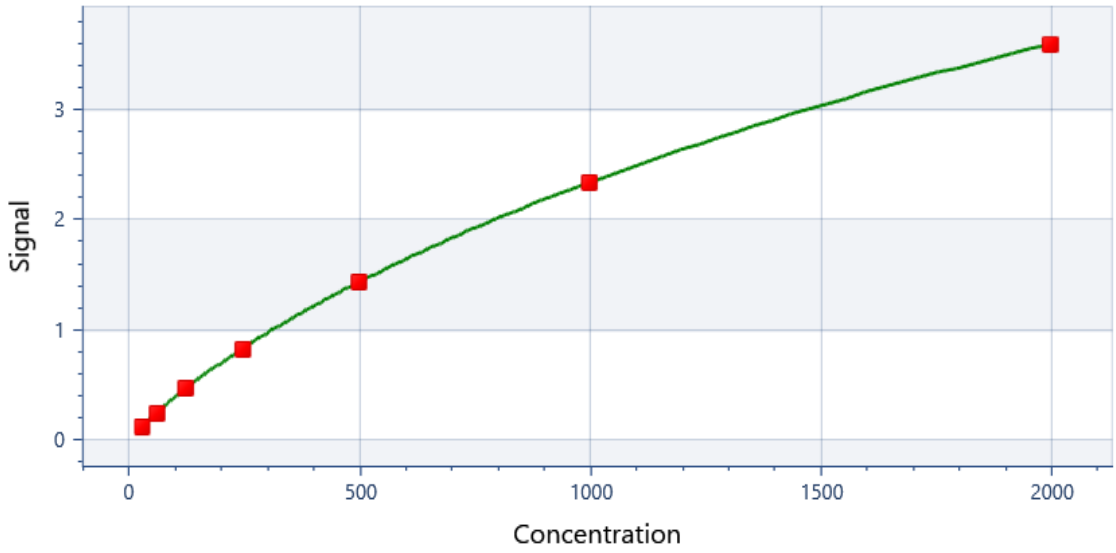


Figure 3.14 – Example of a sTACI (**A**) and sBCMA (**B**) standard curves generated using 2-fold serial dilutions of the respective standards.

APRIL ELISA

The serum samples were processed, and the APRIL standard was reconstituted and diluted by 2-fold serial dilutions to produce the standard curve, as described in the previous ELISA methods (Figure 3.15). The ELISA plate was washed twice with a squirt bottle and thorough decanting of contents between washes as described above. 100 µL of each standard was added and 100 µL of Sample Diluent was added to the blank wells. The serum was diluted 'in-well' 2-fold by adding 50 µL of Sample Diluent to the sample wells, followed by 50 µL of serum and 50 µL of Biotin-

Conjugate. The plate was incubated at room temperature for 2 hours on a microplate shaker set at 500 rpm.

The plate was washed six times before 100 µL of diluted Streptavidin-HRP was added. The plate was incubated at room temperature for 1 hour on a microplate shaker. The plate was then washed again six times before 100 µL of TMB Substrate Solution was added to all wells. Following a 20-minute incubation in the dark, the enzyme reaction was stopped by quickly pipetting 100 µL of Stop Solution into each well.

The optical density of each well was determined within 15 minutes. Wavelength readings at 620nm were subtracted from the readings at 450 nm.

The following calculations were carried out on the microplate reader using SkanIt™ Software:

1. Subtract wavelength readings (450nm-620nm)
2. Average duplicate readings
3. Subtract the reading from the blank cells
4. Create a standard curve (5 PL)
5. Multiple sample readings by dilution factor

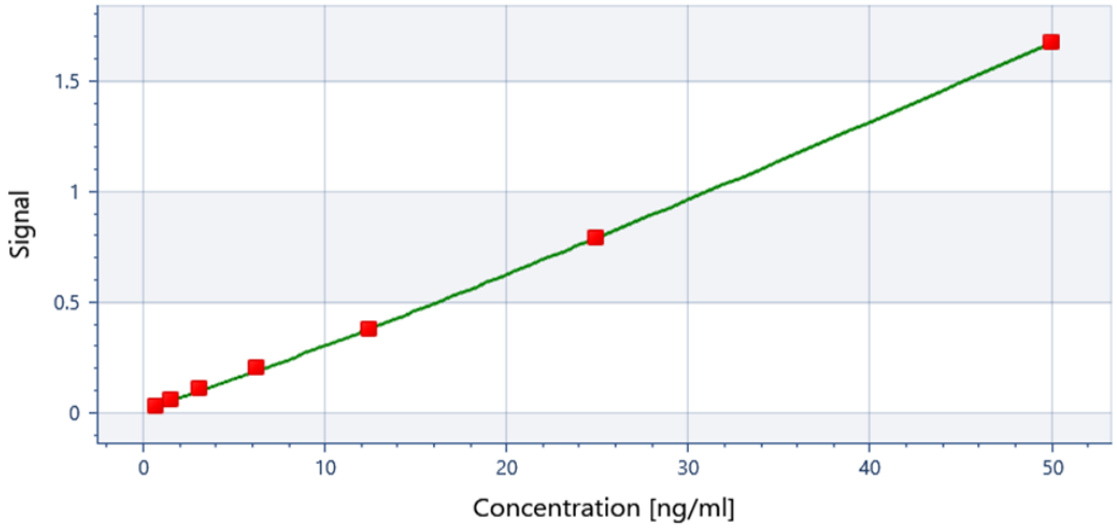


Figure 3.15 – Example of an APRIL standard curve generated using 2-fold serial dilutions of the APRIL standard.

3.5.13 CD19⁺ B cell mRNA sequencing (mRNAseq)

The mRNAseq processing as described in this section was performed by Raf Hussain and Jon Coxhead at the Newcastle University Core Genomics Facility.

Determining the integrity of RNA is a critical step in gene expression analysis, as degraded RNA samples can affect the detection of differentially expressed genes (*Gallego Romero et al. 2014*). To measure the quality and quantity of total RNA in the 65 study samples, electrophoretic separation was performed using a Tapestation™ 4200 machine (Agilent). The quality of RNA was determined using the RIN (RNA Integrity Number) which provides a measure of RNA integrity from completely intact, non-degraded RNA ('10') to total RNA degradation ('0'). Samples were diluted to 10 ng and mRNA sequencing libraries were prepared using the NEBNext® Low Input RNA Library Prep Kit for Illumina® (New England Biolabs) following manufacturer's instructions. This kit works most efficiently with high RIN values of ≥8. RIN values generated for the study samples are presented in Chapter 8: Results 5 – CD19⁺ B cell RNAseq data.

A pooled library was sequenced at 60 million (2 x 100 bp) reads per sample on a NovaSeq 6000 S2 flow cell (Illumina®) following manufacturer's instructions. An overview of this process is presented in Figure 3.16.

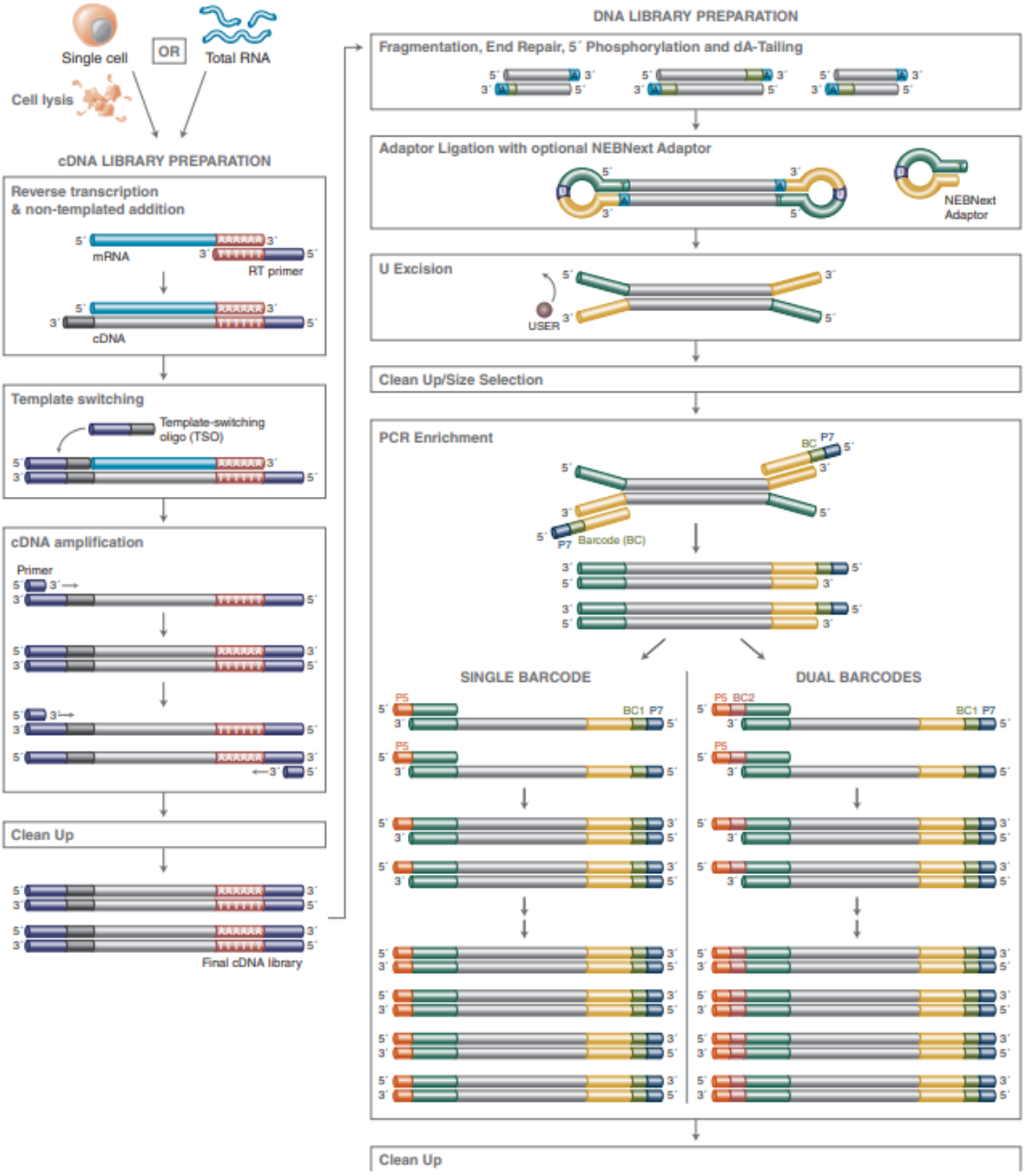


Figure 3.16 – Summary of the library preparation workflow used for the NEBNext® Low Input RNA Library Preparation

3.6 Data analysis

There were two main methods of data analysis used for this study. Firstly, the patient cohort was dichotomised into those that had relapsed or remitted their GD by one year, and this allowed the use of statistical tests that use a binary outcome. However, this method of analysis fails to consider the vast differences that may occur in the

time it takes for a patient to relapse once they have stopped their ATD. Therefore, survival analysis, such as Cox regression analysis, was performed which considers differences in the time between ATD cessation and relapse.

3.6.1 Analysis of clinical, cytokine, KRECs and FACs data

All analyses were performed using R Statistical Software (v4.1.2; R Core Team 2021), with additional packages as listed in Table 3.9.

Package name	Package version	Citation
cluster	2.1.4	Maechler *et al.* (2022)
cutpointr	1.1.1	Thiele and Hirschfeld (2021)
devtools	2.4.3	Wickham *et al*. (2021)
dplyr	1.0.7	Wickham *et al*. (2021)
forestmodel	0.6.2	Kennedy (2020)
ggplot2	3.3.5	Wickham (2016)
ggplotify	0.1.2	Yu G (2023)
ggpubr	0.4.0	Kassambara (2020)
glmnet	4.1.8	Tay *et al.* (2023)
gtsummary	1.5.0	Sjoberg *et al*. (2021)
lme4	1.1-27.1	Bates *et al*. (2015)
lmerTest	3.1-3	Kuznetsova *et al*. (2017)
MASS	7.3-54	Venables and Ripley (2002)
PairedData	1.1.1	Champely (2018)
pheatmap	1.0.12	Kolde R (2019)
pROC	1.18.0	Robin *et al*. (2011)
plotROC	2.2.1	Sachs (2017)
rstatix	0.7.0	Kassambara (2021)
summarytools	1.0.0	Comtois (2021)
survival	3.3-1	Therneau (2022)
survminer	0.4.9	Kassambara *et al*. (2021)
tidyverse	1.3.1	Wickham *et al.* (2019)

Table 3.9 – Additional R packages used for data analysis

Data distribution

Prior to data analysis the distribution of the data was determined and visualised using both density and Q-Q plots, to provide a visual judgment about the shape of the distribution and correlation between a given sample and the normal distribution, respectively. The Shapiro-Wilk's test was used as a significance test comparing the sample distribution to a normal one to ascertain whether the data demonstrated deviation from normality. A P-value > 0.05 implied that the distribution of the data was not significantly different from normal and therefore normality was assumed. If the data was not normally distributed, as in the case of the cytokine and soluble receptors concentrations which were all positively skewed, these values were analysed following natural log-transformation.

TRAb values below the lower limit of detection

A large proportion of the cohort (68%) had undetectable TRAb levels (<1 IU/L) on stopping ATD and therefore for the purposes of analysis the TRAb values that were below the assay LLOD had values imputed using the method of LLOD divided by the square root of 2, which has been reported to provide a good estimate of both the mean and standard deviation (*Hornung and Reed, 1990*).

Binary outcome data

The binary outcome of relapse or remission was used in the analysis of the clinical, cytokine, KRECs and FACs data to perform univariate logistic regression analyses to explore the unadjusted association between variables and outcome. This was followed by multivariate logistic regression analyses comprising variables identified from the univariate analysis. A cut-off P value of ≤0.2 was used for each variable in the univariate analyses to determine which variables were entered into a multivariate analysis. This higher threshold of P value was used to reduce the risk of Type II error and avoid incorrectly disregarding a covariable effect (*Dales and Ury, 1978, Mickey and Greenland, 1989*).

In all multivariate analyses performed in this study, there were several clinically relevant variables included in the model that were not necessarily significantly associated with outcome in this study but are well described in the literature to be associated with GD outcome, namely age, gender, goitre size, smoking status and TRAb titre on stopping ATD (discussed in Chapter 1: Introduction). The allowance to include variables such as these in multivariate analyses, even if they are statistically insignificant in this particular study, reflects the nature of building a model that is clinically relevant and translatable to clinical practice.

Variables that did not contribute to the multivariate model (P= >0.05) were eliminated. The results of the multivariate analysis were visualised using the R programme ‘forestmodel’ to generate a forest plot including all the variables entered into the analysis.

Survival analysis

The association of variables with time to relapse of GD was performed using univariate Cox regression analyses. Consistent with the binary regression analysis,

those with a P value of ≤0.2 were entered into a multivariate Cox regression analysis including the same additional clinically relevant variables.

The fundamental assumption in the Cox model is that the hazards are proportional, which means that the effect of a particular variable remains constant over time. To test the proportional hazards assumption for all covariates entered into Cox analyses, statistical tests and graphical diagnostics using scaled Schoenfeld residuals were undertaken. Using the R package, 'survival', the corresponding set of scaled Schoenfeld residuals for each covariate was correlated with time, to test for independence between residuals and time. Additionally, a global test was performed for the whole model. Proportionality of hazards was supported if there was a non-significant (P=>0.05) relationship between residuals and time.

ROC curve analysis

A ROC curve reflects the ability of a variable to correctly classify subjects regardless of a specific threshold, and therefore this analysis can be useful to determine the most discriminating biomarker for outcome. Following ROC curve analysis, a threshold value needs to be determined to enable use of this variable as a diagnostic test, or in this case, a predictor of outcome. The optimal threshold of the ROC curve is the one that leads to the highest sensitivity and specificity.

For each significant continuous variable identified from the multivariate Cox regression analyses, the R programme 'cutpointr' generated the optimal cut-off biomarker concentration which produced the highest sensitivity and specificity for a particular variable to predict relapse or remission. Confidence intervals and the area under the ROC curve (ROC_{AUC}) were calculated by bootstrapping (2000 replicates) and using the trapezoidal rule (*Lasko et al. 2005*), respectively. ROC curves were produced using the 'plotROC' R package.

The survival distributions of high and low biomarker concentrations were analysed by the log-rank test to determine their predictive utility for time-to-relapse after ATD cessation. This data is presented in Kaplan-Meier plots.

Heatmap analysis

The cytokine and differential gene expression heatmaps were generated using the package 'pheatmap' to visualise hierarchical clustering of samples based on cytokine or gene expression. Agglomerative hierarchical clustering (*Saxena et al. 2017*) was

performed, and Ward's method (*Murtagh et al. 2014*) was observed to be the hierarchical clustering method that identified the strongest clustering structures. To determine the optimal number of clusters, the average silhouette method was performed which determines how well each object lies within its cluster with a high average silhouette width indicating a good quality clustering (*Rousseeuw et al. 1987*).

Prior to heatmap analysis, the normalised gene read counts underwent log2(+1) transformation and scaling of the data was performed to assess relative differences in cytokine or gene expression between samples.

Composite scores

Composite biomarker scores were used to analyse whether a combination of the significant variables associated with GD outcome could have predictive value for relapse. Variables were selected if they were observed to be significant in the binary logistic and/or cox regression analyses. Least absolute shrinkage and selection operator (lasso) regression was performed using the 'glmnet' package in R (*Friedman J et al. 2010*) to select the specific variables that had the strongest relationship with GD outcome. Lasso regression performs both variable selection and regularisation to reduce overfitting by reducing the magnitude of the regression coefficients, thus encouraging simpler and more interpretable models. A relaxed lasso regression was performed on the variables determined significant by the original lasso regression (*Hastie et al. 2020*). The purpose of the relaxed lasso was to refine the shrinkage of the selected variable coefficients to improve the predictive accuracy of the model.

These coefficients generated from the relaxed lasso regression provided a balanced weight for each variable, which was then multiplied by the respective individual biomarker value and added to the other variables to create a composite score.

An example of this calculation for a four-variable biomarker composite score is shown below:

Composite score = 1.7 (BAFF) + 0.83 (sTACI) + 1.62 (MLR) + 0.52 (TRAb)

The composite scoring was also used to compare combinations of genes +/- biochemical biomarkers. An example of this calculation for a six-variable biomarker and gene composite score is shown below:

Composite score = 1.21 (DDX11L17) + 0.098 (POC1B) + 2.537 (SOD2) + 2.55 (BAFF) + 0.61 (sTACI) + 0.57 (TRAb)

The predictive performance of the top biomarker/gene composite scores were then evaluated using ROC curve analysis and the survival distributions of high and low biomarker scores were analysed by the log-rank test.

3.6.2 Analysis of mRNAseq data

The initial bioinformatics analysis for the mRNAseq data (apart from the multivariate Cox regression) was performed by Ann Hedley, Senior Experimental Scientific Officer, Bioinformatics Support Unit, Newcastle University.

FASTQ files were generated from the sequencer output using Illumina's bcl2fastq and quality checks on the raw data were done using FastQC (*Andrews 2018*) and FastQ Screen (*Wingett et al. 2018*). Alignment of the RNAseq paired-end reads was to the GRCh38 (*Zerbino et al. 2018*) version of the human genome and annotation using HISAT2 (*Kim et al. 2015*). Expression levels were determined and statistically analysed by a workflow combining HTSeq (*Anders et al. 2014*) and the R environment (*R Core Team, 2018*), utilising packages from the Bioconductor data analysis suite (*Huber et al. 2015*). Differential gene expression (DGE) analysis was based on the negative binomial distribution using the DESeq2 package (*Love et al. 2014*). Further data analysis and visualisation used R and Bioconductor packages. Software Versions of the R packages used can be found in Chapter 3: Methods.

RNAseq analysis for DGE was initially performed using a P value (<0.05) adjusted for false-discovery rate (FDR), before secondary analysis was undertaken using an unadjusted P value of <0.001 to capture genes with a fold change of ≥ 1.6 which did not reach statistical significance when adjusted for FDR. The unshrunken log2 fold change values were used to detect subtle changes in gene expression for genes that were significantly differentially expressed. Further exploratory secondary analysis was performed using the unadjusted P value at a threshold of <0.001.

Pathway analysis was performed using QIAGEN Ingenuity Pathway Analysis (IPA) (*QIAGEN Inc*. Available at https://digitalinsights.qiagen.com/IPA) (*Krämer et al., 2014*). The differentially expressed genes entered into the pathway analysis were those with an unadjusted P value of <0.01 DGE between relapse and remission groups.

The association between DGE and the time-to-relapse was analysed using a univariate Cox Regression model. Genes which were significantly (unadjusted P value <0.001) associated with time-to-relapse were entered into a multivariate Cox Regression model.

3.7 Ethical approval and study governance

The SPRING study was approved by the NHS HRA Research Ethics Committee (REC) in East Midlands - Leicester South (REC reference: 18/EM/0371). Professor Simon Pearce was the Chief Investigator for the study, and the Newcastle-upon-Tyne Hospitals NHS Foundation Trust acted as Study sponsor. All patients provided written informed consent before recruitment into the study, which was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013). The NIHR BioResource Steering Committee approved the use of the healthy control patient DNA samples from the Newcastle NIHR BioResource Centre.

Chapter 4: Results 1 – Clinical data

In this results chapter, I will describe clinical characteristics of the patient cohort recruited to the SPRING study including patient demographics, biochemical data, and clinical outcome. The clinical variables are compared between the relapse and remission groups and their association with time-to-relapse is analysed by Cox regression.

4.1 Study procedures

4.1.1 Patient recruitment

A total of 68 potential patients were identified for possible recruitment to the SPRING study, all of whom were being considered for cessation of their ATD medication. Patients were recruited from March 2019 to September 2020. Out of the 68 patients identified, 3 declined to take part in the study and therefore 65 patients were enrolled. All these 65 patients met the inclusion criteria and were willing to stop their ATD medication.

4.1.2 Patient outcomes

Of the 65 patients who stopped ATD, 16 (25%) of patients had relapsed by 12 months of study follow-up. One patient was lost to follow-up during this 12-month period, but they had their thyroid function checked 2 years after stopping ATD which showed they remained in remission off ATD. This means that it is very likely that this patient was also in remission at the 12-month timepoint, and therefore they were assigned the outcome of remission. The clinical outcome 12 months after ATD withdrawal was used for the primary analysis in this Chapter.

4.1.3 Long-term clinical outcome

As the length of the project was extended and the patients continued to be followed up routinely in the thyroid clinic or in primary care, outcomes at 2 years after stopping ATD were also available for all the patients. Thyroid function checked after two years off ATD medication was deemed to be indicative for 2-year outcome. Of the remaining 49 patients that hadn't already relapsed by 12 months, a further 8/49 (16%) relapsed within 2 years. Therefore, the total number of patients relapsing by two years was 24 (37%). More patients relapsed within the first year of stopping ATD

compared to the second year (16 vs. 8 patients), which is consistent with the fact that most GD relapses occur in the first 12 months after stopping ATD.

A flow diagram detailing the patient recruitment and outcomes is presented in Figure 4.1.

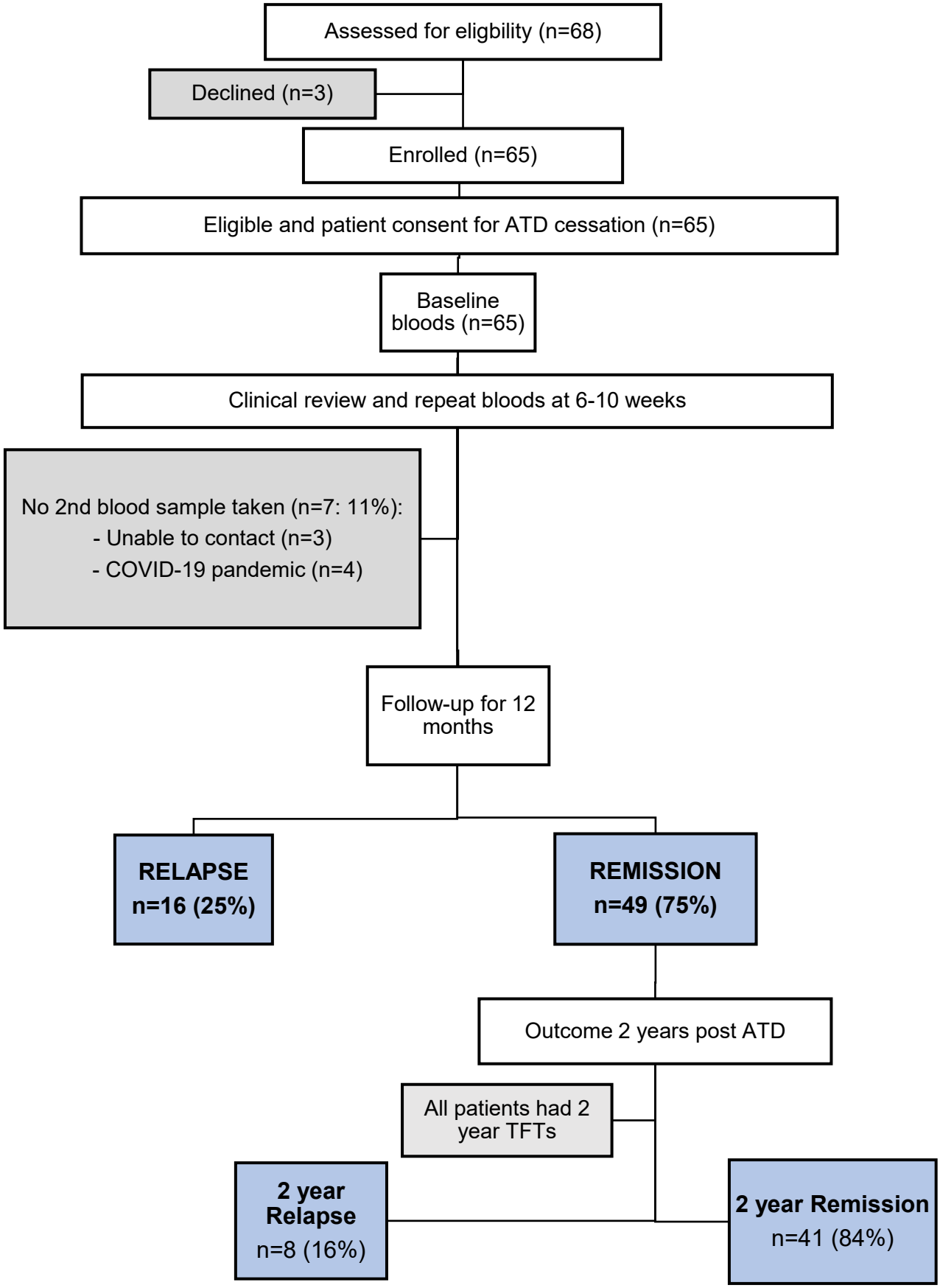


Figure 4.1 – Flow diagram detailing the patient recruitment and outcomes

TFTs: thyroid function tests, ATD: antithyroid drug

4.2 Quality control

4.2.1 Study visits

Once patients had stopped their ATD they were reviewed again at a median of 7 weeks later (IQR 6-10 weeks), however there were 6 patients who had their second blood sample taken between 11-19 weeks due to delays in them returning for a clinic appointment. The distribution of the timing of each patient's 6–10-week follow-up appointment is presented in Figure 4.2.

Following this second appointment, the study patients were seen every 3 months when they had a clinical review and thyroid function checked. All patients, apart from the patient described in 4.1.2, attended for thyroid blood tests at least twice between the 6–10-week appointment and the last appointment at 12 months.

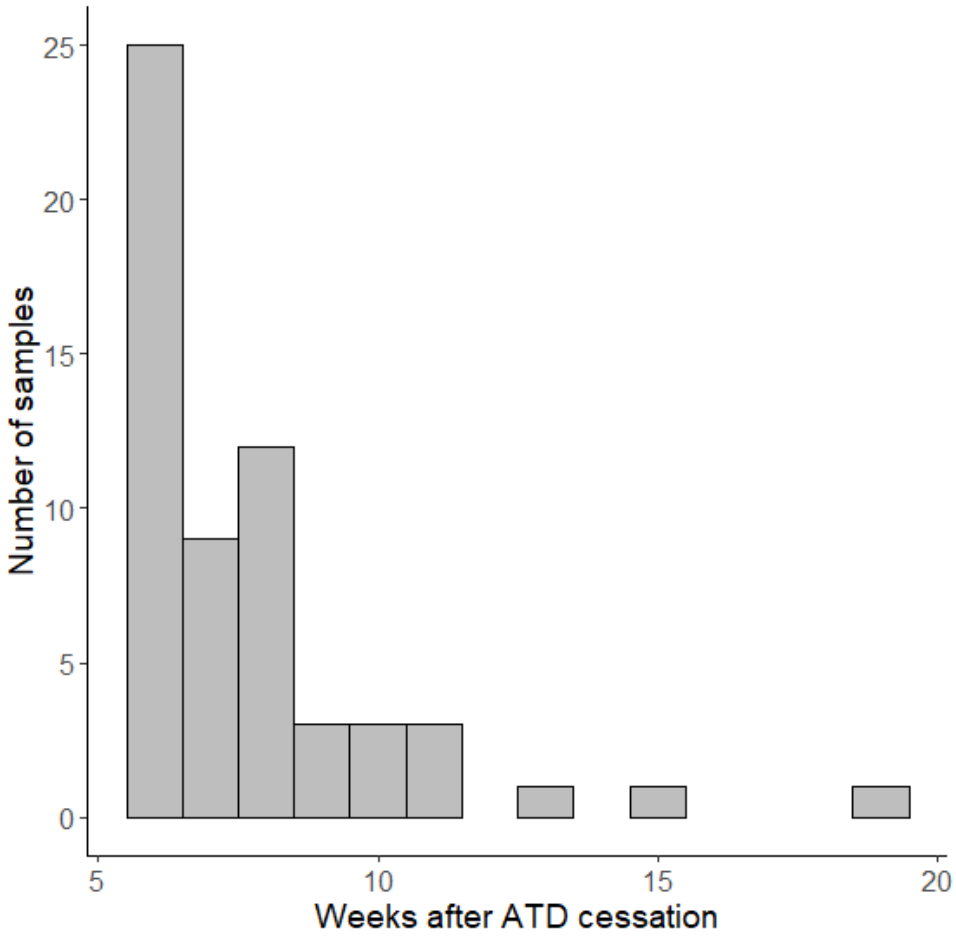


Figure 4.2– The distribution of the timing of each patient's 6–10-week follow-up appointment.

4.2.2 Missing data

Although all 65 patients had baseline blood tests done, 7/65 (11%) did not have the second 6–10-week blood sample taken after stopping ATD medication. In three cases this was because the patient did not return for their follow-up appointment and could not be contacted at the time. For the remaining 4 patients, they were unable to return for their follow-up appointment as they were during the 'lockdown' period of the COVID-19 pandemic and therefore they were not able to attend hospital. Other than this, all other prospective clinical data collected is complete.

In terms of the retrospective clinical data collected from the electronic patient record, FT4 at diagnosis was unavailable for one patient and FT3 at diagnosis was unavailable for two patients. Six patients did not have a TRAb titre available on the specific date of diagnosis. All patients had the start date of their ATD documented in their medical records.

4.2.3 Assay consistency

As patients were recruited from two different hospitals, the assays used to measure thyroid and TRAb concentrations differed between the two sites. Therefore, extra serum was taken for the seven patients recruited at the second site and these samples were processed at the same laboratory as the other samples in the study to ensure consistency of results. The discrepancy between results was particularly notable for TRAb levels, where two patients had markedly different TRAb concentrations depending on the assay, 0.7 vs. 2.8 iU/L and 1.4 vs. 6.6 iU/L, with the higher results found at the central laboratory used for the study.

4.3 Descriptive analysis

4.3.1 Patient demographics and clinical characteristics

The demographic details of all patients recruited to the SPRING study are presented in Table 4.1.

Overall, the 65 patients that stopped ATD are largely representative of the adult population with GD. The median age of patients was 49 years and 52 (82%) of the patients were female. Nearly a quarter of patients had had a previous relapse and were on their second course of ATD. Most patients (65%) were being managed with ATD using a 'block and replace' approach, which is the commonest regimen used to

treat adult GD in the Newcastle thyroid clinic. The duration of ATD was variable (12 – 105 months), but all patients had received a minimum of 12 months of ATD treatment before the ATD was stopped, as per the study inclusion criteria.

Out of the GD cohort, 21% described themselves as a 'current smoker' which is slightly higher than the prevalence of smoking described in the North-East of England by the Department of Health and Social Care in 2020 (15.7%) (*Department of Health and Social Care 2023*). There is minimal ethnic diversity demonstrated in the patient cohort, which is representative of the data from the Office for National Statistics where 90.6% of the North-East England population are described to be White British (*Office for National statistics 2022*).

Prior to stopping ATD, 44 (68%) of the patients had tested negative for TRAbs (<1 IU/L), and of those which tested positive 13 (62%) had TRAb levels <2 IU/L. Patients with detectable TRAbs had ATD withdrawn in cases where they had already decided to stop as they no longer wanted to take medical therapy, or in two cases where they were thought to have negative TRAb levels as they tested negative on the TRAb assay at the second site, which when later repeated at the central laboratory site was found to be positive. Having a low or negative TRAb concentration prior to stopping ATD is consistent with the established management guidelines for GD and therefore this patient cohort represents most of the adult GD patient population at the timepoint of stopping ATD treatment.

Demographic / Clinical variable	Value
Total number of patients who stopped ATD	65
Age (years): median (IQR) [range]	49 (39 – 56) [20 – 80]
Female: n (%)	53 (82%)
Ethnicity: n (%)	
White	60 (92%)
Black African	1 (2%)
Asian	4 (6%)
Smoking history: n (%)	
Current smoker	14 (21%)
Previous smoker	16 (25%)
Never smoked	35 (54%)
Family history of thyroid disease in 1st degree relative: n (%)	34 (52%)
Previous history of Graves' disease relapse: n (%)	15 (23%)
Thyroid function at diagnosis (where available): median (IQR) [range]	
TSH (miU/L)	<0.05 [<0.05 – 0.3]
FT4 (pmol/L)	41.8 (33.1 – 58.2) [18.7 – >100]
FT3 (pmol/L)	17.8 (12.9 – 25.5) [6.6 – >46.5]
TRAb (IU/L)	6.5 (3.6 – 11.1) [1.7 – >100]
Time to normal TSH after ATD started (months): median (IQR) [range]	5.9 (3.4 – 12.3) [1.4 – NA]
ATD regimen: n (%)	
BR	42 (65%)
DT	23 (35%)
Duration of ATD (months): median (IQR) [range]	17 (14.3 – 23.4) [12 – 105]
Thyroid function when stopping ATD: median (IQR) [range]	
TSH (miU/L)	1.99 (1.2 – 3.8) [< 0.05 – 36.1]
FT4 (pmol/L)	16.5 (14.5 – 18.5) [6.5 – 32.9]
FT3 (pmol/L)	4.3 (4 – 4.8) [3.2 – 11]
TPOAb (IU/ml)	24 (12 – 68.3) [8 – 301]
TRAb (IU/L)	< 1 (0 – 1.2) [<1 – 6.6]
Goitre size on stopping ATD: n (%)	
0	21 (32%)
1	33 (51%)
2	11 (17%)
Thyroid function 6-10 weeks after stopping ATD: median (IQR) [range]	
TSH (miU/L)	1.22 (0.6 – 1.7) [<0.05 – 3.5]
FT4 (pmol/L)	16.5 (14.7 – 17.7) [11.4 – 27]
FT3 (pmol/L)	4.8 (4.5 – 5.5) [3.8 – 9.5]
TPOAb (IU/ml)	27.5 (12.8 – 87.8) [<5 – 294]
TRAb (IU/L)	< 1 (0 – 0.8) [<1 – 5.8]

Table 4.1 – Demographics and clinical characteristics of patients who stopped ATD.

4.3.2 Distribution of Graves' disease relapse events up to 12 months

Of the 65 patients who stopped ATD, 16 (25%) relapsed their GD within 12 months after ATD withdrawal. The median (IQR) time to relapse was 187 (83 – 238) days and ranged from 42 – 344 days (Figure 4.3). The probability of remaining in remission 12 months after stopping ATD is 75% (95% confidence interval 0.66 – 0.87). The relapse-free survival of all patients after stopping ATD is shown in Figure 4.4.

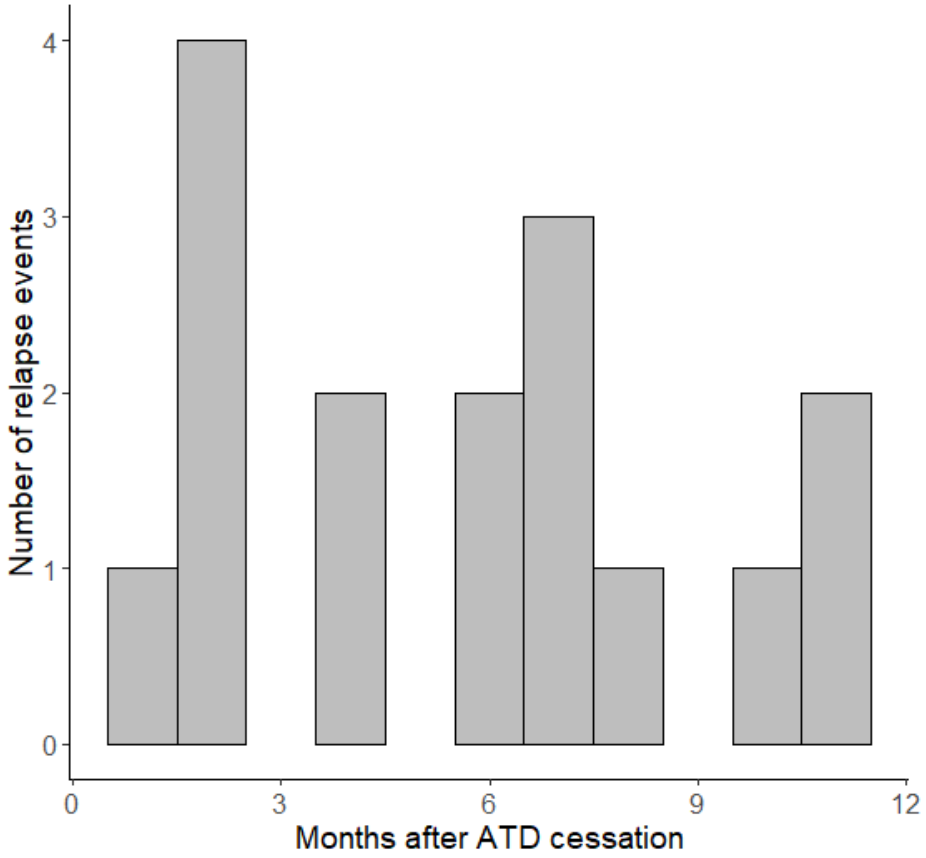
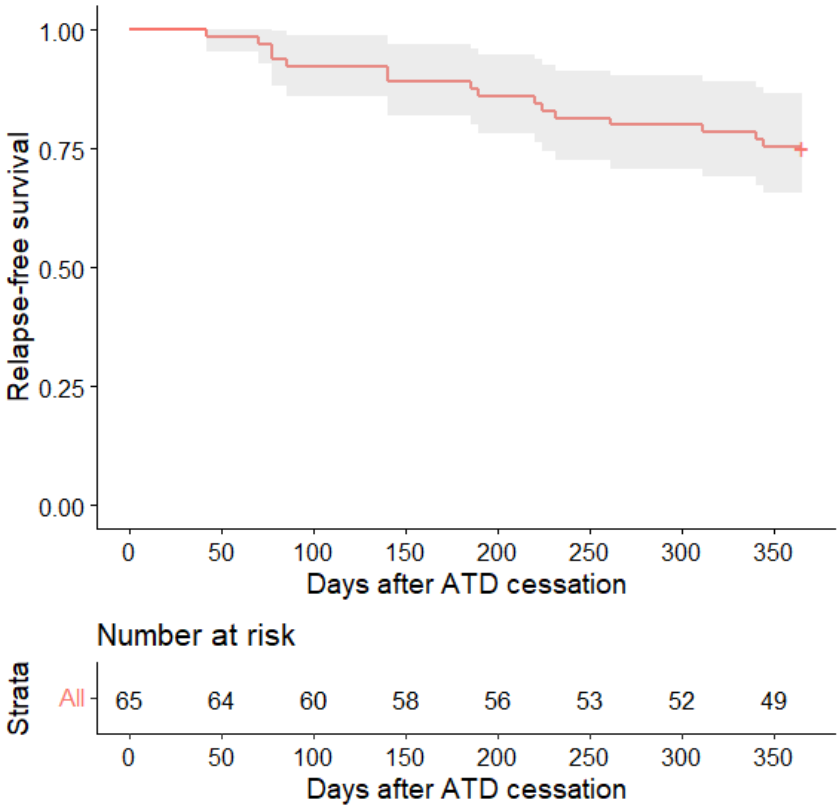


Figure 4.3 – The distribution of relapse events by months after ATD withdrawal.


1.00
0.75
0.50
0.25
0.00
Relapse-free survival
0
50
100
150
200
250
300
350
Days after ATD cessation

Number at risk
Strata
All
65
64
60
58
56
53
52
49
0
50
100
150
200
250
300
350
Days after ATD cessation

Figure 4.4 – Kaplan-Meir plot of relapse-free survival of patients after stopping ATD. Out of the 16 patients that relapsed, 10 (63%) had a positive TRAb level (≥1 IU/L) at the time of stopping ATD treatment. The median TRAb level when stopping ATD treatment in patients that relapsed was 1.15 (<1 – 6.6 IU/L), compared to those that remitted who mostly had negative TRAb levels <1 (<1 – 3.8 IU/L). Those that relapsed by 12 months had significantly higher TRAb concentrations when stopping ATD (P=0.015) (Figure 4.5).

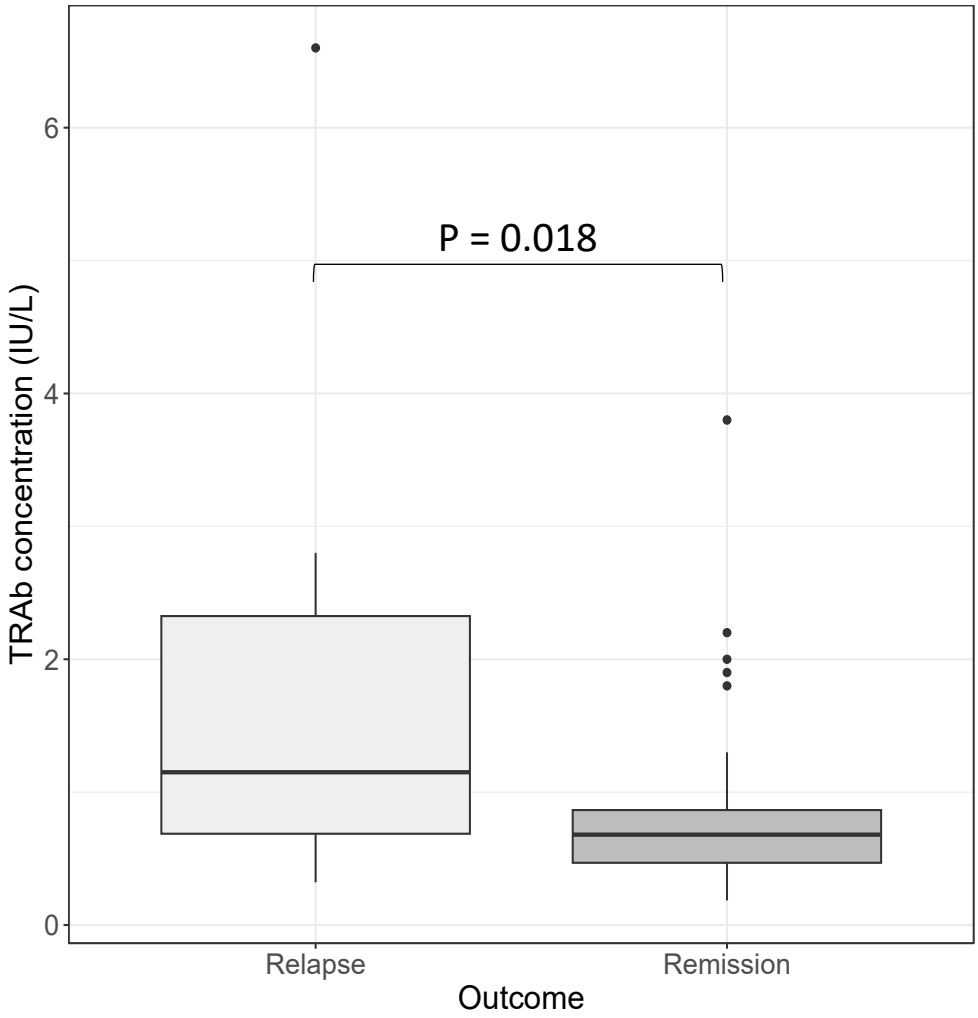


Figure 4.5 – TRAb concentrations in the relapse and remission groups when stopping ATD.

4.4 Comparison of the clinical and demographic variables between the relapse and remission groups

The clinical characteristics and demographics of patients who stopped ATD were compared between the relapse and remission outcome groups, at the timepoint of stopping ATD treatment (Timepoint 1) and 6-10 weeks later (Timepoint 2). Statistical

analysis of differences in clinical and demographic variables between the two outcome groups was assessed by univariate binary logistic regression at diagnosis and Timepoints 1 and 2 (Table 4.2).

(a)

Variable (At diagnosis)	Relapse group (n=16)	Remission group (n = 49)	B	ORrelapse	95% CI ORrelapse	Univariate P value
Median TRAb (IU/L)	8.1	6.2	0.03	1.03	0.99 – 1.06	0.117
Median FT3 (pmol/L)	20.3	16.7	0.03	1.03	0.97 – 1.09	0.301
Median FT4 (pmol/L)	52.4	40	0.01	1.01	0.99 – 1.04	0.337
Median Age	51.5	48	0.02	1.02	0.97 – 1.064	0.425
Gender Male n (%)	4 (25)	8 (16)	0.54	1.70	0.44 – 6.67	0.440

(b)

Variable (Timepoint 1)	Relapse group (n=16)	Remission group (n = 49)	B	ORrelapse	95% CI ORrelapse	Univariate P value
Median TRAb (IU/L)	1.15**	0.63**	0.97	2.64	1.31 – 6.46	**0.018***
Median MLR	0.4	0.29	3.85	46.9	1.8 – 1255	**0.021***
ATD regimen DT n (%)	3 (19)	20 (41)	-1.09	0.33	0.08 – 1.33	0.119
Goitre size n (%)						
0	3 (19)	18 (37)				
1	9 (56)	24 (49)	0.81	2.25	0.53 – 9.52	0.270
2	4 (25)	7 (14)	0.81	3.43	0.60 – 19.4	0.163
Median NLR	2.94	2.52	0.23	1.26	0.82 – 1.92	0.28
Median PLR	243.7	210	0.004	1	0.99 – 1.01	0.28
Median time to normal TSH (months)	5.98	5.98	0.03	1.03	0.96 – 1.01	0.42
Median duration of ATD (months)	18.4	16.8	0.01	1.01	0.98 – 1.04	0.464
Median TPOAb (IU/ml)	41.5	22.5	-0.003	0.99	0.99 – 1	0.505
Median FT3 (pmol/L)	4.2	4.3	0.16	1.18	0.71 – 1.95	0.525
Median FT4 (pmol/L)	15.9	16.6	0.05	1.05	0.89 – 1.22	0.563
Median TSH (miU/L)	2.29	1.94	-0.03	1.05	0.89 – 1.13	0.718
Current smoker n (%)	3 (19)	11 (22)	-0.22	0.79	0.19 – 3.31	0.754
Previous GD relapse n (%)	4 (25)	11 (22)	0.14	1.15	0.30 – 4.29	0.834

(c)

Variable (Timepoint 2)	Relapse group (n=16)	Remission group (n = 49)	B	ORrelapse	95% CI ORrelapse	Univariate P value
Median TSH (miU/L)	0.32	1.4	-2.17	0.11	0.03 – 0.39	**0.0006***
Median FT3 (pmol/L)	5.5	4.7	1.16	3.19	1.43 – 7.14	**0.004***
Median TRAb (IU/L)	0.86**	0.63**	0.90	2.46	1.27 – 5.92	**0.02***
Median MLR	0.38	0.3	4.9	148	3.9 – 7235	**0.01***
Median FT4 (pmol/L)	17.5	15.9	0.21	1.24	1.02 – 1.51	**0.030***
Median TPOAb (IU/ml)	33	23	-0.0008	0.99	0.99 - 1	0.838

Table 4.2 – Association of clinical variables and patient demographics with GD relapse by univariate logistic regression, at **(a)** diagnosis of GD **(b)** Timepoint 1: the time of ATD withdrawal **(c)** Timepoint 2: 6-10 weeks after ATD withdrawal.

Out of these clinical variables, there are a total of 6 variables that are significantly ($P<0.05$) associated with GD relapse, however only 2 of these variables are known prior to ATD withdrawal, namely the TRAb concentration and the monocyte-lymphocyte ratio (MLR). It is these variables that would be most useful clinically to help predict outcome, as they are values that could be used to inform a clinical decision on timing of ATD withdrawal. Although age, gender, goitre size and smoking are factors described in the literature to be associated with outcome of GD, there was no association observed in this study. Further analysis splitting the goitre variable into two groups on the presence or absence of a goitre (Grade 0 vs. Grade 1 and 2) remained non-significant ($P=0.19$). However, as these clinical factors have been reasonably well-established to be associated with GD outcome, these variables alongside TRAb and MLR at Timepoint 1 were entered into a multivariate binary logistic regression model (Table 4.3).

Variable	B	ORrelapse	95% CI ORrelapse	Multivariate P value
TRAb on stopping ATD	1.05	2.87	1.2 – 8	**0.027***

MLR	3.82	45.69	1.4 – 3267	**0.04***
Goitre size				
1	1	2.7	0.4 – 20.6	0.28
2	1.92	6.8	0.80 – 77.7	0.09
Gender (Male)	0.94	2.56	0.45 – 14.2	0.27
Age	-0.004	0.99	0.94 – 1.06	0.90
Current smoker	0.08	1.09	0.19 – 5.3	0.92

Table 4.3 – Association of clinical variables with occurrence of GD relapse following ATD withdrawal, using a multivariate binary logistic regression model.

Following multivariate binary logistic regression analysis, two variables at baseline (Timepoint 1) were significantly ($P<0.05$) associated with the occurrence of a GD relapse, namely the TRAb concentration and MLR. The relative contributions of each clinical variable to this multivariate model are presented in Figure 4.6.

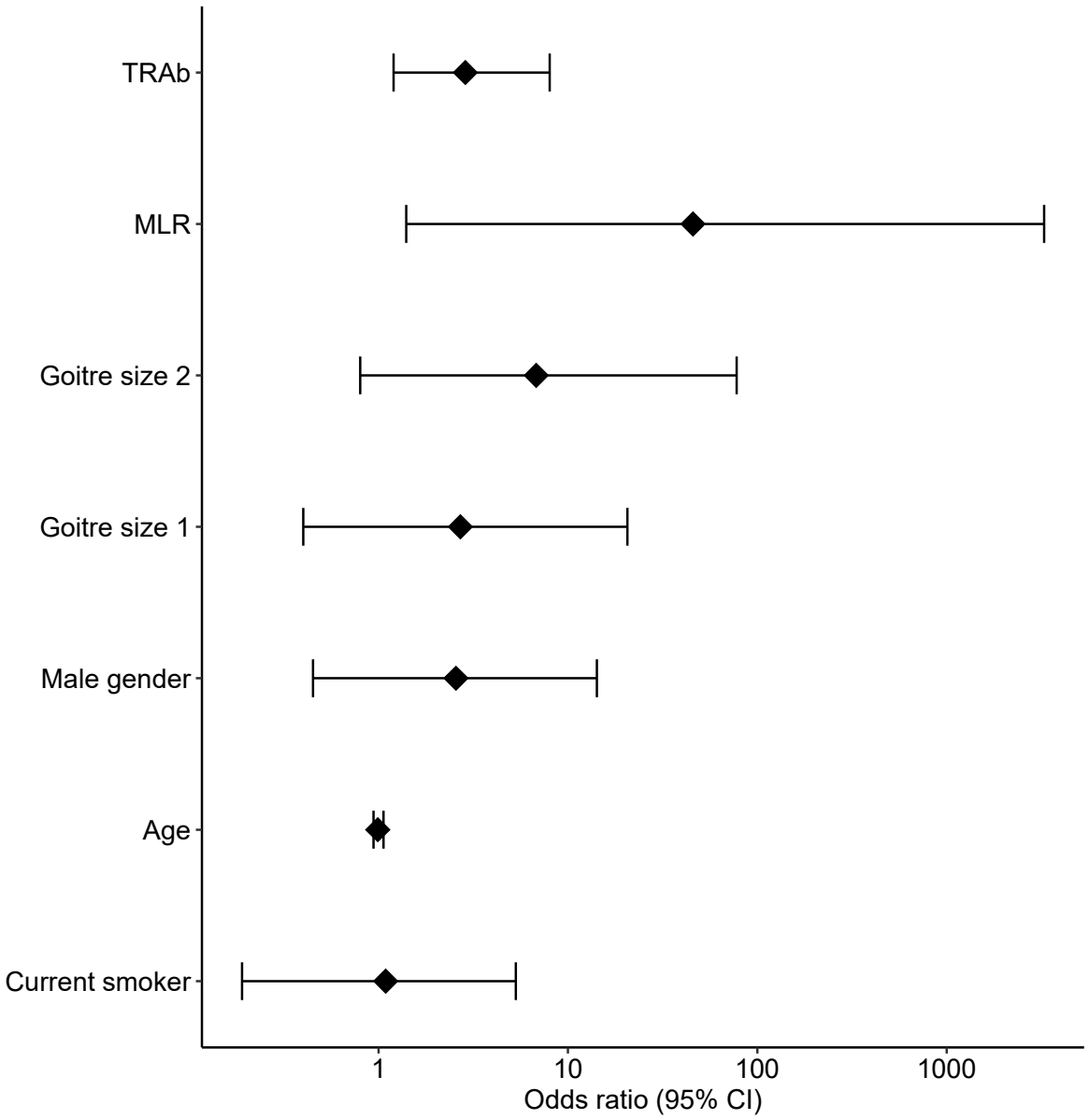


Figure 4.6 – The association of baseline clinical variables with the occurrence of GD relapse following ATD withdrawal, using a multivariate binary logistic regression model. **4.5 Monocyte-lymphocyte ratio (MLR)**

The MLR is the absolute monocyte count divided by the absolute lymphocyte count. It is described as a biomarker of inflammation and was studied in relation to GD relapse as it has been observed to be elevated in many acute and chronic diseases including autoimmune disease, and in some cases associated with disease activity and outcome. The results of the binary logistic regression model described above suggest an elevated MLR is associated with GD relapse.

The MLR at baseline and 6-10 weeks later was higher in the relapse patients (median 0.4 vs. 0.29, P= 0.02; median 0.38 vs. 0.3, P=0.01, respectively) (Figure 4.7).

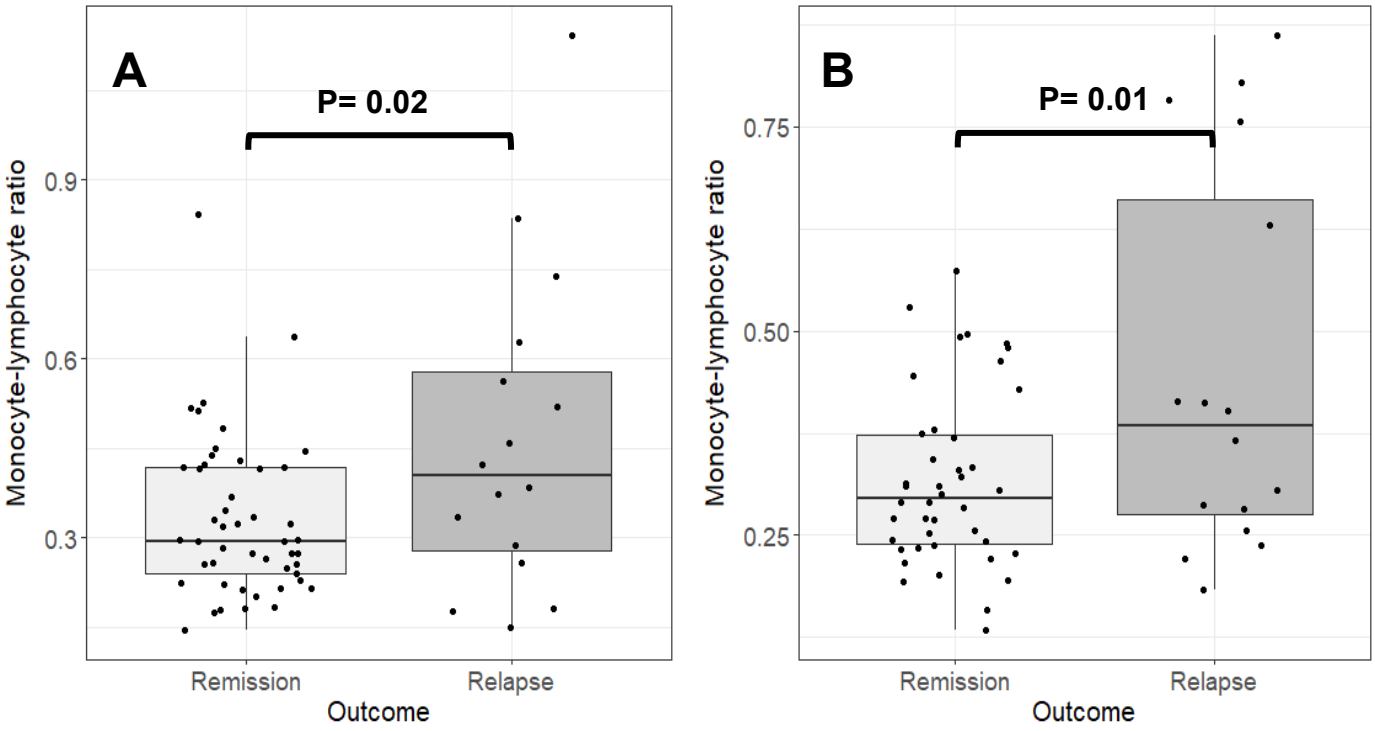


Figure 4.7 – Association of the monocyte-lymphocyte ratio and outcome in GD at baseline (**A**) and 6-10 weeks later (**B**).

Further analysis was undertaken to identify what was driving the elevated MLR by comparing the absolute monocyte and lymphocyte counts between the relapse and remission patients. There was no overall difference in the absolute lymphocyte count between the relapse and remission patients at baseline or 6-10 weeks (P=0.29, P=0.34, respectively). However, there was an observed increase in the absolute monocyte count in the relapse patients at 6-10 weeks (P=0.025), although this was not significant at baseline (P=0.07) (Figure 4.8).

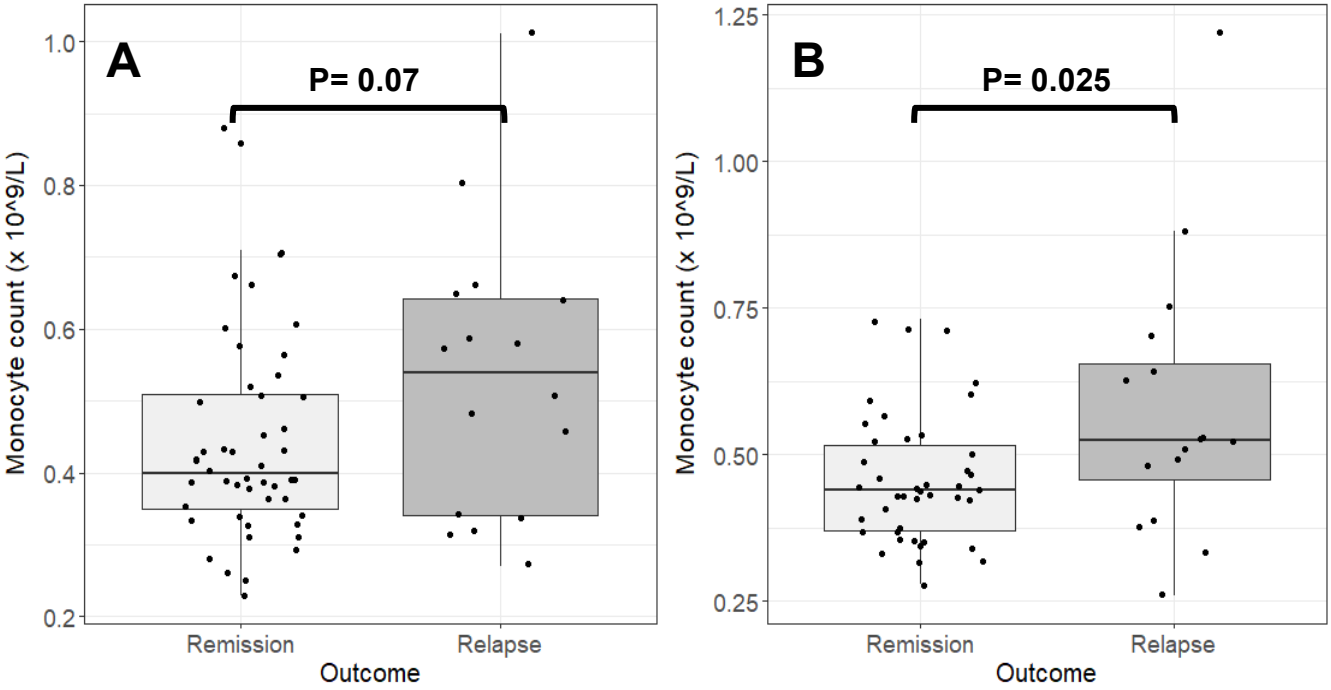


Figure 4.8 – Association of the absolute monocyte count and outcome in GD at baseline (**A**) and 6-10 weeks later (**B**).

4.6 Application of the 'GREAT' predictive score for Graves' disease outcome

As discussed in Chapter 1: Introduction, there is currently only one validated predictive score available to predict outcome of GD. Retrospective analysis was undertaken using the 'GREAT' scoring criteria on the specific diagnostic variables of SPRING study patients (Table 4.4).

GREAT score (class)	Relapse group (%) (n=16)	Remission group (%) (n = 49)
0-1 (I)	4 (25)	16 (33)
2-3 (II)	6 (37.5)	24 (49)
4-6 (III)	6 (37.5)	9 (18)

Table 4.4 – The association of the different classes of GREAT score with outcome in GD patients.

This demonstrated that there was a relapse rate of 20% for both Class I patients and Class II patients and 40% for Class III patients. Thus, a higher 'GREAT' score predicted a higher risk of relapse after ATD withdrawal if in Class III but it did not differentiate between the low or intermediate class I and II scores. The difference between the groups was not statistically significant on chi square (P=0.29) or Kruskal-Wallis testing (P=0.28). This disparity compared to other studies that have demonstrated good validity of the score may be partly due to the retrospective nature of the analysis including differences in clinical documentation of goitre size at diagnosis and a shorter follow-up period to define relapse (12 months vs. >2 years). In addition, the 'GREAT+' score which includes genetic risk alleles demonstrated an improved predictive value for relapse when compared to just the clinical factors alone (*Vos XG et al. 2016*), and these were not measured in this study.

4.7 Survival analysis

The association between the clinical variables at baseline and time to relapse following ATD withdrawal was assessed using univariate Cox regression analysis (Table 4.5).

Variable (Timepoint 1)	B	HRrelapse	95% CI HRrelapse	Univariate P value
MLR	3.8	43	4.43 – 420	**0.001***
TRAb (IU/L)	0.38	1.5	1.12 – 1.92	**0.005***
ATD regimen DT	-0.97	0.38	0.11 – 1.34	0.13
Goitre size				
1	0.7	2.02	0.55 – 7.5	0.29
2	1.1	3.03	0.68 – 13.5	0.15
NLR	0.24	1.3	0.88 – 1.8	0.19
PLR	0.004	1	0.99 – 1.01	0.23
FT3 (pmol/L)	0.31	1.4	0.82 – 2.28	0.23
FT4 (pmol/L)	0.07	1.1	0.92 – 1.30	0.37
Time to normal TSH (months)	0.02	1.02	0.96 – 1.09	0.46
Duration of ATD (months)	0.008	1	0.99 – 1.03	0.46
TPO (IU/ml)	-0.003	1	0.99 – 1.00	0.53
TSH (miU/L)	-0.03	0.97	0.84 – 1.12	0.69
Current smoker	-0.25	0.78	0.22 – 2.7	0.70
Previous GD relapse	0.09	1.1	0.35 – 3.40	0.87

Table 4.5 – Association of clinical variables at baseline with timing of GD relapse following ATD withdrawal, using a univariate Cox regression model.

The assumptions of the Cox Proportional Hazards Model were tested for each variable using Schoenfeld Residuals to assess proportionality of hazards (as described in Chapter 3: Methods. The proportional hazard assumption was supported by a non-significant relationship between residuals and time for each variable, apart from FT4 at baseline (P=0.025; Figure 4.9). However, this variable was not significantly associated with time to GD relapse and the global Schoenfeld test for the multivariate Cox regression model was not significant (P=0.20).

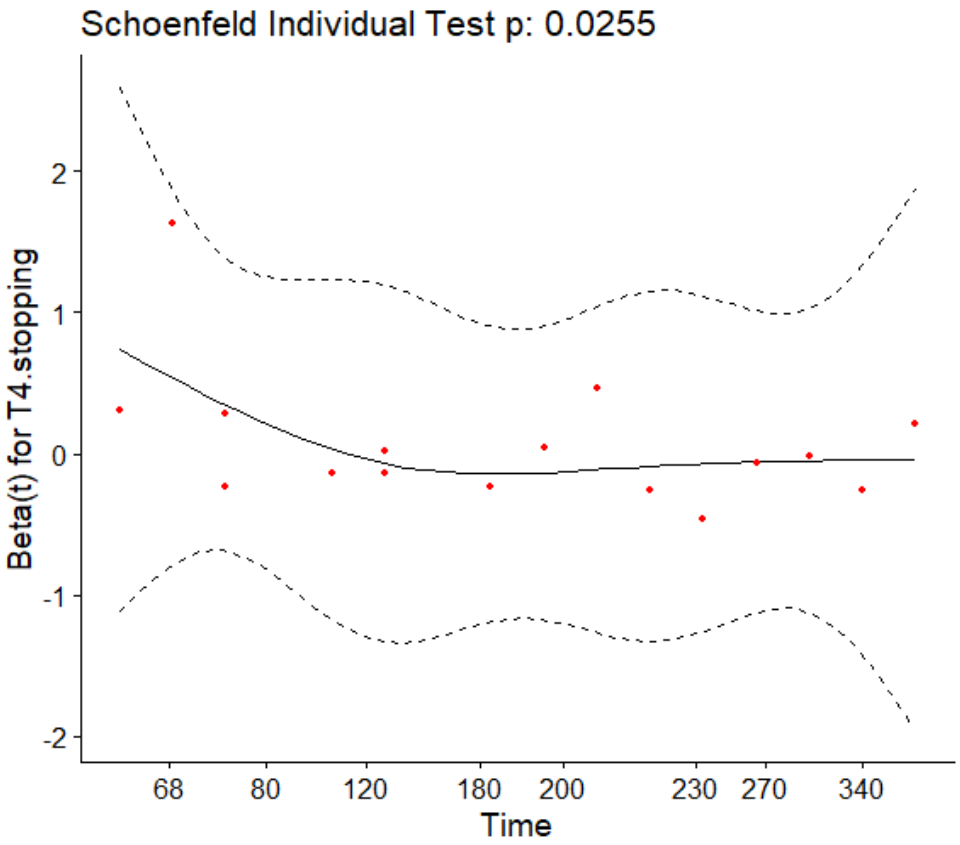


Figure 4.9 – Correlation of the scaled Schoenfeld residuals for FT4 against the transformed time (Beta(t)).

Out of these clinical variables assessed in the univariate Cox regression analysis, there were two variables significantly (P<0.05) associated with time to GD relapse, namely the TRAb concentration and the MLR at the time of stopping ATD. Both these biomarkers were also significantly associated with outcome on the binary logistic regression analysis. Therefore, these two variables alongside the other variables included in the previous multivariate binary logistic regression analysis (age, gender, goitre size, smoking, TRAb level prior to stopping ATD) were entered into a multivariate Cox regression model (Table 4.6).

Variable	B	HRrelapse	95% CI HRrelapse	Multivariate P value
MLR	3.69	40.1	2.18 – 739	**0.013***
Goitre size				
1	0.77	2.15	0.5 – 9.3	0.31

2	1.82	6.6	1.1 – 39.7	**0.039***
TRAb on stopping ATD	0.16	1.17	0.80 – 1.71	0.41
Gender (Female)	-0.57	0.57	0.14 – 2.24	0.42
Age	0.02	1.02	0.97 – 1.07	0.48
Current smoker	0.05	1.05	0.27 – 4.08	0.95

Table 4.6 – Association of baseline clinical variables with timing of GD relapse following ATD withdrawal, using a multivariate Cox regression model.

Following multivariate Cox regression analysis, two variables at baseline were significantly (P<0.05) associated with a faster time to GD relapse, namely the MLR and the presence of a size 2 goitre (Figure 4.10).

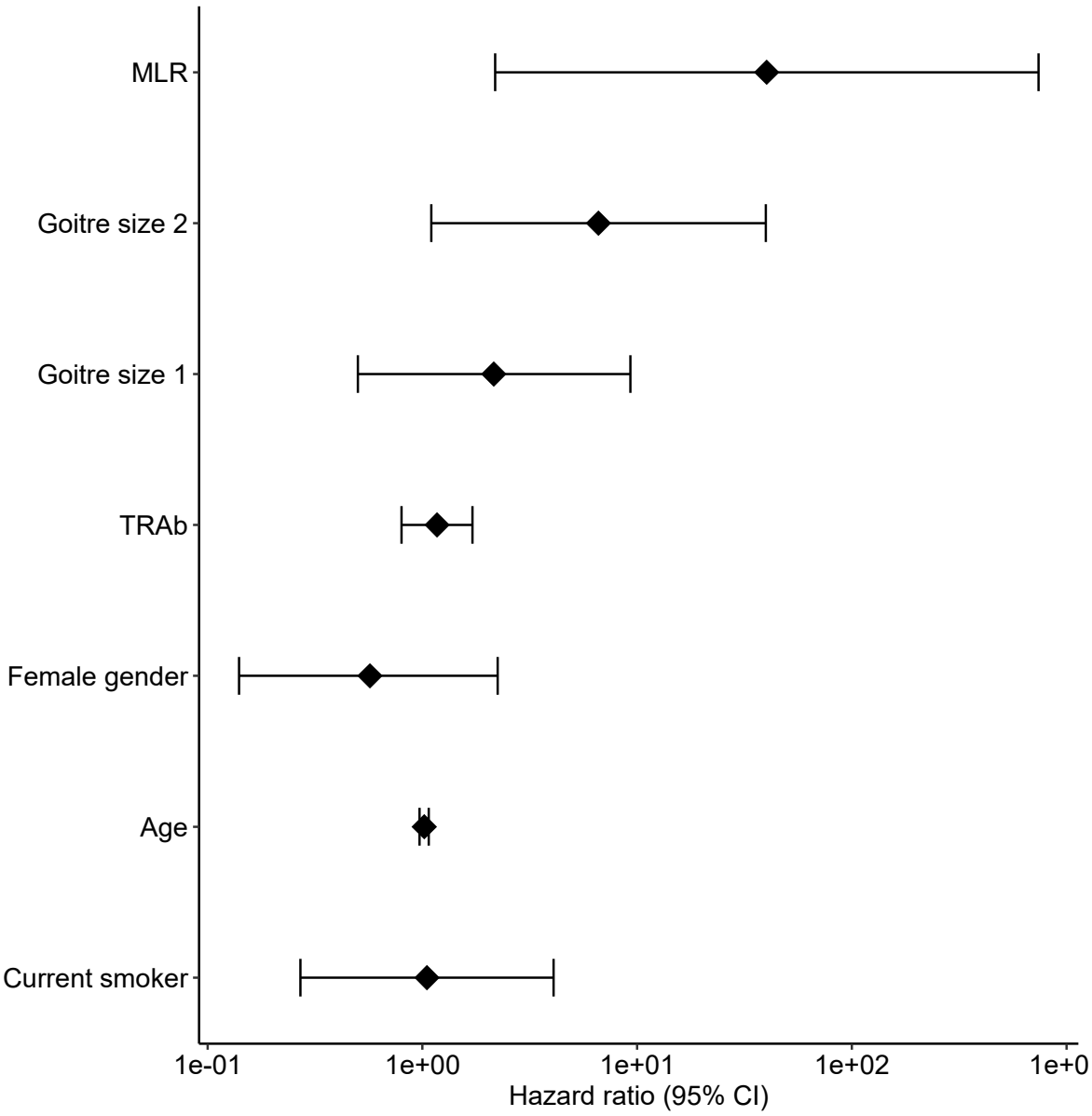


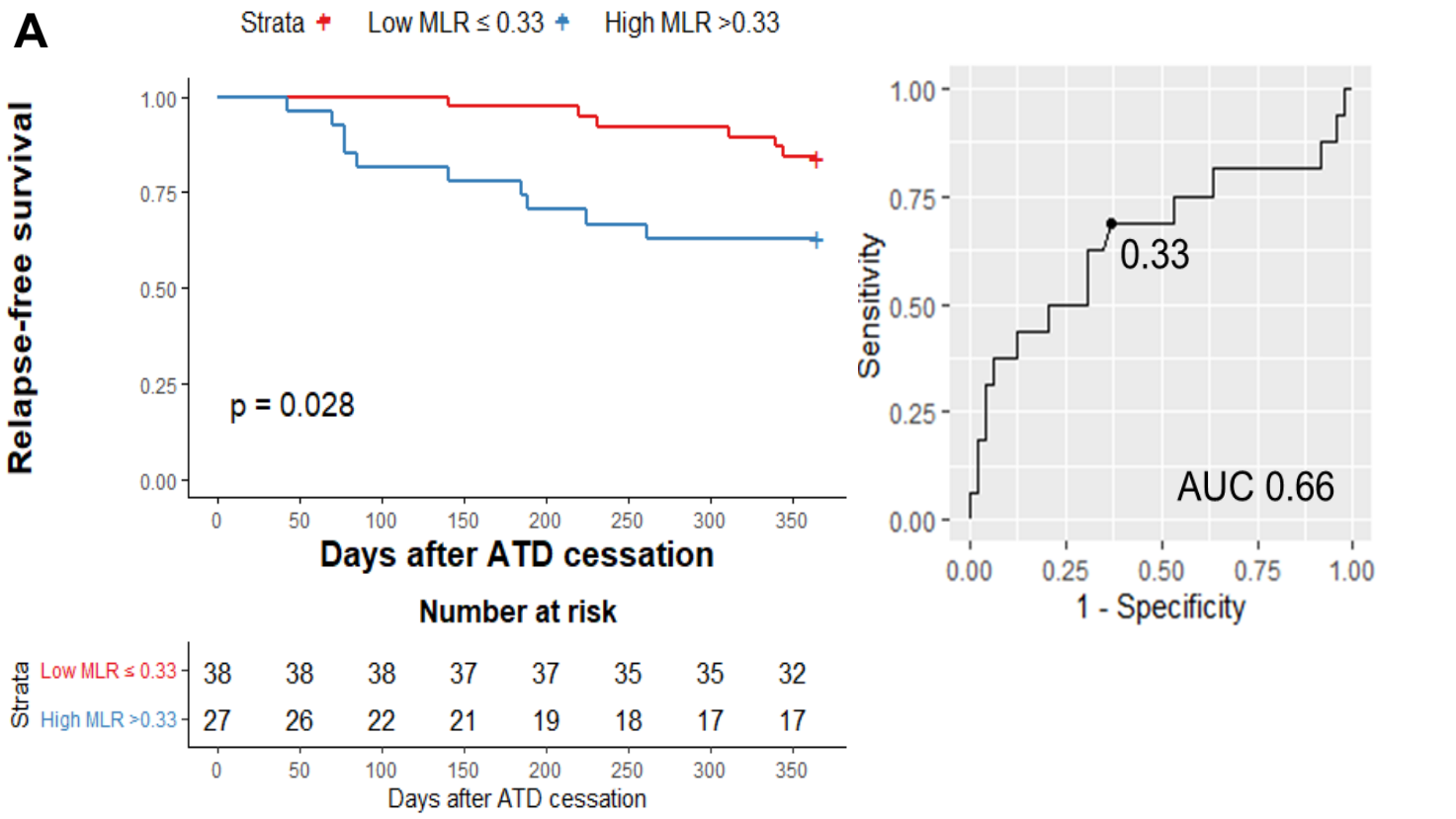
Figure 4.10 – The association of baseline clinical variables with timing of Graves' disease relapse following ATD withdrawal, using a multivariate Cox regression model.

4.6.1 The predictive utility of the monocyte-lymphocyte ratio, goitre size and TRAb concentration

The survival distributions were analysed for the variables observed as significant on multivariate Cox regression analysis (baseline MLR and goitre size). TRAb titre was also included as it had been significant on the univariate analysis and had a P value of <0.2.

ROC curve analysis was undertaken to determine the optimal TRAb and MLR thresholds to predict the outcome of relapse vs. remission. The R programme ‘cutpointr’ generated the optimal cut-off which produced the highest sensitivity and specificity for the prediction of relapse or remission (as described in Chapter 3: Methods).

The study cohort was dichotomised into high/low TRAb or MLR as described previously (Chapter 3: Methods). The survival distributions were analysed by the log-rank test and are presented in Kaplan-Meier plots, with the corresponding ROC curve (Figure 4.11).


A
Strata Low MLR ≤ 0.33 High MLR >0.33
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p = 0.028
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Low MLR ≤ 0.33 38 38 38 37 37 35 35 32
High MLR >0.33 27 26 22 21 19 18 17 17
0 50 100 150 200 250 300 350
Days after ATD cessation
Sensitivity
0.33
AUC 0.66
1 - Specificity

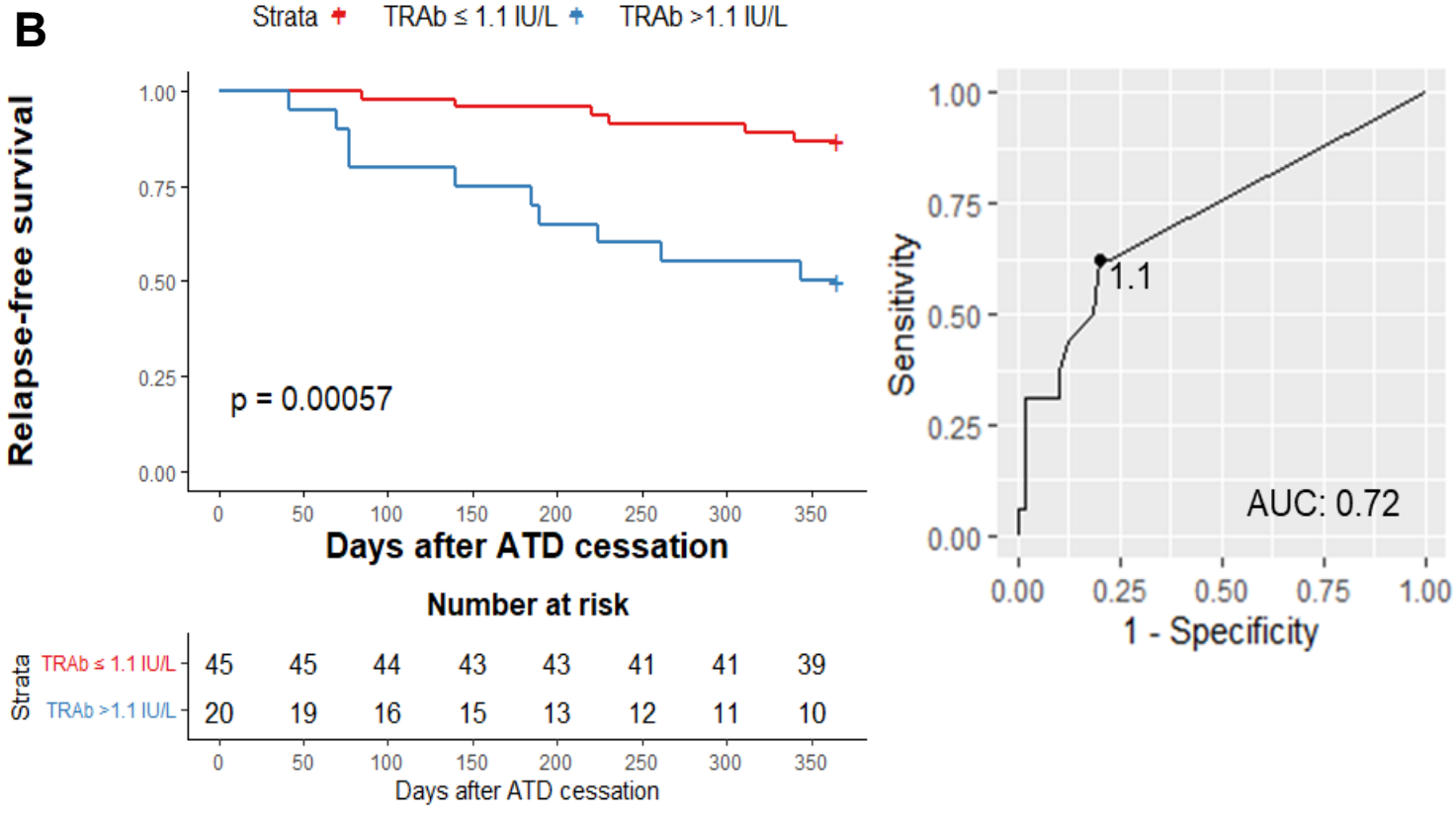

B
Strata TRAb ≤ 1.1 IU/L TRAb >1.1 IU/L
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p = 0.00057
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
TRAb ≤ 1.1 IU/L 45 45 44 43 43 41 41 39
TRAb >1.1 IU/L 20 19 16 15 13 12 11 10
0 50 100 150 200 250 300 350
Days after ATD cessation
Sensitivity
1.1
AUC: 0.72
1 - Specificity

Figure 4.11 – Kaplan-Meier plots of survival following ATD cessation stratified by high (blue) or low (red) level of MLR or TRAb. **A**The difference in survival distributions for the presence of an increased MLR (>0.33) and TRAb (>1.1 IU/L) met statistical significance (P=0.028 and P=0.00057, respectively). For both biomarkers, this difference was particularly apparent beyond the first 2 months after stopping ATD. There was no significant difference between the survival distributions for goitre size (P=0.33) (Figure 4.12). Furthermore, pairwise comparisons adjusted for multiple test correction (Benjamini & Hochberg) did not reveal any significant difference in survival distribution between any of the goitre size groups (Goitre size 0 and 1, P= 0.42; Goitre size 1 and 2, P=0.50; Goitre size 0 and 2, P=0.41).

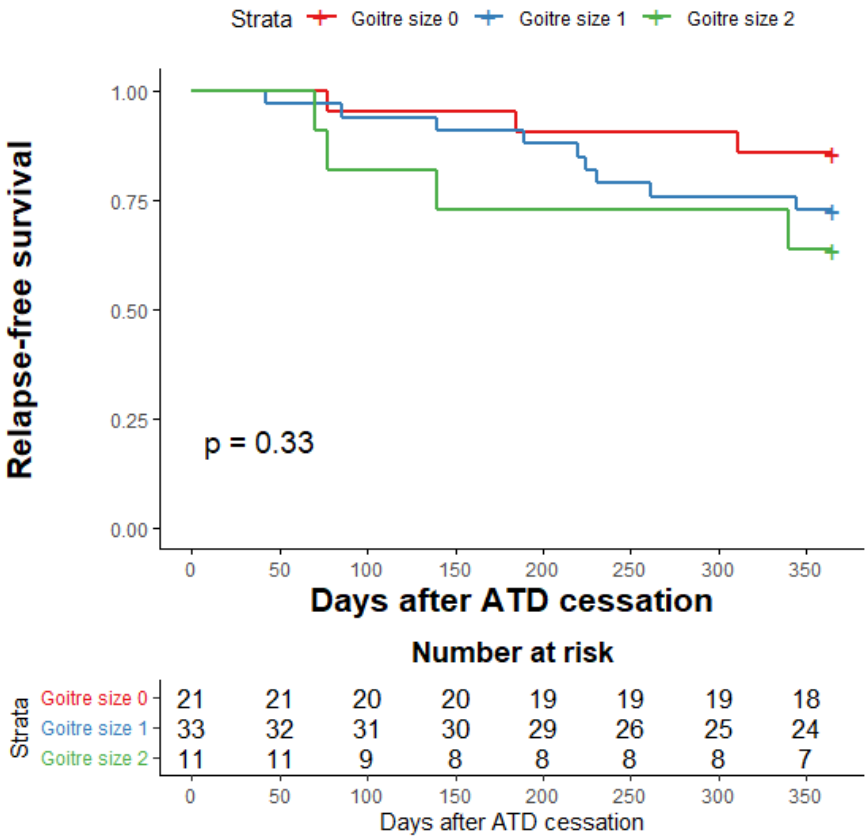


Figure 4.12 – Kaplan-Meier plot of survival following ATD cessation stratified by goitre size; 0 (red), 1 (blue), and 2 (green).

ROC_{AUC} values for both MLR and TRAb thresholds to predict relapse, including their sensitivity, specificity, and positive and negative predictive values, is presented in Table 4.6.

Baseline variable(s)	ROCAUC (95% CI)	Optimal Threshold	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95%CI)
MLR	0.66 (0.49-0.81)	0.33	**0.69** (0.50 – 0.85)	0.63 (0.46 – 1)	0.38 (0 – 1)	0.86 (0.68 – 0.93)
TRAb	0.72 (0.60-0.84)	1.1 IU/L	0.63 (0.33 – 0.82)	**0.80** (0.72 – 0.98)	**0.5** (0.43 – 0.86)	**0.87** (0.79 – 0.89)

Table 4.7 – ROC curve analysis of the predictive utility of MLR and TRAb biomarker thresholds to predict GD relapse following cessation of ATD.

4.7 Discussion

4.7.1 Graves' disease patient characteristics

The clinical and demographic characteristics of GD patients recruited to this study were largely representative of the typical population affected by Graves' disease. Over three-quarters of the cohort were female (82%) with a median age of 49 years, which is consistent with the usual patients affected by GD reported in the literature (*Hussain et al. 2017*). Most of the patients (92%) were of white ethnicity and therefore further studies would need to be undertaken to validate the findings in other ethnicities. However, the results from this study should be applicable to most Caucasian GD patients seen in clinical practice.

4.7.2 Graves' disease relapse

In this study, of the 65 GD patients who had their ATD withdrawn, 16/65 (25%) had relapsed by 12 months, with 24/65 (37%) relapsing by 24 months. This is consistent with the literature for patients such as these that primarily had low/negative TRAb titres, with previous studies in those with low TRAb concentrations demonstrating relapse rates of 22% and 36% 12 months and 24 months after ATD cessation, respectively (*Tun et al. 2016*).

4.7.3 Timing of Graves' disease relapse

The median time to patients relapsing their GD in this study was 187 days (6 months). Only 8 patients relapsed their GD in the second year after stopping ATD compared to 16 patients in the first 12 months. This is consistent with previous studies including a systematic review that found around two thirds of GD patients

experience relapse within 6-18 months after stopping ATD (*Struja et al. 2017*). However, there are many factors that may influence time to relapse including whether the TRAb titre was negative at the time of stopping ATD and the frequency of patient follow-up to detect a relapse. Patients in this study were followed up 3 monthly after the first two blood samples had been taken, and only one patient contacted in between appointments with symptoms suggestive of relapse necessitating an earlier blood test. Therefore, documented time to relapse could vary depending on whether the individual patient proactively attends for a blood test if symptoms develop prior to their next appointment. However, the consistent and regular frequency of patient follow-up in this prospective study means that there was a greater accuracy in capturing the time of relapse for most patients.

The timing of the second blood sample, most of which were taken at 6-10 weeks (60/65; 92%), aimed to capture the patient's immune activity prior to a relapse developing to try and gain insight into the factors that could be driving or initiating a GD relapse. However, 4/65 (6%) patients had relapsed their GD before 10 weeks and therefore had relapsed by/at the time the second blood sample was taken. Therefore, any changes in the biomarkers of these patients may have occurred as a result of the ongoing immune activity or intercurrent thyrotoxicosis and may not be useful in the context of predicting relapse.

4.7.4 Baseline predictors of Graves' disease relapse

The most clinically useful predictors of GD relapse would be those available at the time of stopping ATD when a clinical decision is made about whether it is the right time to stop treatment. There were two clinical variables associated with relapse at the time of stopping ATD, namely the TRAb titre and the MLR. Both these variables were also associated with time-to-relapse of GD on univariate cox regression analysis, however only the MLR and the presence of a size 2 goitre at the time of stopping ATD were significantly associated with time to relapse on the multivariate cox regression analysis. All these three variables are easily measurable in clinical practice, either through simple venepuncture (TRAb, MLR) or through routine clinical examination (goitre size).

TSH-receptor autoantibodies (TRAb)

This study found that the presence of an elevated TRAb titre at the time of stopping ATD had the strongest association with the likelihood of a patient relapsing their GD and was also associated with a poorer relapse-free survival. This is consistent with most previous studies that observed TRAb titre prior to stopping ATD as the most significant factor in predicting GD relapse (Chapter 1: Introduction). Indeed, most of the patients in this study that had relapsed by 12 months had positive TRAb titres at the time of stopping ATD. This explains the current guidance that advises using the TRAb titre prior to stopping ATD to determine which patients could consider ATD withdrawal (*Kahaly et al. 2018*).

The presence of circulating TRAb suggests ongoing disease activity and therefore it is logical that TRAb levels are directly associated with the likelihood of a patient relapsing. However unfortunately, over half of GD patients will experience a relapse 4 years after stopping ATD, despite having had low TRAb levels at the time of stopping ATD (*Tun et al. 2016*). Therefore, it is important to explore other measurable variables that may contribute to a more accurate prediction of relapse.

Monocyte-lymphocyte ratio (MLR)

This study found that an elevated MLR was associated with both relapse and time-to-relapse, as well as a poorer relapse-free survival. The MLR has been proposed as a biomarker of inflammation for both acute and chronic disease states, and previous studies have observed an association with disease activity and outcome in autoimmune disease, including SLE and RA (*Suszek et al. 2020, Du et al. 2017*).

Although monocytes are key components of the innate immune system, they also have a role in modulating the adaptive immune response by activating T and B lymphocytes through their capacity as antigen-presenting cells and their ability to produce proinflammatory cytokines (*Kapellos et al. 2019*). Proinflammatory cytokines produced by monocytes include those previously observed to be elevated in GD, such as BAFF and IL-6 (Chapter 1: Introduction). Furthermore, circulating monocyte levels have been observed to be higher in GD patients, associated with active disease, and positively correlated to TRAb titres (*Turan 2019, Chen et al. 2021*). Monocytes have also been demonstrated to infiltrate the thyroid gland in GD (*Chen et al. 2021*). In the SPRING study although the absolute monocyte count 6-10 weeks after stopping ATD was associated with relapse, this association was not observed in

the sample taken at baseline prior to stopping ATD. However, this may be due to monocytic infiltration within the thyroid gland reducing circulating peripheral monocytes.

The ratio of absolute monocyte to lymphocyte count is described as a biomarker of chronic systemic inflammation. An elevated MLR was associated with relapse and time-to-relapse in this study, which indicates it may be a biomarker of active GD and could reflect persistent inflammation within the thyroid gland, that if present at the time of stopping ATD, results in a higher likelihood of relapse. Given the previous association observed with MLR and autoimmune disease activity, the MLR may be directly correlated to the severity of inflammation which is why it may be associated with the time it takes patients to relapse. The wide confidence interval observed with the MLR as a variable likely reflects the small sample size in this study and therefore the findings would need to be validated on a larger scale. However, the MLR is an accessible and cost-effective biomarker, the measurement of which could easily be translated into routine clinical practice.

Goitre size

The presence of a goitre is a common clinical manifestation of GD where circulating TRAb stimulate the TSHR on thyroid follicular cells to promote thyroid hormone synthesis, resulting in diffuse enlargement of the thyroid gland. Indeed, in this study even after completing at least 12 months of ATD, nearly three quarters of patients were observed to have a palpable goitre. The presence of a large goitre at diagnosis and at the time of ATD withdrawal are well established risk factors for GD relapse (Chapter 1: Introduction). A size 2 goitre (goitre visible with the head in a normal position) at the time of ATD withdrawal was significantly associated with time-to-relapse on multivariate cox regression analysis. This is consistent with a previous study of 133 GD patients with mild to moderate goitres, which observed that those who had non-visible goitres after ATD treatment tended to have lower rates of relapse compared to those with a visible goitre (8.9% vs. 44.4%) (*Liu et al. 2015*).

A quarter of patients in this study that had a size 2 goitre at the time of ATD withdrawal had relapsed by 12 months, with over 50% of these patients relapsing by 24 months. The persistent enlargement of the thyroid gland could indicate ongoing stimulation by TRAb, although the majority (73%) of patients with a size 2 goitre in this study also had negative TRAb concentrations. However, this could reflect the

fact that there may be circulating TRAb undetectable by current laboratory assays. Although in this study all the patients were clinically examined by one practitioner, the subjective nature of routine clinical examination and the presence of various different classifications of goitre size could potentially impact the usefulness of assessing goitre size in a standardised way.

4.8 Summary

The clinical and demographic features of patients in this study are consistent with the usual population of patients affected by GD and therefore the findings should be largely generalisable to most GD patients. Furthermore, the rates of relapse and time-to-relapse were consistent with previously reported data in the literature.

This study emphasises the importance of TRAb concentration at the time of ATD withdrawal in predicting GD relapse, but also highlights that TRAb alone is inadequate in its predictive utility. Similarly, goitre size is a more subjective measure and although associated with time-to-relapse was not able to adequately predict GD relapse. The MLR has not previously been studied in association with outcome in GD, but this study found it to be an independent biomarker for predicting relapse and time-to-relapse. Additionally, the relationship between MLR and inflammation may provide insight into the immunopathology driving active and relapsing GD.

All the three clinical variables discussed are easily measurable and cost-effective, as venepuncture and clinical examination are undertaken routinely in GD assessment. The use of a biomarker score encompassing different clinical variables may help to standardise and improve the accuracy of predicting relapse.

Chapter 5: Results 2 – Kappa-deleting recombination excision circles (KRECs)

5.1 Introduction

Kappa-deleting recombination excision circles (KRECs) are circular DNA molecules generated during B cell maturation in the bone marrow which remain present in B cells that have recently emigrated from the bone marrow, and therefore they provide a quantitative measure of B cell output. They can also provide insight into the replicative activity of B cells in the peripheral compartment. The aim of investigating humoral immune activity in GD using KRECs as a marker of B cell output and proliferation was to improve our understanding of the immunopathology driving GD, provide mechanistic insight into the relapse process and assess its utility as a potential biomarker of relapse.

In this chapter, I present analysis of the KRECs data from SPRING study patient samples taken at baseline (on ATD withdrawal) and again 6-10 weeks later, aiming to explore differences in humoral activity between the relapse and remission patients. Furthermore, healthy control samples and additional hyper-and euthyroid GD patients were combined with the SPRING study patients to further examine the association between KRECs with GD and thyroid status.

The KRECs analysis was also used in conjunction with the cytokine and flow cytometry data to provide insight into the contribution of these different factors to the underlying humoral immune activity observed in GD (Chapter 9: Integrative analysis).

5.2 Quality control

KRECs were quantified by real-time quantitative polymerase chain reaction (RT-qPCR) using a comparator triple-insert plasmid (containing TREC and KREC signal joint (SJ) fragments and the T-cell receptor alpha constant gene (TCRAC) reference gene) supplied by Sottini and colleagues (*Sottini et al. 2014*). Both the standard curve and samples were run in triplicate, and 10% of samples were repeated to assess the inter- and intra-assay coefficient of variation. Negative controls and standard curves were run on each plate and the PCR efficiency was maintained between 92-100% with a slope value between $-3.5 \geq -3.3$.

5.2.1 Sample collection and processing

The samples used in this experiment were from the blood samples collected at the timepoint of stopping ATD therapy and again 6-10 weeks later. The EDTA sample was collected at these timepoints and stored at $^{-}80^{O}C$ before undergoing batch DNA extraction.

DNA was available for all baseline samples, however some patients (7/65;11%) did not have the second 6–10-week blood sample taken either because they could not be contacted (3/65) or because they were unable to attend due to the COVID-19 pandemic (4/65).

Further analysis of KREC levels was undertaken using 67 samples of previously stored DNA from hyperthyroid and euthyroid GD patients recruited to an observational study of autoimmune disease (Pathogenesis and Treatment of Endocrinopathy Study). Hyperthyroidism was defined as having a suppressed TSH with an elevated FT3 (> 6.8pmol/L) and/or FT4 (> 22pmol/L)). Euthyroidism was defined as a FT3 and FT4 within the normal reference range (FT3 3.1 – 6.8pmol/L and FT4 10 – 22pmol/L).

In addition, DNA was also extracted from 140 healthy control samples provided by the NIHR BioResource Centre, Newcastle-upon-Tyne. Therefore, in total there were 330 separate samples from which KREC levels were quantified (including 123 samples from the SPRING study patients – baseline and 6-10 week samples).

The quantity of KRECs is reported as number of KRECs per 10^6 cells and the replicative history of the B cells is reported as a ratio using the SJ and CJ Ct values. The calculations used for this are outlined in Chapter 3: Methods.

5.3 Demographic and clinical data of the healthy control samples and hyperthyroid/euthyroid Graves' disease samples

The clinical and demographic characteristics of the SPRING study patients are presented in Chapter 4: Results 1 – Clinical. The demographic data of the healthy control samples used as part of this analysis is presented in Table 5.1. The age of the healthy control samples was closely aligned to the SPRING study patients (median 51 years vs. 46 years), although there were less female healthy controls (62% vs 82%).

Demographic variable	Healthy controls	SPRING patients
Total number of subjects	140	65
Age (years): median (IQR) [range]	51 (39 – 57) [20 – 70]	49 (39 – 56) [20 – 80]
Female: n (%)	87 (62%)	53 (82%)

Table 5.1 – The demographic characteristics of the healthy control samples compared to the SPRING patients.

The demographic and clinical characteristics of the 67 GD patients (from the Pathogenesis and Treatment of Endocrinopathy Study as detailed above) used as part of this analysis is presented in Table 5.2. As expected, most of these GD patients were female which was comparable to the GD SPRING study patients (87% vs. 82%).

Demographic / Clinical variable	Value
Total number of hyperthyroid patients	46
Total number of euthyroid patients	21
Age (years): median (IQR) [range]	42 (32 – 51) [20 – 92]
Female: n (%)	59 (87%)
Thyroid function at time of sample (hyperthyroid patients):	
median (IQR) [range]	
FT4 (pmol/L)	36 (28 – 54) [12.9 – 100]
FT3 (pmol/L)	15.5 (10 – 30) [6.2 – 50]
TRAb (IU/L) – 34/46 (74%) available	9.5 (5 – 28) [1.7 – 100]
TPO Ab (IU/ml) – 11/46 (24%) available	67 (8 – 318) [2 – 600]
Thyroid function at time of sample (euthyroid patients):	
median (IQR) [range]	
FT4 (pmol/L)	15 (13 – 18) [10.8 – 21.7]
FT3 (pmol/L)	4.7 (4.2 – 5.5) [3.8 – 5.8]
TRAb (IU/L) – 5/18 (28%) available	0 (0 – 3.6) [0 – 85.9]
TPO Ab (IU/ml) – 2/18 (11%) available	98 (81 – 116) [63 – 134]

Table 5.2 – The demographic characteristics and thyroid function results of the hyper- and euthyroid Graves' disease patients, excluding the SPRING study patients.

5.4 KREC quantity and clinical characteristics

5.4.1 Association of KRECs with demographic and clinical characteristics

To evaluate the association between age and gender with KRECs, the total sample population was used (330 samples). There was no significant difference in age between the GD and control cohorts (46 years vs. 51 years; P = 0.38), but there were more females represented in the GD cohort (84% vs. 62%; P = 0.00004). However, there was no association found between KRECs and gender (P=0.50) or age (P=0.20) (Figure 5.1).

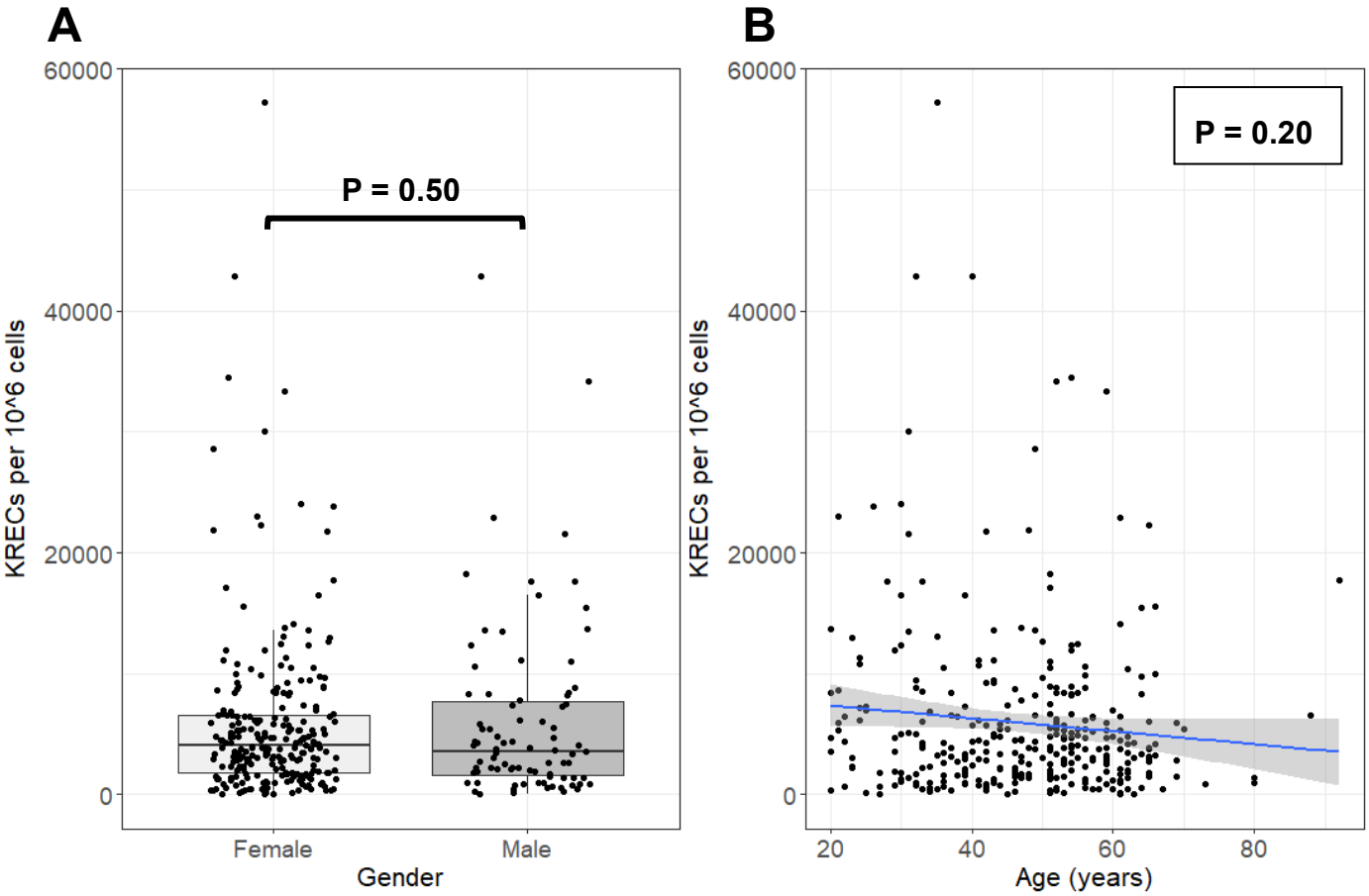


Figure 5.1 – The association of gender (**A**) and age (**B**) with KRECs per 10^6 cells. There was no association observed between KRECs and smoking status (P=0.64) or presence of GO (P=0.83) in the 65 GD SPRING study patients where this data was available.

5.4.2 Infective/vaccination status

Infective status and recent vaccination both have the potential to influence dynamic B cell activity. In the SPRING study patients, intercurrent illness was objectively determined by measuring the level of the acute inflammatory marker CRP, which demonstrated no association with KREC level at either baseline or 6-10 weeks later (P=0.95, P=0.94, respectively). None of the SPRING study patients had received a vaccination within one month of having a sample taken. CRP levels and vaccination status were not available for the additional GD patient samples or the healthy controls.

5.5 KRECs in Graves' disease vs. healthy controls

The association of KREC levels with disease status was evaluated by comparing GD patients with healthy controls, followed by comparisons between both hyperthyroid and euthyroid GD patients with healthy controls (Figure 5.2).

There was a significantly higher quantity of KRECs per 10^6 cells in GD patients compared to controls (median 4952 vs. 2571; P= 1.5×10^{-9}). Furthermore, a Kruskal-Wallis test revealed a significant difference between the 3 groups (controls, euthyroid, hyperthyroid; P= 4.5×10^{-14}). Pairwise comparisons using Wilcoxon rank sum test with multiple test correction (false discovery rate: Benjamini & Hochberg (*Benjamini et al. 1995*)) revealed a significant difference in KRECs per 10^6 cells between controls vs. euthyroid (median 2571 vs. 4092; P= 5.7×10^{-4}) and controls vs. hyperthyroid patients (median 2571 vs. 8614; P= 2.9×10^{-12}). There was also a significant difference observed between hyperthyroid vs. euthyroid GD patients (median 8614 vs. 4092; P= 8.8×10^{-9}). This indicates that elevated thyroid hormones are associated with higher KREC levels, but also that regardless of thyroid hormone status, those with GD are more likely to have elevated KREC levels.

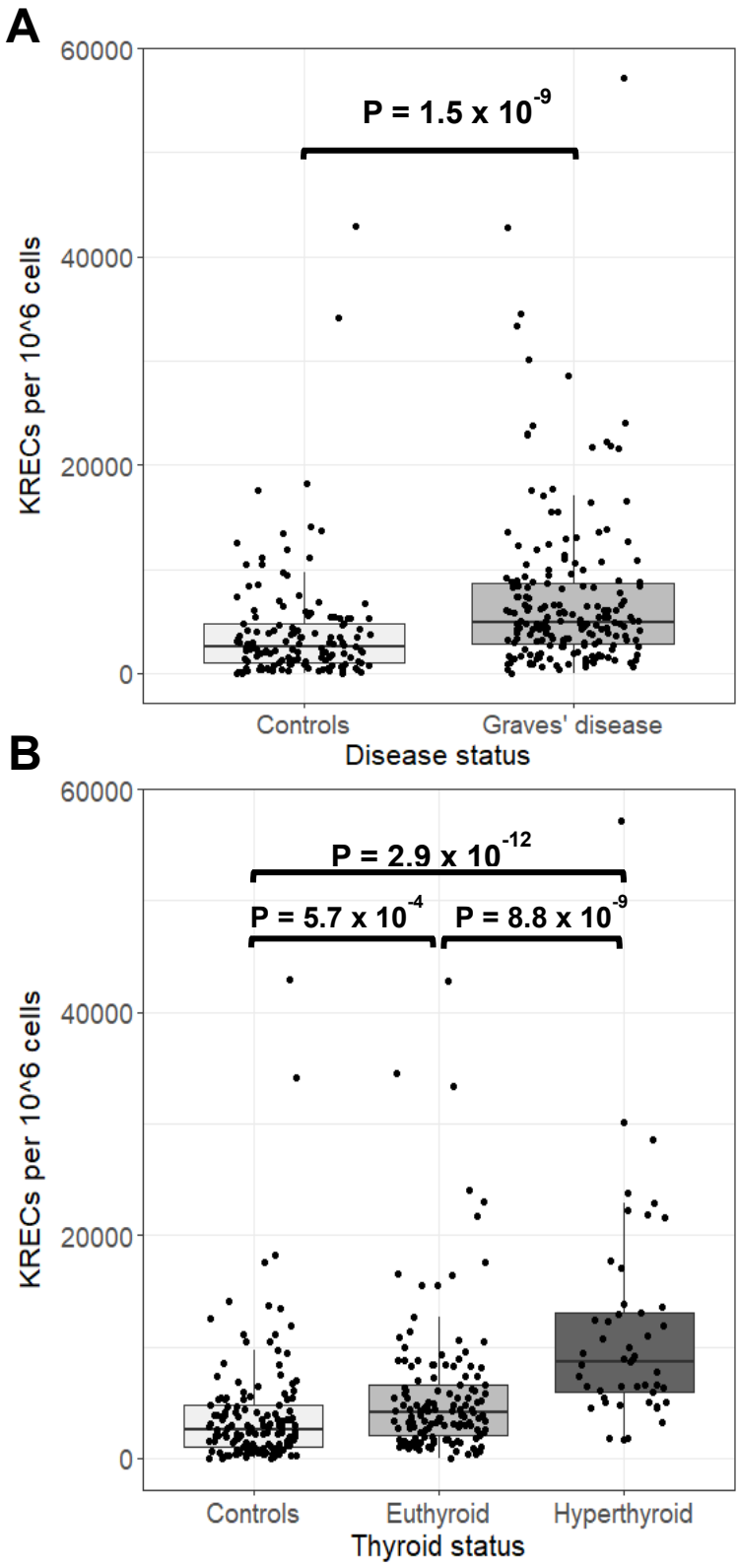


Figure 5.2– Association of KREC levels in controls vs. Graves' disease patients **5.6 Association of KRECs with thyroid status in Graves' disease**

To further evaluate the association between thyroid status and KRECs, all 190 GD samples (123 from the SPRING study and 67 from the Pathogenesis and Treatment of Endocrinopathy Study) where thyroid biochemistry was available (FT4 (all samples), FT3 (187/190; 98%), TPOAb (134/190; 71%) and TRAb titre (168/190; 88%) were entered into the correlation analysis. There was a positive correlation observed between KRECs and the log-transformed thyroid hormones and autoantibodies (FT4 ($P = 1.99 \times 10^{-5}$, rs = 0.30), FT3 ($P = 1.77 \times 10^{-7}$, rs = 0.37), TPOAb (P = 0.007, rs = 0.23) and TRAb ($P = 1.34 \times 10^{-5}$, rs = 0.33)) (Figure 5.3).

The results of a univariate linear regression analysis to evaluate whether thyroid status significantly predicted KREC quantity are presented in Table 5.3. All the thyroid hormones and antibodies measured were significantly associated with KREC quantity on the univariate analysis.

Thyroid hormone/antibody	B	SE B	95% CI B	t	Adj R^2	F	Univariate P value
ln(FT4)	38.7	11.3	16.4 – 61	3.4	0.05	11.7	**7.5×10^{-4}***
ln(FT3)	42.9	9.4	24.4 – 61.5	4.6	0.10	20.9	**9×10^{-6}***
ln(TRAb)	38.8	5.2	16.6 – 37.3	5.1	0.14	25.8	**7.51×10^{-7}***
ln(TPO Ab)	9.3	3.9	1.6 – 17	2.4	0.03	5.7	**0.019***

Table 5.3 – Association of the log-transformed (ln) thyroid hormones and autoantibodies with quantity of KRECs in Graves' disease patients, using a univariate linear regression model.

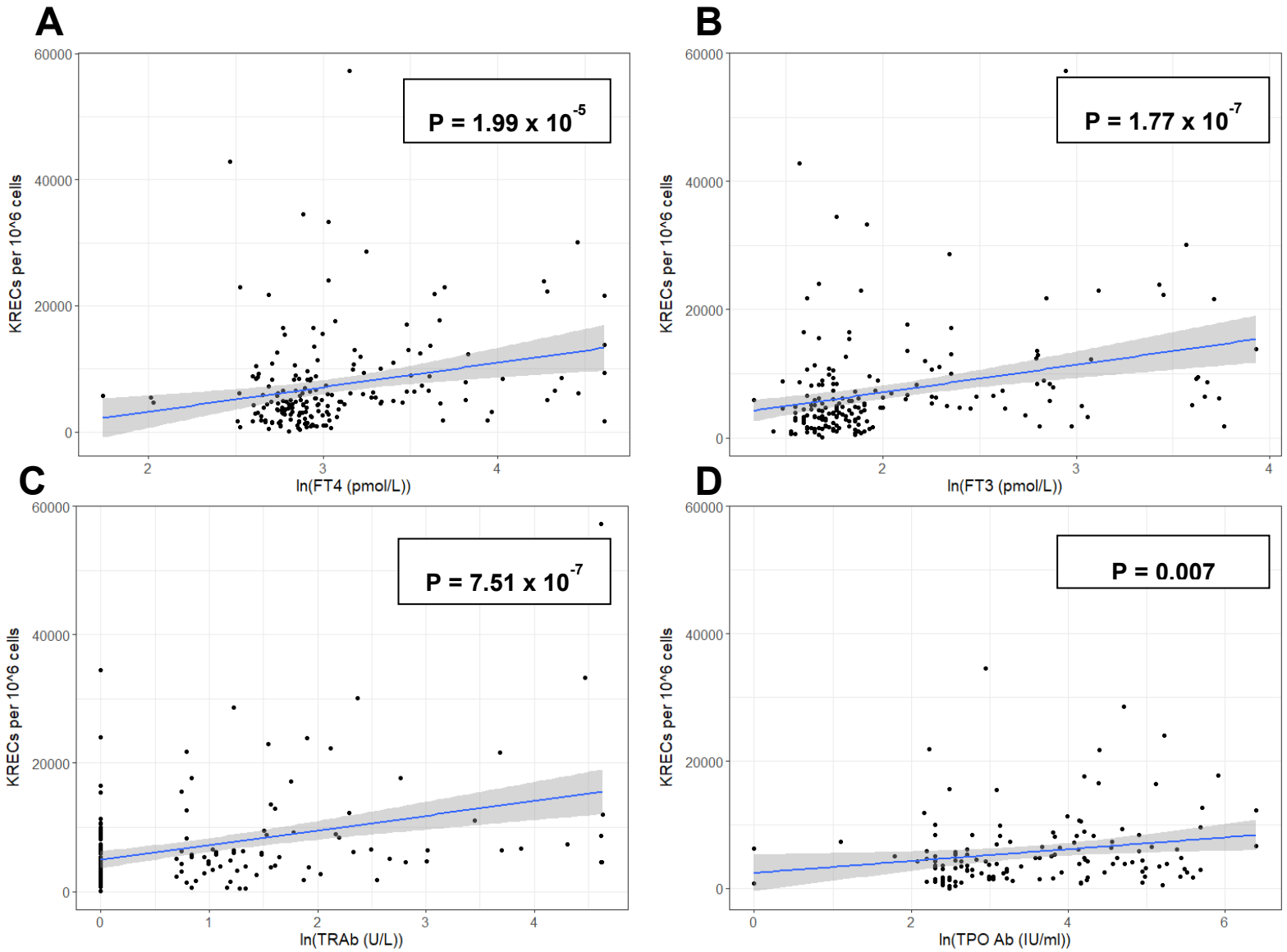


Figure 5.3 – Association between log-transformed (ln) thyroid hormones and serum autoantibodies (**A** FT4, **B** FT3, **C** TRAb, **D** TPO Ab) and KRECs per 10^6 cells.

A multiple linear regression analysis including all four measures of thyroid status did not reach overall statistical significance (P=0.11), although TPO Ab remained statistically significant in this model (Table 5.4). However, given the positive correlations that co-exist between both FT3 and FT4 thyroid hormones and autoantibodies it is likely that this multivariate model suffers from multicollinearity, which makes it challenging to distinguish between any individual effects of the thyroid hormones/antibodies on KRECs.

Thyroid hormone/antibody	B	SE B	95% CI B	t	Multivariate P value
ln(FT4)	7.9	24	-39 – 55.6	0.33	0.70
ln(FT3)	13.7	20	-26.4 – 53.9	0.68	0.50
ln(TRAb)	-8.6	7.8	-16.4 – 14.7	-0.11	0.91
ln(TPO Ab)	9.5	4.1	1.4 – 17.8	2.3	**0.02***

Table 5.4 – Association of the log-transformed (ln) thyroid hormones and autoantibodies with quantity of KRECs in Graves' disease patients, using a multiple linear regression model.

Variance inflation factor (VIF) is used for detecting multicollinearity, which measures the correlation and strength of correlation between the independent variables in a regression model. A value of VIF between 1-5 demonstrates a moderate degree of correlation which was observed with all the variables, more notable for FT4 and FT3 with values exceeding 2 (FT4 2.05, FT3 2.002, TRAb 1.3, TPO 1.01) (Figure 5.4).

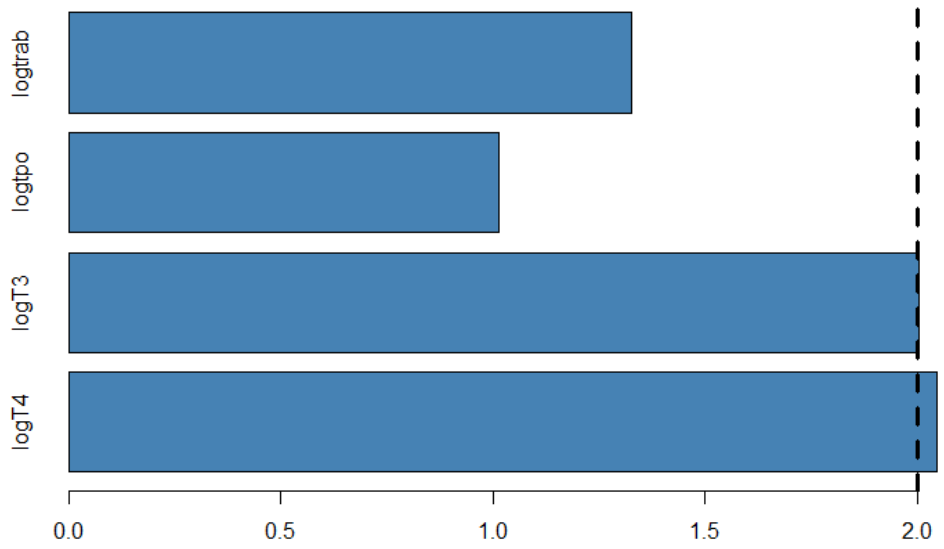


Figure 5.4 – Barplot presenting the observed Variance inflation factor values (VIF) for all the variables in the multiple regression model.

Given that FT4 and FT3 are highly correlated in this model, FT4 (which has the highest VIF) was removed from the regression model as its information is captured by including FT3 in the model. Following the removal of FT4, the VIF values for all the variables decreased (FT3 1.2, TRAb 1.2, TPO Ab 1.008) and this model demonstrated that TPOAb continued to be an independent predictor of KREC quantity (P=0.024) (Table 5.5). Within the 71% of patients that had a TPOAb measured, only 2 didn't have a measurable value – one in each of the hypothyroid and hyperthyroid group. After removal of these patients from the analysis, alongside those that didn't have a TPOAb measured, TPOAb remained significantly associated with KREC quantity (P=0.024)

Thyroid hormone/antibody	B	SE B	95% CI B	t	Multivariate P value
ln(FT3)	17.8	16	13.9 – 49.5	1.1	0.27
ln(TRAb)	-3.3	7.6	-15.4 – 14.8	-0.04	0.97
ln(TPOAb)	9.5	4.1	1.4 – 16.7	2.3	**0.024***

Table 5.5 – Association of the log-transformed (ln) thyroid hormones with quantity of KRECs in Graves' disease patients, using a multiple linear regression model, excluding FT4.

5.7 KRECs and Graves' disease outcome

The association between KRECs and the outcome of GD at 12 months was studied in the SPRING study patients both at the time of ATD withdrawal (baseline) and 6-10 weeks later. There was no significant association between KRECs per 10^6 cells at baseline and outcome (P=0.87), however those with higher KREC levels at 6-10 weeks were more likely to relapse (P=0.04) (Figure 5.5).

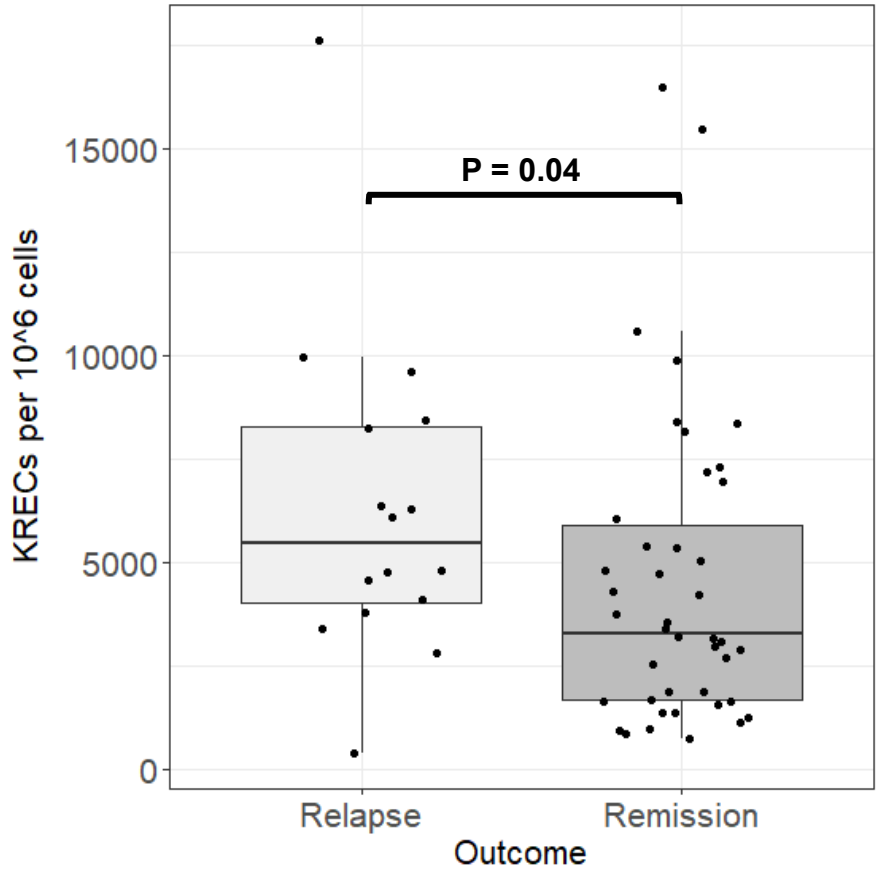


Figure 5.5 – Association of KRECs per 10^6 cells 6-10 weeks after stopping ATD with outcome at one year.

A multivariate logistic regression analysis was undertaken to investigate the ability of KRECs at 6-10 weeks to predict outcome in GD, including other factors already known to be associated with relapse of GD as used in previous analyses: age, gender, TRAb titre, smoking and goitre size (Chapter 4: Results 1 – Clinical). This regression model did not demonstrate KRECs at 6-10 weeks to be independently

associated with outcome (P=0.07), although as reported previously, high TRAb titre remained significantly associated with relapse in this model (P=0.02).

As previously discussed in this Chapter, higher thyroid hormone levels are positively correlated to KRECs and it was also observed that patients at 6-10 weeks who had higher FT3, FT4 and TRAb were more likely to relapse (Chapter 4: Results 1 – Clinical). Indeed, FT3 was positively correlated to KRECs per 10^6 cells in these 16 relapsing patients (P=0.018, rs 0.58) (Figure 5.6). There was no association observed in these patients between KRECs and TRAb (P = 0.13) or FT4 concentrations (P = 0.13).

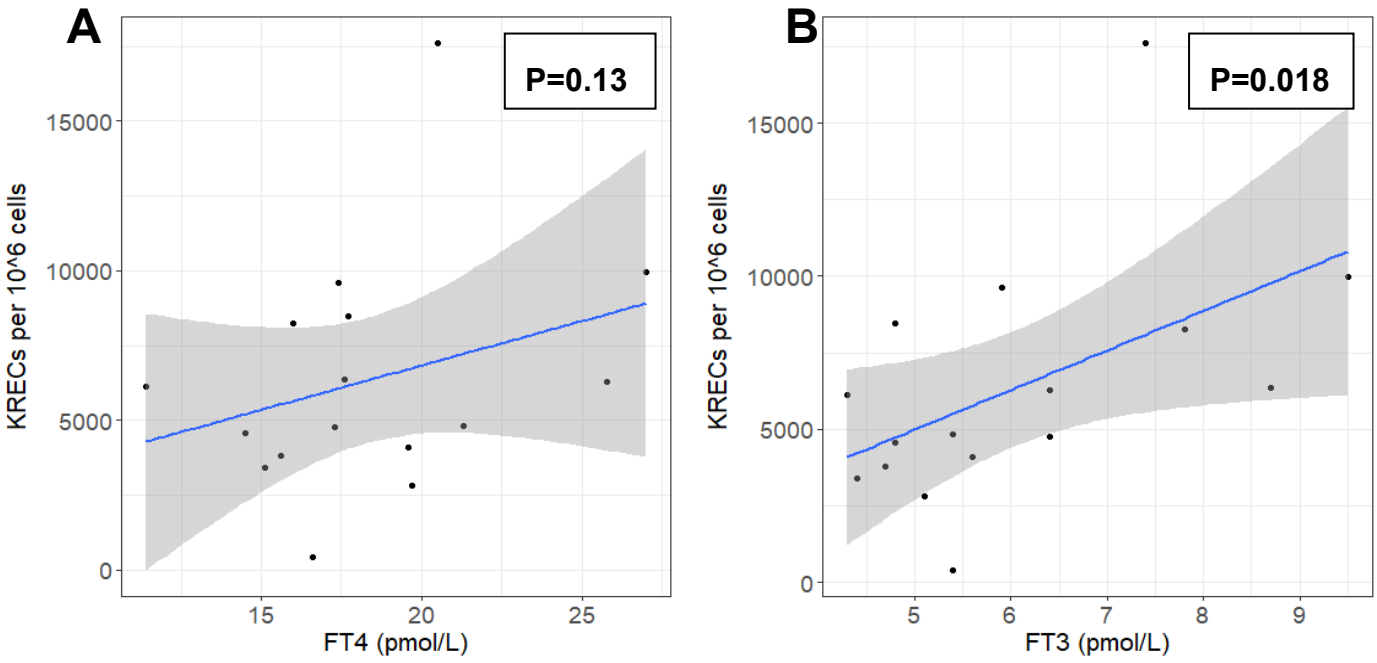


Figure 5.6 – Association between thyroid function and KRECs per 10^6 cells in the 16 relapsing patients at 6-10 weeks.

Of the 16 patients that relapsed at 12 months and 49 patients that remitted, there were 15/16 and 42/49 paired samples available, respectively, from both the timepoint of ATD withdrawal and 6-10 weeks later. Of the relapsing patients, the majority (12/15; 80%) either had an increase (10/15; 67%) or stable (2/15; 13%) KREC levels from withdrawal of ATD to 6-10 weeks later (Figure 5.7A) compared to those that remitted where only 17/42 (40%) had an increase or stable KREC levels (Figure 5.7B). This change in KRECs between the two outcome groups was nominally

significant on chi-square analysis (P = 0.046). A change of less than 50 KREC copies from baseline to 6-10 weeks was used for the purposes of defining ‘stable’ KREC levels. There was a trend towards a quicker relapse in the patients that had an increasing or stable KREC trajectory following withdrawal of ATDs compared to those with falling KRECs (mean 165 vs. 243 days) (P=0.24).

To evaluate whether KREC levels at baseline or 6-10 weeks were associated with the time it took patients to relapse their GD, a univariate cox regression analysis was performed which did not show any association between KREC levels and time to relapse (P=1).

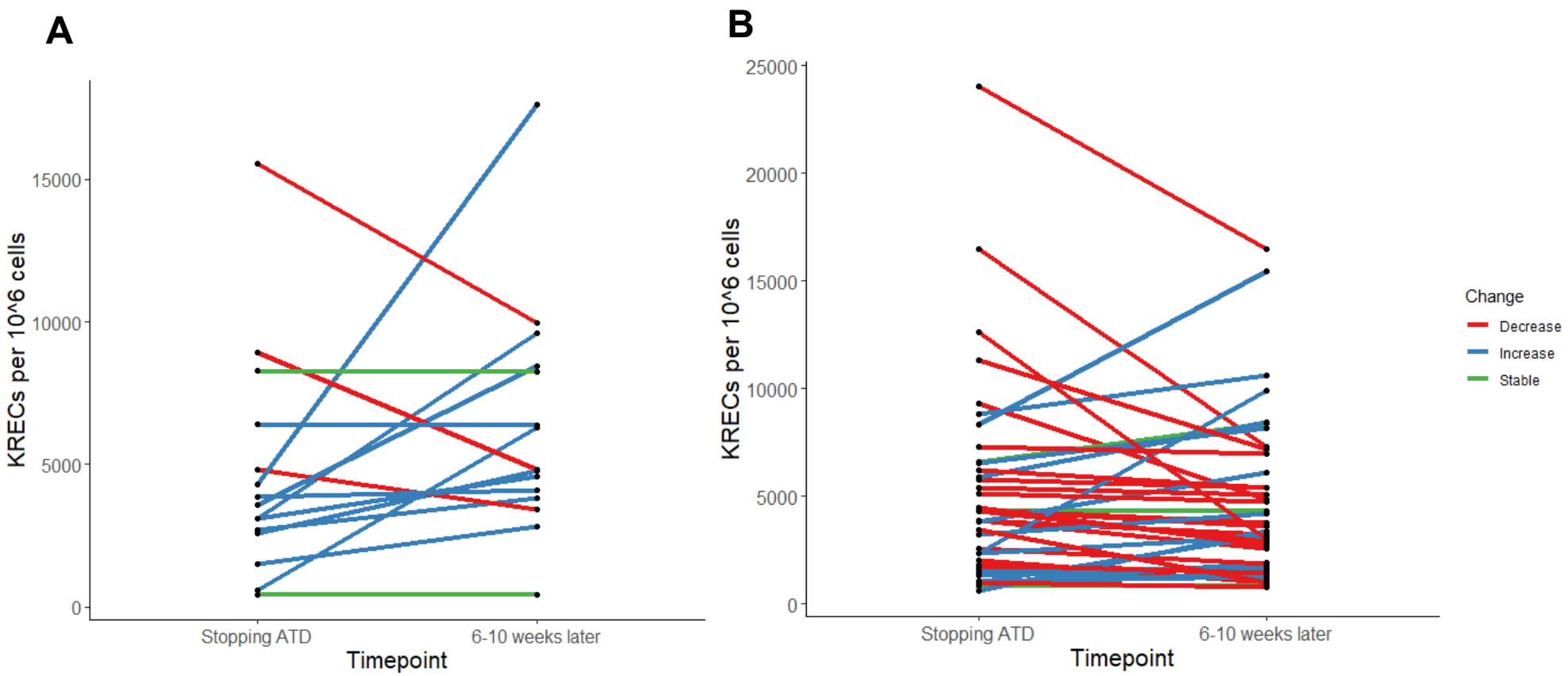


Figure 5.7 – KRECs per 10^6 cells at the timepoint of stopping ATD and 6-10 weeks later in 15 of the Graves' disease patients that relapsed (**A**) and 42 of the patients that remitted (**B**) at 12 months.

5.8 Longitudinal paired KRECs data

The difference in KREC levels in the SPRING GD patients at baseline (prior to ATD withdrawal) and 6-10 weeks later was investigated, however there was no significant change in KRECs between the two timepoints (P=0.75). This cohort was split by outcome to determine if there was any difference in longitudinal KREC levels between the relapse and remission patients. Although there was an increase in KRECs observed in the relapse group from baseline to 6-10 weeks compared to a slight decrease in the remission group, this difference was not significant (P=0.12) (Figure 5.8).

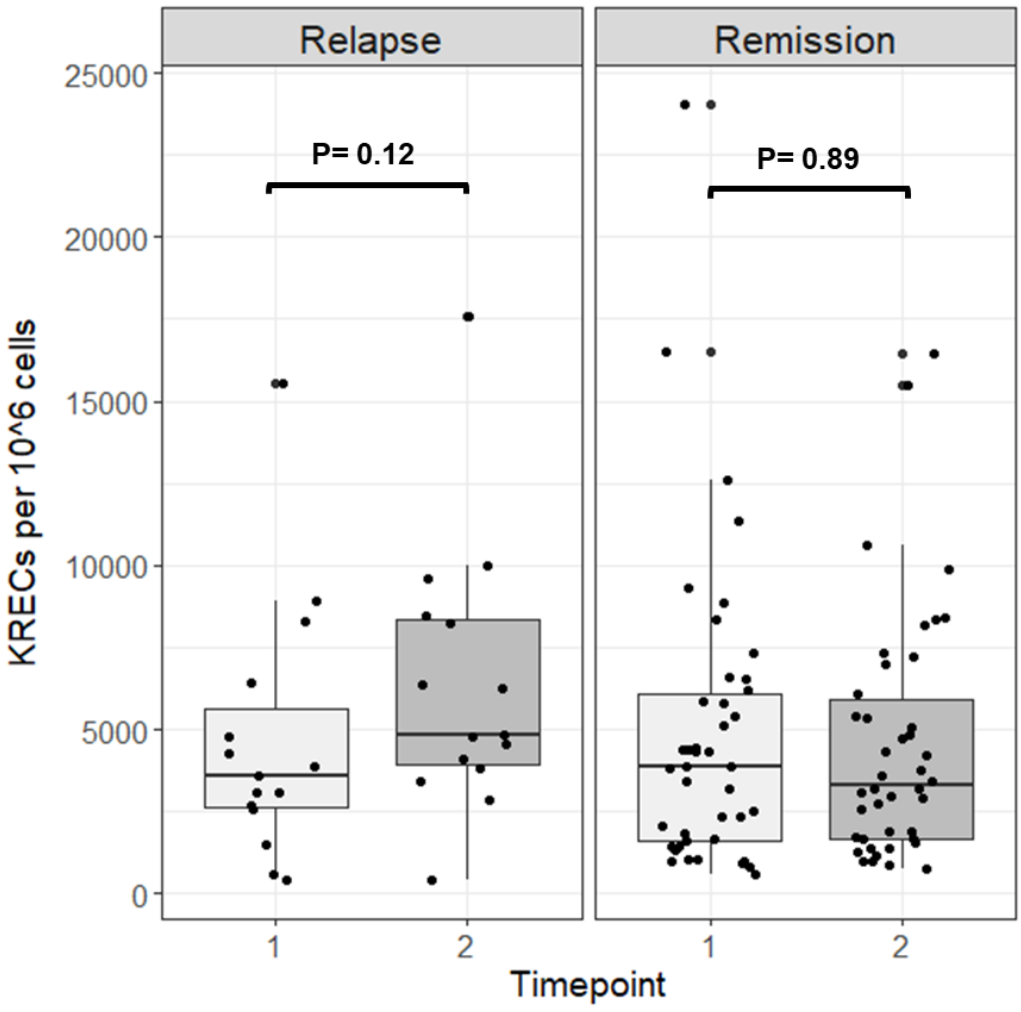


Figure 5.8 – KRECs per 10^6 cells at ATD withdrawal (Timepoint 1) and 6-10 weeks later (Timepoint 2), divided into relapse and remission groups.

Longitudinal data examining changing KREC levels following the diagnosis and treatment of GD was also investigated for 10 of the SPRING GD patients, where KREC levels were available from both diagnosis (sample storage from previous study) and at ATD withdrawal when they were later recruited into the SPRING study. The clinical and demographic characteristics of these 10 patients is presented in Table 5.6. Paired analysis showed a significant decline in KRECs per 10^6 cells from diagnosis to the timepoint of stopping ATD (median 8119 vs. 3628, P= 0.0019) (Figure 5.9).

Demographic / Clinical variable	Value
Total number of paired samples	10
Age (years): median (IQR) [range]	46.5 (33 – 49.5) [27 – 66]
Female: n (%)	8 (80%)
Ethnicity: n (%)	
White	10 (100%)
Black African	0 (0%)
Asian	0 (0%)
Smoking history: n (%)	
Current smoker	3 (30%)
Previous smoker	4 (40%)
Never smoked	3 (30%)
Thyroid function at diagnosis: median (IQR) [range]	
TSH (miU/L)	<0.05 [<0.01 – <0.05]
FT4 (pmol/L)	56.6 (40.3 – 71.2) [25 – 76]
FT3 (pmol/L)	24.9 (16.7 – 30.4) [9 – 35]
TRAb (IU/L)	6.9 (5.3 – 8.7) [2 – 26]
Duration of ATD (months): median (IQR) [range]	16 (14 – 16) [12 – 17]
Thyroid function when stopping ATD: median (IQR) [range]	
TSH (miU/L)	1.7 (0.6 – 2.8) [0.1 – 10]
T4 (pmol/L)	15.2 (14.4 – 17) [14 – 19]
T3 (pmol/L)	4.5 (4.2 – 5.1) [4 – 6]
TRAb (IU/L)	< 1 (0 – 1.1) [<1 – 2.8]
KRECs per 10^6 cells at diagnosis: median (IQR) [range]	8119 (4733 – 22120) [1832 – 28597]
KRECs per 10^6 cells at stopping ATD: median (IQR) [range]	3628 (1454 – 4406) [20 – 15540]

Table 5.6 – Demographic and clinical characteristics of the ten SPRING patients with available KREC levels at diagnosis and on stopping ATD. At 6-10 weeks after ATD withdrawal, there was an increase in KREC levels observed in the two patients in this cohort that relapsed their GD within 12 months after ATD withdrawal (Figure 5.9).

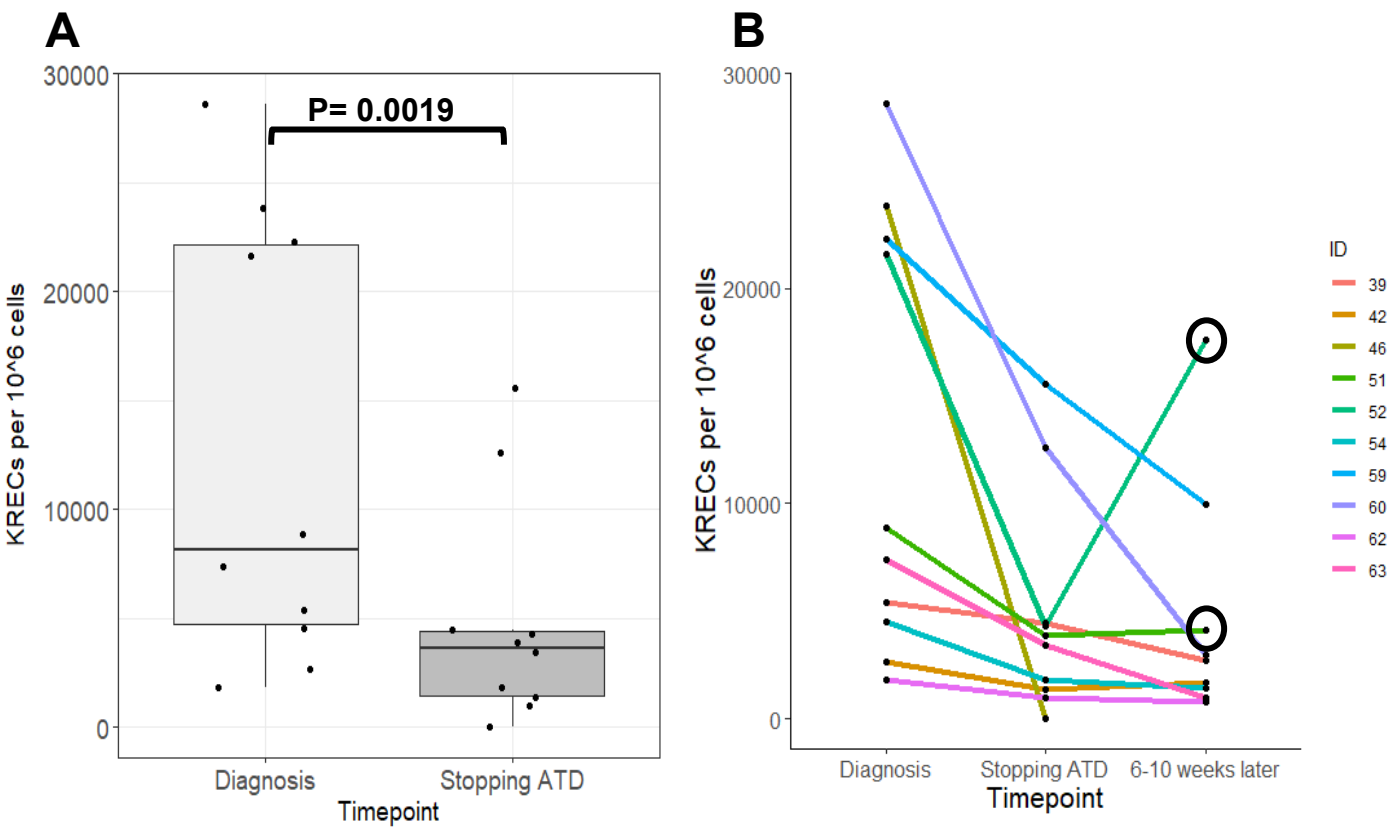


Figure 5.9 – Average number of KRECs per 10^6 cells at the timepoint of Graves' disease diagnosis and stopping ATD in ten Graves' disease patients (**A**) and individual longitudinal change in KRECs per 10^6 cells between three timepoints; diagnosis, stopping ATD and 6-10 weeks later (**B**). **5.9 Peripheral B cell replicative history**

As described in Chapter 3: Methods, measuring KRECs allows us to determine the total number of average cell divisions that B cells have undergone since their emigration from the bone marrow, and this reflects B cell proliferation in the peripheral compartment.

To determine whether there was any association between B cell proliferation and outcome in GD, the average number of B cell divisions was compared between the relapse and remission SPRING GD patient groups at baseline and 6-10 weeks, but this did not reveal any significant difference (P=0.16, P=0.84, respectively).

Unexpectedly, when healthy controls were compared to GD patients there was a significantly increased number of B cell divisions observed in healthy controls (median 1.5 vs. 0.6; P= $<2 \times 10^{-16}$) (Figure 5.10A). Pairwise comparisons using

Wilcoxon rank sum test with multiple test correction (false discovery rate: Benjamini & Hochberg (*Benjamini et al. 1995*)) revealed a significant difference in average number of B cell divisions between controls vs. euthyroid (median 1.45 vs. 0.74; P= $<2 \times 10^{-16}$) and controls vs. hyperthyroid patients (median 1.45 vs. 0.039; $<2 \times 10^{-16}$) (Figure 5.10B). There was also a significant difference observed between hyperthyroid vs. euthyroid GD patients (median 0.74 vs. 0.039; P= 6.2×10^{-7}) (Figure 5.10B). Furthermore, a Kruskal-Wallis test revealed that overall there was a significant difference between the 3 groups (controls, euthyroid, hyperthyroid; P= $<2.2 \times 10^{-16}$). These findings suggest that healthy controls have the highest peripheral B cell proliferative activity, followed by euthyroid and then hyperthyroid GD patients.

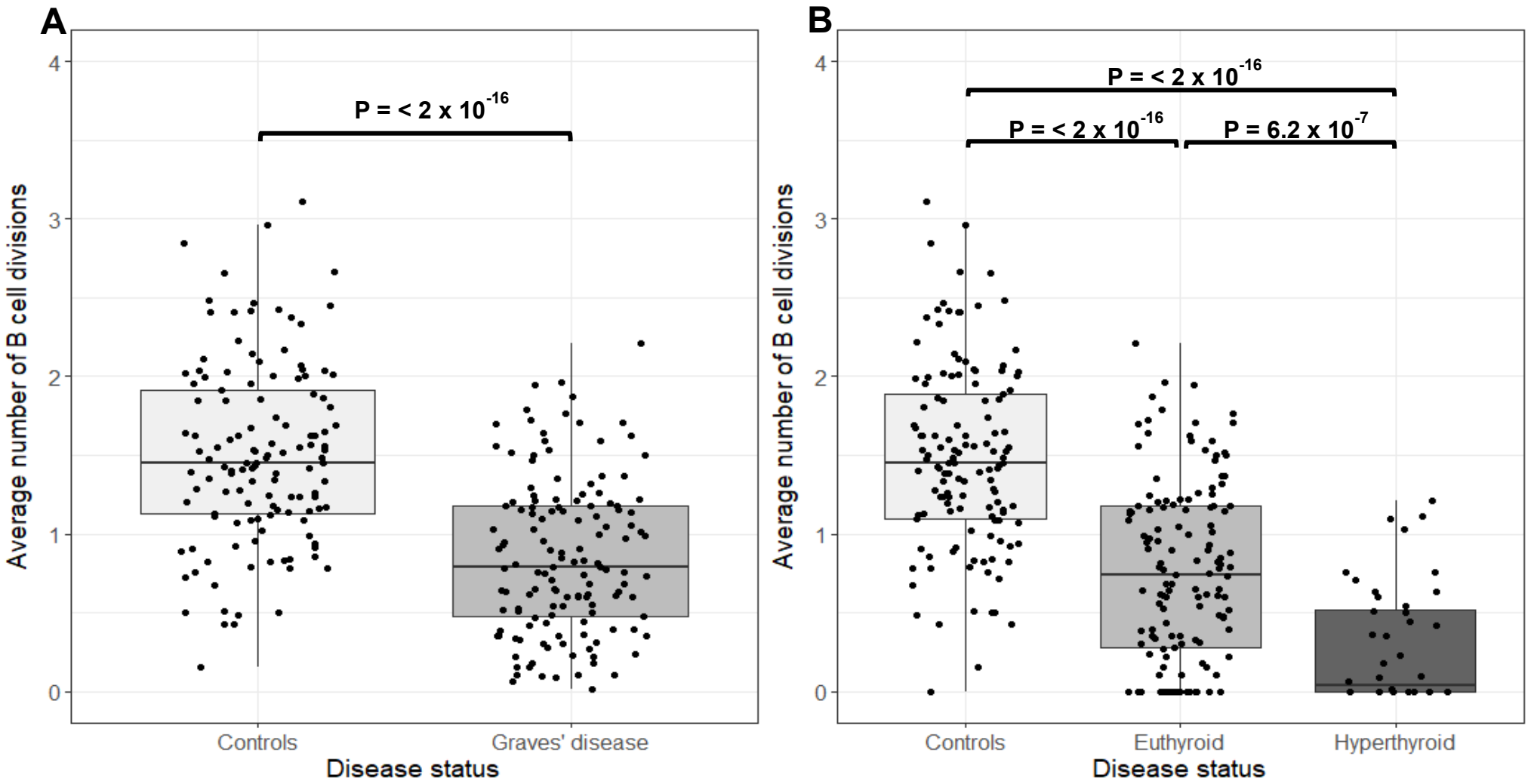


Figure 5.10 – Association of average number of B cell divisions in controls vs. Graves' disease (**A**) and controls vs. euthyroid vs. hyperthyroid Graves' disease patients (**B**)

5.10 Coding joint (B cell quantity) in Graves' disease vs. controls

The genomic coding joints (CJ) are found in all B cells and therefore represent number of circulating B cells. The coding joint cycle threshold (Ct) values between GD patients and controls were statistically different with the GD patients having lower average Ct values (lower Ct values suggests a greater number of circulating B cells) (P=0.02) (Figure 5.11). There was no difference in CJ Ct values between the controls and hyperthyroid or euthyroid GD patients (P=0.06).

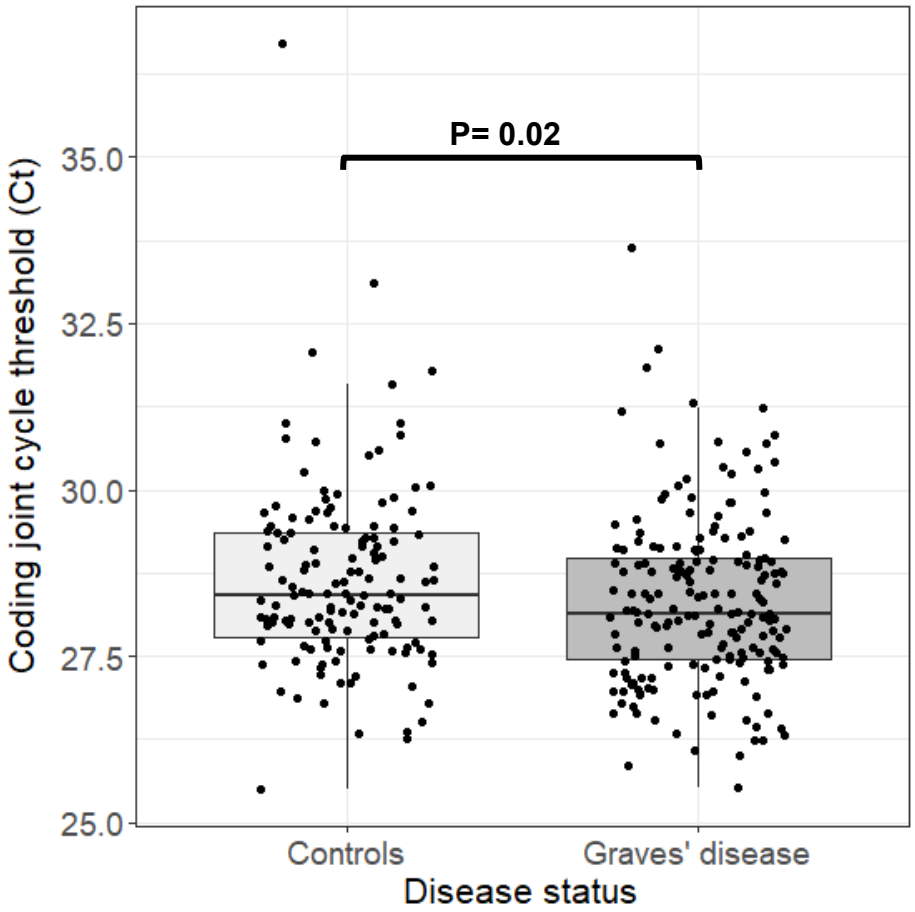


Figure 5.11 – Association of coding joint cycle threshold (Ct) values in controls vs. Graves' disease patients

5.11 Discussion

Although the importance of B cells is well-established in the immunopathology of GD, the specific mechanisms driving autoimmunity and relapse remain largely unknown. Despite their obvious relevance in autoimmune disease, KRECs are most widely studied alongside TRECs in the field of primary immunodeficiencies to detect B and

T cell developmental defects (*Dasouki et al. 2020*). However, in addition to the association of KRECs with disease status and activity in the B cell-mediated autoimmune condition ITP (*Levy-Mendelovich et al. 2017*), a recent proof-of-concept study demonstrated that KRECs, when used as a measure of oligoclonal peripheral B cell expansion, could predict response to the B-cell depleting agent Rituximab in children with the autoimmune disease, juvenile dermatomyositis (JDM) (*Ochfeld et al. 2022*). Although KRECs have not previously been studied in GD, an increased concentration of TRECs has been observed in patients with GD and associated with thyroid status (*Strawa et al. 2014*).

5.11.1 KRECs and Graves' disease patient characteristics

There was no association observed between KRECs and age or gender in the GD patients or healthy controls. This is consistent with the literature in adults reporting that unlike their T cell equivalent, TRECs, KRECs generally remain stable after 18 years of age and show no association with gender (*Sottini et al. 2014*). There is some evidence that the B cell compartment can be modified by smoking and result in dysregulation of the humoral immune response (*Qiu et al. 2017*), however the lack of association observed between KRECs and smoking status of the GD patients in this study did not support any smoking-related variability in B cell activity.

The ability of KRECs to reflect humoral immune activity indicates there is the potential for dynamic changes in KRECs to occur due to external factors that influence the activity of B cells, such as infection or vaccination. However, there is limited literature on the magnitude, or not, of change in KRECs during intercurrent illness or after vaccination. Recent data examining KREC levels in patients with acute COVID-19 infection revealed median KREC levels of 7050 per 10^6 cells (*Khadzhieva et al. 2021*), which is higher than the levels of both the GD patients (4952 per 10^6 cells) and healthy controls (2571 per 10^6 cells) observed in this study. Nevertheless, as indicated by the lack of association between KRECs and circulating CRP observed in this study, there appeared to be no relationship between KRECs and detectable intercurrent inflammation/infection of the SPRING study patients. Furthermore, none of the SPRING patients received a vaccination within 4 weeks of their samples being taken so vaccination is unlikely to have contributed to any changes observed in humoral activity in these patients.

5.11.2 KRECs in Graves' disease vs. healthy controls

In this study, elevated KRECs were observed in GD patients compared to healthy controls, suggesting an association with GD and increased B cell output from the bone marrow. Indeed, the lower average CJ Ct value observed in the GD patients compared to controls is consistent with this finding, as a lower CJ Ct value indicates increased circulating B cells. Elevated circulating B cells have previously been described in both untreated and treated GD patients compared to healthy controls (*Iwatani et al. 1989, Bossowski et al. 2003*). Iwatani et al. demonstrated that the highest quantity of circulating B cells was observed in individuals that were hyperthyroid, however even those that had received ATD and were euthyroid continued to have higher circulating B cells compared to controls, suggesting that thyroid status is not the sole factor in determining B cell output.

5.11.3 KRECs and thyroid status in Graves' disease

My study demonstrated that thyroid status was associated with KREC levels, as those with hyperthyroid GD had elevated KRECs compared to euthyroid GD patients. The association with thyroid status is also reflected in the notable decline in KREC levels observed in the untreated GD patients from diagnosis to ATD withdrawal, likely reflecting the fact that these patients are being rendered biochemically euthyroid by the ATD, and therefore KRECs are declining as the GD becomes increasingly inactive. Indeed, both FT3 and FT4 thyroid hormones and TRAb and TPO autoantibodies were observed to be positively associated with KREC levels.

There is growing evidence of 'bidirectional crosstalk' between thyroid hormones and the immune system, with thyroid hormones proposed as direct regulators of the adaptive immune response and essential for primary B cell development and activation *(Torimoto et al. 2022, Wenzek et al. 2022, Liu et al. 2022, Jaeger et al. 2021).* Specifically, FT4 concentrations have been demonstrated to be positively associated with circulating B lymphocytes and induce the proliferation of B cell subsets, including the naïve B cells (*Jaeger et al. 2021*). Furthermore, immunophenotyping studies investigating the impact of thyroid hormones on immune homeostasis have demonstrated the enrichment of thyroid hormone sensitive gene expression in B-cell functional pathways, highlighting the role of thyroid hormones in regulating B lymphocyte function (*Jaeger et al. 2021*). The increased B cell output, as quantified by KRECs, observed in this study of GD patients may reflect bone marrow

hyperactivity that could be partially secondary to hyperthyroidism. Indeed, thyroid hormones are reported to be positively correlated to transitional and naïve B cells *(Van der Weerd et al. 2013),* and hyperthyroidism has been associated with the presence of hypercellular bone marrow, lymphocytosis, and splenomegaly, with a direct effect demonstrated on bone marrow activity (*Axelrod et al. 1951, Foster et al. 1999, Bloise et al. 2014*).

However, the euthyroid GD patients still had significantly higher KREC levels compared to controls, suggesting that regardless of thyroid hormone status, those with GD are more likely to have elevated circulating KREC levels. This indicates that KREC levels are not only influenced by thyroid status, and they may highlight additional factors that have a role in disease pathogenesis. These findings provide insight into the influence of thyroid status on B cell activity and may facilitate our understanding of the impact of thyroid hormones on clinical outcome in GD.

5.11.4 KRECs and Graves' disease outcome following ATD treatment

It is well-established that higher thyroid hormone and TRAb concentrations at the time of ATD withdrawal are associated with a greater risk of relapse in GD (*Sjolin et al. 2019*). Indeed, this study demonstrated that KREC levels at 6-10 weeks were not only associated with relapse by 12 months but were also positively correlated to FT3 concentration. This indicates that elevated KRECs are, at least partially, reflecting the rising FT3 levels seen in patients that are beginning to relapse their GD. Furthermore, individuals that had falling KREC levels 6-10 weeks after ATD withdrawal were more likely to remit their GD, highlighting the potential association between KRECs and GD activity. However, the multivariate regression model adjusted for other factors associated with relapsing GD did not reveal KRECs as an independent biomarker of outcome in GD.

5.11.5 Peripheral B cell replicative history

Unexpectedly, this study demonstrated that those with hyperthyroid GD had the lowest level of peripheral B cell proliferation, followed by the euthyroid GD group, and then the healthy control group which had the highest B cell proliferative activity. The individual B cell subpopulations have differing levels of replicative activity, with immature B cells demonstrating substantially lower proliferative activity compared to their mature counterparts that have exited the highly replicative germinal centres (*van*

Zelm MC et al. 2007, Tangye SG et al. 2003). Specifically, the CD5+ naïve B lymphocytes have been demonstrated not to undergo proliferation in the peripheral B cell compartment (*van Zelm MC et al. 2007).* This could therefore explain why individuals with active hyperthyroid or inactive euthyroid GD, whom may have an expanded immature B cell population relative to healthy controls, might demonstrate lower levels of peripheral B cell replicative activity. It may also be that those with GD that have a greater B cell output and potentially more circulating autoreactive B cells, have lower levels of peripheral proliferation in order to maintain B cell homeostasis *(Lee et al. 2020*).

An overall summary of the findings of this study including the relationship between GD and thyroid status with KREC levels and B cell output/proliferation is presented in Figure 5.12.

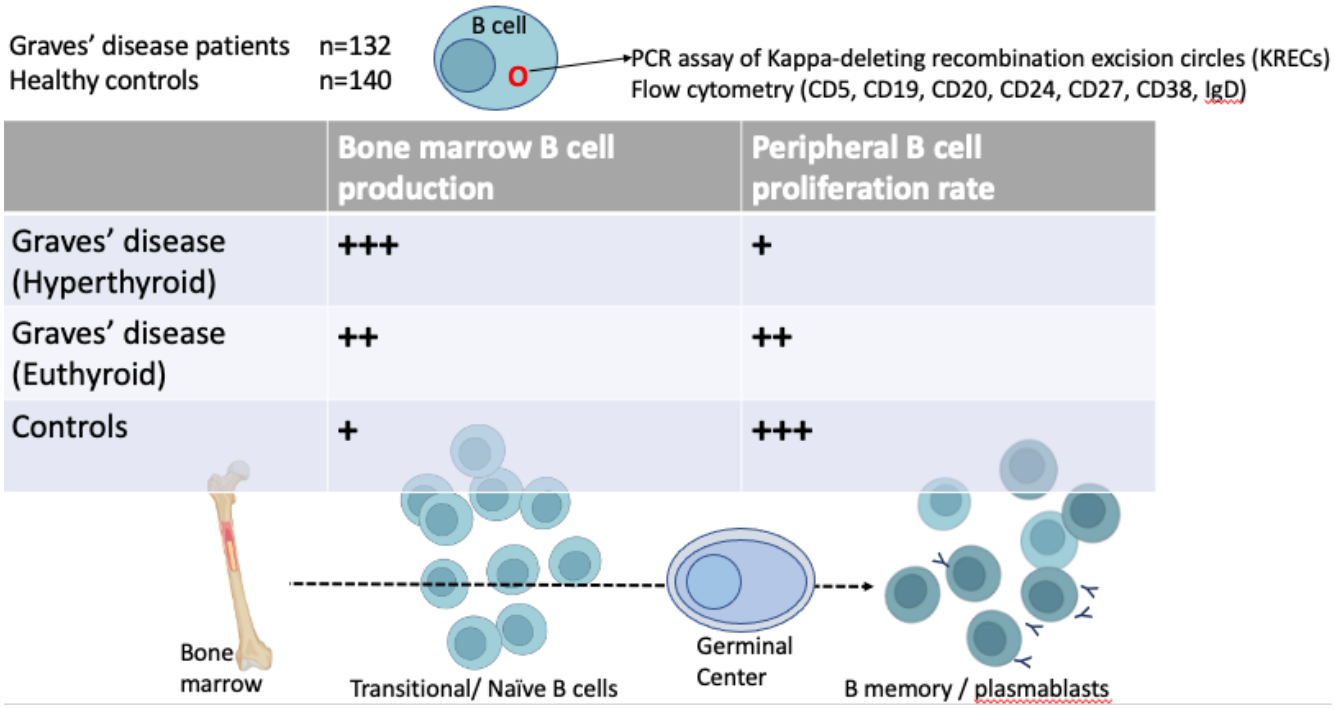


Figure 5.12 – Summary of the findings of the KREC study examining the expansion of the immature B lymphocyte compartment in Graves' disease (*Lane LC et al. 2023*).

Further analysis to explore these observed differences is detailed in Chapter 9: Integrative Analysis, which analyses the KRECs in conjunction with the flow cytometry data of B cell subpopulations.

5.12 Limitations

As this study was performed in a relatively small cohort of GD patients, the findings should be validated on a larger scale. Wider analysis of B cell subpopulations at different stages of disease activity would help to establish the role of immature B cell subsets in relation to GD activity. To further explore the underlying mechanism of KRECs and peripheral B cell replicative activity in GD, analysis of KRECs on sorted B cell subpopulations may provide greater insight into the functional roles of specific B cell subsets in the immunopathology of GD.

5.13 Summary

This Results Chapter presents a robust association between KRECs and GD, with the findings highlighting the importance of B cells in the pathogenesis of GD and the influence of thyroid status on B cell activity. The association between KRECs and thyroid status indicate a potential role for KRECs as a marker of disease activity and outcome in GD, although the true value of KRECs may lie in their ability to provide insight into the humoral immune activity of GD, rather than as a predictive biomarker of relapse.

Chapter 6: Results 3 – Cytokine and Chemokine analysis

6.1 Introduction

Circulating cytokines and chemokines have a key role in modulating B cell activity and although they are described in the pathogenesis of GD, their potential role in disease relapse is currently unknown. They can be easily measured by venepuncture and following a simple centrifugation step to produce serum, patient samples can be stored frozen for periods of time, providing an appealing potential biomarker that could translate to clinical practice and inform patient management.

In this Chapter I present the analysis of the cytokine and chemokine data from samples taken at baseline (on stopping ATD medication) and again 6-10 weeks later. The aim of this work was to measure the B cell cytokines/chemokines at these timepoints to see if they were able to differentiate between those patients who relapsed or remitted one year after stopping ATD treatment. Furthermore, this analysis was used in conjunction with the KRECs and FACs data to better understand the contribution of these factors to the underlying humoral immune activity in GD (Chapter 9: Integrative analysis).

6.2 Quality control

Commercially available solid phase ELISA kits were used for the detection of five cytokines and one chemokine (Chapter 3: Methods 3.5.10). These validated kits are based on the two-site sandwich principle in which two highly specific antibodies are used to detect the target analyte. Each cytokine/chemokine measurement was performed in triplicate, with negative controls on every plate. 10% of samples were repeated.

6.2.1 Sample collection and processing

The samples used in this experiment were from the first blood sample collected at the timepoint of stopping ATD therapy and the second blood sample collected 6-10 weeks later. The fresh sample underwent centrifugation, and the separated serum was aliquoted and stored at $^{-}80^{O}C$ for future batch analysis.

Serum was available for all baseline samples, however as discussed previously, 7/65 (11%) of patients did not have the second 6–10-week blood sample. Cytokines and chemokines can degrade if left to stand at room temperature for prolonged periods of time. Therefore, the time from venepuncture to centrifugation was recorded for all the baseline and 41/65 (63%) of the 6–10-week serum samples collected, with a median value of 60 (IQR: 45 – 90, range 30 – 180) minutes. There was no significant difference between the relapse and remission groups and time to serum centrifugation (P=0.52) (Figure 6.1).

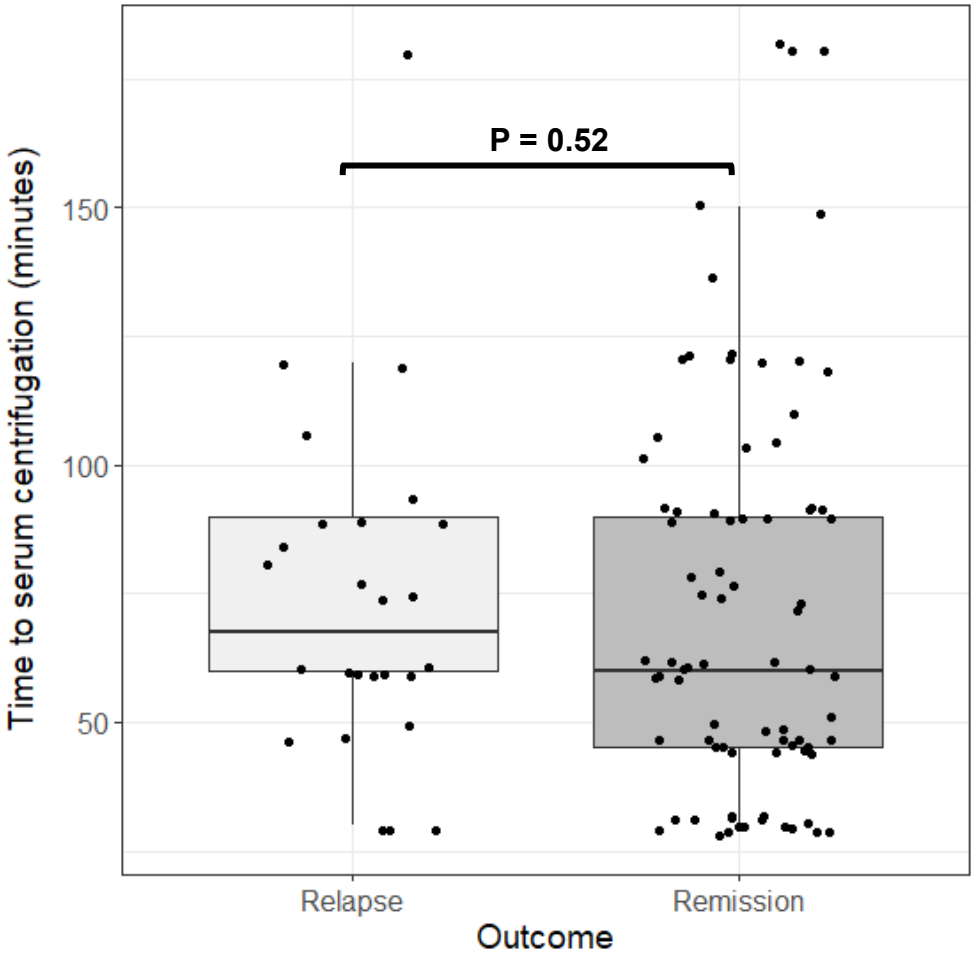


Figure 6.1 – Time to commencement of serum centrifugation from the time of sample collection in both outcome groups.

6.2.2 Coefficient of variation of ELISA kits

The reproducibility of results obtained within and between ELISA plates is important to ensure that the results generated throughout an experiment are accurate and enable the comparison of results between samples.

This reproducibility is measured as a coefficient of variation (CV) from the mean value. Two types of precision were considered in this experiment: the intra-assay CV and the inter-assay CV. Intra-assay CV describes the reproducibility between wells within an assay, whereas the inter-assay CV is the reproducibility between assays.

The manufacturer stated average intra- and inter-assay CVs of the ELISA kits are presented in Table 6.1 (where available), alongside the CVs generated in this experiment for each cytokine and chemokine.

Cytokine/chemokine	**Manufacturer**		**This experiment**	
	Mean Intra-assay CV (%)	*Mean Inter-assay CV (%)*	*Mean Intra-assay CV (%)*	*Mean Inter-assay CV (%)*
BAFF	5.7	10.5	4.8	11.7
APRIL	8.1	7.1	8	11.6
CXCL13	3.4	9.2	4.8	5.2
sTACI	NA	NA	5.7	9.5
sBCMA	NA	NA	5.5	11.7
IL-6	4.1	6.5	6.2	14

Table 6.1 – The manufacturer ELISA kits stated mean coefficient of variations (CV) compared to those generated in this experiment.

The CVs generated from this experiment meet the acceptable limits for ELISA intra- and inter-assay CVs, which are <10% and <20%, respectively (*Reed et al. 2002, Tighe et al. 2015*).

6.2.3 Limits of detection of ELISA kits

The assay range including the upper (ULOD) and lower limits of detection (LLOD) for each ELISA kit is presented in Chapter 3: Methods 3.5.10. In total, there were 1474 measurements of cytokines/chemokines recorded (including the duplicate results and excluding negative controls and repeats). Out of these measurements, four sample concentrations (all APRIL) exceeded the ULOD, whereas forty samples (9/40 APRIL, 31/40 sTACI) had a concentration below the LLOD.

When the sample was above the ULOD, the sample was re-diluted to a greater degree to achieve a value that was on the standard curve. When the sample concentration was below the LLOD the sample was assigned the concentration of the

stated LLOD for the kit. All the BAFF, CXCL13, IL-6, sBCMA and 92.7% of APRIL and 75% of sTACI had sample concentrations that were above the LLOD (Figure 6.2).

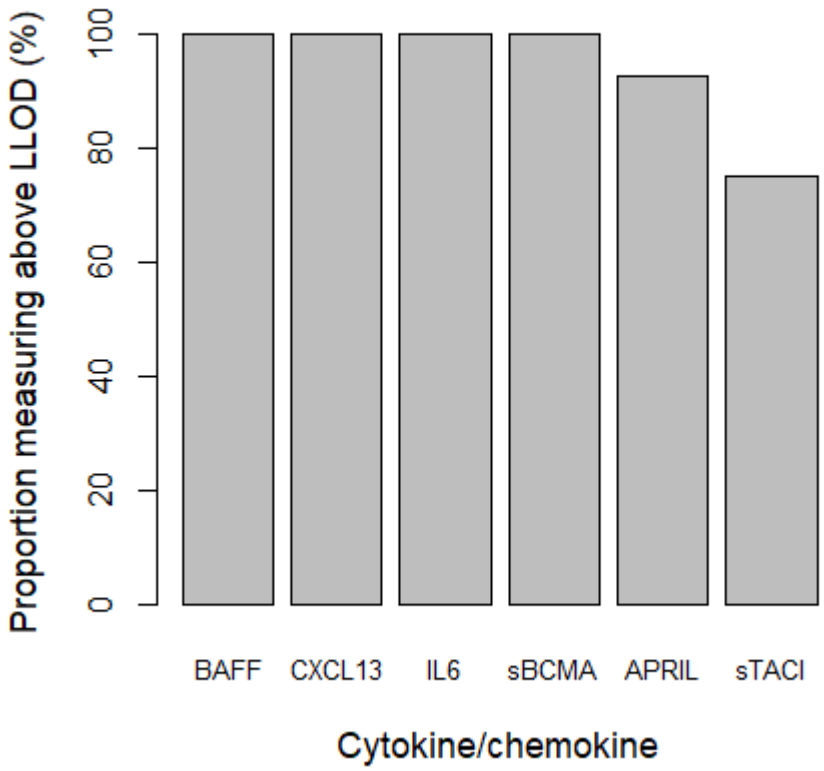


Figure 6.2 – The proportion of cytokine and chemokine samples which measured above the LLOD.

6.2.4 Logarithmic data transformation

The cytokine and chemokine values generally demonstrated a positively skewed dataset, and therefore the data was standardised by natural log transformation. An example of this logarithmic transformation is presented in Figure 6.3.

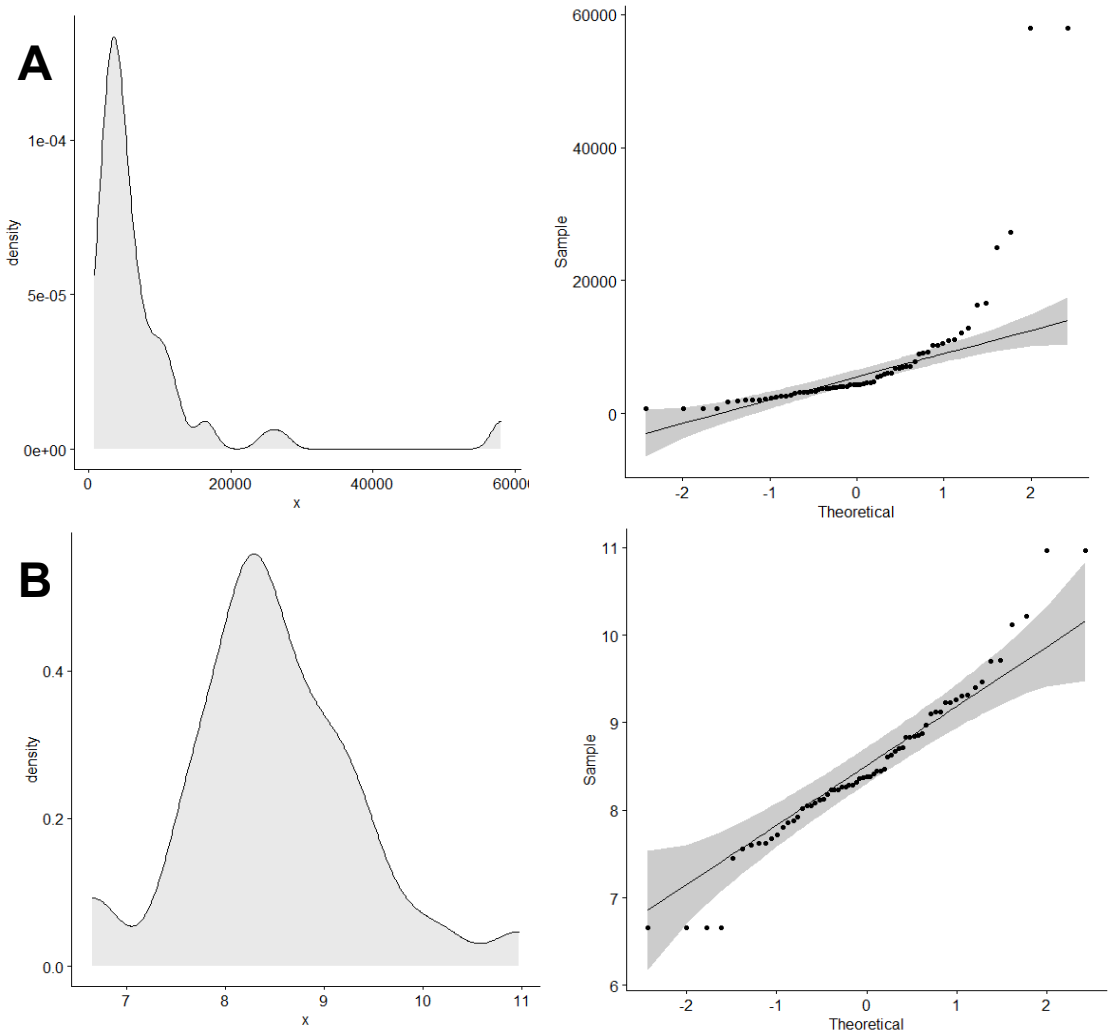


Figure 6.3 – Density plots and quantile-quantile (Q-Q) plots of baseline APRIL concentration before (**A**) and after (**B**) natural logarithmic data transformation

6.3 Cytokines/chemokines and clinical characteristics

Analysis was performed to investigate for any relationships between the cytokines and age, gender, or smoking status. None of the cytokines or chemokines measured were associated with gender at any timepoint. However, although none of the cytokines at baseline were associated with age, APRIL and sTACI concentrations at 6-10 weeks, were positively and negatively correlated with age (P=0.02, rs=0.32; P=0.04, rs= -0.27), respectively (Figure 6.4).

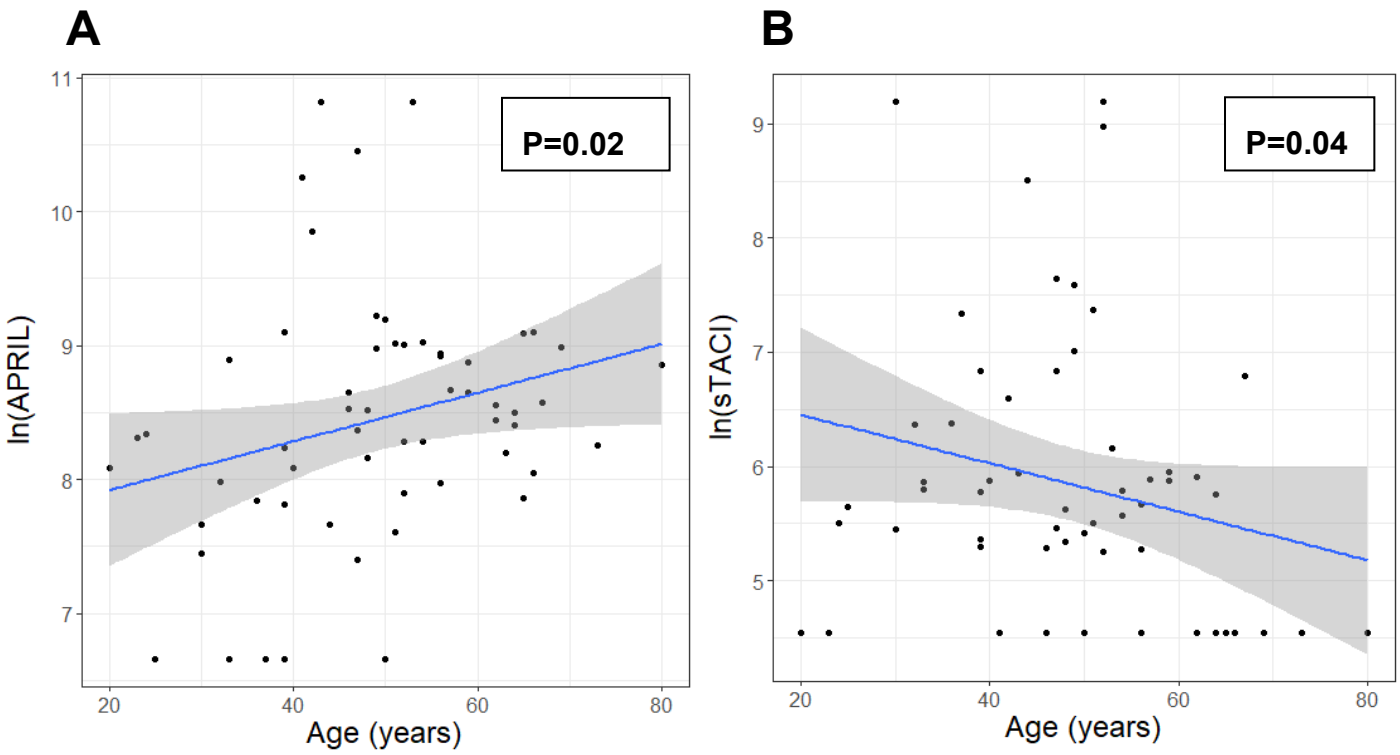


Figure 6.4 – Correlations between log transformed (ln) APRIL (**A**) and sTACI (**B**) with age 6-10 weeks after stopping ATD.

The relationship between serum cytokine concentrations and smoking status was explored, with a significant increase in serum APRIL concentrations found in smokers compared to non-smokers (P=0.002) (Figure 6.5). None of the other cytokines or chemokine were associated with smoking status.

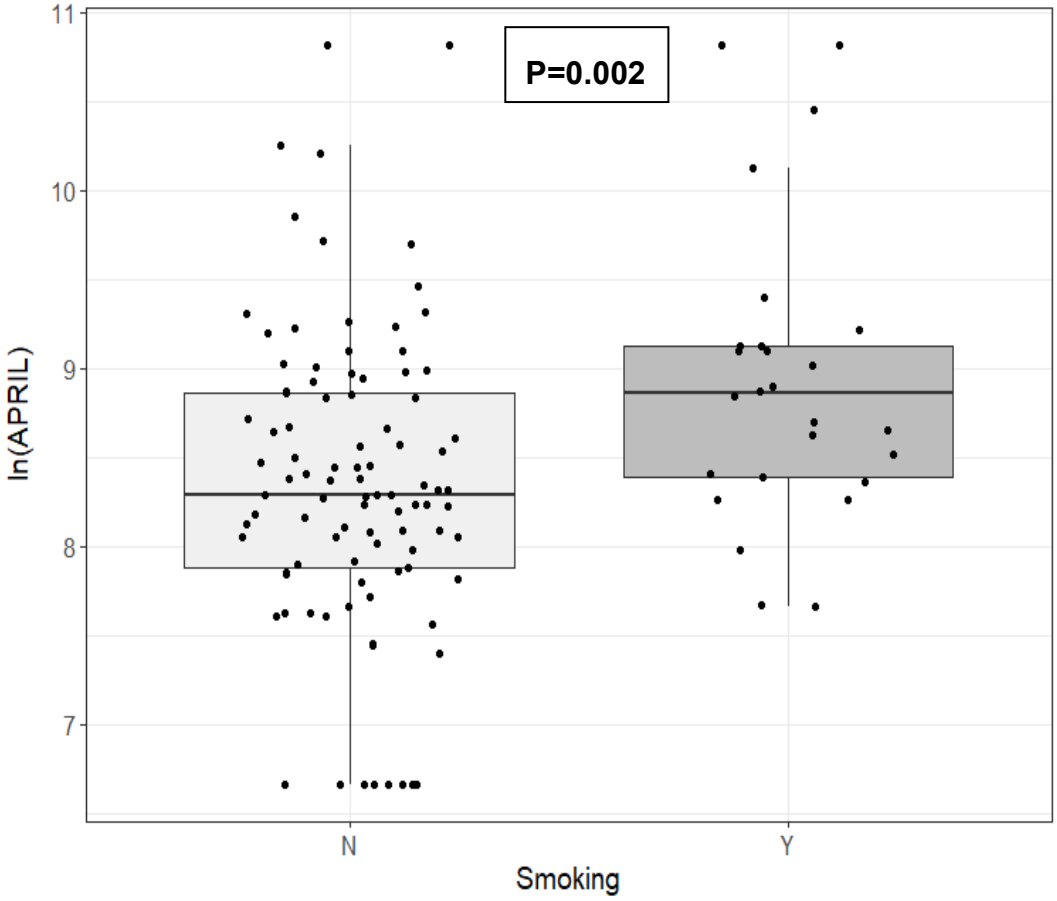


Figure 6.5 – Smoking status and log transformed (ln) APRIL concentrations

6.4 Association of cytokines/chemokines with thyroid hormones

The association between circulating thyroid hormones and cytokine/chemokine concentration was studied. There was a positive correlation between BAFF concentrations and FT4 at baseline (P=0.008, rs=0.32) and FT3 (P=0.014, rs=0.32) at 6-10 weeks after stopping ATD (Figure 6.6), however BAFF wasn't observed to be correlated with FT3 at baseline (0.42) nor FT4 at 6-10 weeks (0.069). When the data from baseline and 6-10 weeks later was analysed together, BAFF was positively correlated with both FT4 (P=0.018, rs=0.28) and FT3 (P=0.019, rs=0.21). There was no correlation between TRAb (P=0.16, P=0.97) or TPOAb (P=0.46, P=0.82) and BAFF at either timepoint.

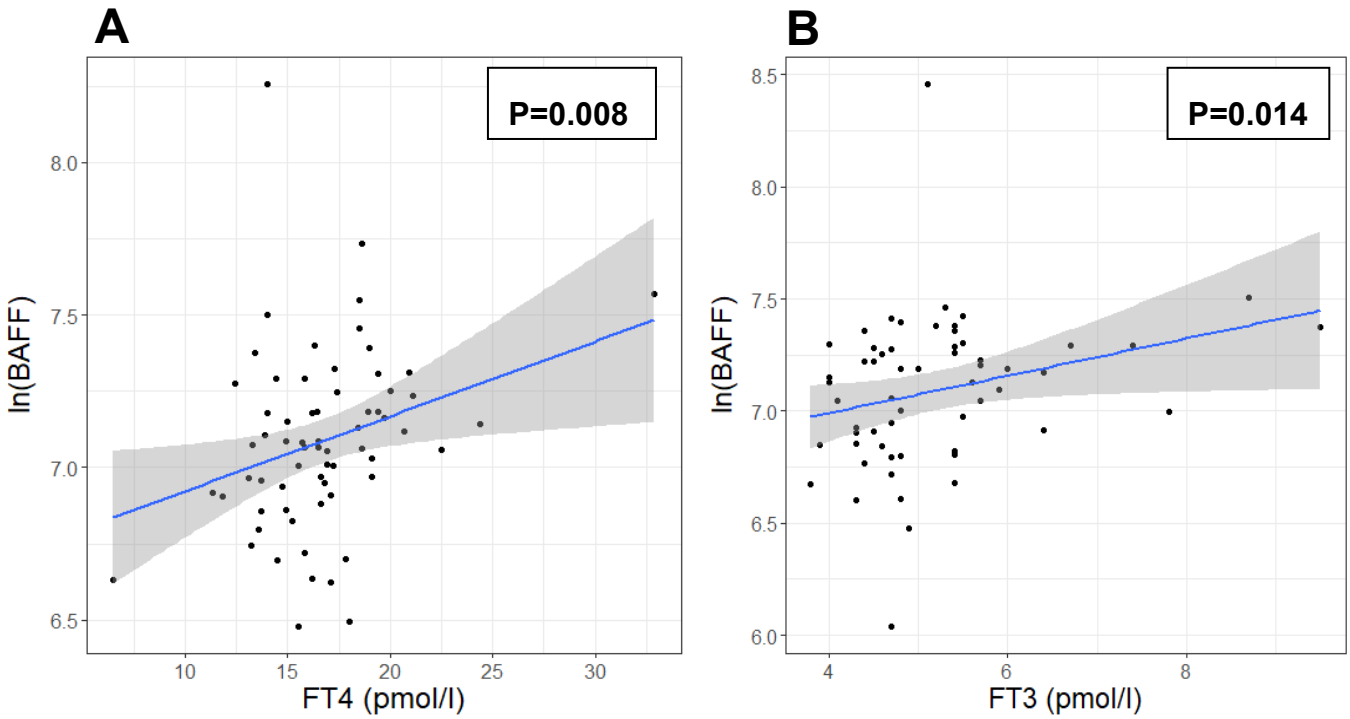
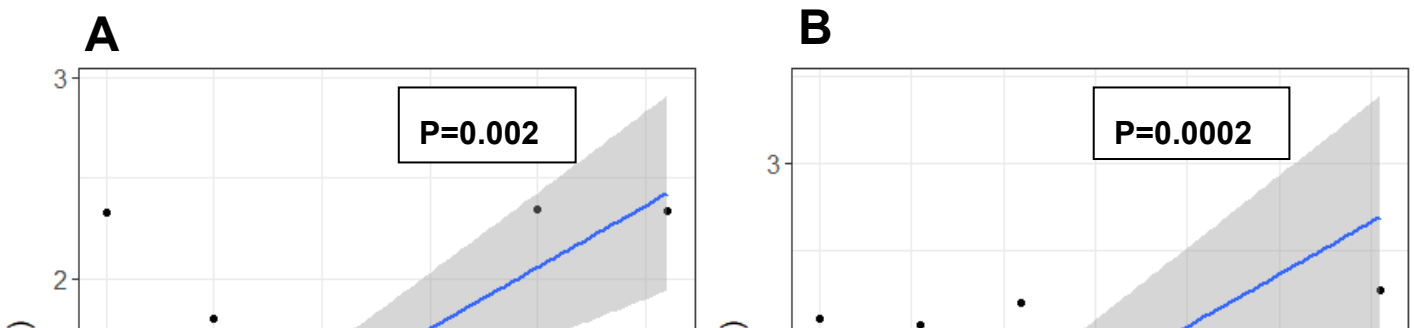


Figure 6.6 – Correlations between log transformed (ln) BAFF concentration and FT4 at baseline (**A**) and FT3 at 6-10 weeks (**B**). **6.5 Association of cytokines/chemokines with CRP**

Given the dynamic nature of B cell activity with potential intercurrent illness/infection, a CRP was taken from the patients at both timepoints to consider any underlying infective illness that may have been modulating the B cell activity and affecting cytokine production independently of autoimmune activity. Most patients had a negative CRP (<1 mg/L) at both baseline (54/65; 83%) and 6-10 weeks later (46/58; 79%). The IL-6 cytokine was found to be positively associated with CRP at both timepoints (P=0.002, rs=0.37; P=0.0002, rs=0.47), including after those with negative CRP values had been excluded (P=0.02, rs=0.68, P=0.03, rs=0.60) (Figure 6.7). None of the other cytokines or chemokine were associated with CRP concentration.



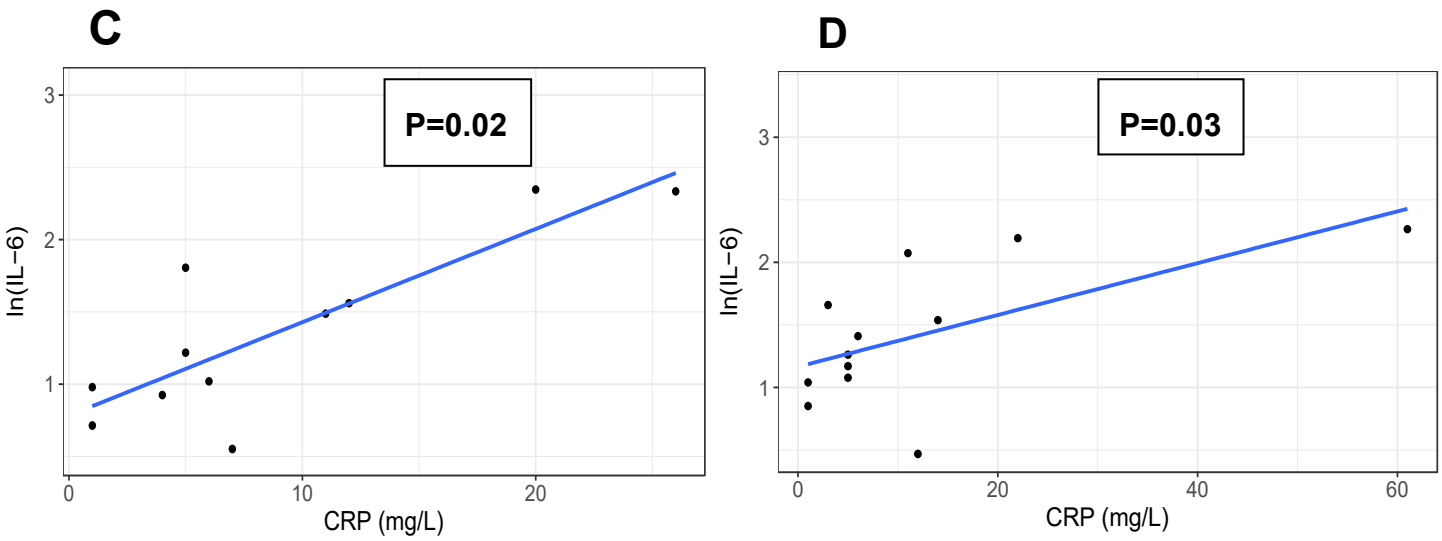


Figure 6.7 – Correlations between log transformed (ln) IL-6 concentration and CRP at baseline (**A**) and at 6-10 weeks (**B**), and including only those with a detectable CRP at baseline (**C**) and 6-10 weeks (**D**). **6.6 Association between the cytokines/chemokines**

Analysis was undertaken to determine the relationships between the circulating cytokines and chemokine concentrations at baseline and 6-10 weeks (Table 6.2).

Circulating BAFF and CXCL13 were positively correlated at baseline and 6-10 weeks later (P= 0.02, P= 0.04, respectively), and IL-6 was positively correlated with both BAFF and CXCL13 at both timepoints (P=0.03, P=0.01; P=0.03, P=0.047, respectively). However, it was the soluble BAFF receptors, sBCMA and sTACI, that

were observed to demonstrate the strongest significant positive correlation with each other at both timepoints (P= $6.3x10^{-5}$, P= $2.1x10^{-5}$, respectively) (Figure 6.8).

Log transformed B cell cytokine/chemokine			BAFF	CXCL13	sBCMA	APRIL	sTACI	IL-6
BASELINE	ln(BAFF)	rs P	.	0.28 **0.02***	-0.29 **0.02***	0.22 0.08	-0.12 0.31	0.27 **0.03***
	ln(CXCL13)	rs P		.	0.18 0.13	-0.06 0.62	-0.07 0.60	0.31 **0.01***
	ln(sBCMA)	rs P			.	-0.10 0.38	0.47 **$6.3x10^{-5}$***	0.03 0.80
	ln(APRIL)	rs P				.	0.008 0.94	0.007 0.95
	ln(sTACI)	rs P					.	0.06 0.62
	ln(IL-6)	rs P						.
6-10 WEEKS	ln(BAFF)	rs P	.	0.26 **0.04***	-0.15 0.24	-0.20 0.11	-0.31 **0.02***	0.29 **0.03***
	ln(CXCL13)	rs P		.	0.22 0.09	0.04 0.76	0.02 0.87	0.26 **0.047***
	ln(sBCMA)	rs P			.	0.08 0.52	0.53 **$2.1x10^{-5}$***	0.50 -0.09
	ln(APRIL)	rs P				.	-0.002 0.98	0.09 0.50
	ln(sTACI)	rs P					.	0.01 0.91
	ln(IL-6)	s P						.

Table 6.2 – Spearman's correlation of each log transformed (ln) B cell cytokine/chemokine.

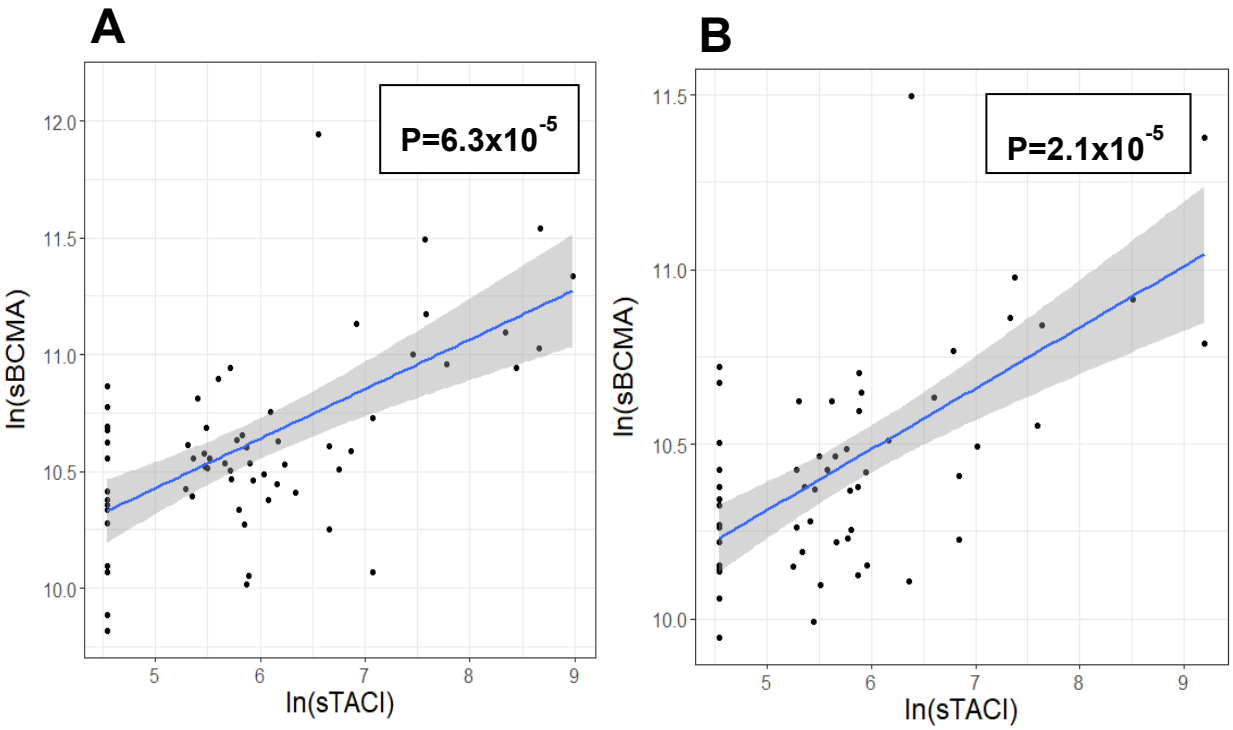


Figure 6.8 – Correlations between log transformed (ln) sBCMA and sTACI concentrations at baseline (**A**) and 6-10 weeks (**B**). **6.7 Comparison of the cytokine/chemokine concentrations between the relapse and remission groups**

The cytokine and chemokine concentrations were compared between the relapse and remission outcome groups, at the baseline timepoint of stopping ATD treatment (Timepoint 1) and 6-10 weeks later (Timepoint 2). Statistical analysis of differences in cytokine and chemokine concentrations between the two outcome groups was evaluated by univariate binary logistic regression analysis at both timepoints (Table 6.3).

(a)

(b)

Log transformed cytokine (Timepoint 2)	B	ORrelapse	95% CI ORrelapse	Univariate P value
BAFF	2.08	7.9	0.97 – 63.8	0.05
sTACI	0.14	1.15	0.73 – 1.8	0.54
sBCMA	0.96	2.6	0.7 – 9.4	0.14
BAFF	1.02	2.7	0.46 – 16.7	0.26
CXCL13	1.27	3.6	0.45 – 28.7	0.23
CXCL13	1.4	3.9	0.5 – 31	0.20
IL-6	0.46	1.58	0.73 – 3.6	0.24
sBCMA	1.5	4.7	0.8 – 26	0.08
APRIL	0.34	1.4	0.74 – 2.7	0.30
IL-6	0.11	1.11	0.46 – 2.67	0.80
APRIL	0.02	1.02	0.5 – 1.9	0.94

Table 6.3 – Association of cytokines and chemokines with Graves' disease relapse by univariate binary logistic regression, at **(a)** Timepoint 1: the time of ATD withdrawal **(b)** Timepoint 2: 6-10 weeks after ATD withdrawal.

On univariate analysis, it was only the sTACI concentration at the time of stopping ATD that was significantly associated with relapse in GD (P=0.02), although this did not remain significant after Benjamini-Hochberg multiple test correction (P=0.12). The distributions of the cytokines at baseline by outcome are summarised in Figure 6.9.

Cytokines at baseline with P=<0.2 on univariate analysis (sTACI, sBCMA and BAFF) were entered into separate multivariate binary logistic regression analyses to include other variables that are widely reported in the literature to be related to GD outcome and used in previous multivariate analyses in this study (age, gender, goitre size, smoking, TRAb level prior to stopping ATD) (Table 6.4).

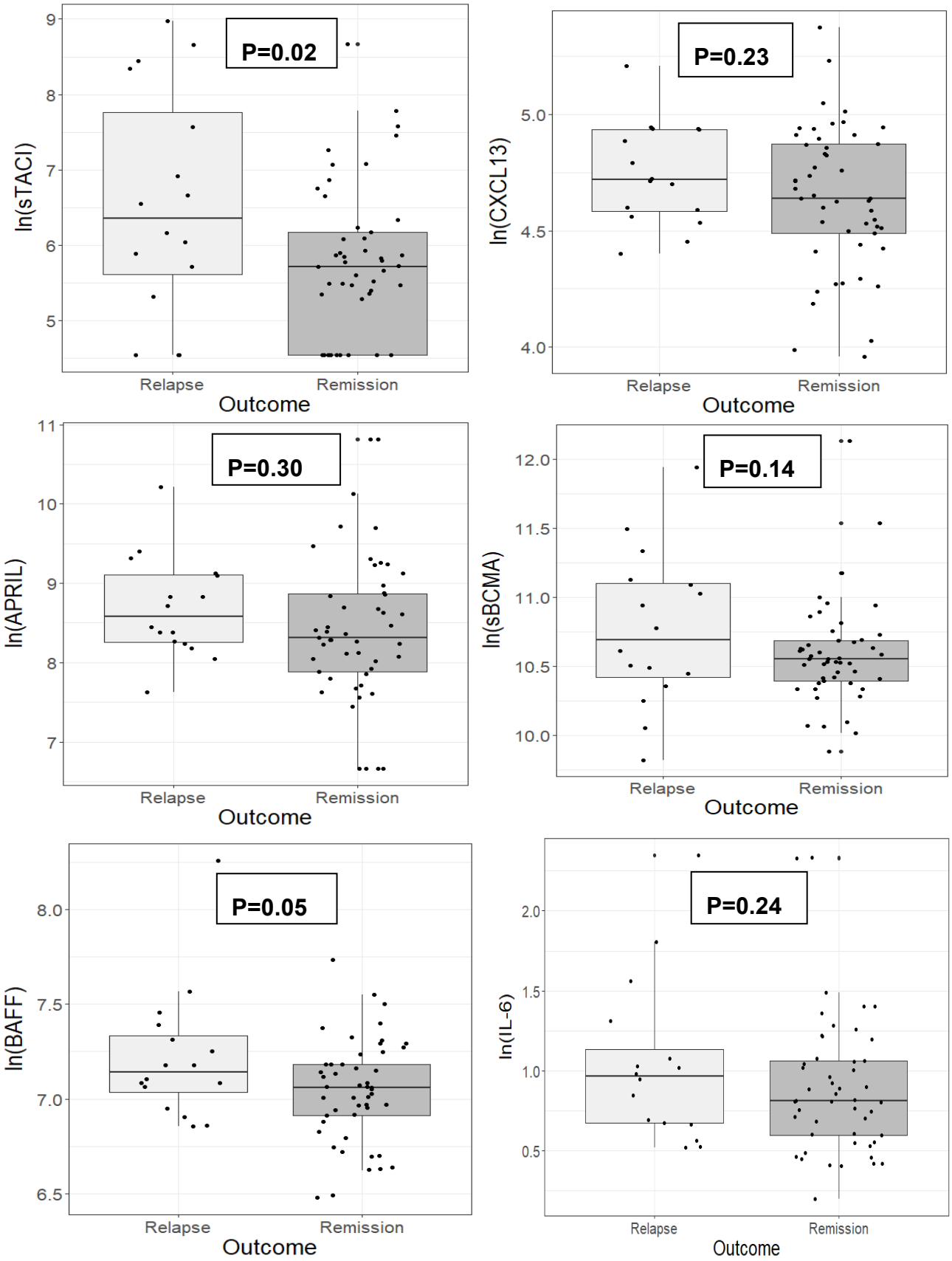


Figure 6.9 – Association of the log transformed (ln) cytokines at baseline in the univariate logistic regression analysis and their relation to outcome in Graves' disease

(a)

Baseline variable	B	ORrelapse	95% CI ORrelapse	Multivariate P value
ln(sTACI)	0.85	2.3	1.3 – 4.6	**0.007***
TRAb	1.2	3.3	1.5 – 9.9	**0.01***
Goitre size				
1	0.5	1.6	0.28 – 11.3	0.59
2	1.64	5.2	0.57 – 58.7	0.16
Gender (Male)	1.2	3.5	0.54 – 23	0.18
Age	0.007	1	0.94 – 1.09	0.83
Current smoker	-0.02	0.97	0.15 – 5.1	0.97

(b)

Baseline variable	B	ORrelapse	95% CI ORrelapse	Multivariate P value
ln(sBCMA)	1.82	6.1	1.3 – 39.8	**0.036***
TRAb	1.23	3.4	1.4 – 10.5	**0.016***
Goitre size				
1	0.97	2.6	0.45 – 18	0.28
2	2.11	8.3	0.88 – 108	0.07
Gender (Male)	1.5	4.7	0.74 – 33.3	0.10
Age	0.001	1	0.94 – 1.08	0.78
Current smoker	0.03	1.03	0.18 – 4.8	0.97

Table 6.4 – Association of baseline cytokines with occurrence of Graves' disease relapse following ATD withdrawal, using two multivariate binary logistic regression models, including **(a)** sTACI **(b)** sBCMA.

Following multivariate binary logistic regression analyses, although BAFF was not independently associated with relapse (P=0.08), both the baseline sTACI (P=0.007) and sBCMA (P=0.036) individually were independently associated with the occurrence of a GD relapse alongside TRAb titre at time of stopping ATD (Figure 6.10). Baseline sTACI on multivariate analysis remained significant (P=0.04) after Benjamini-Hochberg multiple test correction, unlike sBCMA (P=0.20). Goitre size remained non-significant on multivariate analysis with both sTACI (P=0.12) and sBCMA (P=0.09) when the cohort was split into two groups for analysis based on the presence of a goitre (Grade 0 vs. Grade 1 and 2).

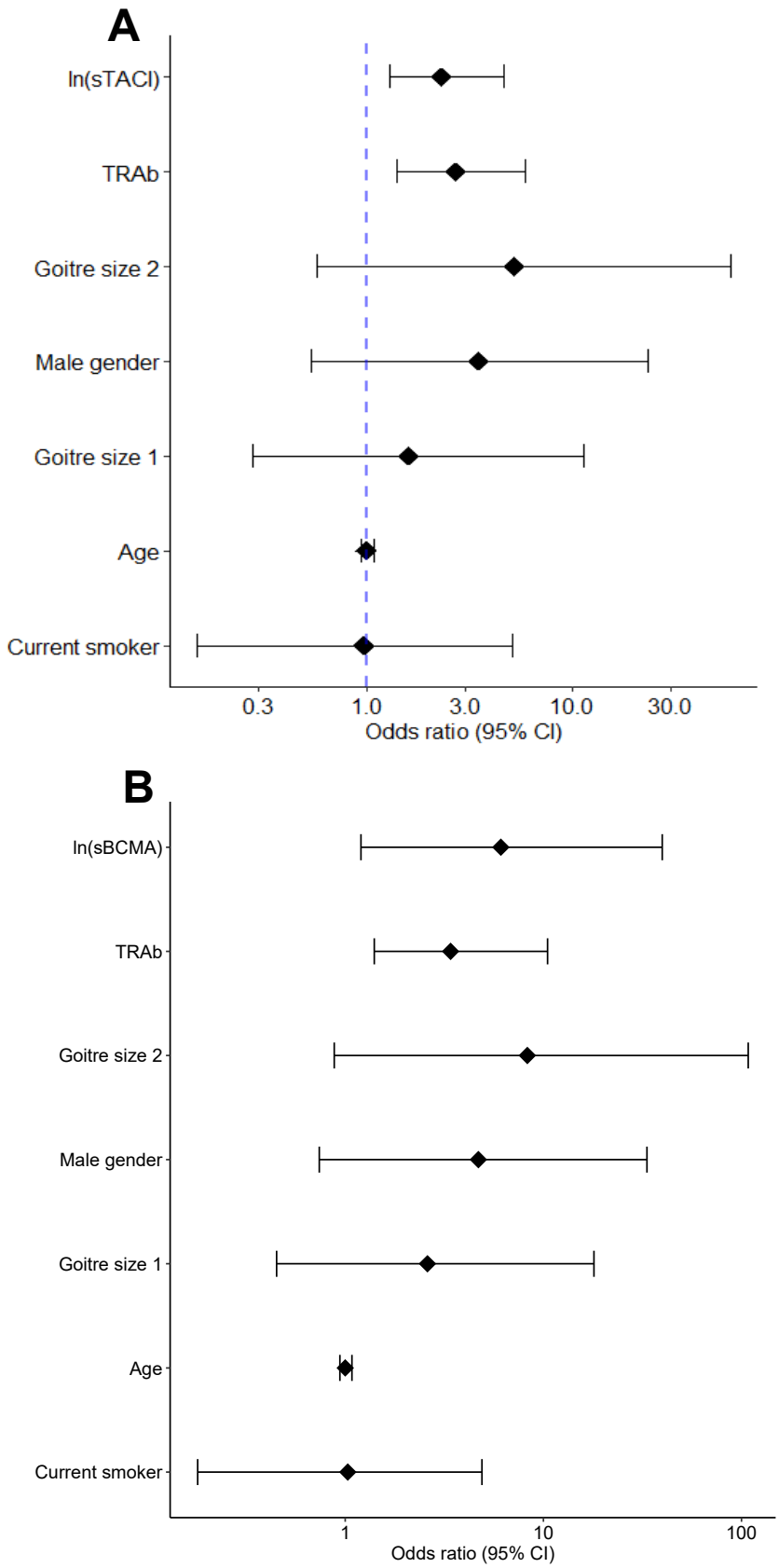


Figure 6.10 – Forest plots summarising the association of baseline sTACI **(A)** and sBCMA **(B)** with the occurrence of Graves' disease relapse following ATD withdrawal, using a multivariate binary logistic regression model.

The logistic regression analysis of sTACI was repeated to exclude the 25% of values that were below the LLOD and had been assigned the manufacturer LLOD concentration in case this had affected the results. However, once these patients were excluded, sTACI remained associated with GD outcome on both univariate (P=0.016) and multivariate analysis (P=0.024).

6.8 Heatmap analysis of cytokine expression and outcome

Using the hierarchical clustering method as set out in Chapter 3: Methods, a heatmap was generated to visualise clustering of samples based on cytokine expression. Following hierarchical clustering there were four clusters identified within the heatmap (Figure 6.11).

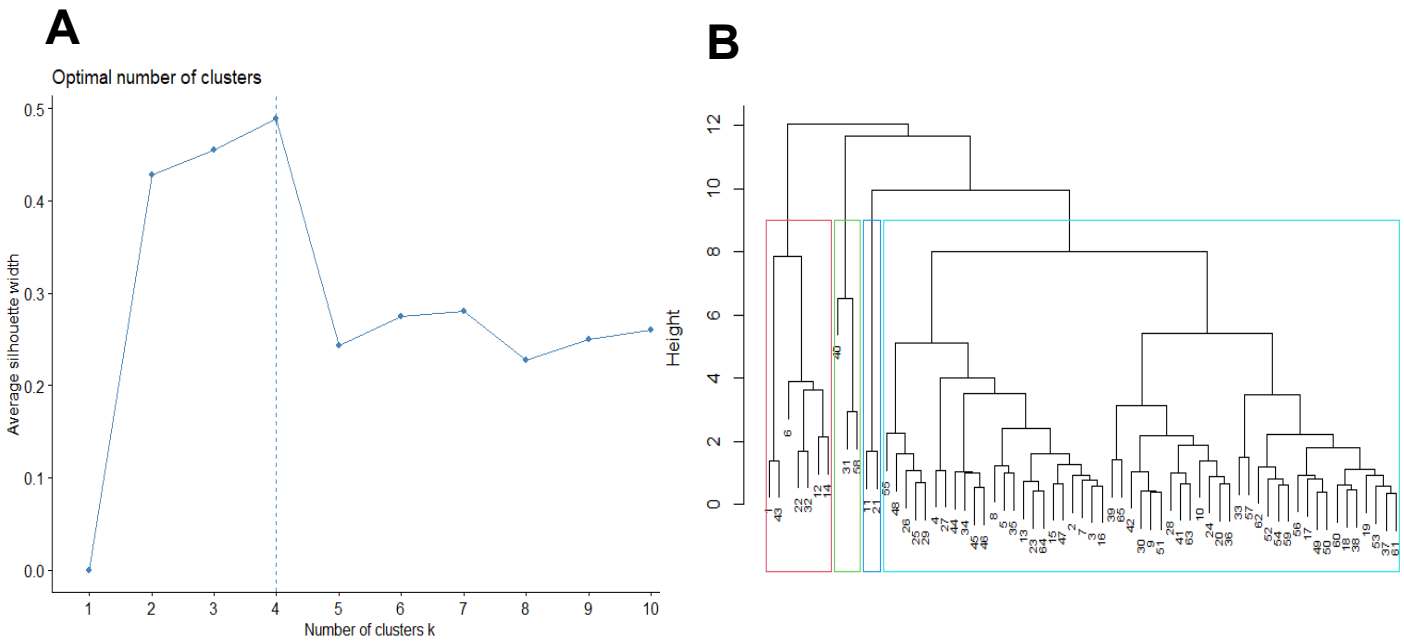


Figure 6.11 – **A**: Silhouette plot used to determine the optimal number of clusters in the dataset **B**: Cluster dendrogram produced following hierarchical clustering to identify four separate clusters

The clustered heatmap of cytokine/chemokine expression of all 65 patients at the time of stopping ATD is presented in Figure 6.12. This clustered heatmap shows four separate clusters, including one (labelled 'Group 1') which demonstrates a group of patients (mainly containing relapsing patients (5/7; 71%)) with higher sTACI and sBCMA expression, consistent with the findings from the multivariate analysis that demonstrated an association with both sTACI and sBCMA and outcome (Figure 6.10). Apart from one relapsing patient that is in a separate cluster (related to isolated high levels of BAFF), the other relapsing patients all fall within the same cluster as remission patients (labelled 'Group 2') and there is no discernible difference in cytokine expression visualised within this group of patients.

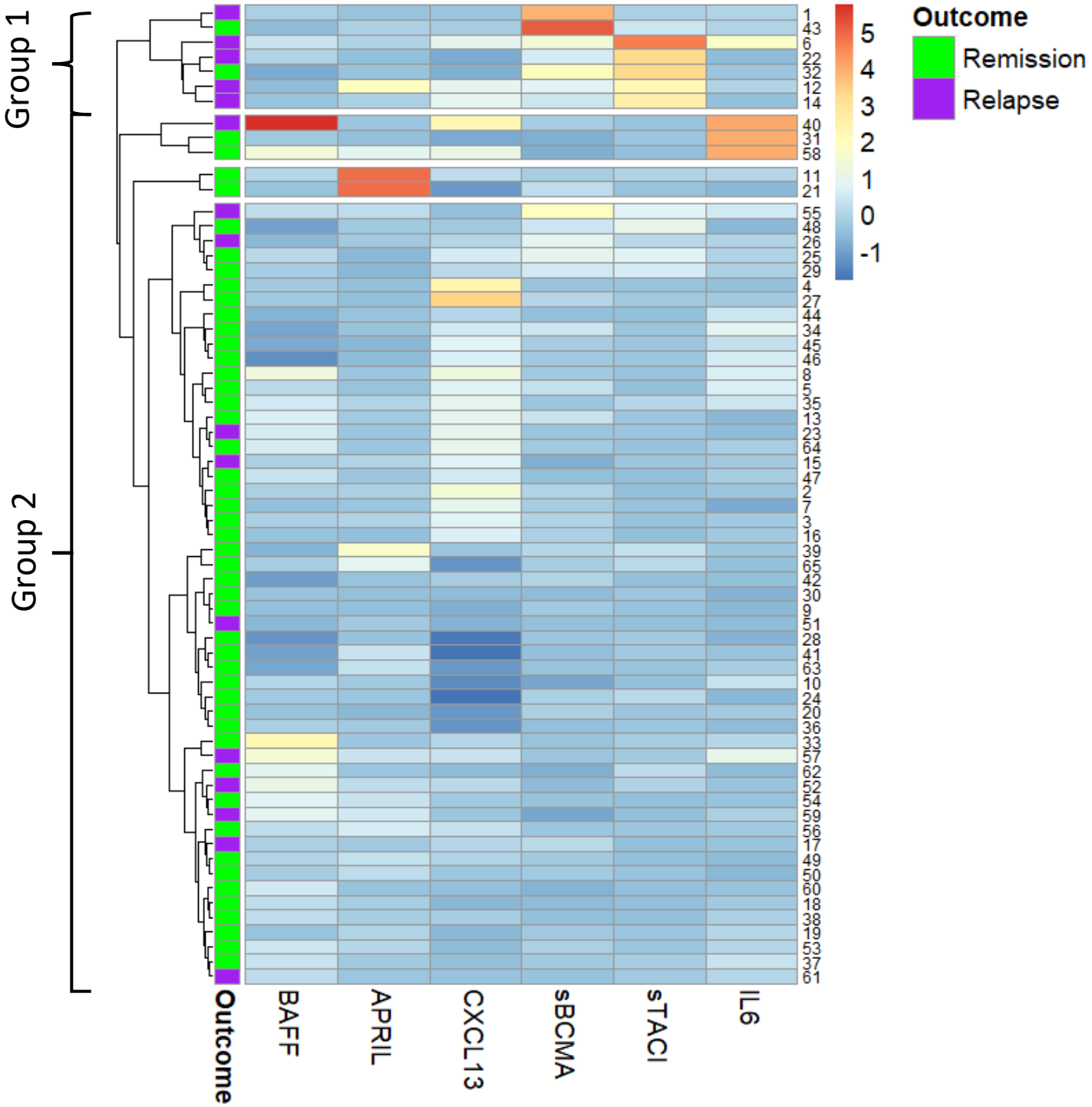


Figure 6.12 – Clustered heatmap of cytokine/chemokine expression at the time of stopping ATD.

The two groups identified on Figure 6.12 included five (Group 1) and eleven (Group 2) relapsing patients respectively. Additional analysis was undertaken to further characterise these two groups of relapsing patients (Chapter 9: Integrative analysis).

A heatmap was also generated to examine any clustering of cytokine/chemokine expression based on the time it took patients to relapse after stopping ATD. However, there was no clear association of cytokine/chemokine expression with timing of relapse (Figure 6.13).

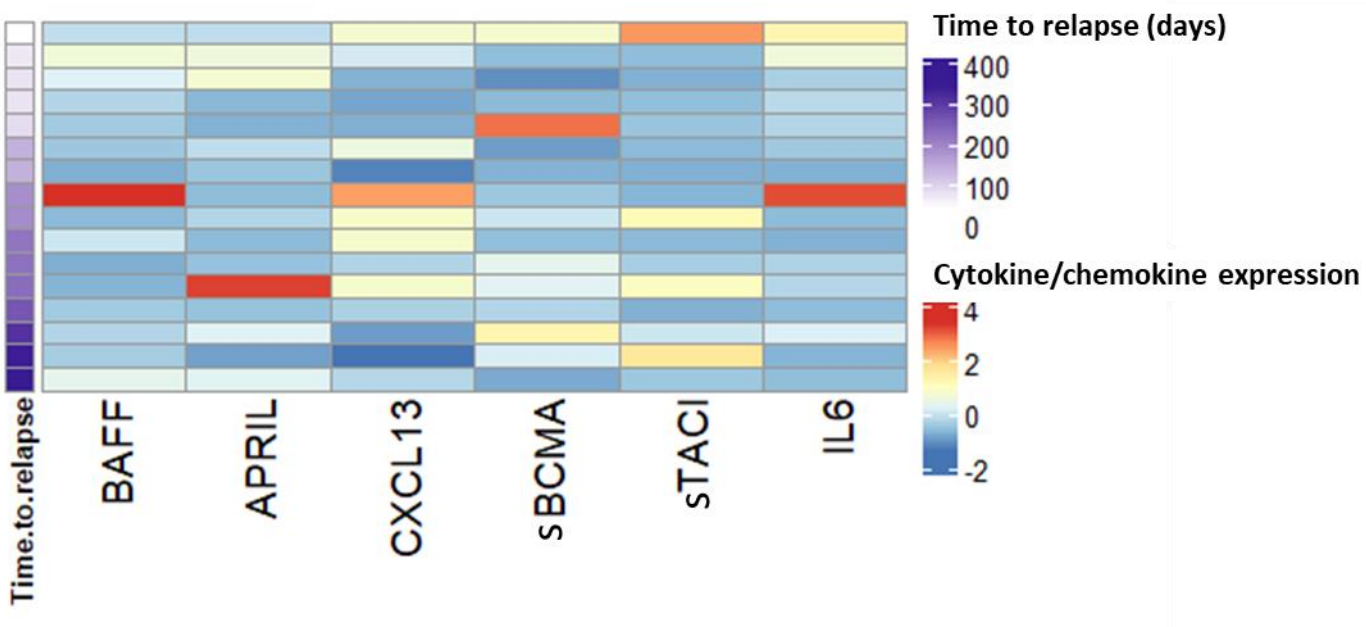


Figure 6.13 – Heatmap of cytokine/chemokine expression including only relapsing patients at the time of stopping ATD. **6.9 Longitudinal analysis: Baseline to 6-10 weeks after ATD cessation**

Paired analysis was undertaken to investigate the change in cytokines/chemokine from stopping ATD to 6-10 weeks later. There was a decline in sBCMA concentrations from baseline to 6-10 weeks later (P=0.01), but when split into relapse and remission patient groups this decline was only observed as significant in the remission patients (P= 0.004) (Figure 6.14). For all the other cytokines/chemokine, there was no significant change in concentrations in the 6–10-week period following ATD withdrawal.

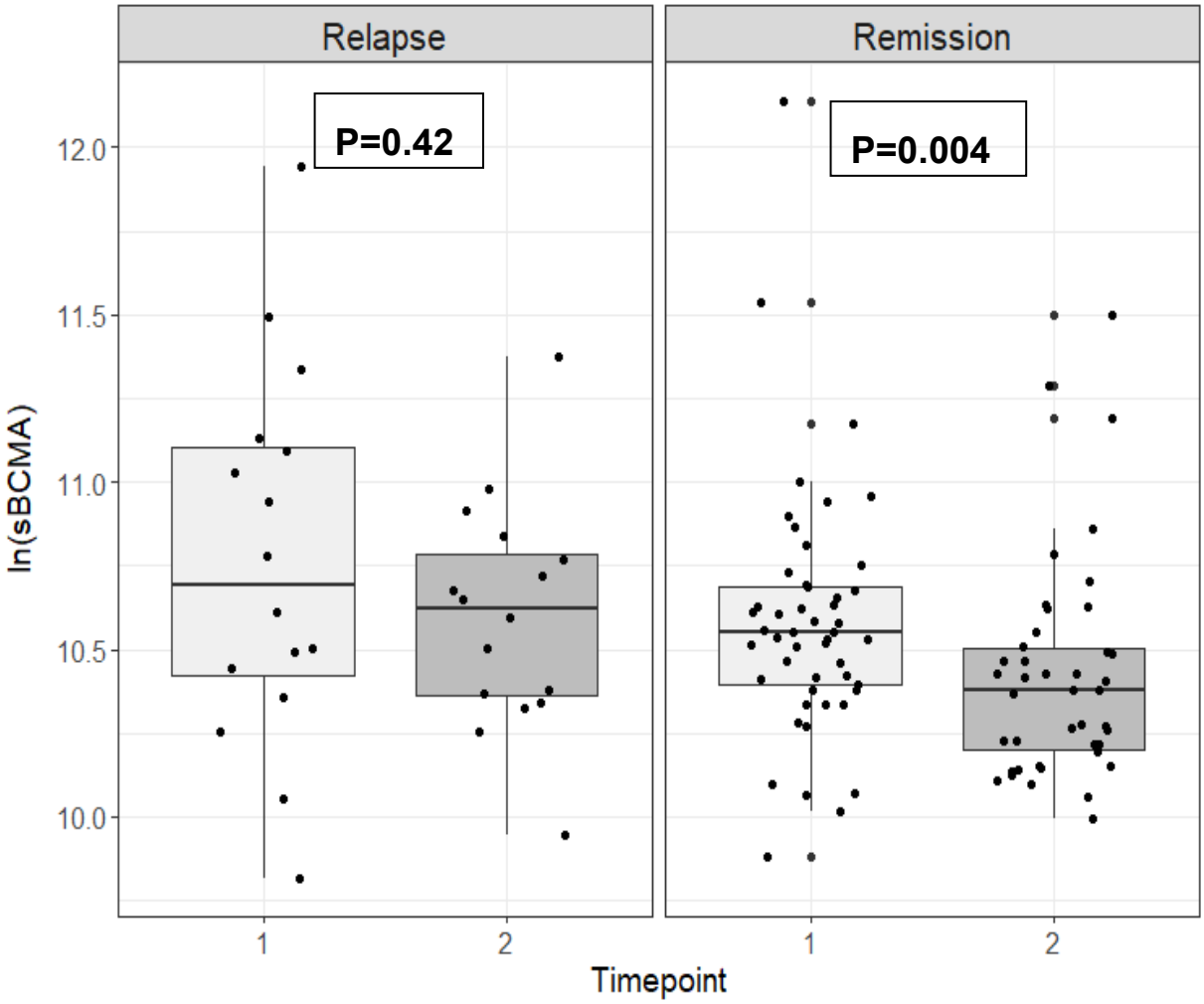


Figure 6.14 – Log-transformed (ln) sBCMA concentrations at ATD withdrawal (Timepoint 1) and 6-10 weeks later (Timepoint 2), divided into relapse and remission groups.

6.10 Survival analysis

The association between the cytokines/chemokine concentration and time-to-relapse following ATD withdrawal was assessed using univariate Cox regression analysis (Table 6.5). The assumptions of the Cox Proportional Hazards Model were tested for each variable using Schoenfeld Residuals to assess proportionality of hazards (as described in Chapter 3: Methods). The proportional hazard assumption was supported by a non-significant relationship between residuals and time for each cytokine/chemokine.

(a)

Baseline log-transformed cytokine/chemokine	B	HRrelapse	95% CI HRrelapse	Univariate P value
ln(sTACI)	0.46	1.6	1.08 – 2.3	**0.019***
ln(BAFF)	1.7	5.2	1.3 – 21	**0.023***
ln(sBCMA)	0.78	2.2	0.8 – 5.9	0.12
ln(IL-6)	0.43	1.5	0.8 – 2.9	0.18
ln(CXCL13)	1.1	2.9	0.5 – 16	0.2
ln(APRIL)	1.3	1.1	0.78 – 2.2	0.29

(b)

6-10 weeks log-transformed cytokine/chemokine	B	HRrelapse	95% CI HRrelapse	Univariate P value
ln(sTACI)	0.12	1.1	0.76 – 1.7	0.54
ln(BAFF)	0.98	2.7	0.64 – 11	0.18
ln(sBCMA)	1.2	3.2	0.9 – 10	0.06
ln(IL-6)	0.09	1.1	0.5 – 2.22	0.79
ln(CXCL13)	1.1	3.1	0.6 – 15	0.16
ln(APRIL)	0.02	1	0.6 – 1.7	0.94

Table 6.5 – Association of cytokines/chemokines at baseline **(a)** and 6-10 weeks **(b)** with timing of Graves' disease relapse following ATD withdrawal, using a univariate Cox regression model.

Cytokines at baseline with P=<0.2 on univariate cox regression analysis (sTACI, sBCMA, IL-6 and BAFF) were entered into separate multivariate cox regression analyses. These models included additional variables used in the previous logistic regression analysis, which are known to be associated with GD outcome (age, gender, goitre size, smoking, TRAb level prior to stopping ATD) (Table 6.6).

(a)

Baseline cytokine	B	HRrelapse	95% CI HRrelapse	Multivariate P value
ln(sTACI)	0.62	1.86	1.2 – 2.9	**0.007***
TRAb	0.55	1.7	1.1 – 2.6	**0.009***
Goitre size				
1	0.3	1.4	0.3 – 6	0.69
2	1.29	3.6	0.6 – 21	0.14
Gender (Male)	0.73	2	0.5 – 8	0.29
Age	0.01	1	0.95 – 1.08	0.70
Current smoker	-0.06	0.94	0.2 – 3.8	0.93

(b)

Baseline cytokine	B	HRrelapse	95% CI HRrelapse	Multivariate P value
ln(sBCMA)	1.1	3	0.9 – 10.6	0.09
TRAb	0.37	1.5	1.01 – 2.07	**0.04***
Goitre size				
1	0.46	1.6	0.4 – 6.9	0.54
2	1.6	5.1	0.9 – 30	0.07
Gender (Male)	1.6	2.5	0.7 – 9.4	0.17
Age	0.007	1	0.95 – 1.07	0.71
Current smoker	-0.07	0.92	1.07 – 2.05	0.97

(c)

Baseline cytokine	B	HRrelapse	95% CI HRrelapse	Multivariate P value
ln(BAFF)	1.8	5.8	1.3 – 26.3	**0.02***
TRAb	0.4	1.5	1.03 – 2	**0.03***
Goitre size				
1	0.76	2.1	0.43 – 8.9	0.33
2	1.73	5.6	0.9 – 34	0.06
Gender (Male)	0.97	2.7	0.7 – 10.6	0.17
Age	0.02	1	0.96 – 1.06	0.53
Current smoker	-0.55	0.57	0.15 – 2.2	0.42

(d)

Baseline cytokine	B	HRrelapse	95% CI HRrelapse	Multivariate P value
ln(IL-6)	0.76	2.14	1.02 – 4.5	**0.04***
TRAb	0.41	1.5	1.1 – 2.2	**0.03***
Goitre size				
1	0.73	2.08	0.46 – 9.4	0.34
2	1.85	6.4	1.10 – 36.9	**0.038***
Gender (Male)	0.76	2.13	0.5 – 8.8	0.30
Age	0.01	1	0.96 – 1.06	0.56
Current smoker	-0.35	0.70	0.18 – 2.7	0.61

Table 6.6 – Association of the cytokines **(a)** sTACI **(b)** sBCMA **(c)** BAFF **(d)** IL-6, with timing of Graves' disease relapse following ATD withdrawal, using multivariate Cox regression models.

The multivariate Cox regression models demonstrated that sTACI, BAFF and IL-6 were independently associated with time to GD relapse, and in each model the TRAb concentration at the time of stopping ATD was the only other factor consistently associated with time-to-relapse. These three cytokines that remained significantly

associated with time-to-relapse on individual multivariate Cox regression models were then all entered together into a multivariate cox regression analysis (Table 6.7).

Baseline variable	B	HRrelapse	95% CI HRrelapse	Multivariate P value
ln(sTACI)	0.83	2.23	1.35 – 3.89	**0.002***
TRAb	0.65	1.9	1.2 – 3	**0.006***
ln(BAFF)	1.42	4.13	0.63 – 27	0.14
ln(IL-6)	0.69	2	0.75 – 5.34	0.16
Goitre size				
1	0.75	2.1	0.39 – 11.6	0.39
2	1.38	3.9	0.5 – 26.9	0.16
Gender (Male)	0.51	1.7	0.37 – 7.5	0.50
Age	0.009	1.009	0.95 – 1.08	0.77
Current smoker	-0.30	0.74	0.21 – 2.9	0.71

Table 6.7 – Association of the cytokines (sTACI, IL-6 and BAFF) together, with timing of Graves' disease relapse following ATD withdrawal, using a multivariate Cox regression model.

Following this analysis, baseline sTACI, and TRAb titre remained independently associated with a shorter time-to-relapse (P=<0.05). The results of the multivariate cox regression analysis are summarised in Figure 6.15.

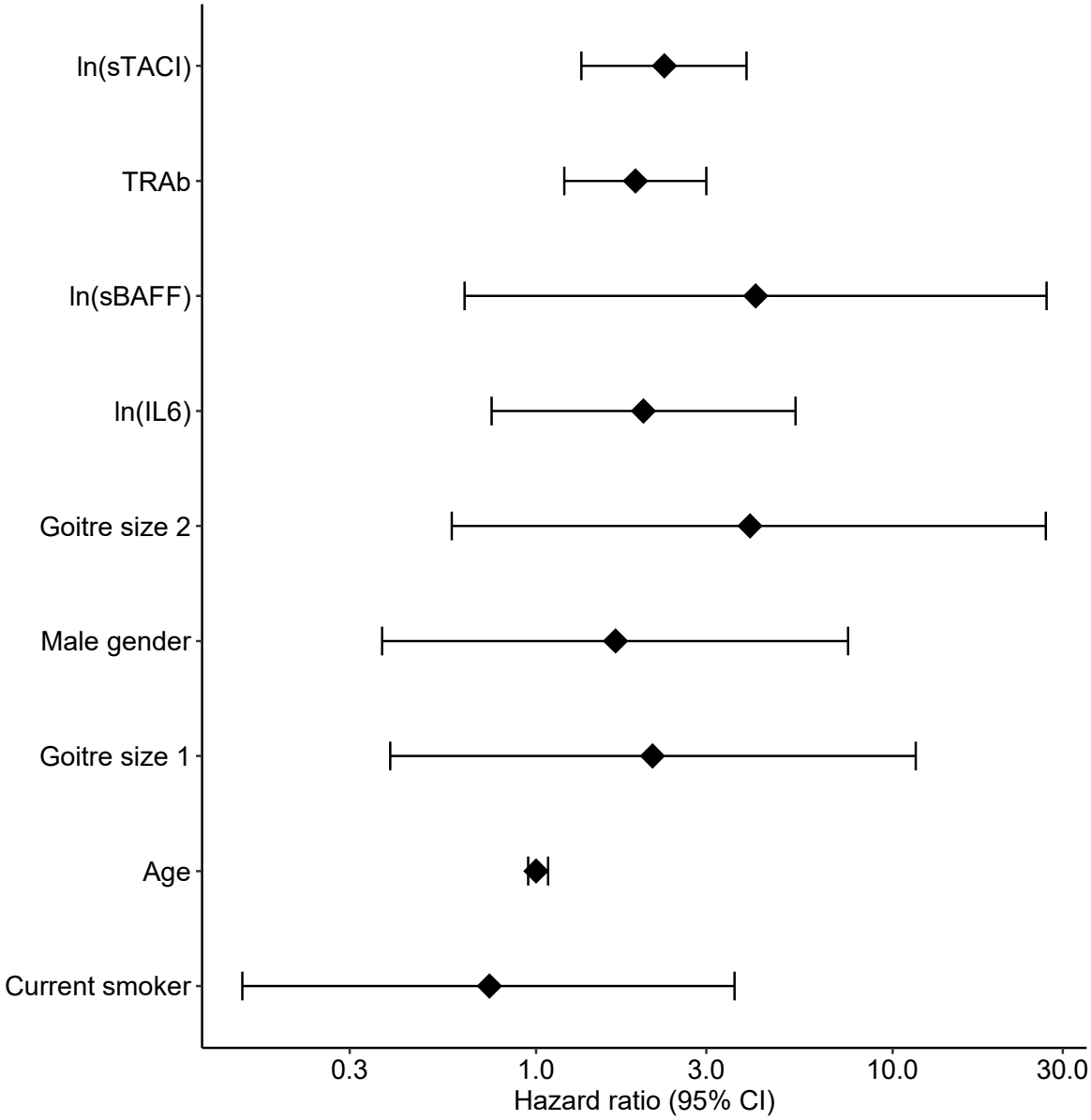
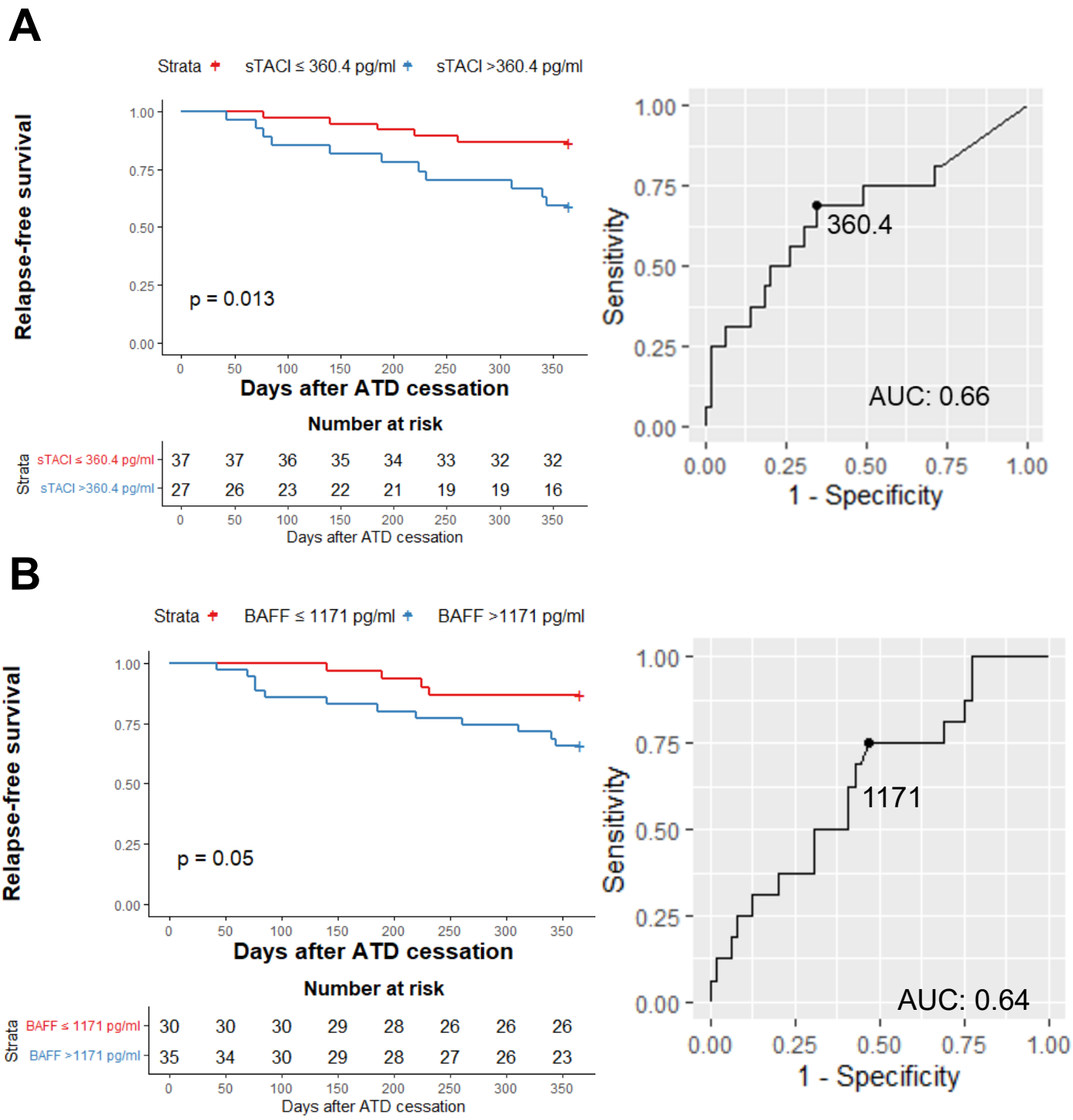


Figure 6.15 – Forest plot summarising the association of baseline sTACI**,** IL-6 and BAFF with the timing of Graves' disease relapse following ATD withdrawal, using a multivariate cox regression model. **6.11 Cytokine ROC curve analysis**

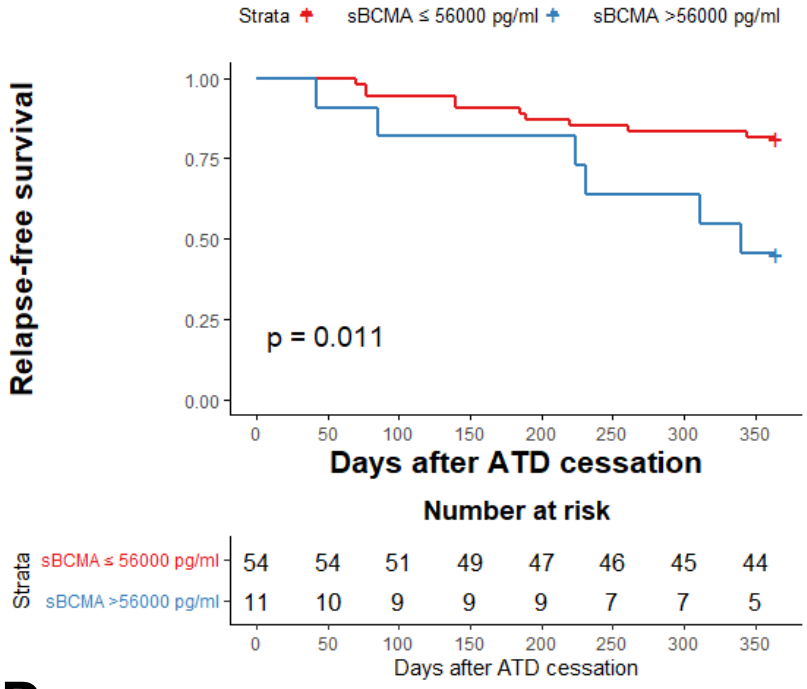
ROC curve analysis was undertaken to assess the performance of different cytokine/chemokine biomarker(s) in distinguishing between the outcome of relapse vs. remission. The R programme 'cutpointr' generated the optimal cytokine concentration cut-off which produced the highest sensitivity and specificity of a specific cytokine for the prediction of relapse or remission (Chapter 3: Methods).

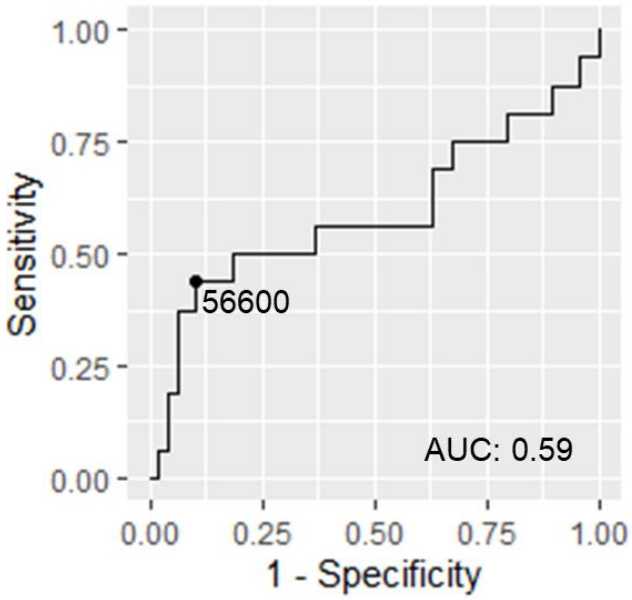
Given their observed significance in the multivariate logistic and/or cox regression model, baseline sBCMA, sTACI, BAFF and IL-6 levels were analysed. The study

cohort was dichotomised into high/low cytokine concentrations as described previously (Chapter 3: Methods). The survival distributions of these cytokines were analysed by the log-rank test and are presented in Kaplan-Meier plots, including the ROC curves for each cytokine in Figure 6.16.

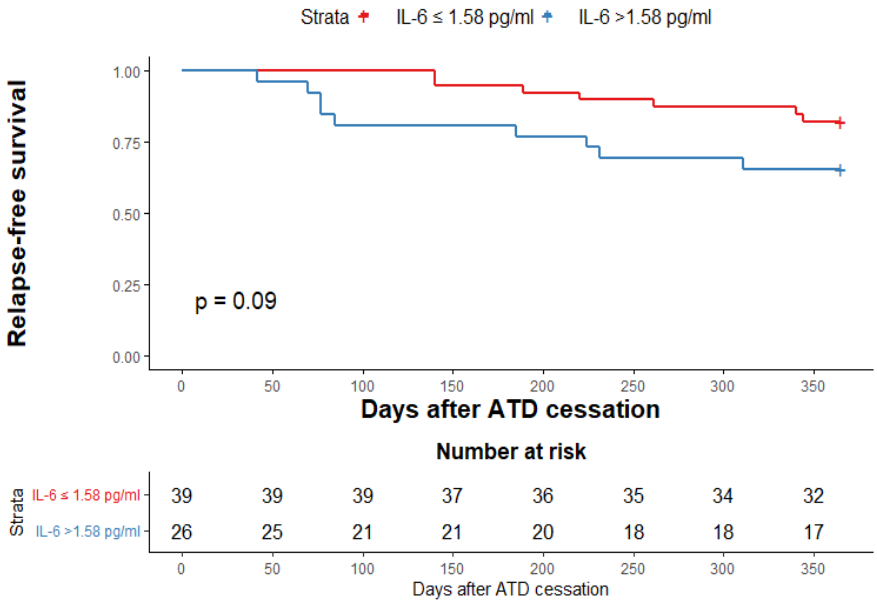


C





D



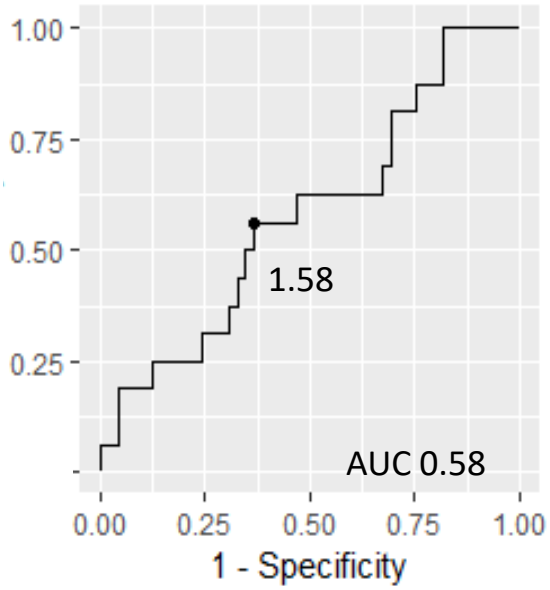


Figure 6.16 – Kaplan-Meier plots of survival following ATD cessation stratified by high (blue) or low (red) level of cytokine. **A**: sTACI; **B**: BAFF; **C**: sBCMA, **D**: IL-6.

ROC_{AUC} values for each cytokine to predict relapse, including their sensitivity, specificity, and positive and negative predictive values, is presented in Table 6.8.

Baseline variable(s)	ROCAUC (95% CI)	Optimal Threshold	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95%CI)
sTACI	**0.66** (0.52-0.86)	360.4 pg/ml	0.68 (0.5 – 0.8)	0.65 (0.47 – 1)	0.39 (0.32 – 0.8)	0.86 (0.77 – 0.90)
BAFF	0.64 (0.49-0.77)	1171 pg/ml	**0.75** (0.37 – 1)	0.53 (0.20 – 0.93)	0.34 (0.29 – 0.8)	**0.87** (0.77 – 1)
sBCMA	0.59 (0.43-0.76)	56000 pg/ml	0.44 (0.27 – 0.75)	**0.90** (0.73 – 0.98)	**0.58** (0.31 – 0.68)	0.83 (0.78 – 0.91)
IL-6	0.58 (0.43-0.71)	1.58 pg/ml	0.56 (0.27 – 0.76)	0.63 (0.10 – 1)	0.33 (0.22 – 0.80)	0.82 (0.76 – 1)

Table 6.8 – ROC curve analysis of the predictive utility of sTACI, BAFF, sBCMA and IL-6 biomarker thresholds to predict Graves' disease relapse following cessation of ATD.

The difference in survival distributions for the presence of an increased sTACI (>360.4 pg/ml) and sBCMA (>56000 pg/ml) met statistical significance. For sTACI, this difference was particularly apparent beyond the first 2 months after stopping ATD, whereas with sBCMA, the survival difference was more notable from around 7 months after stopping ATD. In contrast, there was no significant difference in survival distributions for BAFF or IL-6, although this was on the borderline of statistical significance in the former (P=0.05). ROC curve analysis was also performed to assess the performance of these cytokines when used in combination with each other and other variables (Chapter 9: Integrative analysis).

6.12 Discussion

Cytokines form a crucial part of the autoimmune response and have an important role in modulating adaptive immunity (*Moudgil KD et al. 2011*). Circulating peripheral cytokines can interact synergistically to drive inflammation and autoimmunity,

creating a pro-inflammatory state often described in autoimmune disease (*Cassese G et al. 2003, Wahren-Herlenius et al. 2013).* Several different cytokines, including IL-6 studied in this Chapter, have been implicated in GD pathogenesis and suggested as potential biomarkers for disease activity (*Pedro AB et al. 2011*).

Despite their relevance as a reflection of autoimmune activity, the use of cytokines as a predictive tool for GD outcome following the cessation of ATD has not previously been studied. In this Chapter, the relationship between cytokine/chemokine activity and outcome of GD was explored. The association between inflammatory status and thyroid hormones/autoantibodies with cytokines was also investigated.

6.12.1 Cytokines and Graves' disease patient characteristics

Age, gender, and smoking status are all variables that have been described to influence an individual's cytokine dynamics (*Ter Horst R et al. 2016*). There is a growing body of evidence regarding the presence of sex-based differences in forming an immune response, through genetic or hormonal mediators, that result in variations to the activity, distribution and production of cytokines and chemokines (*Klein et al. 2016*). Indeed, oestrogen has been implicated as a factor modulating BAFF expression in GD (*Cheng CW et al. 2021*). However, my study found no relationship between any of the cytokine concentrations and gender of the patients. Regardless, as male gender is associated with an increased risk of relapse in GD, gender was included as a factor in all multivariate analysis of the cytokines.

Immunosenescence associated with increasing age has also been suggested to trigger cytokine dysregulation resulting in a pro-inflammatory state (*Rea et al. 2018*). Although my study did find a weak positive correlation between APRIL and patient age, this finding was only apparent at the 6-10 week sample and conversely, at this timepoint there was also a negative correlation found with age and sTACI. Despite these unclear findings, as age is a clinical variable already known to be associated with likelihood of relapse in GD, it was included as a factor in all multivariate analyses of the cytokines.

Smoking status is another factor known to be associated with relapse in GD. Exposure to cigarette smoke is described to affect both the innate and adaptive immune response and directly affect cytokine secretion, with current smokers demonstrating an increased inflammatory response to stimulation (*Saint-André et al. 2024*). My study found that out of the cytokines measured it was only APRIL that was

found to be associated with smoking status, with significantly elevated levels demonstrated in those described as current smokers. Serum APRIL has previously been found to be elevated in smokers compared to non-smokers (*Gümüş P et al. 2014*). As well as age and gender, smoking status was also included as a factor in all multivariate analyses of the cytokines.

6.12.2 Cytokines and thyroid status

The proposed bidirectional 'crosstalk' between thyroid hormones and the immune system suggest that thyroid status can directly affect immune homeostasis, and conversely components of the immune system may interfere with the regulation of systemic thyroid hormone levels (*Wenzek C et al. 2022*). Serum BAFF in patients with GD has been described as elevated compared to healthy controls and positively correlated with FT4, FT3 and TRAb (*Liu S et al. 2022, Cheng CW et al. 2021*). Elevated levels of FT3 have been described to induce overexpression of BAFF in thyroid follicular cells affecting adaptive immunity by promoting the differentiation of autoreactive B cells (*Liu S et al. 2022*). In my study BAFF was found to be positively correlated with both FT3 and FT4, although there was no correlation with the thyroid autoantibodies. Therefore, it is possible that patients with higher thyroid hormone concentrations which are potentially driving BAFF secretion, may be more likely to demonstrate dysregulated autoreactive B cell differentiation affecting GD activity.

6.12.3 Cytokines and inflammatory status

Cytokines have a diverse and pleotropic role in orchestrating antigen-specific immune responses and inflammatory reactions (*Kany S et al. 2019*). They are used in clinical practice to monitor inflammatory disease activity and have been shown to correlate with disease activity in GD (*Pedro AB et al. 2011)*. Elevated levels of IL-6 and IL-6 receptor have been reported in the serum of GD patients (Salvi et al. 1996), and *Pedro et al.* reported that untreated GD patients had higher IL-6 levels compared to controls (3.0 vs. 5.0 pg/ml), which fell significantly once euthyroid on during ATD treatment. Most of the patients in my study were euthyroid at the time the sample was taken, as they were either still taking or had recently finished ATD which may explain why the IL-6 was not elevated in most of these patients.

Although IL-6 was not associated with GD outcome in my study, it was found to be positively correlated with CRP at both measured timepoints. Consistent with this, IL-6

is known to stimulate the secretion of the acute phase reactant CRP and has been implicated in activating and maintaining an inflammatory response in both acute and chronic inflammation in autoimmunity (*Hirano T 2021*). Cytokines including IL-6 have been demonstrated to be produced by both intrathyroidal lymphocytes and thyroid follicular cells (*Weetman AP et al. 1997*) and have been proposed as a target for novel immunomodulatory therapies in GD and GO (*Lane et al. 2020*, *Fallahi P et al. 2021).* Secretion of IL-6 by thyroid follicular cells has been associated with the degree of lymphocytic infiltrate within the thyroid (*Grubeck-Loebenstein B et al. 1989*). ATD have been shown to inhibit IL-6 production from the thyroid follicular cells and reduce thyroid lymphocytic infiltrate, highlighting the potential immunomodulatory role of ATD in enabling remission of GD and possibly explaining the low levels of IL-6 seen in patients in my study on ATD (*Weetman AR et al. 1992).*

The elevated levels of IL-6 associated with CRP in my study may suggest that these patients were experiencing an acute intercurrent illness, however in only one case did a patient subjectively report a recent history of illness and there was no history of recent vaccinations (< 4 weeks) that could account for these elevated levels. As both CRP and IL-6 can be raised in both acute and chronic inflammatory states, elevated levels in these patients may reflect lymphocytic infiltrate and inflammation of the thyroid gland.

6.12.4 Relationships between circulating cytokines

Cytokines are often secreted together by the same immune cells and work synergistically or antagonistically to produce an immune response. The cytokines measured in this study are all closely related to B cell activity. BAFF was found to be positively correlated to both CXCL13 and IL-6 at both timepoints in my study, which is consistent with previous findings in autoimmune disease (*Salazar-Camarena DC et al. 2020, Traianos EY et al. 2020*). BAFF, CXCL13 and IL-6 all have roles in enhancing B cell activity (Chapter 1: Introduction). CXCL13 forms a chemokine axis with the receptor CXCR5 expressed on B cells, and promotes chemotaxis, germinal centre formation and differentiation of B cells into memory and plasma cells (*Pan Z et al. 2022*). CXCL13 and CXCR5 mRNA levels in thyroid tissue have been demonstrated to directly correlate with the number of focal lymphocytic infiltrates and germinal centres in the thyroid *(Aust G et al. 2004).* Similarly, IL-6 is produced by activated B cells and thyroid follicular cells and has been demonstrated as critical for

germinal centre formation (*Arkatkar T et al. 2017*). The production of IL-6 in the thyroid may promote ectopic germinal centre formation and contribute to the thyroid as a site for autoantibody synthesis in GD (*Weetman AP et al. 1997)*.

BAFF has been demonstrated to induce both IL-6 production and enhance the chemotactic response of B cells to CXCL13, specifically notable in increasing the chemotaxis of memory B cells (*Yoshimoto K et al. 2011, Badr G et al. 2008*). Therefore, the positive association of both these cytokines with BAFF may reflect their apparent synergistic role in promoting autoreactive B cell proliferation and survival through enhanced ectopic germinal centre activity in the thyroid.

Both BAFF and APRIL regulate B cell homeostasis by binding to their receptors BAFFR, TACI and BCMA (*Meinl E et al. 2021*). The soluble forms of the latter two receptors measured in this study, sTACI and sBCMA, are cleaved and shed from the membrane of activated late-stage B cells (memory/plasma cells) (*Meinl E et al. 2021).* In my study, both these soluble receptors were found to be strongly positively correlated with each other at both timepoints, which would be consistent in the context of their likely joint secretion in an activated B cell environment. Indeed, previous studies have also reported this positive correlation between sTACI and sBCMA (*Thaler FS et al. 2017).* Once secreted, sBCMA and sTACI are reported to act as decoy receptors blocking BAFF and APRIL (*Meinl E et al. 2021)*, which may explain why there was some negative correlation found in my study between BAFF and sBCMA and sTACI, although this wasn't consistent at both timepoints and isn't a consistent finding in previous studies (*Alfaro R et al. 2022*).

6.12.5 Cytokines and Graves' disease outcome following ATD treatment

As discussed previously, the cytokines measured in this study have key roles in modulating the humoral immune response and therefore may be important biomarkers of outcome in GD. In multivariate analysis, both sTACI and sBCMA at the time of stopping ATD were independently associated with relapse of GD. As expected, an elevated TRAb level at the time of stopping ATD was also consistently associated with relapse in all multivariate analysis.

On longitudinal analysis, sBCMA only fell significantly in the remission patients but remained high in the relapse patients at 6-10 weeks after stopping ATD, again suggesting a potential role of sBCMA in GD relapse. In addition, survival analysis also demonstrated that those with a high sBCMA or sTACI at the time of stopping

ATD had a poorer relapse free survival. This suggests that alongside TRAb, sTACI and sBCMA concentration at the time of stopping ATD are potential biomarkers of relapse in GD. Indeed, although not previously studied in GD, both sTACI and sBCMA have been found elevated in other autoimmune conditions including SLE, RA and MS and suggested as potential biomarkers of disease activity (*Salazar-Camarena DC et al. 2020, Hoffmann FS et al. 2015, Rodríguez-Carrio J et al. 2018, Laurent, S et al. 2015, Vincent FB 2019*). *Salazar-Camarena DC et al.* demonstrated that sBCMA was elevated in SLE patients compared with healthy controls (49.03 vs. 25.60 ng/mL; $p < 0.05$) and active SLE patients who achieved low disease activity demonstrated decreased sBCMA (53.30 vs 35.30 ng/mL; $p < 0.05$).

Circulating levels of sTACI and sBCMA may be a direct reflection of B cell activation considering they are both cleaved and shed once B cells become activated. Therefore, the presence of elevated sTACI and sBCMA observed in relapsing GD patients in my study suggests that an activated B cell state is associated with relapse and highlights the potential use of these soluble receptors as biomarkers of outcome. A proposed model of mechanistic insight into how these soluble receptors could be biomarkers for relapse in GD is presented in Figure 6.17.

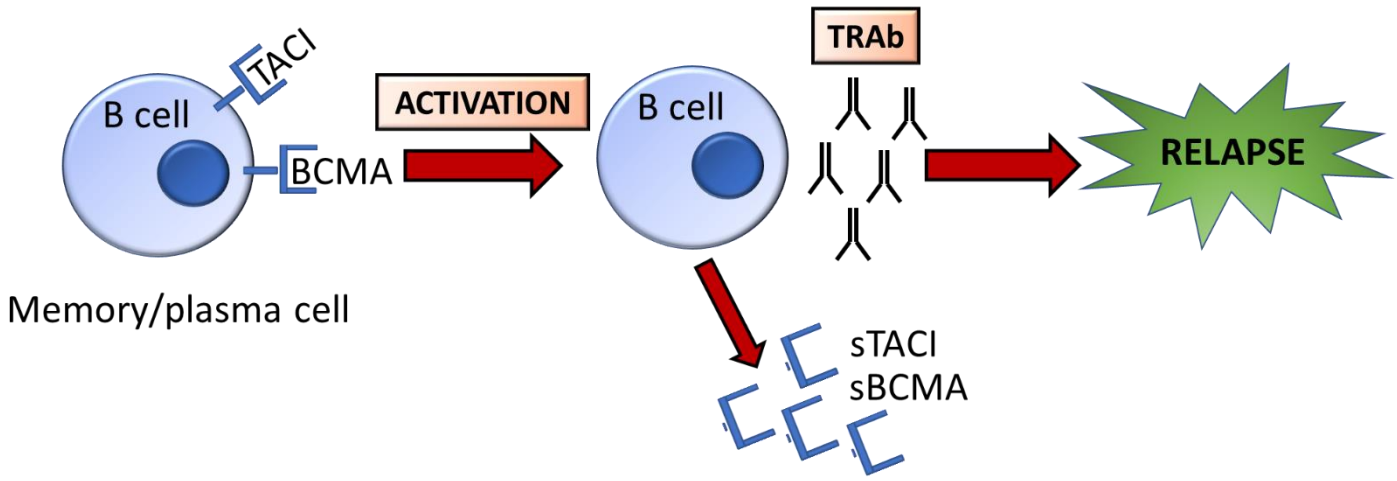


Figure 6.17 – A proposed model of how sTACI and sBCMA could be biomarkers of Graves' disease relapse.

Briefly, membrane bound TACI and BCMA receptors which are expressed on late-stage B cells (memory and plasma cells) are shed once these cells become activated. These B cells differentiate into antibody secreting cells producing, in the case of GD, TRAb. Therefore, the soluble receptors cleaved from these activated B cells could reflect an increase in antibody secreting B cell activity – which may explain why elevated sBCMA and sTACI are associated with GD relapse.

6.12.6 Cytokines and time-to-relapse of Graves' disease following ATD treatment

Alongside BAFF and IL-6, both sTACI and TRAb were also associated with time-to-relapse as elevated levels of these cytokines at the time of stopping ATD were associated with a shorter time-to-relapse. As discussed previously, it is possible that circulating sTACI and TRAb are reflective of a state of B cell activation, the severity of which could be directly linked to how quickly a patient may relapse their GD. Although the absolute concentrations of BAFF and IL-6 weren't independently associated with the binary outcome of GD, it may be that BAFF-induced IL-6 is contributing to ectopic germinal centre formation in the thyroid thereby enhancing TRAb production and resulting in a faster time-to-relapse. In addition, decreased methylation levels of the IL-6 gene have been associated with intractable GD, hypothesised to result from increased IL-6 production *(Hirai, N et al. 2019).*

Considering the role of IL-6 in acute and chronic inflammation (*Hirano T 2021)*, and the previous association found in my study with another biomarker of systemic inflammation, MLR, and the speed of relapse (Chapter 4: Results 1 Clinical), an elevated IL-6 may also reflect severity of inflammation within the thyroid which could potentially be driving a faster relapse.

6.13 Limitations

As my study was performed on a relatively small cohort of patients, most of whom were female, it may be difficult to detect differences in sex-specific cytokine responses that may have provided mechanistic insight into known sex-specific differences in relapse rates. Therefore, to address this issue the study would need to be repeated on a larger scale using an independent patient cohort, including more male participants, to validate the findings. In addition, repeating the cytokine measurements on a more frequent basis in the first year after stopping ATD may provide greater insight into the pathophysiological cytokine dynamics indicating pending relapse.

Regarding the measurements of cytokines, it is worth noting that the peripheral compartment of cytokines may not accurately reflect the cytokines present within the thyroid tissue. Ideally, additional cytokine analysis of thyroid tissue in conjunction with peripheral analysis would provide more in-depth insight into the pathophysiology of cytokine dynamics and relapse. However, in terms of the objective of this study which is to find a suitable biomarker(s) to inform clinical management, peripheral cytokines are more accessible and easier to process for routine clinical practice.

The assay technology meant that although most cytokine measurements were within the limits of detection, 25% and 8% of the sTACI and APRIL measurements, respectively, were below the LLOD. Due to the relatively small cohort of patients, these were assigned the lowest detectable concentration as per the manufacturers kit for the purposes of including them in the analysis. Therefore, although it may be that these limited number of measurements do not provide precise quantitative data the results of the univariate and multivariate analysis remained significant even with these cases excluded.

As with any immunological studies, cytokines are dynamic and may change in relation to factors other than that related to autoimmune disease, such as intercurrent

illness and vaccinations. CRP level and subjective questioning regarding illness and recent vaccinations was explored with patients to try and account for this.

The specific B cell cytokines studied in this thesis were chosen due to their well-established, synergistic role in B cell immunity and their specific relevance in autoimmunity. However, there are a number of additional B cell cytokines demonstrated to be associated with GD activity such as IL-1β, soluble IL-2R, IL-5 and tumour necrosis factor-*α* (*Pedro AB et al. 2011*) that would also be relevant to study and may contribute additional information about humoral immunity in GD relapse. Furthermore, tolerogenic interleukins, such as IL-10 and IL-35, have been implicated in GD (*Saeed, M et al. 2021, Takeoka, K et al. 2004*), and may be relevant in the context of their anti-inflammatory role.

6.14 Summary

This Results Chapter presents the findings of the cytokine analysis that suggests an association between B-cell related cytokines with thyroid hormones, inflammation, and outcome in GD. Cytokines may work in synergy to promote lymphocytic infiltrate and inflammation within the thyroid, resulting in ectopic germinal centre formation, autoantibody synthesis and ultimately GD relapse. The association of baseline sTACI and sBCMA at the time of stopping ATD with relapse not only provides mechanistic insight into the presence of an activated B cell environment, but also indicates a potential role for these cytokines as predictive biomarkers of outcome in GD.

Chapter 7: Results 4 – Flow Cytometry

7.1 Introduction

The purpose of undertaking flow cytometry in this study was to determine the purity of the isolated $CD19^+$ B cells and to undertake immunophenotyping to identify the $CD19^+$ B lymphocyte subpopulations. Each B cell subpopulation has a distinct role in the humoral immune response and therefore it is important that B cell phenotypes are examined alongside the other humoral immune markers, as well as the $CD19^+$ transcriptome, to gain an in-depth understanding of the pathophysiological mechanisms underlying B cell activity. Additionally, it is important to consider the purity of the isolated $CD19^+$ B cell population as contaminating T cells and monocytes may affect interpretation of the RNA sequencing data.

The peripheral B cell compartment has been described in active GD (Chapter 1: Introduction), but detailed phenotypic studies on peripheral blood B cell subpopulations related to outcome in GD are lacking. Variations in certain B cell subpopulations may signify a more 'active' state of humoral immunity which could reflect an impending relapse and provide mechanistic insight into the relapse process.

In this Results Chapter, I present the analysis of flow cytometry data from samples taken at the timepoint of ATD withdrawal. The aim of this work was to compare the composition of peripheral B cell subpopulations between patients who relapsed or remitted one year after stopping ATD. Furthermore, the flow cytometry data was used in combination with the KRECs and cytokine data to provide insight into the contribution of these different factors to the underlying humoral immune activity observed in GD (Chapter 9: Integrative analysis).

7.2 Quality control

7.2.1 Sample collection and processing

The samples used in this flow cytometry experiment were the blood samples collected at the timepoint of stopping ATD. The isolated PBMCs and $CD19^+$ B cells were stored overnight at 4^oC and were processed the following day, normally by 24 hours of sample collection, median 24 (23 – 27 hours) (Figure 7.1). As noted on Figure 7.1, there was one sample that was delayed by 24 hours due to a mechanical

fault in the flow cytometry machine. However, this sample still demonstrated high cell viability (92.7%) and therefore it was analysed with the other samples. There was no significant difference between the relapse and remission groups and time to flow cytometry (P=0.81).

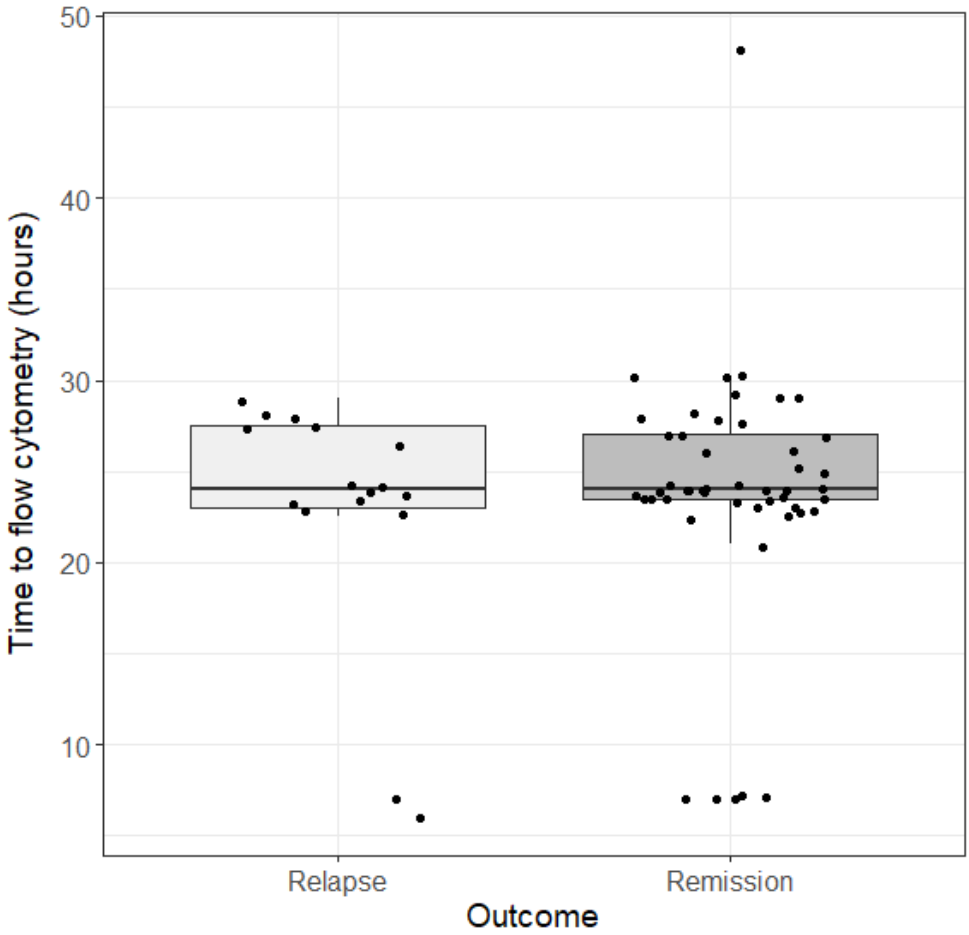


Figure 7.1 – Time to commencement of flow cytometry processing from the time of sample collection in both outcome groups.

7.2.2 Flow cytometry compensation and controls

Controls are essential in flow cytometry to reliably distinguish cell data from background fluorescence and non-specific binding. Dead cells can bind non-specifically to antibodies and have a high level of autofluorescence, which can lead to false positives and make it challenging to detect rare cell populations. Therefore, a viability dye (Zombie Aqua™) was included in the B cell subset flow cytometry panel to exclude dead cells. Overall cell viability was satisfactory with a median of

90.3% (IQR: 86.3 – 92.4%, range 77.1 – 96.3%) of live CD19^{+} B cells. The cell viability did not differ between the two outcome groups (P=0.4) (Figure 7.2).

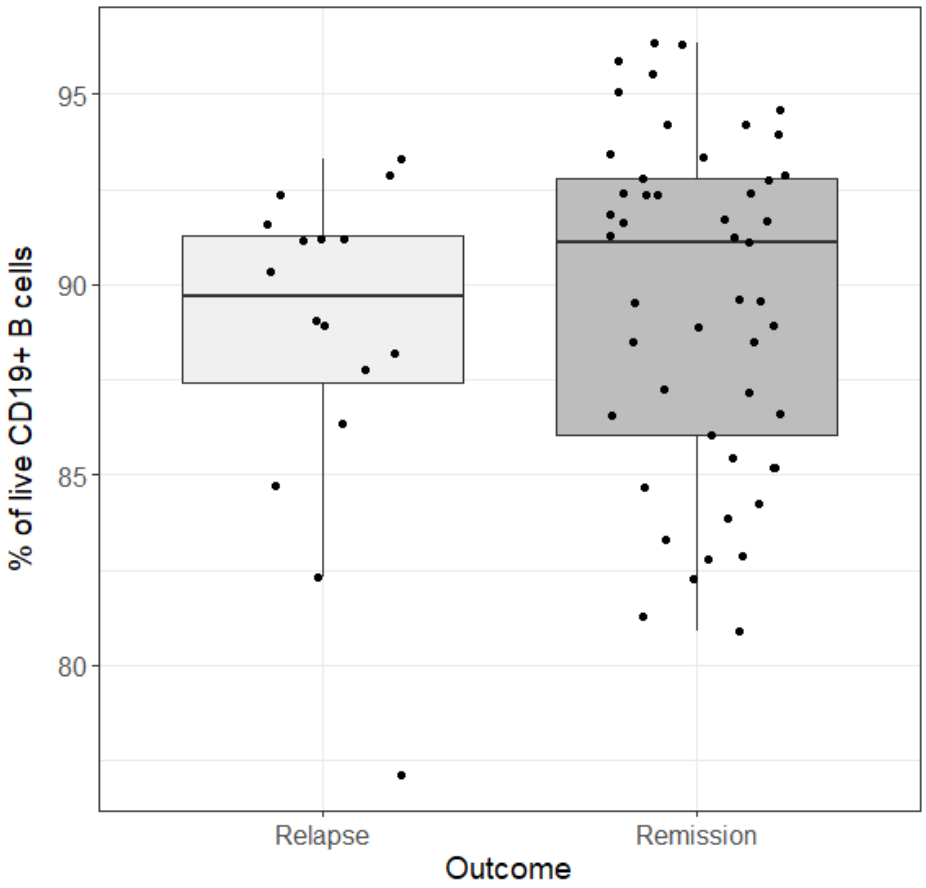


Figure 7.2 – Percentage of live CD19^{+} B cells determined by flow cytometry in both relapse and remission outcome groups.

Compensation control beads (BD™ CompBeads) were used to optimise the fluorescence compensation settings for the multicolour B cell phenotyping. The purpose of this was to establish compensation corrections for spectral overlap for the specific combination of fluorochrome-labelled antibodies used in this experiment.

Fc receptors found on B cells bind to antibodies via the constant Fc domain, rather than the antigen-specific Fab domain, resulting in antibodies binding to unwanted targets *(Andersen MN et al. 2016).* This can lead to false positives and therefore Human IgG immunoglobulin was used in this experiment to block the Fc receptors.

7.3 Peripheral blood mononuclear cell (PBMC) recovery

Isolated PBMCs were used to determine the initial quantities of $CD3^+CD4^+$ and $CD3^+CD4^-$ T cells and $CD19^+$ B cell populations. The cell populations as a median percentage of the total lymphocyte count were 41.4 (IQR 32 – 50%), 20 (IQR 15 – 25%) and 12 (IQR 6 – 16%), respectively (Figure 7.3). There was no significant difference in the absolute quantities of the $CD3^+CD4^+$ and $CD3^+CD4^-$ T cells (P=0.46, P=0.41, respectively) or $CD19^+$ B cells (P=0.40) between the relapse and remission groups.

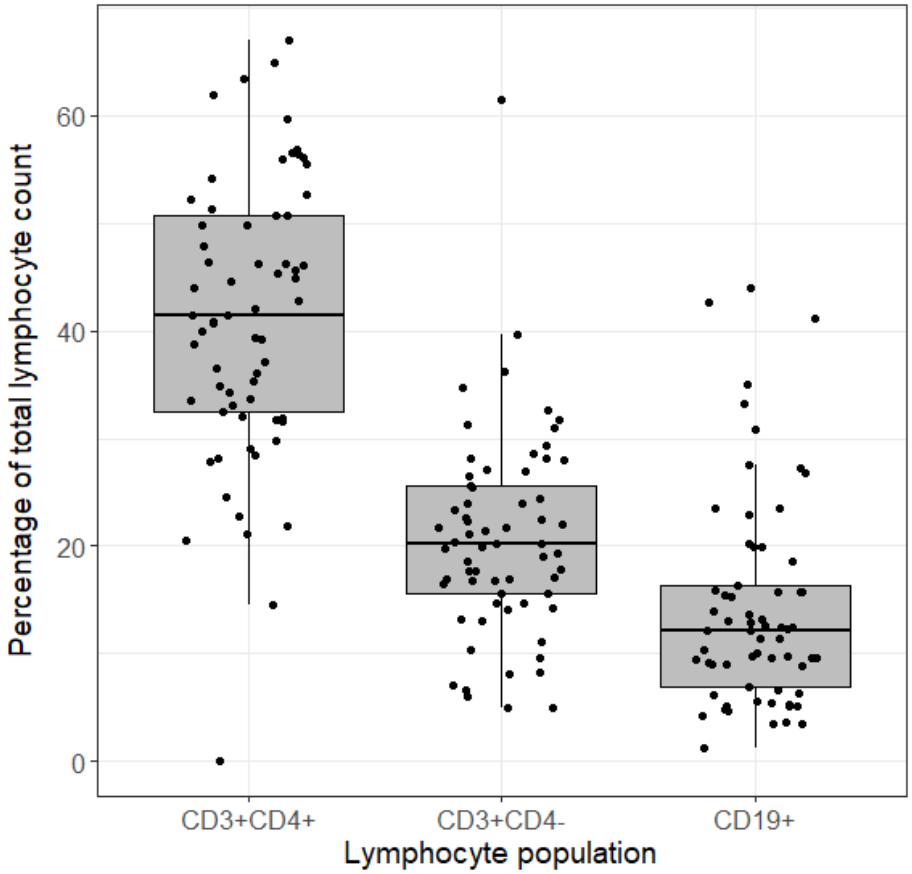


Figure 7.3 – Quantity of $CD3^+CD4^+$ and $CD3^+CD4^-$ T cells and $CD19^+$ B cells, as a percentage of the total lymphocyte count.

The PBMCs were manually gated and then analysed using the unsupervised clustering and dimension reduction programme (FCS Express™ 7) to generate distinct clusters based on cell surface marker expression, including the T and B cells and monocytes (Figure 7.4). As expected, the largest proportion of the lymphocyte

population was the $CD3^+CD4^+$ T cells, followed by the $CD3^+CD4^-$ T cells, the $CD19^+$ B cells and lastly, the $CD14^+$ monocytes (*Kokuina E et al. 2019*).

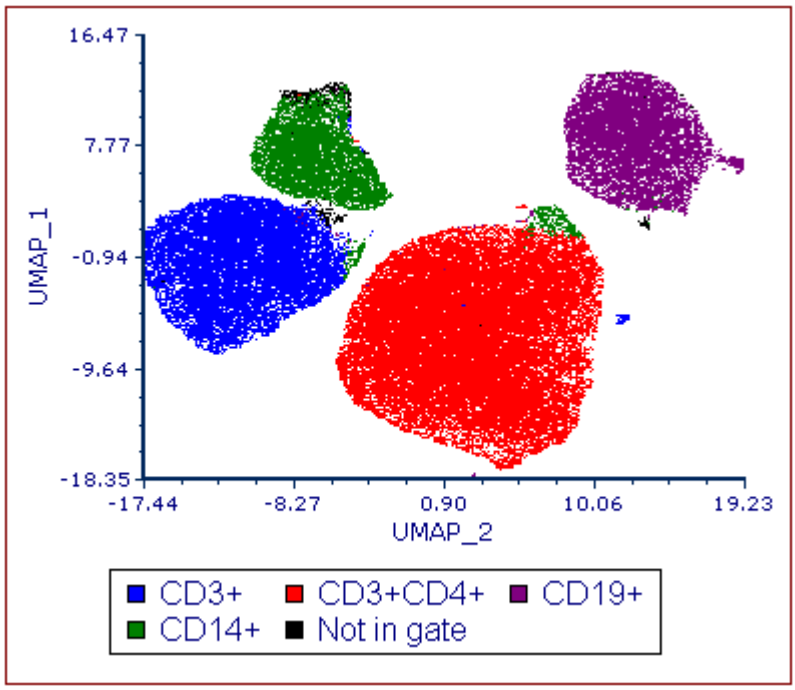


Figure 7.4 – UMAP (Uniform Manifold Approximation and Projection) plot created using the concatenated isolated PBMCs from all patient groups, demonstrating clustering of the main cell populations.

7.4 $CD19^+$ B cell purity

Prior to B cell immunophenotyping, the isolated $CD19^+$ B cells were assessed for purity and presence of contaminating T cells and monocytes. The median percentage of $CD19^+$ B cells in the isolated B cell sample was 96.1 (IQR 93 – 96%). There was no significant difference in B cell purity between the relapse and remission groups (P=0.84) (Figure 7.5). The cell isolation kit used is designed to provide >90% purity of positively selected B cells (*Miltenyi et al., 1990*), however three samples, all in the remission group, had <90% purity. Overall, the median percentage of contaminating T cells and monocytes was acceptable at 2.3 (IQR 1.4 – 4%) and 0.5 (0.2 – 0.9%), respectively.

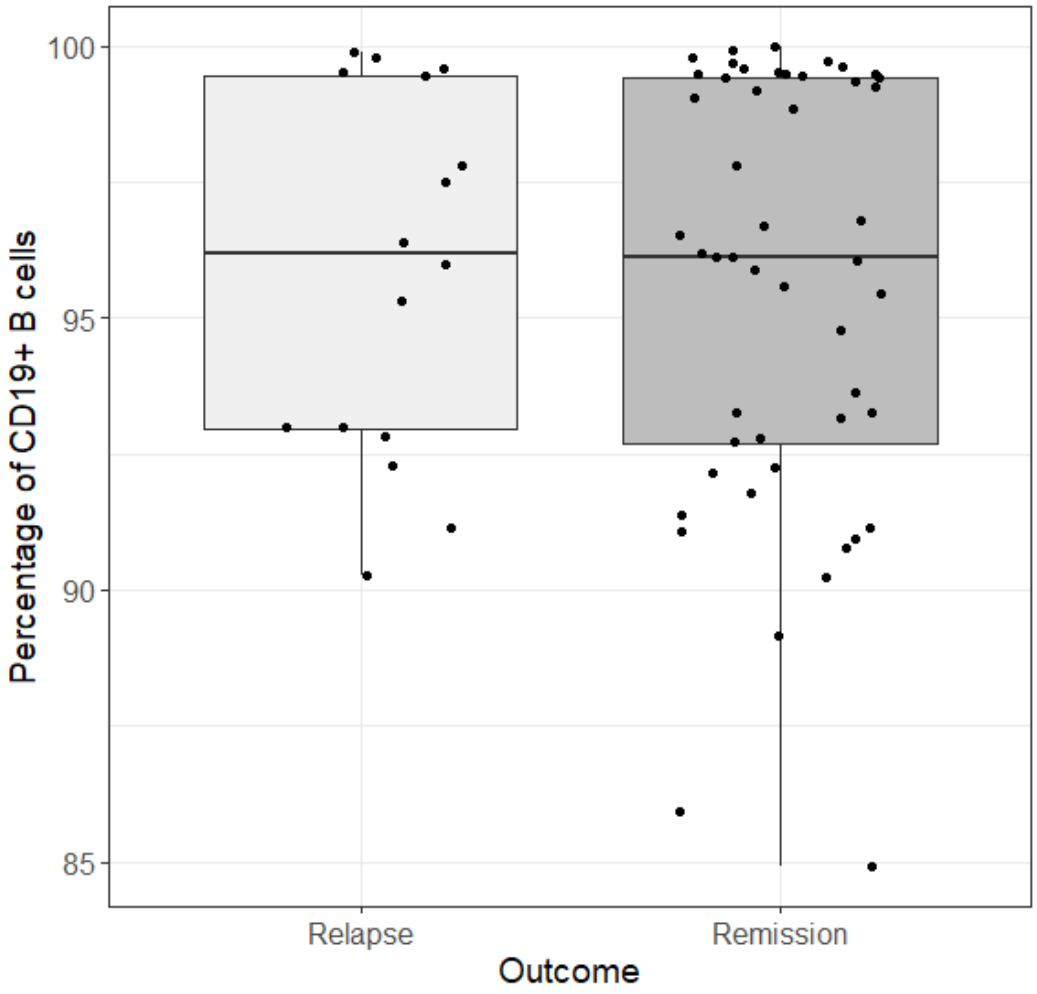


Figure 7.5 – Percentage purity of $CD19^{+}$ B cells present in the isolated B cell samples in both relapse and remission outcome groups.

7.5 B cell immunophenotyping

The B lymphocyte subpopulations studied correspond to different stages of maturation, activation, and differentiation. They encompass the progressive stages of B cell lineage beginning with the exit of immature B cells from the bone marrow to the development of mature, activated B cells (Chapter 1: Introduction).

The different B cell subpopulations and their phenotypes studied in this flow cytometry experiment are presented in Table 7.1.

There were two different phenotypes of B regulatory cells studied. The true origin and phenotype of B regulatory cells remains controversial with no robust B regulatory cell-specific transcription factor identified and most studies suggesting they are

unlikely to originate from a unique cell lineage and instead differentiate from various B cell subsets upon environmental stimulation (*Rosser EC et al. 2015*). However, the phenotypes used in this study are the common phenotypes previously identified to produce the anti-inflammatory cytokines, IL-10 and transforming growth factor beta, and have been previously studied in autoimmune disease including GD (*Iwata Y et al. 2011, Karim MR et al. 2019, Stożek K et al. 2020*)

B cell subset	Phenotype
Transitional	$CD19^{+}$ $CD38^{++}$ $CD24^{++}$ $CD5^{+}$ $CD27^{-}$
Naïve	$CD19^{+}$ IgD^{+} $CD27^{-}$
Plasmablast	$CD19^{+}$ $CD27^{++}$ $CD38^{++}$ $CD20^{-}$
Memory (switched)	$CD19^{+}$ $CD27^{+}$ $IgD^{-}CD38^{low}$
Memory (unswitched)	$CD19^{+}$ $CD27^{+}$ $IgD^{+}CD38^{low}$
Double negative	$CD19^{+}$ $CD27^{-}$ IgD^{-}
B regulatory cells 1	$CD19^{+}$ $CD27^{+}$ $CD24^{++}$
B regulatory cells 2	$CD19^{+}$ $CD5^{+}CD1d^{+}$

Table 7.1 – The phenotypes of B cell subpopulations studied by flow cytometry.

7.5.1 B cell subpopulations and clinical characteristics

Analysis was performed to investigate for any relationships between age, gender, and smoking status with B cell subpopulations. Age was only associated with double negative B cells, where there was a weak positive correlation found (P=0.02; rs=0.28) (Figure 7.6).

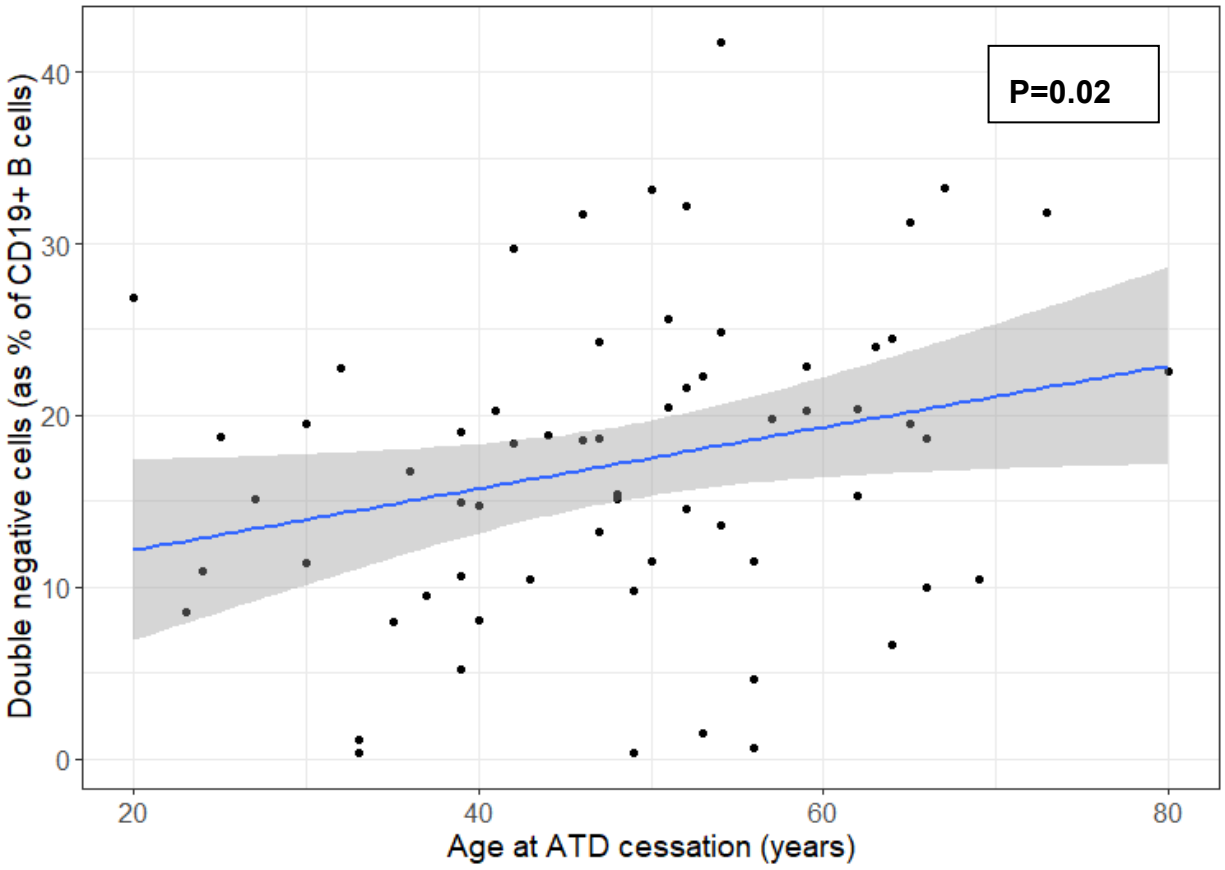


Figure 7.6 – Correlation between the double negative B cells and age at ATD cessation

None of the B cell subpopulations were associated with gender of the patient. Smoking was associated with an increased percentage of circulating Breg1 cells (P=0.04) (Figure 7.7).

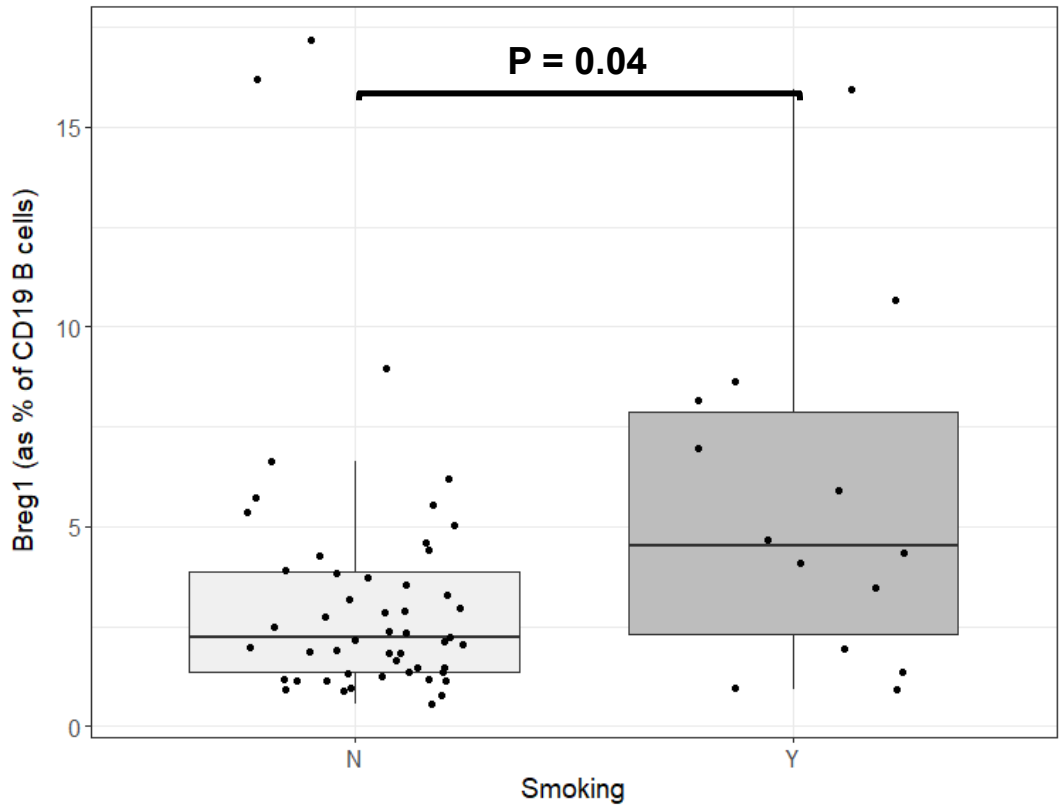


Figure 7.7 – Percentage of Breg1 cells present in the isolated B cell samples in the smoking and non-smoking cohorts

7.5.2 B cell subpopulation analysis

Correlation analysis was undertaken to determine the relationships between the peripherally circulating B cell subpopulations (Table 7.2).

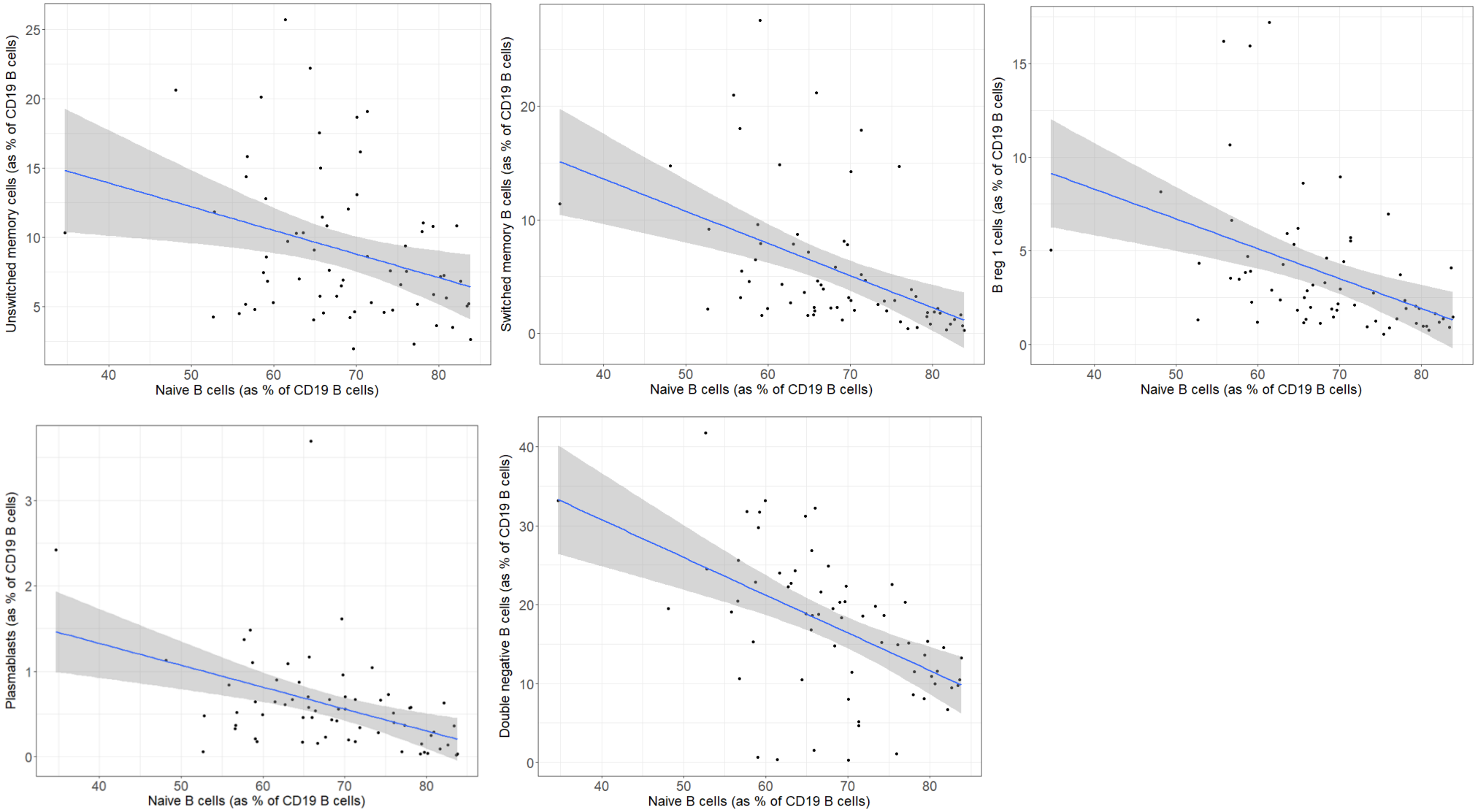
B cell subpopulation		Transitional	Naïve	Unsw memory	Sw memory	Plasma-blast	Double negative	B reg cells 1	B reg cells 2
Transitional	rs	.	0.3	-0.11	-0.19	0.04	-0.05	-0.33	0.04
	P		**0.01***	0.37	0.13	0.78	0.67	**0.007***	0.76
Naïve	rs		.	-0.29	-0.61	-0.46	-0.55	-0.54	-0.02
	P			**0.01***	**$5.9x10^{-8}$***	**0.0001***	**$1.2x10^{-6}$***	**$2.9x10^{-6}$***	0.87
Unswitched memory	rs			.	0.28	0.34	-0.36	0.48	-0.06
	P				**0.02***	**0.005***	**0.003***	**$4.6x10^{-5}$***	0.63
Switched memory	rs				.	0.53	0.04	0.64	-0.12
	P					**$4.5x10^{-6}$***	0.75	**$1.3x10^{-8}$***	0.33
Plasmablast	rs					.	0.12	0.28	-0.07
	P						0.36	**0.02***	0.57
Double negative	rs						.	-0.1	0.05
	P							0.43	0.64
B reg cells 1	rs							.	0.08
	P								0.54
B reg cells 2	rs								.
	P								

Table 7.2 – Spearman's correlation of each B cell subpopulation. Spearman's rho correlation coefficients (rs) and P value (P) are presented for each of the correlations.

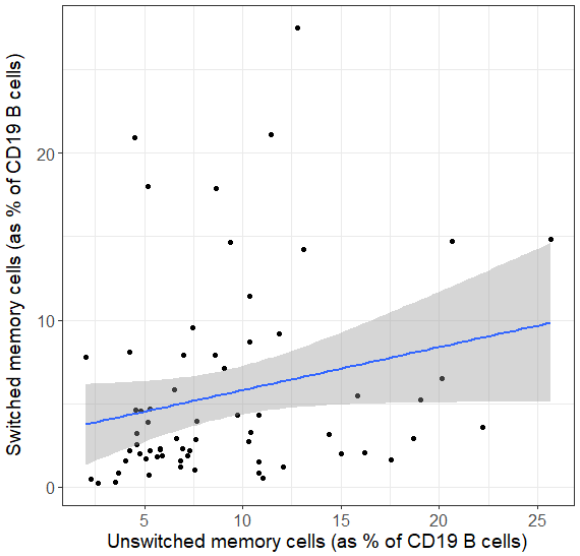
There were 14 significant correlations between the B cell subpopulations, with a mixture of positive and negative associations (Figure 7.8). The naïve B cell subpopulation was negatively correlated to all other B cell subsets (Figure 7.8(**A**)), apart from the B regulatory 2 phenotype. Both switched and unswitched memory cell populations were positively correlated to each other, plasmablasts and B regulatory 1 cells (Figure 7.8(**B,C**)). In addition to the correlations described above, the B regulatory 1 cells were positively correlated to plasmablasts and negatively correlated to transitional cells (Figure 7.8(**D**)).

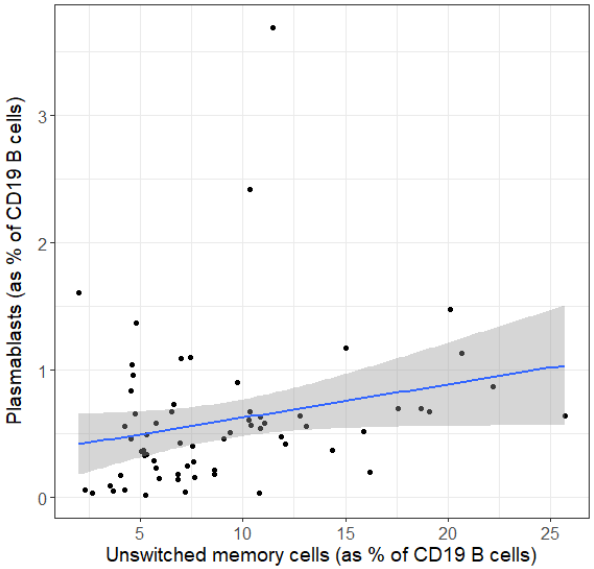
On multiple test correction (Benjamini & Hochberg), many of the correlations described remained significant. Namely, the naïve B cells with switched memory cells ($P=4.2 \times 10^{-7}$), plasmablasts ($P=0.0008$), double negative ($P=9.6 \times 10^{-5}$), and Breg1 cells ($P=2.3 \times 10^{-4}$). In addition, the unswitched memory cells remained correlated to the plasmablasts ($P=0.04$), double negative ($P=0.024$) and Breg 1 cells ($P=0.0004$). The switched memory cells also remained significantly correlated to the plasmablasts ($P=0.0004$) and Breg 1 cells ($P=1.04 \times 10^{-7}$).

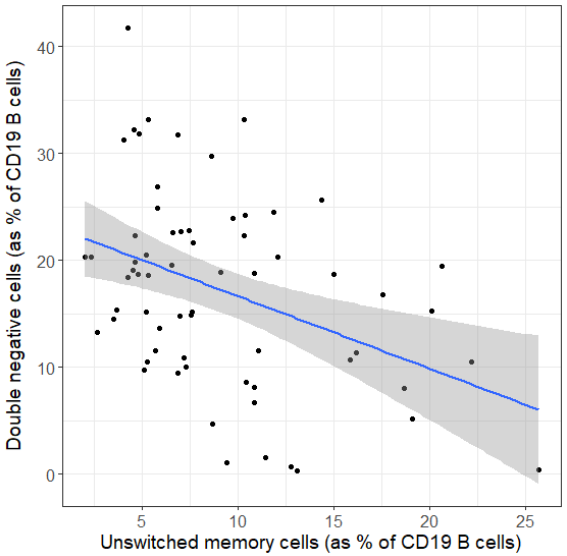
A *Naïve B cells*

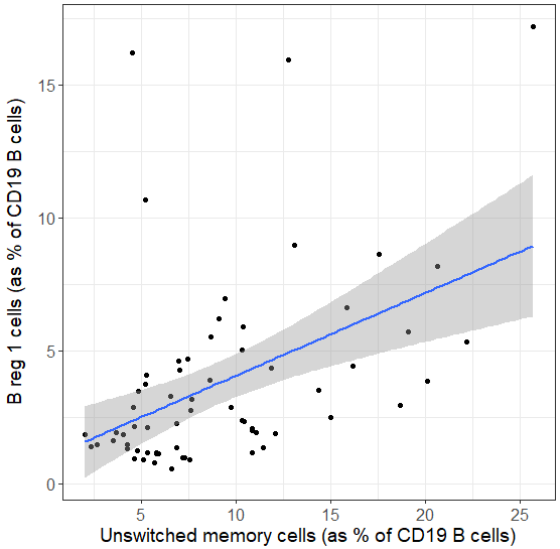


B *Memory B cells (unswitched)*

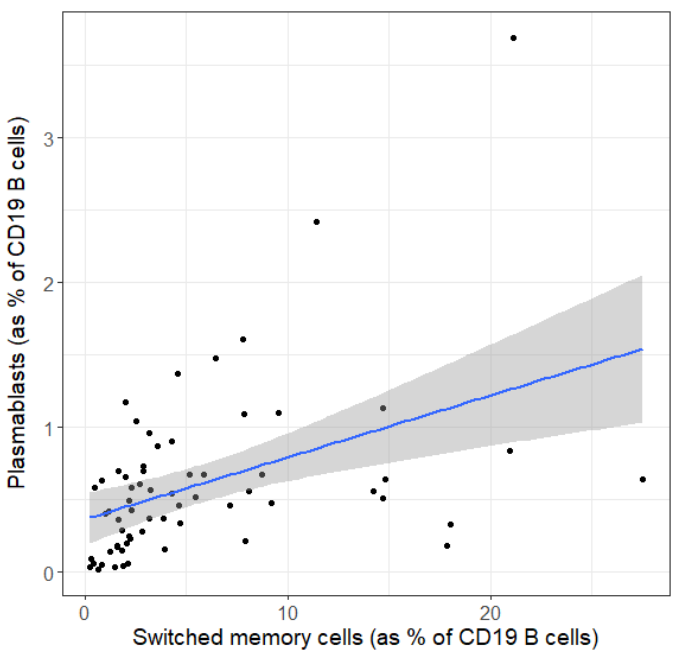


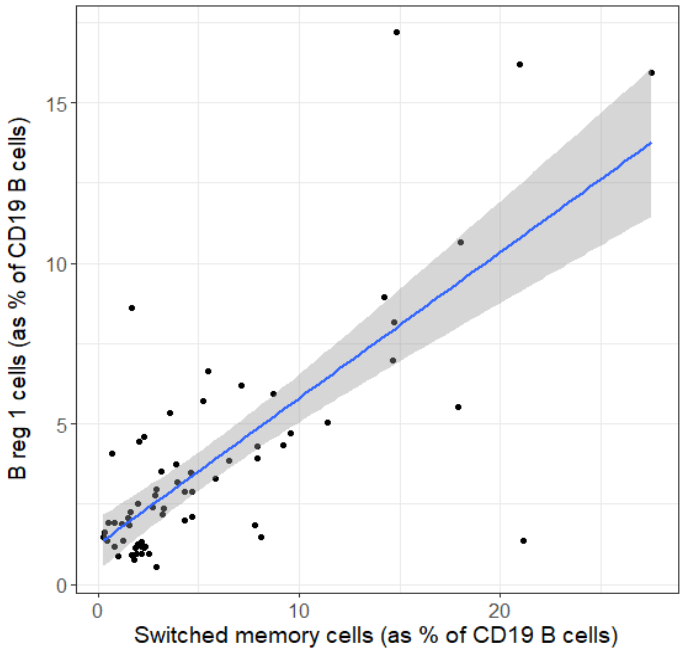




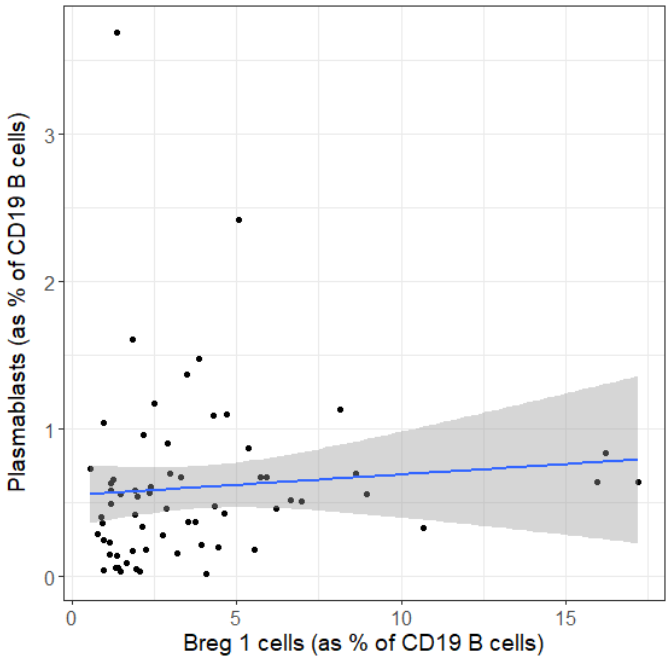


C *Memory B cells (switched)*





D *B regulatory 1 cell*



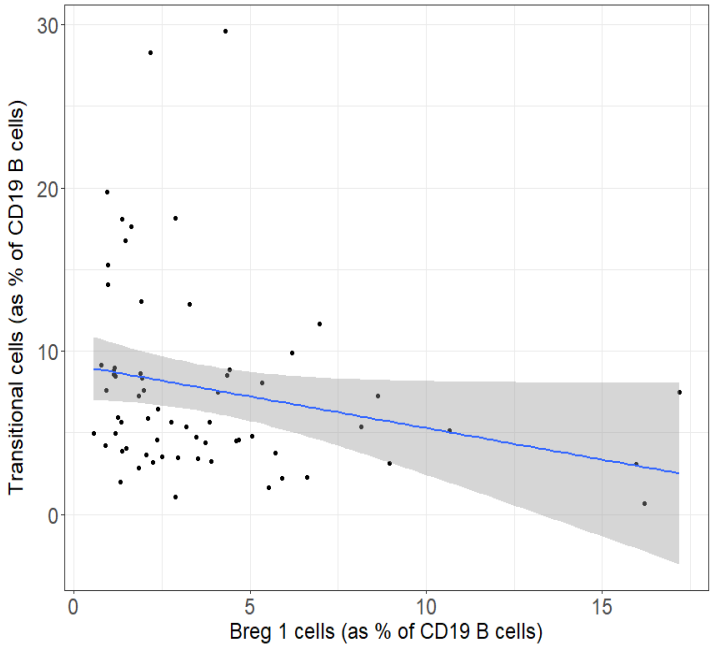


Figure 7.8 – Correlations between the different B cell subpopulations as determined by flow cytometry. (**A**) Naïve (**B**) Memory cells (unswitched) (**C**) Memory cells (switched) (**D**) B regulatory 1 cell phenotype.

7.5.3 B cell subpopulations with thyroid status and C-reactive protein (CRP)

There was no correlation found between the absolute B cell number or the different B cell subpopulations with thyroid hormone or autoantibody concentrations (FT4, FT3, TRAb, TPO Ab) at the time of ATD withdrawal. Furthermore, none of the B cell subpopulations correlated with CRP.

7.5.4 B cell subpopulations in the relapse and remission groups

The isolated $CD19^+$ B cells included differing individual quantities of the various subpopulations, presented in Table 7.3. As expected, overall, the naïve B cells comprised the vast majority of the $CD19^+$ B cell compartment with a median percentage of 68 (IQR 62 – 76%), whilst other cell populations such as plasmablasts were rare, with a median percentage of 0.5 (IQR 0.2 – 0.7%).

The association between B cell subpopulations and outcome of GD was assessed by univariate binary logistic regression analysis. The largest difference between B cell subpopulations and outcome was in the transitional cell population, with a larger percentage of transitional cells present in the relapse group (8.8% vs. 5.4%). However, this did not reach statistical significance in the univariate analysis (P=0.16) (Figure 7.9).

B cell subpopulation	Relapse group (n=16) (% of $CD19^+$ cells)	Remission group (n = 49) (% of $CD19^+$ cells)	B	ORrelapse	95% CI ORrelapse	Univariate P value
Median Transitional	8.8	5.4	-0.06	0.94	0.86 – 1.02	0.16
Median B reg cells 2	3.49	3.75	-0.15	0.86	0.66 – 1.11	0.25
Median double negative	17.08	18.56	0.02	1.01	0.95 – 1.08	0.62
Median memory (unswitched)	6.89	7.56	0.02	1.02	0.91 – 1.14	0.71
Median plasmablast	0.54	0.49	-0.12	0.88	0.35 – 2.24	0.79
Median naïve	69.1	68.4	-0.007	0.99	0.94 – 1.05	0.80
Median memory (switched)	3.07	3.15	-0.007	0.99	0.90 – 1.09	0.88
Median B reg cells 1	2.6	2.5	-0.002	0.99	0.85 – 1.17	0.98

Table 7.3 – The B cell subpopulations as a percentage of the total $CD19^+$ B cell population in both outcome groups, and their association with Graves' disease outcome by univariate logistic regression.

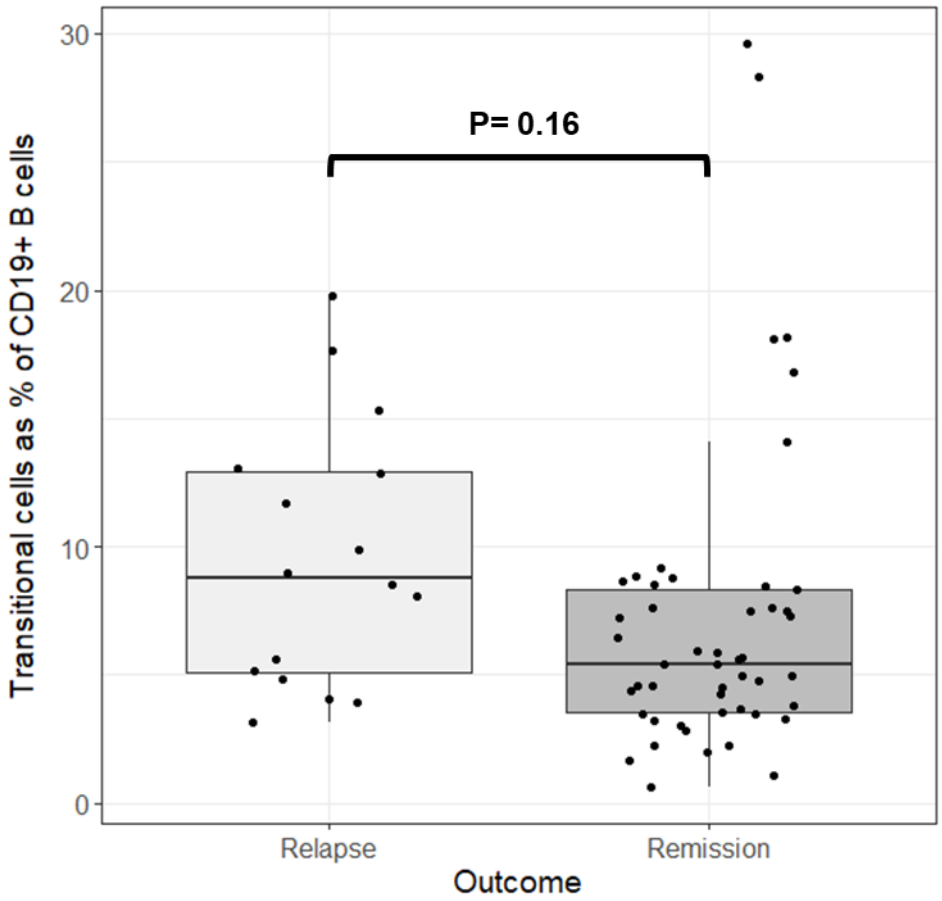
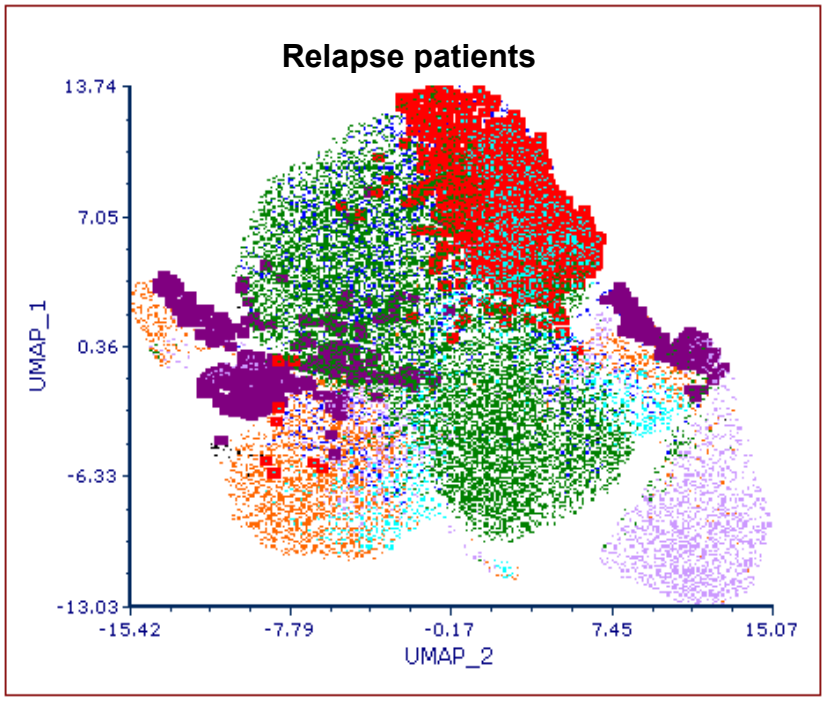


Figure 7.9 – Transitional cells as a percentage of the CD19⁺ B cells present in both relapse and remission outcome groups.

The B cell subpopulations were manually gated and then presented using the unsupervised clustering and dimension reduction programme (UMAP) (FCS Express™ 7) to generate distinct clusters based on cell surface marker expression comparing the relapse and remission groups (Figure 7.10). There is no clear visible distinction between the cell clusters of the relapse and remission groups.

A



B

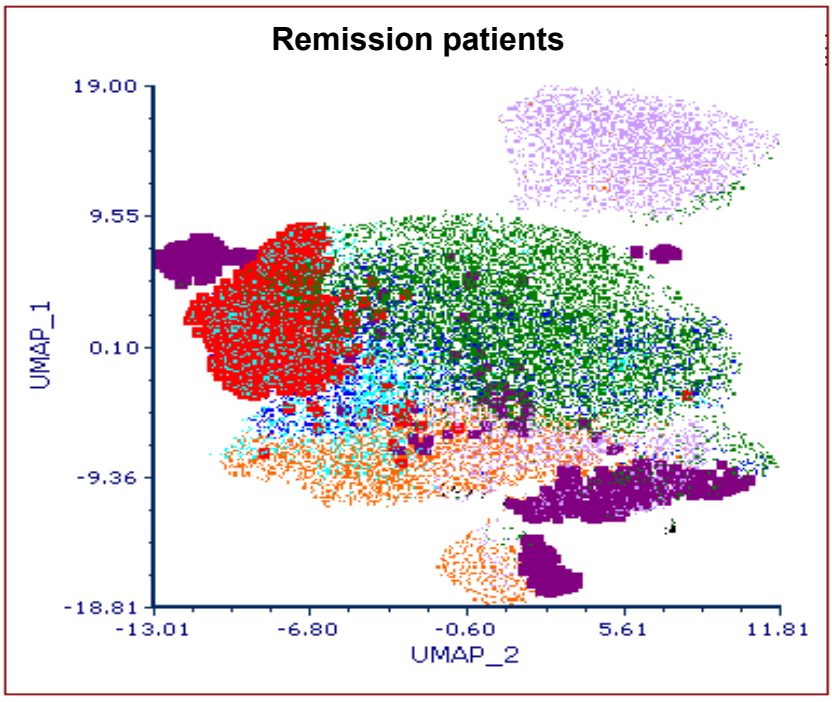


Figure 7.10 – UMAP (Uniform Manifold Approximation and Projection) plot created using the concatenated isolated B cells split into the relapse (**A**) and remission (**B**) patient groups, demonstrating clustering of the main cell populations.

7.6 Survival analysis

The association between the B cell subpopulations at baseline and time-to-relapse following ATD withdrawal was assessed using univariate Cox regression analysis (Table 7.4). Using Schoenfeld Residuals, there was no significant departure from proportional hazards observed for any of the B cell subpopulations.

B cell subpopulation	B	HRrelapse	95% CI HRrelapse	Univariate P value
Transitional	0.05	1.1	0.99-1.12	0.12
B reg cells 2	0.04	1	0.97-1.13	0.26
Double negative	-0.014	0.99	0.93-1.04	0.62
Memory (unswitched)	-0.019	0.98	0.89-1.09	0.72
Plasmablasts	0.096	1.1	0.51-2.39	0.81
Naïve	0.01	1	0.96-1.06	0.71
Memory (switched)	-0.001	1	0.92-1.08	0.98
B reg cells 1	-0.008	0.99	0.86-1.14	0.90

Table 7.4 – Association of B cell subpopulations at baseline with timing of Graves' disease relapse following ATD withdrawal, using a univariate Cox regression model.

Although in the univariate cox regression analysis none of the B cell subpopulations were significantly associated with time-to-relapse, there was a trend towards the negative correlation of transitional cells and the number of days it took a patient to relapse by Spearman's correlation (P=0.14; rs= -0.38) (Figure 7.11). Therefore, the transitional cell population was entered into a multivariate cox regression analysis to include age, gender, smoking status, goitre size and TRAb at ATD withdrawal (Table 7.5).

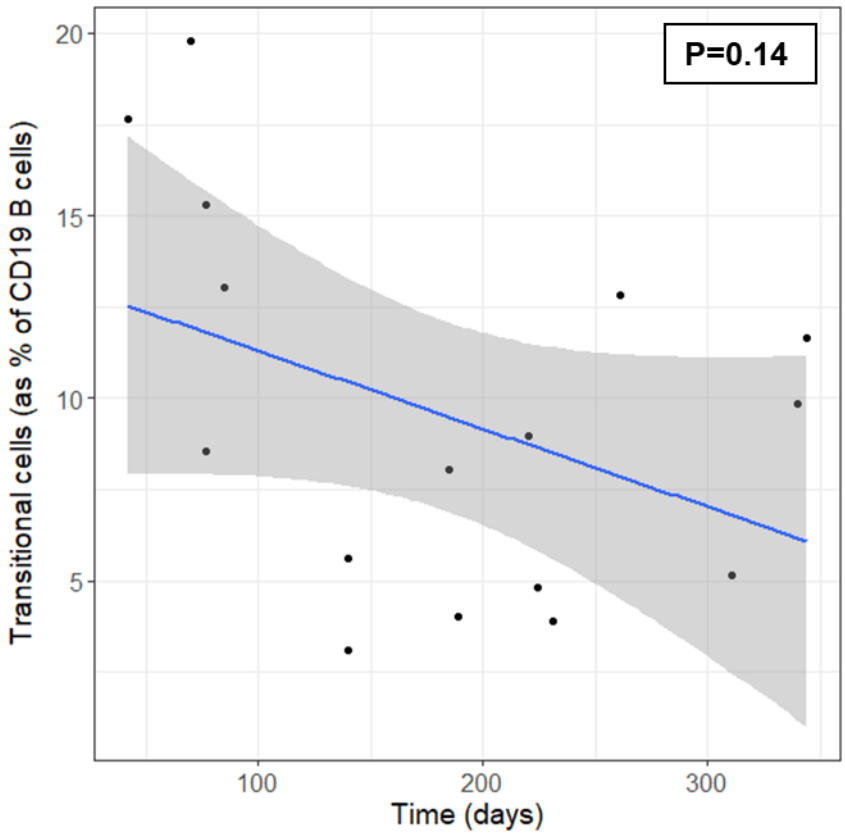


Figure 7.11 – Correlation between transitional B cells (as percentage of $CD19^+$ B cells) and time to relapse (in days).

Variable	B	HRrelapse	95% CI HRrelapse	Multivariate P value
TRAb on stopping ATD	0.35	1.4	1.01 – 1.99	**0.04***
Gender (Male)	0.90	2.46	0.62 – 9.6	0.20
Transitional cells	0.05	1.05	0.97 –1.14	0.19
Goitre size				
1	0.50	1.6	0.39 – 6.9	0.50
2	1.50	4.5	0.83 – 24.8	0.08
Current smoker	-0.3	0.75	0.20 – 2.82	0.68
Age	0.02	1.01	0.97 – 1.07	0.50

Table 7.5 – Association of baseline clinical variables and transitional B cells with timing of Graves' disease relapse following ATD withdrawal, using a multivariate Cox regression model. Out of the variables entered into the multivariate cox regression analysis, it was only the TRAb level prior to ATD withdrawal that was significantly associated with the time it took patients to relapse (Figure 7.12).

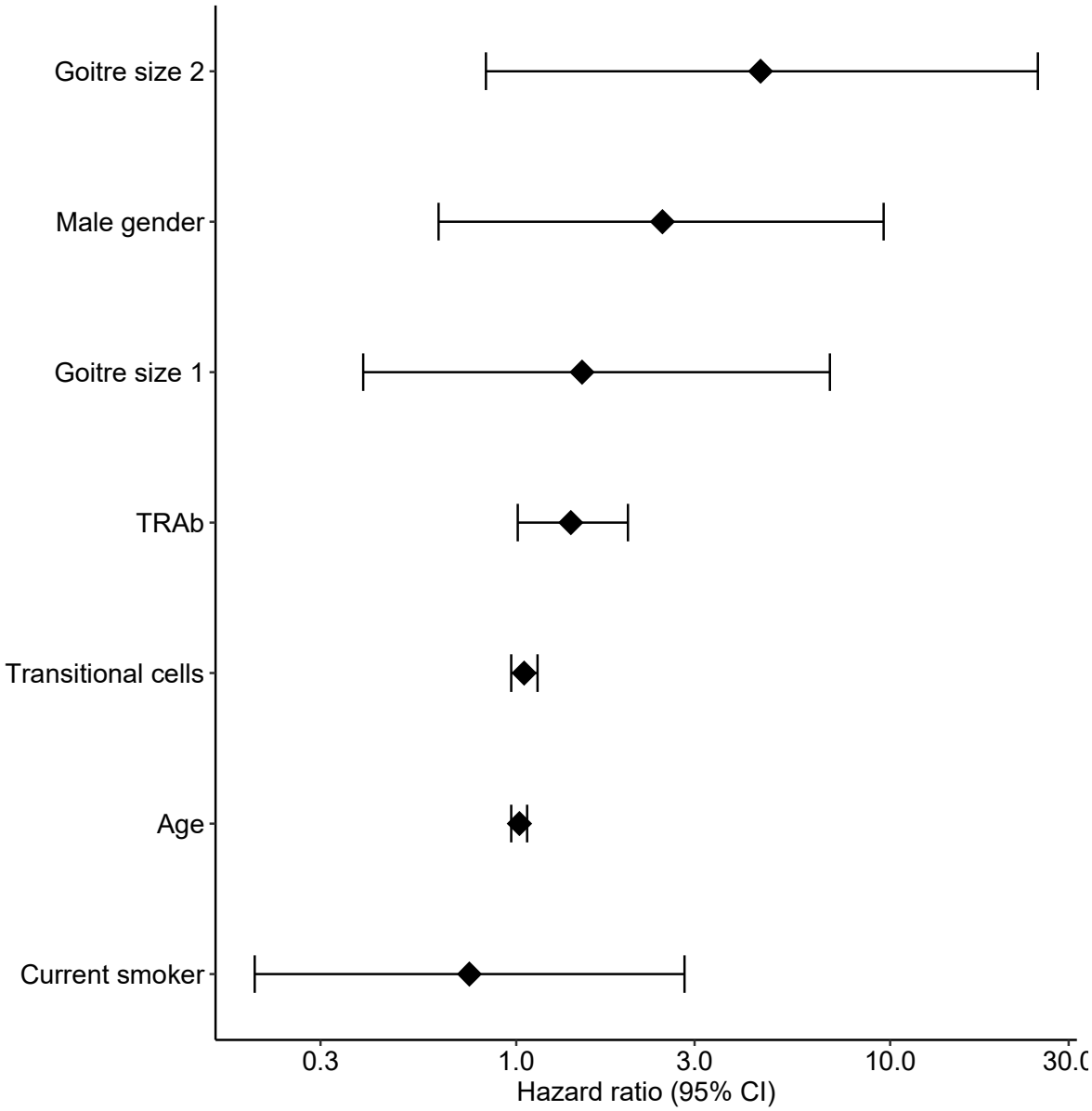


Figure 7.12 – Forest plot presenting the association of baseline clinical variables and transitional B cells with timing of Graves' disease relapse following ATD withdrawal, using a multivariate Cox regression model.

7.7 Discussion

B cell subsets have distinct roles in modifying the humoral immune response, either through secretion of pro-inflammatory cytokines or alternatively, as in the case of B regulatory cells, contributing to the maintenance of peripheral tolerance through the secretion of anti-inflammatory cytokines. The phenotype and frequency of various B

cell subpopulations has been studied in GD, although there is a paucity of data investigating their relationship with clinical outcome.

In this Chapter, humoral immune activity was explored in GD using flow cytometry to gain insight into the circulating B cell subpopulations at the time of ATD withdrawal and investigate a potential pathophysiological link. The association between the B cell subpopulations with patient's characteristics, thyroid status and clinical outcome was explored.

7.7.1 Peripheral B cell compartment in Graves' disease

The composition of the peripheral B cell compartment generally involves naïve B cells as the largest component (58-72%) with plasmablasts occurring in the lowest frequency (<5%) (*Morbach et al. 2010*). This is consistent with the findings in my study where naïve B cells and plasmablasts contributed on average 68% and 0.5% to the total $CD19^+$ B cell pool, respectively. The relative rarity of plasmablasts and the very low frequency of plasma cells within the peripheral blood compartment, means that direct analysis of these precursors to antibody-secreting cells by flow cytometry is challenging.

The absolute quantity of circulating B cells has been implicated in the immunopathology of GD, with increased total B cells observed in both treated and untreated GD compared to healthy controls (*Liu HY et al. 2022, Liu Y et al. 2022, Iwatani et al. 1989, Bossowski et al. 2003*). Although there were no healthy controls included in my study, the median percentage of B cells (12%) (as a percentage of the total lymphocyte count) in the GD patients was comparatively lower than the median percentage of B cells in healthy controls (14.4%), euthyroid (17.7%) and hyperthyroid (22.9%) GD patients described in a previous study (*Iwatani et al. 1989)*. However, elevated total B cells is not a consistent finding in GD, and other studies suggest that rather than just absolute B cell numbers, it is defects in the functional activity of specific B cell subsets that may be involved in the pathogenesis of GD (*Cao Y et al. 2022, Van der Weerd K et al. 2013).*

7.7.2 B cell subpopulations and Graves' disease patient characteristics

Age, gender, and smoking status are all factors that have been described to influence the composition of the peripheral blood lymphocyte compartment (*Morbach H et al. 2010, Feng R et al.* 2022*, Pinto TNC et al. 2024).* Ageing has been reported

to alter the composition of the B cell repertoire and contribute to the development of autoimmunity (*Mouat IC et al. 2022*). Specific 'Age-associated B Cells' (ABCs) with the phenotype, $CD19^+$ $BAFFR^+$ $CD11b^+$ $CD11c^+$ $CD21^-$ ($T\text{-}bet^+$), are a unique population of class-switched, antigen-specific memory cells reported to accumulate with age and have been found to be expanded in autoimmune disease (RA and SLE) (*Adlowitz DG et al. 2015, Wang S et al. 2018*).

The composition of the ageing B cell compartment has been reported to involve a reduction in the naïve B cells and a gradual increase in the proportion of switched and non-switched memory B cells (*Morbach H et al. 2010).* However conversely, other studies report a decrease in the percentage of switched memory B cells, and a significant increase in the percentage of naïve and double negative ($CD27^-IgD^-$) B cells (*Frasca D et al. 2017*). Indeed, in my study although there was no relationship found with age and naïve or memory cells, there was a positive correlation found between age and double negative B cells. Double negative B cells have been proposed to be an intermediate cell state between ABCs and plasma cells and are shown to accumulate in chronic inflammation and autoimmunity (*Jenks SA et al 2020*). These cells have demonstrated an ability to secrete antibody with autoimmune reactivity which may explain how ageing could be a factor in the development of autoimmunity (*Frasca D et al. 2019*).

Sex-specific differences have been reported to influence composition of the peripheral B cell compartment, with females reported to have a higher proportion of naïve B cells and decreased switched memory B cells compared to males (*Feng R et al. 2022*). However, studies have also reported no gender-specific differences in B cells (*Kverneland AH et al. 2016*), which is consistent with my study which found no relationship between the total B cell count or any of the B cell subpopulations and gender.

Exposure to cigarette smoke is reported to affect adaptive immunity partly by altering the composition of the B cell compartment, with higher percentages of switched memory B-cells and decreased naïve B-cells found in smokers compared to non-smokers (*Pinto TNC et al. 2024*). However, in my study, smoking was not found to correlate with naïve or memory B cells, but current smokers were found to have increased B regulatory cells ($CD19^+$ $CD27^+$ $CD24^{++}$) compared to non-smokers. There is limited literature on smoking and B regulatory cells, although a previous study did find a decrease in IL-10+ B regulatory cells in smokers (*Jacobs M et al.*

2022). An increase in B regulatory cells has been described to occur in response to a pro-inflammatory environment which is well-described in smokers and could explain the positive association seen in my study (*Rosser EC et al. 2021*). Of note, none of the B cell subpopulations correlated with CRP which is in-keeping with previous reports (*Li J et al. 2021*).

7.7.3 B cell subpopulations and thyroid status

Thyroid hormones are essential for the normal development, differentiation, and production of B cells from the bone marrow (*Torimoto K et al. 2022*). Studies have shown a positive association between FT4 and FT3 with the total B cell count and both naïve and memory B cell subsets (*Jaeger et al. 2021, Bossowski A et al. 2003).* Furthermore, levels of CD5+ B cells have been demonstrated to correlate with thyroid hormones and TRAb in untreated hyperthyroid GD (*Iwatani Y et al. 1989*). However, my study did not demonstrate any association between the thyroid hormones and the different B cell subpopulations. This may reflect the fact that the study patients were on ATD treatment and euthyroid at the time of the baseline sample, with most having a FT4 and FT3 in the normal range and therefore the thyroid hormones may have had minimal impact on B cell homeostasis at this time.

7.7.4 Relationships between B cell subpopulations

Humoral immune homeostasis is maintained through the balance of various B cell phenotypes working synergistically or antagonistically. In my study the naïve B cells were negatively correlated to most of the other B cell subsets which is similar to patterns observed in a previous B cell immunophenotyping study which found strong negative correlations between naïve B cells with plasmablasts (rs -0.82) and switched memory cells (rs -0.77) (*Li J et al. 2021*). This may represent the balance between a mature and immature B cell compartment with the presence of more circulating naïve cells indicating less cells are differentiating into mature B cells such as memory cells and plasmablasts.

Unswitched and switched memory cells were both positively correlated to the plasmablast population, which likely reflects the fact that upon antigen exposure memory cells are activated and differentiate into plasmablasts and therefore an expanded memory cell population may lead to greater production of plasmablasts.

Unswitched and switched memory cells were also found to correlate to Breg1 cells positively and negatively, respectively. Both switched and unswitched memory cells have been previously found to negatively correlate with B regulatory cells (*Li J et al. 2021*). As discussed previously, it is likely that Breg cells differentiate from various B cell subsets upon environmental stimulation, including from the $CD27^+$ memory cells (*Jansen K et al. 2021*). Indeed, $CD27^+$ memory cells have been found to be a major source of IL-10 producing immunoregulatory B cells in RA (*Bankó Z et al. 2017*). Therefore, their relationship with the memory cell populations may be dynamic and change depending on the immune environment.

7.7.5 B cell subpopulations and Graves' disease outcome following ATD treatment

The relationship between B cell subpopulations and outcome in GD has not previously been described. In my study, there was no evidence of an association with absolute numbers of circulating B cells or between any of the B cell subpopulations at ATD withdrawal and outcome in GD. However, it is possible that samples taken at baseline, which are reflecting a state of inactive, euthyroid GD, may not provide insight into the B cell compartment of a relapsing patient that may rapidly change once the ATD is withdrawn.

Indeed, transitional B cells have been demonstrated to be significantly increased in newly diagnosed untreated GD patients when compared to healthy controls (*Van der Weerd K et al. 2013*), and serial analysis of B cells following administration of ATD observed that total B cell numbers are reduced in treated euthyroid compared to treated hyperthyroid GD patients (*Corrales et al. 1996*). In addition, methimazole treatment in paediatric GD has been shown to significantly alter the B regulatory cell population from baseline and have prognostic implications (*Grubczak K et al. 2021*). This indicates that both thyroid status and ATD treatment can affect B cell subpopulations and therefore samples taken prior to ATD, either at initial diagnosis or after ATD withdrawal, might be more informative in terms of understanding a possible pathological role of B cell subsets and their relationship with outcome in GD.

Given the relationship between TRAb levels and GD relapse, it may be expected that circulating plasmablasts, which differentiate into TRAb-secreting plasma cells, would correlate with relapse. However, the lack of association observed in this study between plasmablasts and outcome suggests that the plasma cells secreting TRAbs

are not located in the blood, but instead could be located in the thyroid or bone marrow.

7.8 Limitations

The samples used for analysis in this flow cytometry study were all from the timepoint of ATD withdrawal and therefore they only provide a limited depiction of B cell subset activity in treated euthyroid GD patients. Future analysis of the 6–10-week PBMCs taken after cessation of ATD may provide further insight into the activity of B cell subpopulations in those patients that relapsed. Although this may provide mechanistic insight into GD relapse, the timing of a sample to determine predictive biomarkers of outcome would preferably be taken prior to ATD withdrawal.

B cell subpopulation data from euthyroid healthy control samples would have been valuable to compare with the euthyroid GD cohort to assess for any differences in the peripheral B cell compartment that may provide mechanistic insight into Graves' immunopathology and the potential immunomodulatory influence of ATD treatment on B cell subsets.

As discussed previously the phenotype of B regulatory cells remains controversial and difficult to define. Although most studies describe $CD19^{+}$ $CD27^{+}$ $CD24^{++}$ as the commonest B regulatory phenotype, functional studies or staining for anti-inflammatory cytokines such as IL-10 may have increased the sensitivity in detecting B regulatory cells in this study. In addition, immunophenotyping B cells from thyroid tissue in conjunction with peripheral B cells may provide additional insight into the distribution of B cells in GD.

7.9 Summary

The balance between various B cell subpopulations is crucial to maintain B cell homeostasis. In the context of inflammation and autoimmunity, this balance can be disrupted and reflect an individual's humoral immune response to pathology. This Results Chapter presents the findings of the FACs analysis and highlights a lack of predictive biomarkers of outcome in the peripheral B cell subsets at the time of ATD withdrawal. This may reflect a lack of active humoral immunity in treated euthyroid GD. The significant correlations between different B cell subpopulations provide

insight into the potential B cell dynamics involved in maintaining immune homeostasis.

Chapter 8: Results 5 – CD19$^+$ B cell RNAseq

8.1 Introduction

In the previous results chapters I have discussed various components of the humoral immune system that were studied with the aim of gaining mechanistic insight into the immunological pathways leading to GD relapse. Humoral immunity is a dynamic process and capturing the specific contributing factors involved in the relapse of GD is challenging. CD19$^+$ B cells are fundamental in the pathogenesis of GD (Chapter 1: Introduction).

The transcriptional profiling of CD19$^+$ B cells at the timepoint of stopping ATD aimed to capture the precise variations in gene expression to enable a molecular understanding of the factors determining relapse or remission in GD. Ultimately, this may enable the use of an individual's transcriptomic profile to predict their disease outcome with translational potential for clinical management. Furthermore, in-depth pathway analysis of the most differentially expressed genes (DEG) observed in relapse and remission patients may identify biomarkers and molecular pathways that could uncover immunopathological insights into GD relapse and reveal novel therapeutic targets.

Previous gene expression profiling of B cells in GD has utilised microarray platforms, however next-generation sequencing allows improved sensitivity and an exploratory approach without the need for pre-defined transcripts.

In this results chapter, I will present the CD19$^+$ B cell RNAseq analysis. The aim of this analysis was to:

1. Explore differences in the CD19$^+$ B cell gene expression between the relapse and remission patients at the timepoint of stopping ATD.
2. Analyse whether transcriptomic patterns can predict outcome in GD.
3. Undertake pathway analysis to identify potential novel molecular mechanisms in the relapse of GD.

8.2 Quality control

8.2.1 Sample collection and processing

The samples used in this sequencing experiment were the blood samples collected at the timepoint of stopping ATD therapy. The freshly collected blood sample underwent processing to isolate the PBMCs, followed by positive selection of the $CD19^+$ B cells (Chapter 3: Methods). These $CD19^+$ B cells were then lysed and stored immediately at $^-80^oC$ until RNA batch extraction.

Blood samples were available for all 65 patients at the timepoint of stopping ATD and therefore sequencing data was available for all the study patients. The median time from taking the patient's blood to freezing of the lysate was recorded for all 65 patients and was 5.5 hours (4.5 – 6 (IQR), 3.25 – 7.5 (range)) (Figure 8.1). There was no difference in the time taken from sample collection to freezing of the lysate between the relapse and remission patient samples (P=0.12).

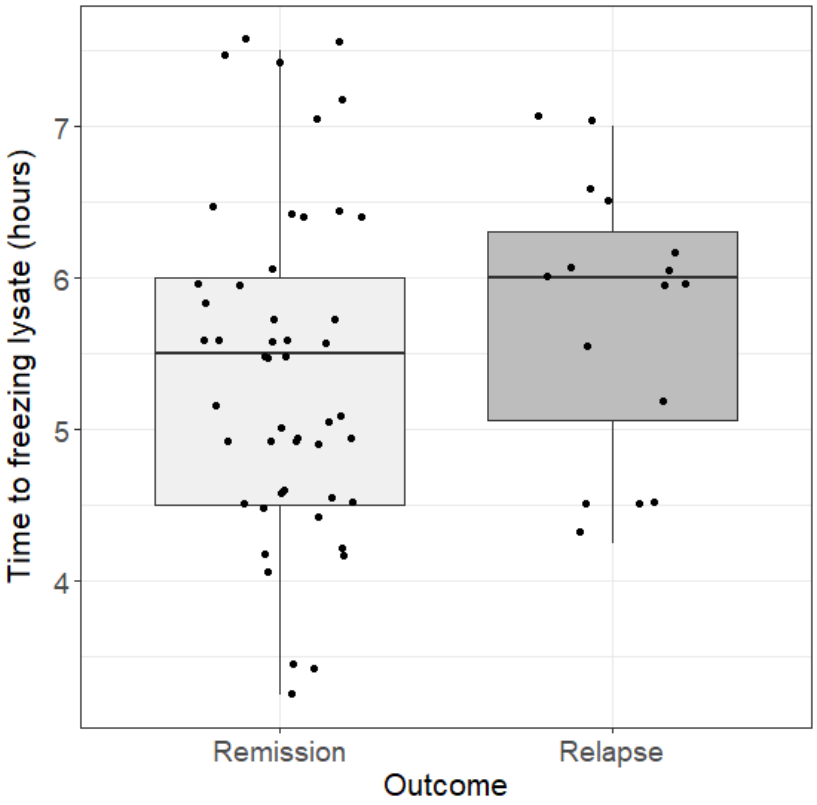


Figure 8.1 – Time to freezing of CD19+ B cell lysate from the time of sample collection in both outcome groups.

8.2.2 Cellular yield and purity

The cellular yield of the CD19⁺ B cells isolated from 100mls of blood was variable between patients, but there was no difference between patient outcome groups (P=0.34) (Figure 8.2).

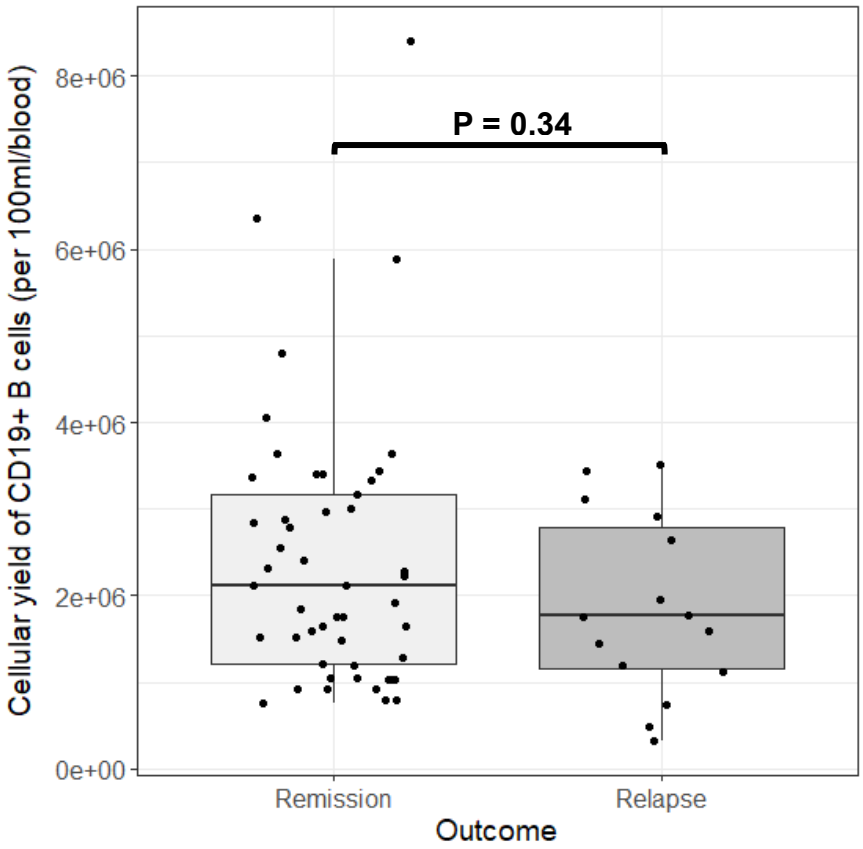


Figure 8.2 – The cellular yield of CD19⁺ B cells isolated in both outcome groups

The purity of the CD19⁺ B cells isolated are important in the context of interpreting the transcriptome. As reported previously in Chapter 7: Results 4 – FACS, the median percentage of CD19⁺ B cells in the isolated B cell sample was 96.1 (IQR 93 – 96%), with low levels of contaminating T cells and monocytes: median percentage 2.3 (IQR 1.4 – 4%) and 0.5 (0.2 – 0.9%), respectively. There was no significant difference observed in B cell purity between the relapse and remission groups (P=0.84).

8.2.3 RNA quantity and integrity

High quality, non-degraded RNA is essential to perform a robust RNA sequencing experiment (Chapter 3: Methods). The quality (integrity) and quantity of RNA was

measured by electrophoretic separation using a Tapestation™ 4200 machine (Agilent) (Figure 8.3A). RNA integrity was assessed using the RIN algorithm (Chapter 3: Methods). Part of the RNA integrity assessment by the Tapestation™ 4200 involves evaluating the 28S to 18S ribosomal RNA ratio, with a 2:1 ratio being considered the benchmark for intact RNA (*Schroeder A et al. 2006*) (Figure 8.3B).

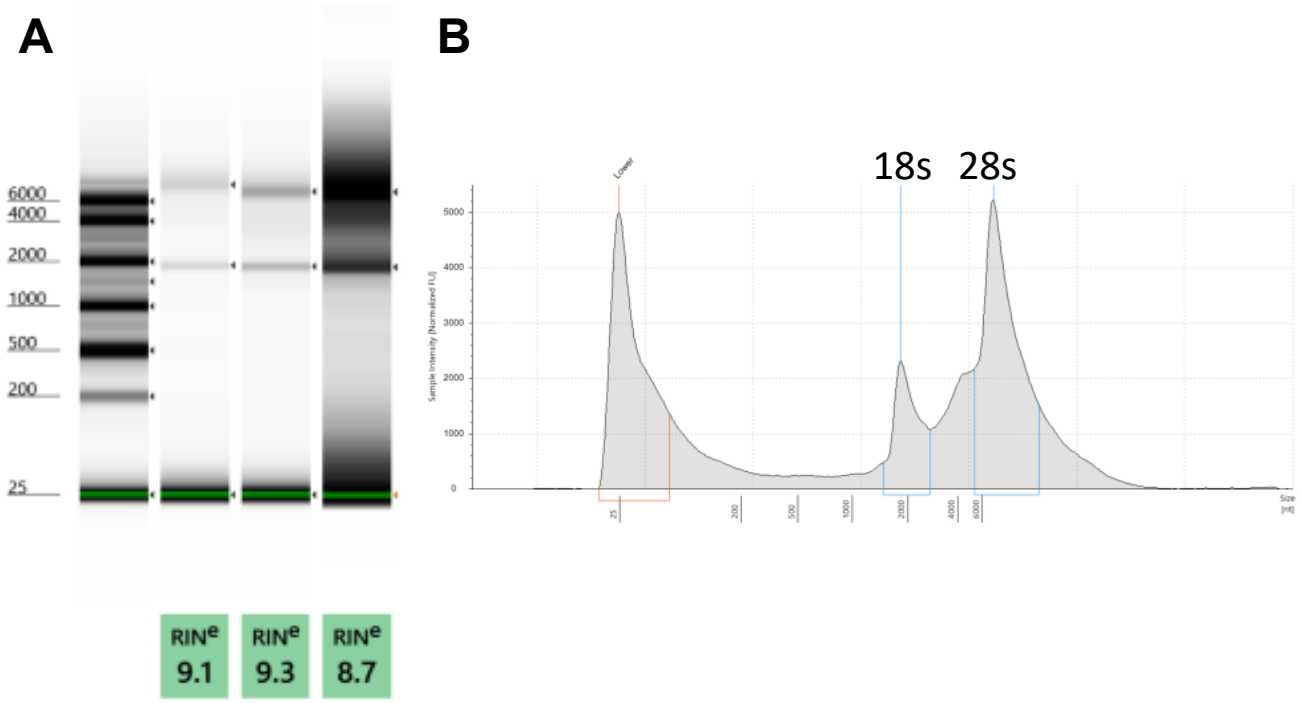


Figure 8.3 – An example of the assessment of RNA integrity for three of the study samples using the Tapestation™ 4200 with RIN^e values generated (**A**). Electrophoretogram demonstrating the 18s and 28s ribosomal peaks used to calculate the 28s/18s ribosomal RNA ratio (**B**).

The median RIN value for the RNA samples was 9.3 (9.2 – 9.5 (IQR), 5.2 – 10 (range)). Illumina® recommends using RNA with a RIN value of ≥ 8 and all these samples, apart from the single one valued at 5.2, had a RIN ≥ 8.9 (Figure 8.4A). This sample with a low RIN value was excluded from analysis. The fragment size distribution of RNA samples has a substantial influence on the library yield and is another determinant of RNA quality. It can be represented as the percentage of RNA fragments above 200 nucleotides by the DV_{200} (Distribution value) quality metric (*Matsubara T et al. 2020*). The median percentage DV_{200} value for the RNA samples was 91.5 (73.1 – 95.6 (IQR), 12.6 – 97.9 (range)) (Figure 8.4B).

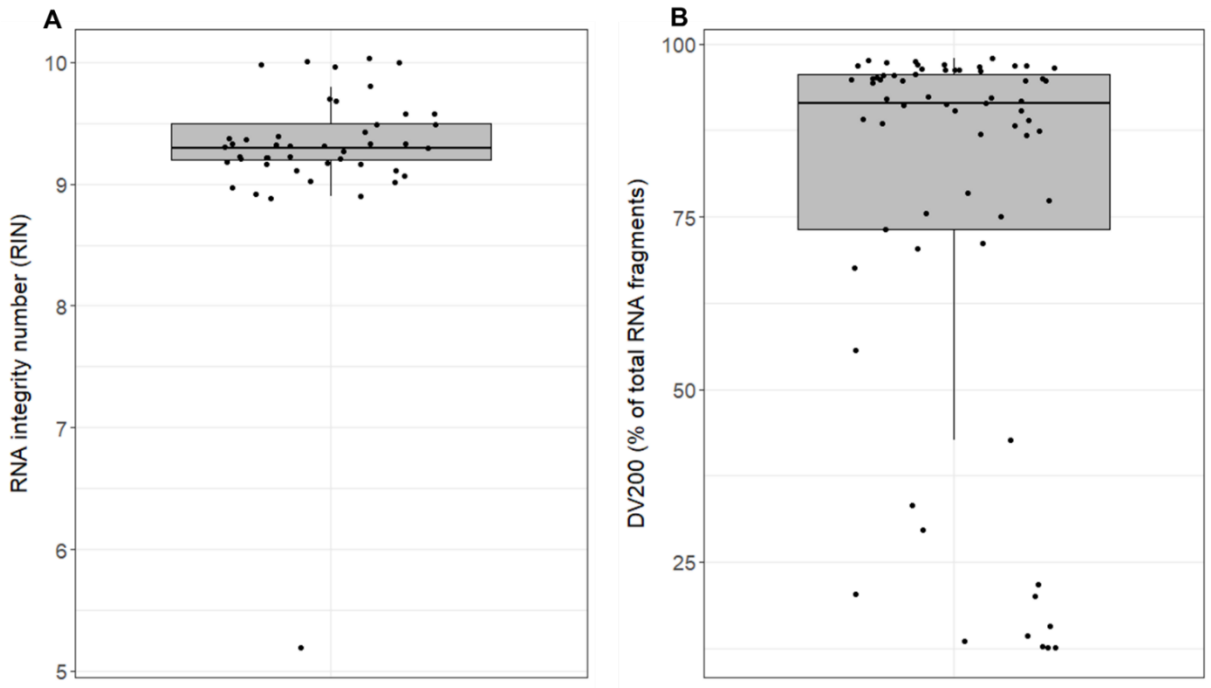


Figure 8.4 – Median RNA integrity number (RIN) for each sample (**A**). DV_{200} (% of total RNA fragments > 200 nucleotides) (**B**)

The median yield of RNA extracted was 20.6 ng/μL (5.6 – 68.6 (IQR), 1.3 – 2410 (range)). There was a positive correlation between the yield of B cells isolated from each patient and the quantity of RNA that was extracted ($P=6.5 \times 10^{-5}$, rs=0.48) (Figure 8.5). As some of the samples contained low quantities of RNA, mRNA sequencing was performed using a low input RNAseq platform (NEBNext® Low Input RNA Library Prep Kit for Illumina®) which accepts quantities of RNA between 2 pg – 200 ng.

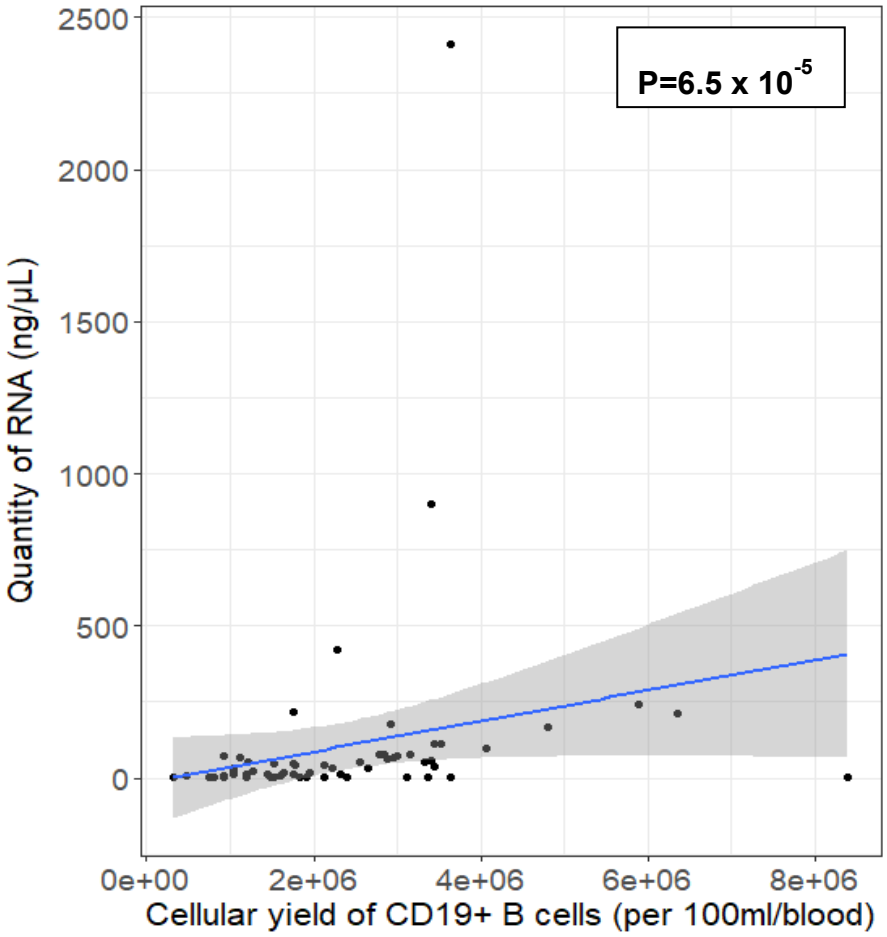


Figure 8.5 – Association between the cellular yield of $CD19^+$ B cells (per 100mls/blood) and the quantity of RNA extracted.

As the RNAseq was performed after all the study samples had been collected and RNA extracted, each sample of frozen lysate was stored for a variable amount of time at $^-80^oC$, with a median storage time of 31 months (24 – 34 (IQR), 23 – 39 (range)). However, there was no difference observed between the length of storage time prior to sequencing and the RIN value (P=0.41). There was also no difference noted in the time taken to freeze the lysate and the RIN value (P=0.34). These findings indicate that the differences in the laboratory cellular processing and storage times did not affect the quality of the RNA.

8.2.4 RNA sequencing depth and power

Sequencing depth describes the number of times that a given nucleotide has been read during an experiment. A higher sequencing depth generates more reads, which increases the statistical power to detect differential expression particularly among genes with lower expression levels. RNA sequencing in this experiment was performed to a depth of 60 million reads per sample, with 100 base pair paired-end reads. As per the power calculation described in Chapter 3: Methods, with 16 samples in the relapse cohort sequenced to a depth of 60 million reads, the study was well powered (99%) to detect effect (fold change) sizes of around 1.9 and adequately powered to detect those around 1.6 (90%), depending on transcript abundance.

8.3 RNA sequencing analysis

The RNAseq analysis was initially performed using a P value (<0.05) adjusted for false-discovery rate (FDR), before secondary analysis was undertaken using an unadjusted P value of <0.001 to capture genes with a fold change of ≥ 1.6 which did not reach statistical significance when adjusted for FDR. Log2 fold changes are reported as these are easier to interpret with a symmetrical distribution in the magnitude of up- and down-regulated genes. Shrinking the log2 fold change may allow for more accurate log2 fold change estimates for low count genes or if high dispersion (variation) present without changing the total number of genes that are differentially expressed (*Love MI et al. 2014*). However, as there was low differential gene expression detected in my dataset, the unshrunken log2 fold change values were used to detect subtle changes in gene expression for genes that were significantly differentially expressed.

The initial bioinformatic analysis including R computer programming scripts, and generation of the PCA/violin plots was performed by Ann Hedley, Senior Experimental Scientific Officer, Bioinformatics Support Unit, Newcastle University.

8.3.1 Relapse vs. remission

Transcriptomic profiling of differentially expressed genes (DEG) between the relapse and remission groups was the primary focus of the RNAseq analysis.

The initial RNA sequencing output using the FDR-adjusted P values demonstrated minimal DEG between the relapse and remission groups, as demonstrated in Principal Component Analysis (PCA) plots (Figure 8.6).

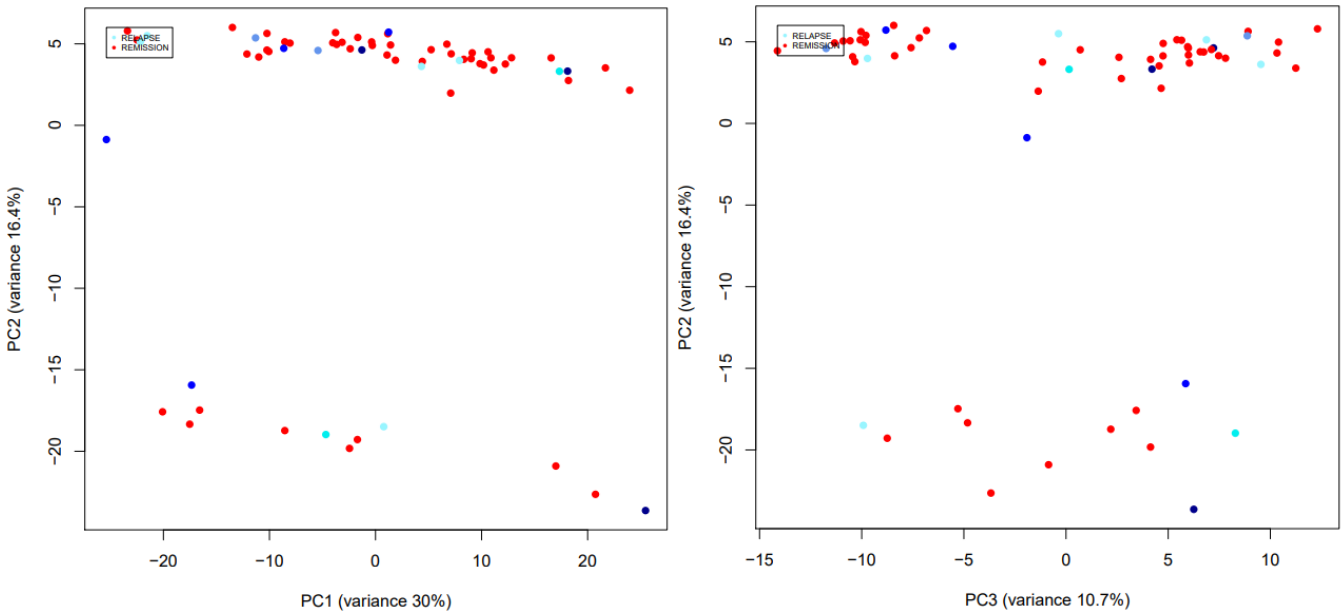


Figure 8.6 – Principal component analysis (PCA) plots showing minimal clustering of the relapse (blue) or remission (red) cohorts.

There were only two genes which were significantly differentially expressed once multiple test correction was applied (FDR-corrected P=<0.05) with a fold change ≥1.6 (Figure 8.7). However, these genes (*DDX3Y* and *USP9Y)* are Y-linked and therefore could potentially be false positives where the model has struggled to control for gender. Once the model was reanalysed split by gender, these genes were no longer differentially expressed. Therefore, further secondary analysis was performed using the unadjusted P value at a threshold of <0.001. Although the standard P value threshold is ideally an FDR-adjusted P=<0.05, the lower unadjusted P value threshold of <0.001 was chosen due to the lack of differential gene expression demonstrated and to permit the analysis of potentially significant genes whilst reducing the rate of false positives.

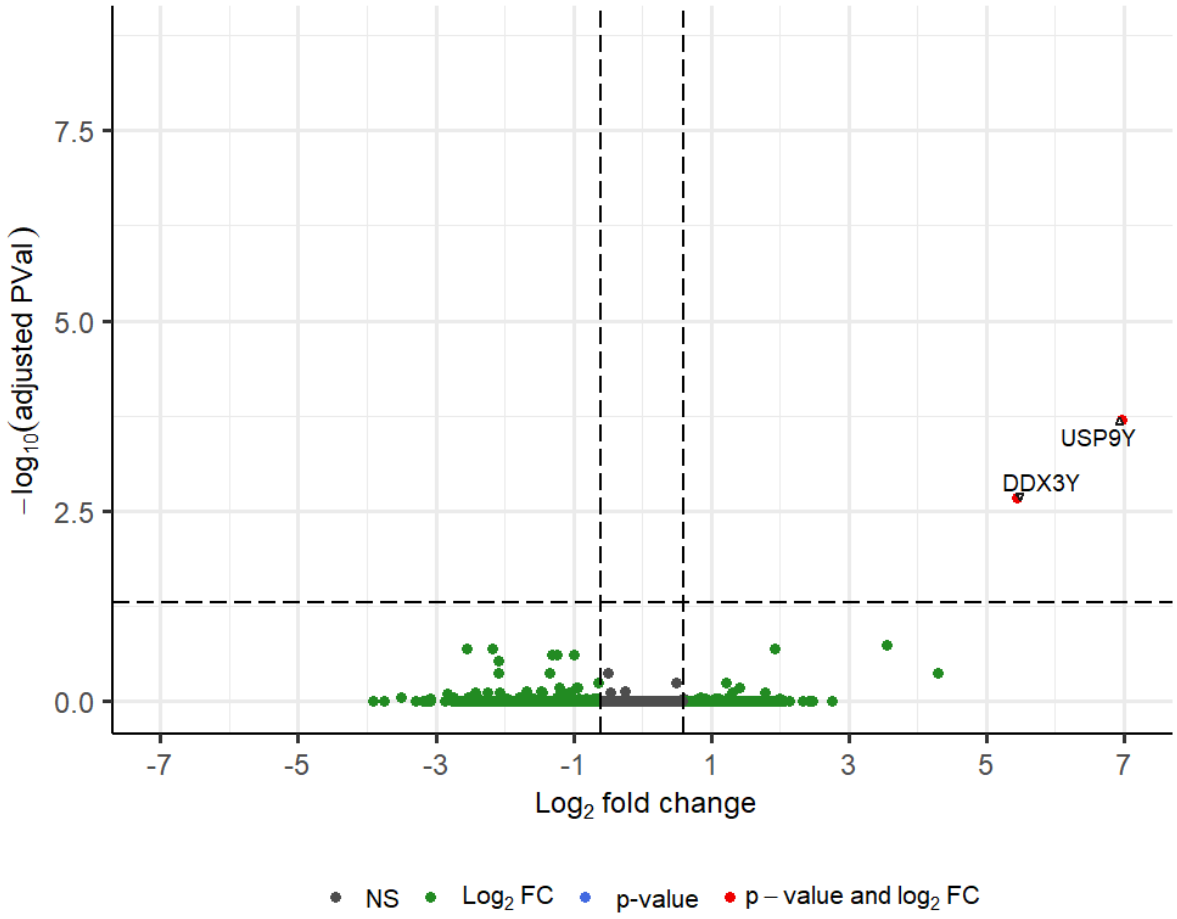


Figure 8.7 – Volcano plot (with FDR-corrected P values and unshrunken Log_2 fold change) presenting baseline differential gene expression in $CD19^+$ B cells between patients that relapsed or remitted following ATD cessation.

Using the unadjusted P value at a threshold of <0.001, there were 33 genes that were significantly differentially expressed between the relapse and remission patient groups. Within these 33 genes, there were 28 genes with a fold change ≥1.6 (Figure 8.8 and Table 8.1).

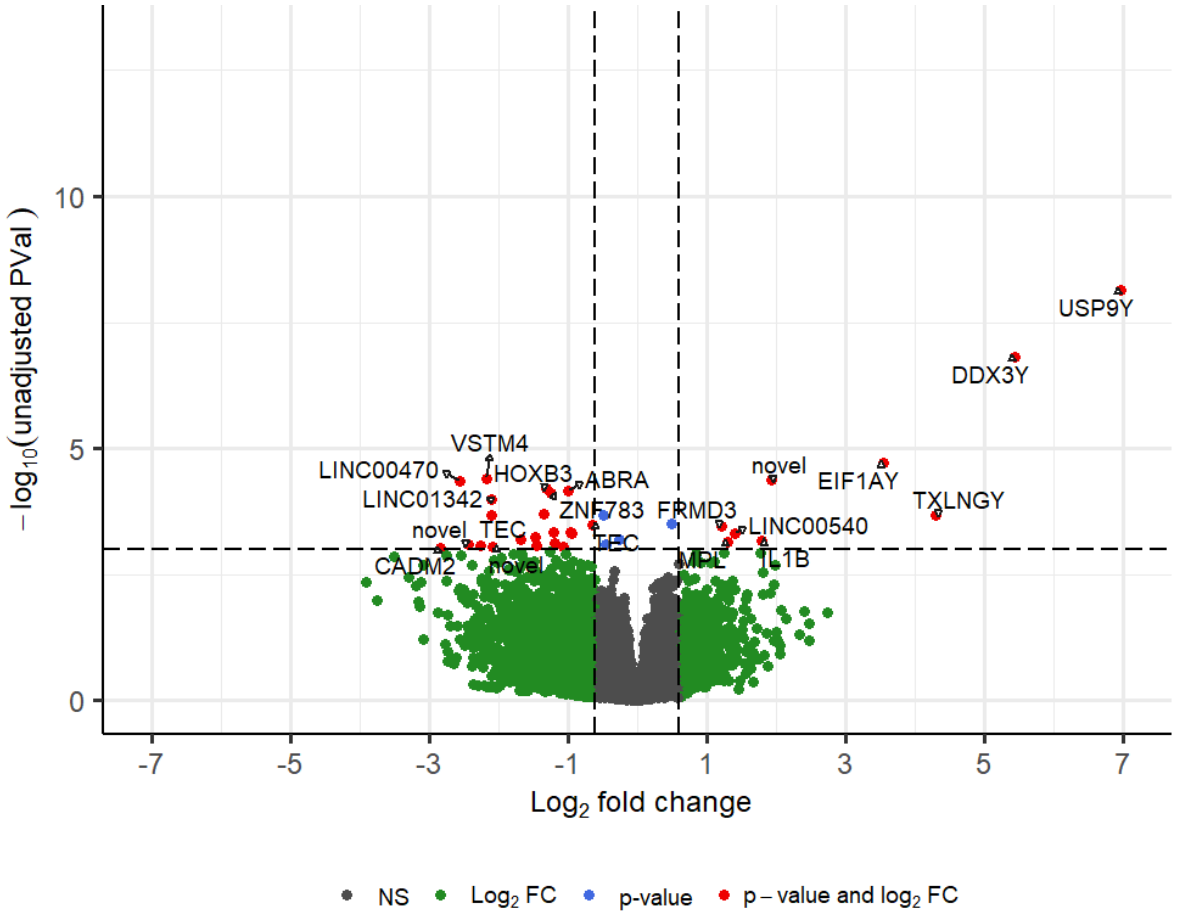


Figure 8.8 – Volcano plot (with unadjusted P values and unshrunken Log_2 fold change) presenting baseline differential gene expression in $CD19^+$ B cells between patients that relapsed or remitted following ATD cessation.

Gene expression	Ensembl gene ID	HGNC symbol	Gene description	Log_2FC	FC	Unadjusted P value
Up-regulated	ENSG00000114374	*USP9Y*	ubiquitin specific peptidase 9 Y-linked	6.9	125.4	7.1×10^{-9}
	ENSG00000067048	*DDX3Y*	DEAD-box helicase 3 Y-linked	5.4	43.6	1.5×10^{-7}
	ENSG00000198692	*EIF1AY*	eukaryotic translation initiation factor 1A Y-linked	3.6	11.8	1.9×10^{-5}
	ENSG00000288253	-	novel transcript	1.9	3.8	4.2×10^{-5}
	ENSG00000131002	*TXLNGY*	taxilin gamma pseudogene, Y-linked	4.3	19.7	0.0002
	ENSG00000112096	*SOD2*	superoxide dismutase 2	0.5	1.4	0.0003
	ENSG00000172159	*FRMD3*	FERM domain containing 3	1.2	2.3	0.0003
	ENSG00000276476	*LINC00540*	long intergenic non-protein coding RNA 540	1.4	2.7	0.0005
	ENSG00000125538	*IL1B*	interleukin 1 beta	1.3	2.5	0.0007
	ENSG00000117400	*MPL*	MPL proto-oncogene, thrombopoietin receptor	0.99	2.00	0.0007
Down-regulated	ENSG00000165633	*VSTM4*	V-set and transmembrane domain containing 4	-2.2	-4.5	3.9×10^{-5}
	ENSG00000132204	*LINC00470*	long intergenic non-protein coding RNA 470	-2.5	-5.9	4.3×10^{-5}
	ENSG00000120093	*HOXB3*	homeobox B3	-1.3	-2.5	6.2×10^{-5}
	ENSG00000174429	*ABRA*	actin binding Rho activating protein	-1	-2	7×10^{-5}
	ENSG00000204946	*ZNF783*	zinc finger protein 783	-1.2	-2.36	7.7×10^{-5}
	ENSG00000223823	*LINC01342*	long intergenic non-protein coding RNA 1342	-2.1	-4.3	0.0001
	ENSG00000215559	*ANKRD20A11P*	ankyrin repeat domain 20 family member A11, pseudogene	-1.09	-2.13	0.0002
	ENSG00000270344	*POC1B-AS1*	POC1B antisense RNA 1	-0.5	-1.4	0.0002
	ENSG00000280216	*TEC*	Tec Protein Tyrosine Kinase	-2.1	-4.3	0.0002
	ENSG00000280037	*TEC*	Uncategorized	-0.6	-1.56	0.0003

	ENSG00000158023	*CFAP251*	cilia and flagella associated protein 251	-0.95	-1.9	0.0004
	ENSG00000269967	-	novel transcript	-1.2	-2.3	0.0005
	ENSG00000164406	*LEAP2*	liver enriched antimicrobial peptide 2	-0.9	-1.9	0.0005
	ENSG00000278931	*ANKRD20A3*	ankyrin repeat domain 20 family, member A3 (ANKRD20A3) pseudogene	-1.5	-2.8	0.0006
	ENSG00000133028	*SCO1*	synthesis of cytochrome C oxidase 1	-0.2	-1.2	0.0006
	ENSG00000232684	*ATP11A-AS1*	ATP11A antisense RNA 1	-1.7	-3.2	0.0006
	ENSG00000233623	*PGAM1P11*	phosphoglycerate mutase 1 pseudogene 11	-1.2	-2.3	0.0008
	ENSG00000144426	*NBEAL1*	neurobeachin like 1	-0.5	-1.4	0.0008
	ENSG00000270104	-	novel transcript	-2.4	-5.4	0.0008
	ENSG00000259648	*HMGB1*	high mobility group box 1 pseudogene	-2.3	-4.8	0.0008
	ENSG00000170074	*FAM153A*	family with sequence similarity 153 member A	-1.5	-2.7	0.0008
	ENSG00000275488	*ZNF337*	zinc finger protein 337 pseudogene	-1.07	-2.1	0.0009
	ENSG00000231859	-	novel pseudogene	-2.08	-4.2	0.0009

Table 8.1 – Differentially expressed genes at baseline between the relapse and remission cohorts, using an unadjusted P value of <0.001.

8.3.2 Controlling for variables

To evaluate whether any variance in gene expression levels could be explained by other co-variates, a secondary analysis of the RNAseq data was performed adjusting for clinical variables including age, gender, ethnicity, smoking status, presence of GO, CRP concentration, thyroid function, and date of sample collection. This analysis is presented as violin plots in Figure 8.9.

Although smoking has the biggest effect on gene expression affecting many genes by a small amount, the mean is only approximately 5% of the total within gene variance. Ethnicity has one, and gender a few, genes where the variance of the gene is close to 100% due to factor effects. Residuals is the variance of gene expression that is not explained by any of the clinical factors entered into the analysis. Therefore, most of the genetic variance (mean 77%) in the dataset cannot be explained by these clinical factors.

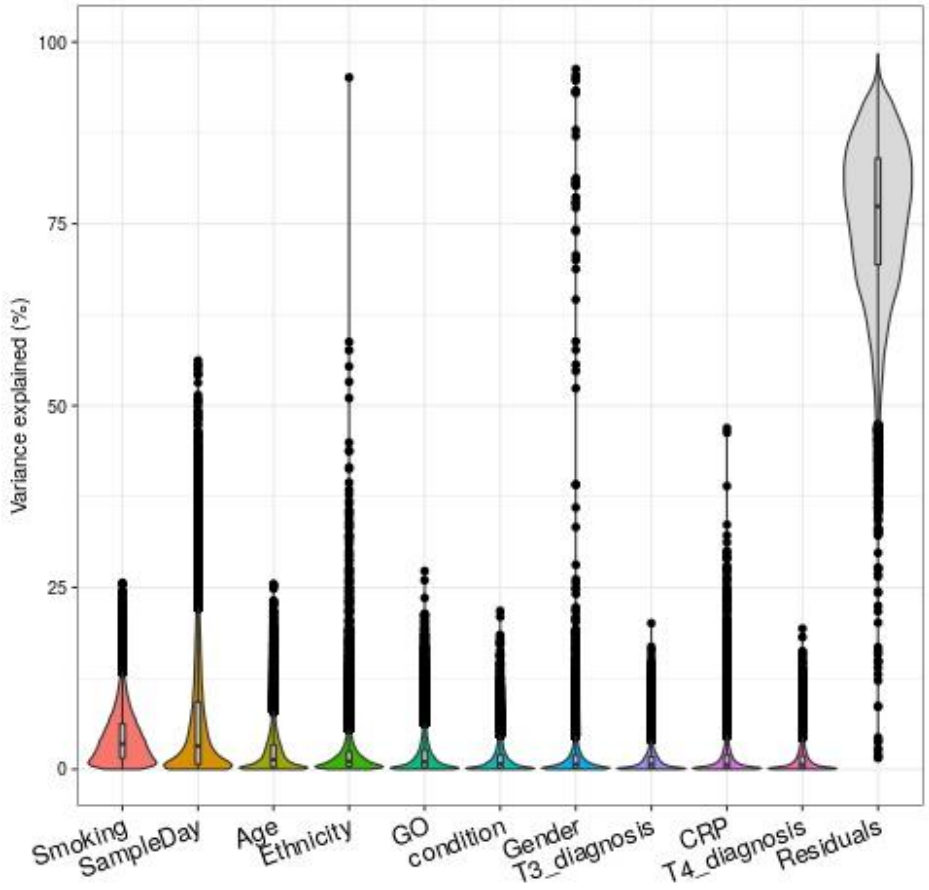


Figure 8.9 – Violin plots showing the variability of gene expression in relation to each clinical factor. *8.3.3 Heatmap analysis of gene expression and outcome*

Using the hierarchical clustering method as set out in Chapter 3: Methods, a heatmap was generated to visualise clustering of samples and genes based on gene expression. A heatmap of the 33 genes that were significantly differentially expressed (unadjusted P value <0.001) between the relapse and remission patient groups is presented in Figure 8.10. Following hierarchical clustering there were two clusters of patient cohorts and three clusters of genes identified within the heatmap (Figure 8.10).

As visualised on the heatmap, the two clusters of patients identified correlates closely with gender (all males patients are in the same cluster) and this is also reflected in a cluster of highly expressed Y-linked genes (*USP9Y*, *EIF1AY*, *DDX3Y*, *TXLNGY*) in these patients. Therefore, the clustering of this heatmap may be representative of the differences in gene expression between genders rather than GD outcome.

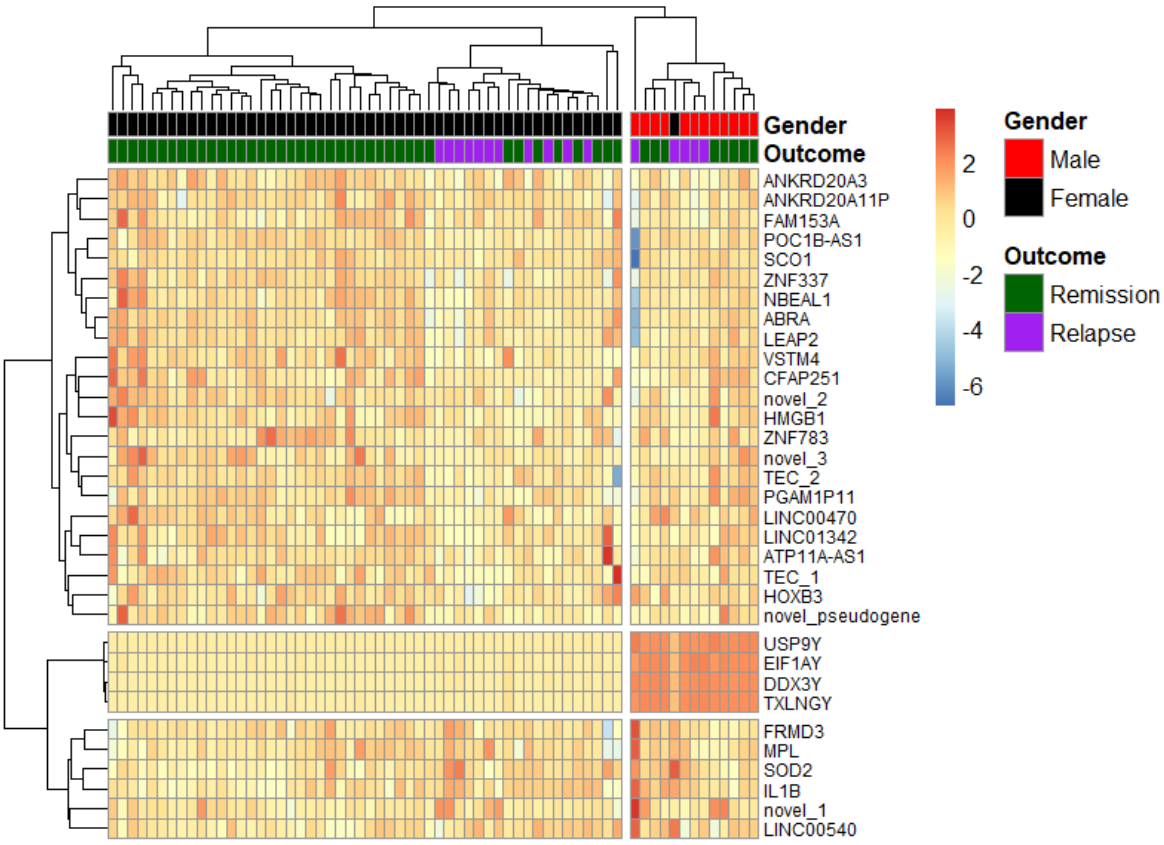


Figure 8.10 – Hierarchical clustered heatmap of $log_2(+1)$ transformed normalised read counts of the 33 genes that were significantly differentially expressed (unadjusted P value <0.001) between the relapse and remission groups.

To further explore the relative differences in gene expression between outcome groups, the highly expressed Y-linked genes found clustered among the male patients were excluded from the analysis. Hierarchical clustering was undertaken as previously discussed, which identified two distinct clusters within the heatmap (Figure 8.11).

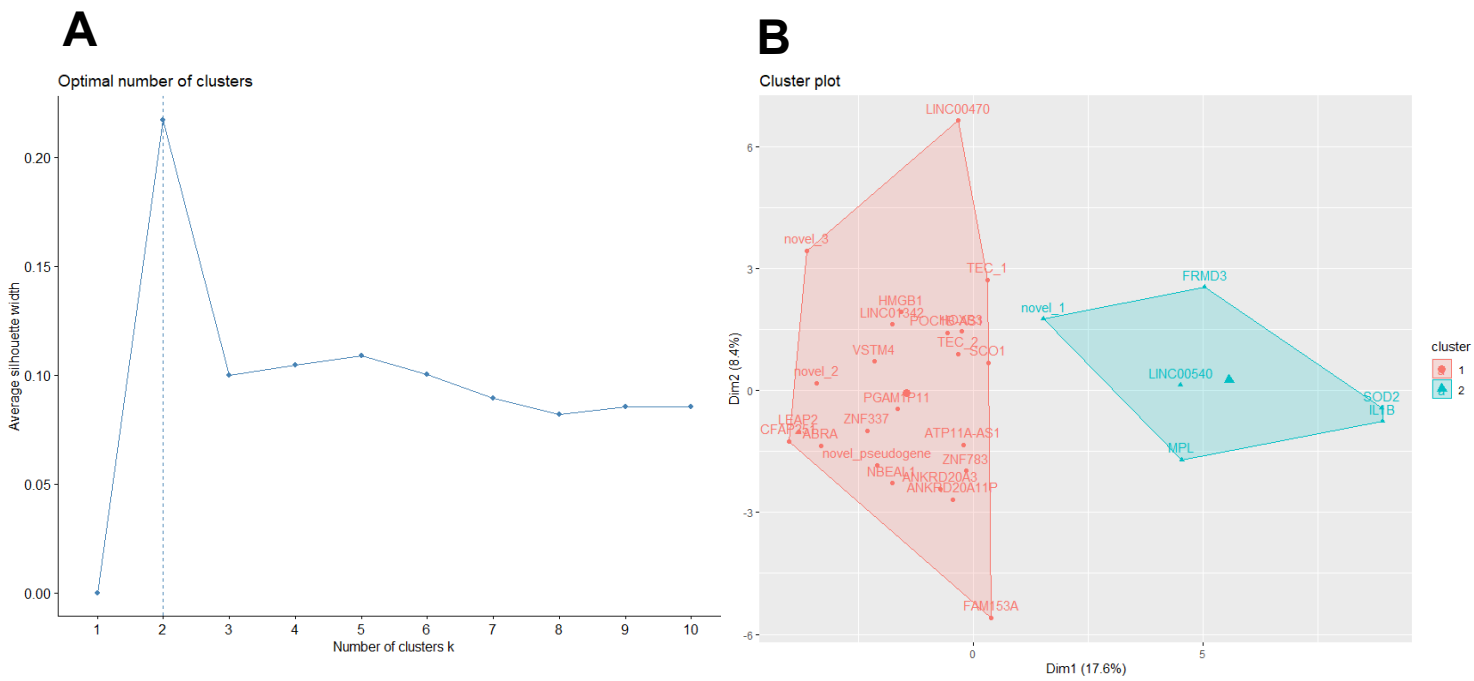


Figure 8.11– **A**: Silhouette plot used to determine the optimal number of clusters in the dataset **B**: Cluster plot produced using principal component analysis (PCA) showing the two separate gene clusters.

The clustering of this secondary heatmap shows gender more evenly spread between the two patient clusters: ‘Group 1’ and ‘Group 2’ (Figure 8.12). Furthermore, all the relapse patients are found within the same cluster (Group 2). There are two separate groups of gene clusters identified, with one cluster containing 6 genes and the other 27 genes. Overall, between the two patient groups, the 6 gene cluster appears more upregulated in ‘Group 2’ compared to ‘Group 1’, whereas the 27 gene group appears more upregulated in ‘Group 1’.

The two groups identified on Figure 8.12 included 33 (Group 1) and 32 (Group 2) patients respectively. Additional analysis was undertaken to further characterise these two groups of patients (Chapter 9: Integrative analysis).

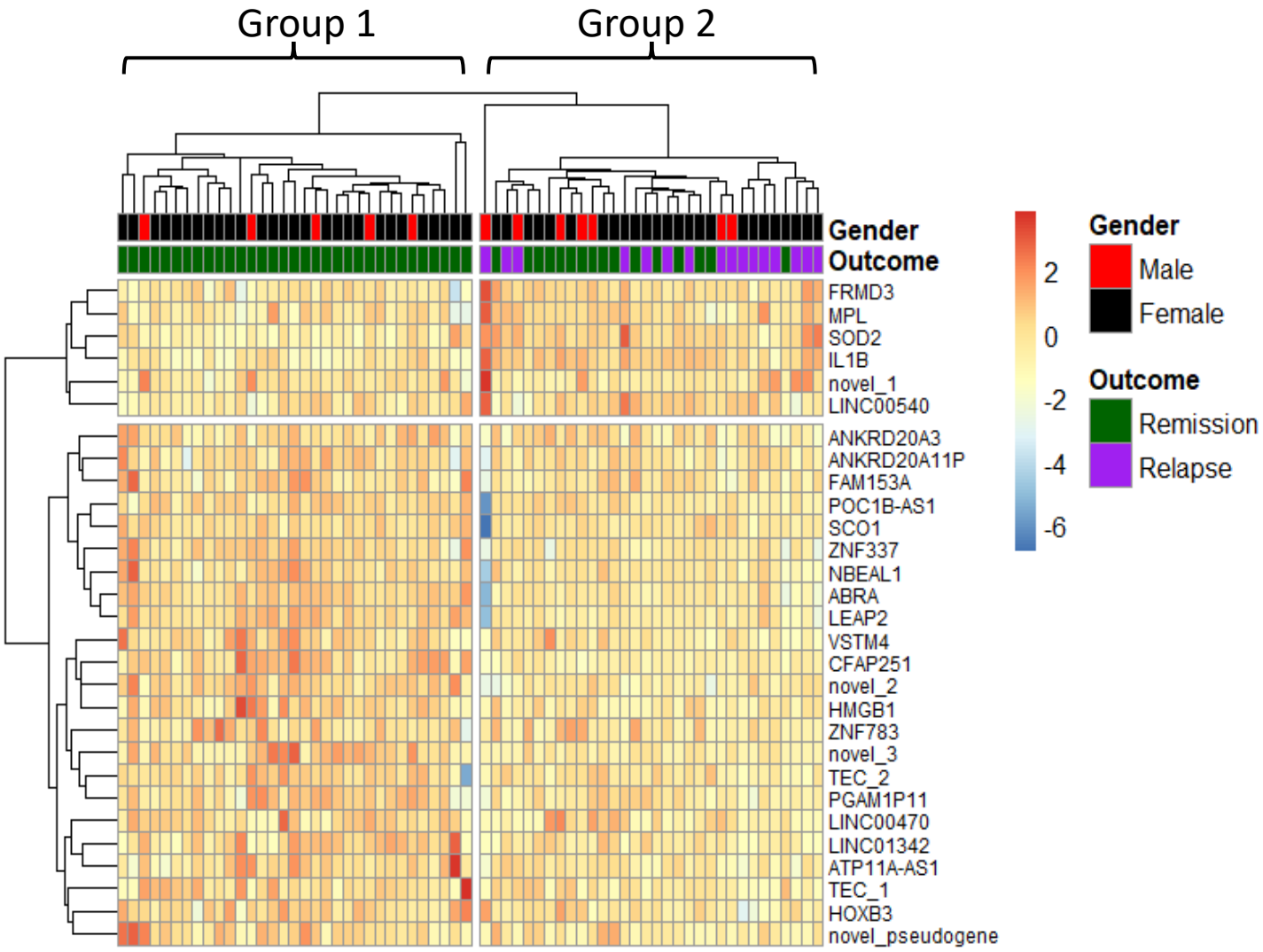


Figure 8.12 – Hierarchical clustered heatmap of $log_2(+1)$ transformed normalised read counts of 29 genes (*USP9Y*, *EIF1AY*, *DDX3Y*, *TXLNGY* excluded) that were significantly differentially expressed (unadjusted P value <0.001) between the relapse and remission groups.

8.3.4 Canonical Pathway Analysis: relapse vs. remission

To investigate the molecular mechanisms underlying GD relapse, pathway analysis of the RNAseq data was performed using Ingenuity Pathway Analysis (IPA™) software (Qiagen) to elucidate significantly involved pathways (https://www.qiagenbioinformatics.com/ products/ingenuity-variant-analysis). All DEG identified between the relapse and remission groups with an unadjusted P value of <0.01 were included in this analysis (258 genes). The DEG were categorised to related canonical pathways based on the Ingenuity pathway knowledge base (IPKB).

IPKB is a database of biological and functional observations generated using modelled relationships between proteins, genes, cells, tissues, and disease (*Krämer A et al. 2014*). The top enriched categories of canonical pathways with a P value <0.05 are presented in Figure 8.13. The significance values (p-value of overlap) for the canonical pathways are calculated by the right-tailed Fisher's Exact Test. This test measures the overlap of observed and predicted regulated gene sets (*Krämer A et al. 2014*).

Ingenuity pathway analysis indicated that ‘Cytokine Signalling’ and ‘Disease-Specific’ pathways were significantly activated prior to GD relapse and contained the greatest number of overlapping genes. The four canonical pathways found within ‘Cytokine Signaling’ and ‘Disease-specific’ categories as predicted by Ingenuity Pathway Analysis, along with their representative DEG are presented in Table 8.2. The remaining significantly enriched canonical pathways and their associated genes are listed in Appendix C. Many of these pathways that were identified included genes involved in the cellular and humoral immune response. The most significantly enriched canonical pathway within both the ‘Cytokine Signalling’ and ‘Disease-Specific’ categories was the ‘Airway Pathology in Chronic Obstructive Pulmonary Disease’ ($^{-}$log P= 3.79). Both *IL-1*β and *IL-13* were found to be associated with all the top enriched categories of canonical pathways.

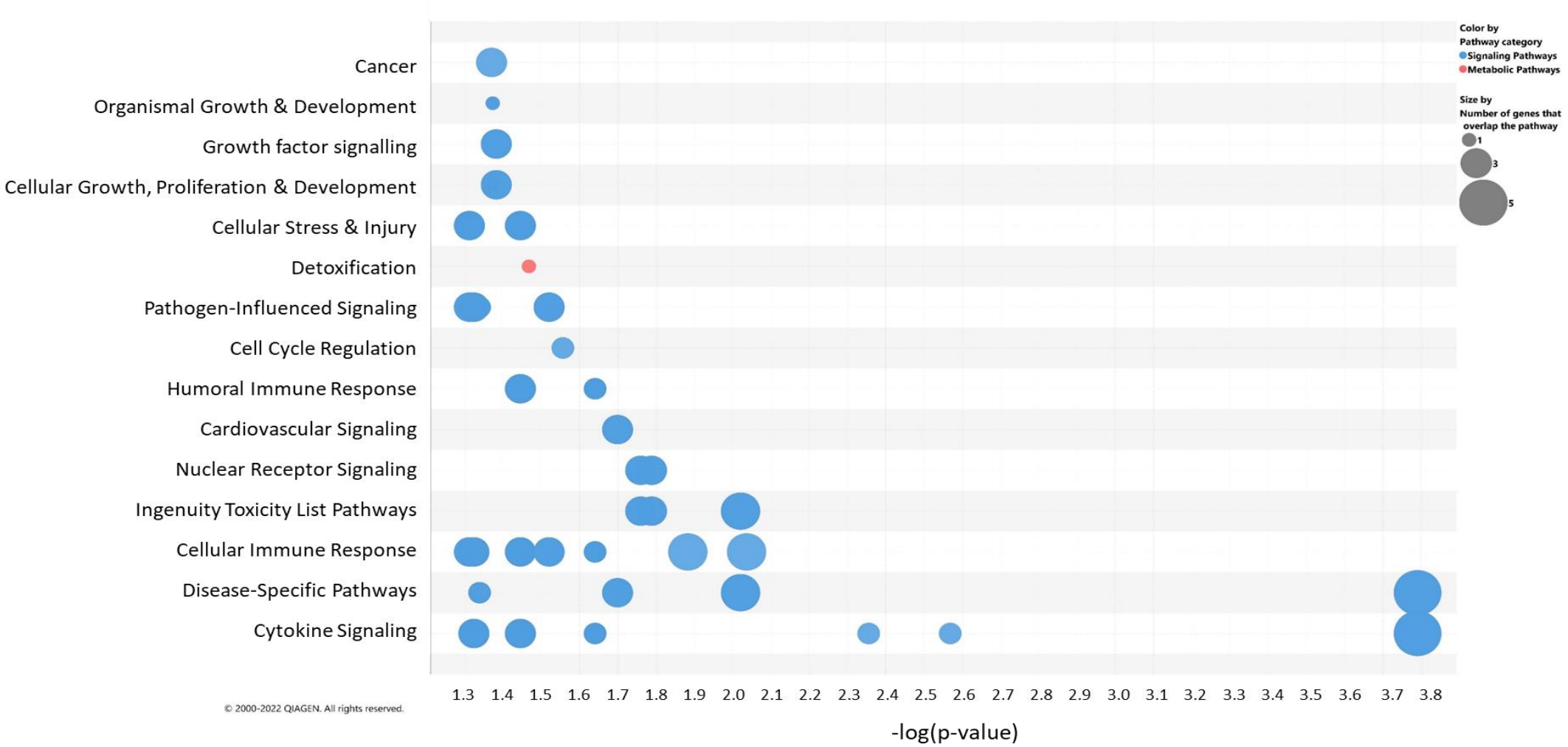


Figure 8.13 – Bubble chart of significant canonical pathways (P <0.05) enriched in the dataset.

Ingenuity Canonical Pathway	Ingenuity Canonical Pathway Category (s)	-Log (P-value)	Representative genes
Airway Pathology in Chronic Obstructive Pulmonary Disease	Cytokine signalling and disease-specific pathways	3.79	*APOD, FGF12, IL-13, IL17D, IL-1β*
Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	Cytokine signalling	2.57	*IL-13, IL-1β*
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	Cytokine signalling	2.36	*IL-13, IL-1β*
Role of Cytokines in Mediating Communication between Immune Cells	Cytokine signalling	1.64	*IL-13, IL-1β*

Table 8.2 – Canonical Pathways and categories as predicted by Ingenuity Pathway Analysis. *8.3.5 Causal analysis (upstream analysis)*

Causal analysis aims to identify upstream transcriptional regulators that can provide biological insight into the observed gene expression changes. Upstream Regulator Analysis is based on expected causal effects between upstream regulators and targets, with the expected causal effects derived from the literature compiled in the IPKB. The Ingenuity Pathway Analysis definition of a transcriptional regulator is any molecule that can affect the expression of other molecules, and therefore upstream regulators can include various types of molecules, such as transcription factors and microRNAs (*Krämer A et al. 2014*).

The top transcriptional regulators (overlap P-value < 0.01) as predicted by Ingenuity Pathway Analysis and their predicted activation states (z-score) are presented in Table 8.3. The overlap P-value measures whether there is a statistically significant overlap between the dataset genes and the genes that are regulated by a transcriptional regulator (calculated using Fisher's Exact Test). The activation z-score is used to infer likely activation states of upstream regulators by assessing the pattern of observed and predicted up/down regulation states *(Krämer A et al. 2014)*.

There were no significant upstream regulators that were predicted to have an inhibitory action.

Upstream Transcriptional Regulator	Molecule Type	Predicted Activation State	Activation z-score	Overlap P-value	Target Genes in Dataset
IL-6	Cytokine	Activated	2.406	0.008	*IL-13, I IL-1β, SOD2, ADAMTS13, CCL3L1, CYP1A2, FCGR1A, GADD45B, IGFBP4,*
miR-3160-3p (miRNAs w/seed GAGCUGA)	Mature microRNA	Activated	2.236	0.0006	*CYP1A2, FAM153A/FAM153B, FOXC1, NPIPA9, TTBK2*
miR-550b-2-5p (and other miRNAs w/seed UGUGCCU)	Mature microRNA	Activated	2.236	0.003	*ABRA, AC004076.7, B3GALT2, HOXB3, MESP1*
REL	Transcription regulator	Activated	2.17	0.0006	*IL-13, IL-1β, SOD2, NR4A3, DCLRE1B, GADD45B,*
IL-1A	Cytokine	Activated	2.167	0.004	*IL-13, IL-1β, SOD2, APOD, CYP1A2*
E. coli B4 lipopolysaccharide	Chemical toxicant	Activated	2.132	0.00717	*IL-13, IL-1β, SOD2, ABCG8, FCGR1A*

Table 8.3 – Transcriptional regulators as predicted by Ingenuity Pathway Analysis.

This causal analysis indicates that the expression pattern in the dataset is consistent with six upstream regulators that are predicted to have downstream activating or inhibitory effects on some of the differentially expressed genes identified in the GD outcome groups. The top predicted transcriptional regulator which had the greatest activation z-score (2.4) was IL-6. The mechanistic network of IL-6 as an upstream regulator is presented in Figure 8.14. The different node shapes represent different types of molecules, as presented in Appendix E.

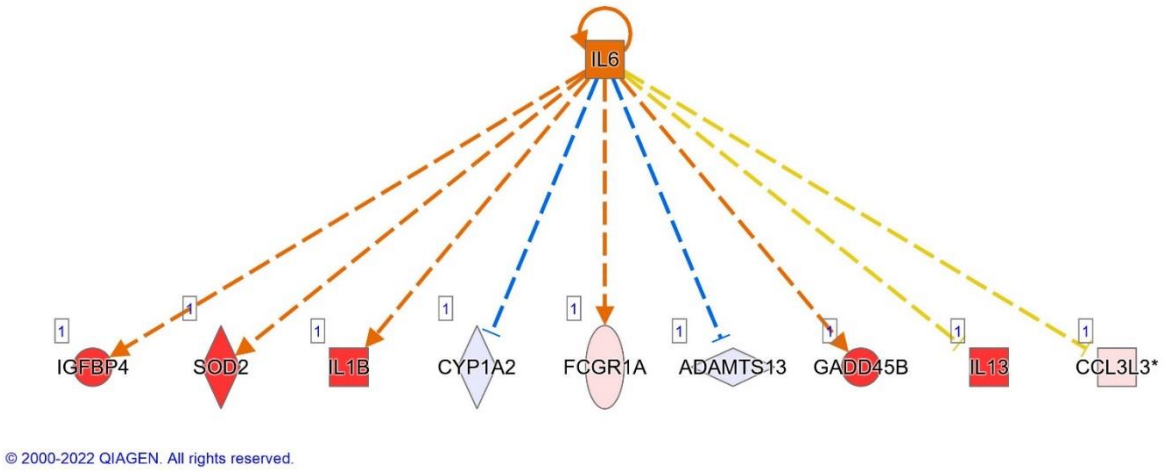


Figure 8.14 – Mechanism of action of IL-6 as an activated upstream regulator. *8.3.6 Disease and Function Analysis (downstream analysis)*

Downstream analysis aims to identify biological functions or diseases that are expected to be increased or decreased given the observed gene expression changes. Biological functions and diseases (focused on cellular function and immune response) predicted to be significantly (P=<0.01) affected by the DEG in the GD dataset are presented in Figure 8.15. The individual P values and the genes involved in each biological function is detailed in Appendix D.

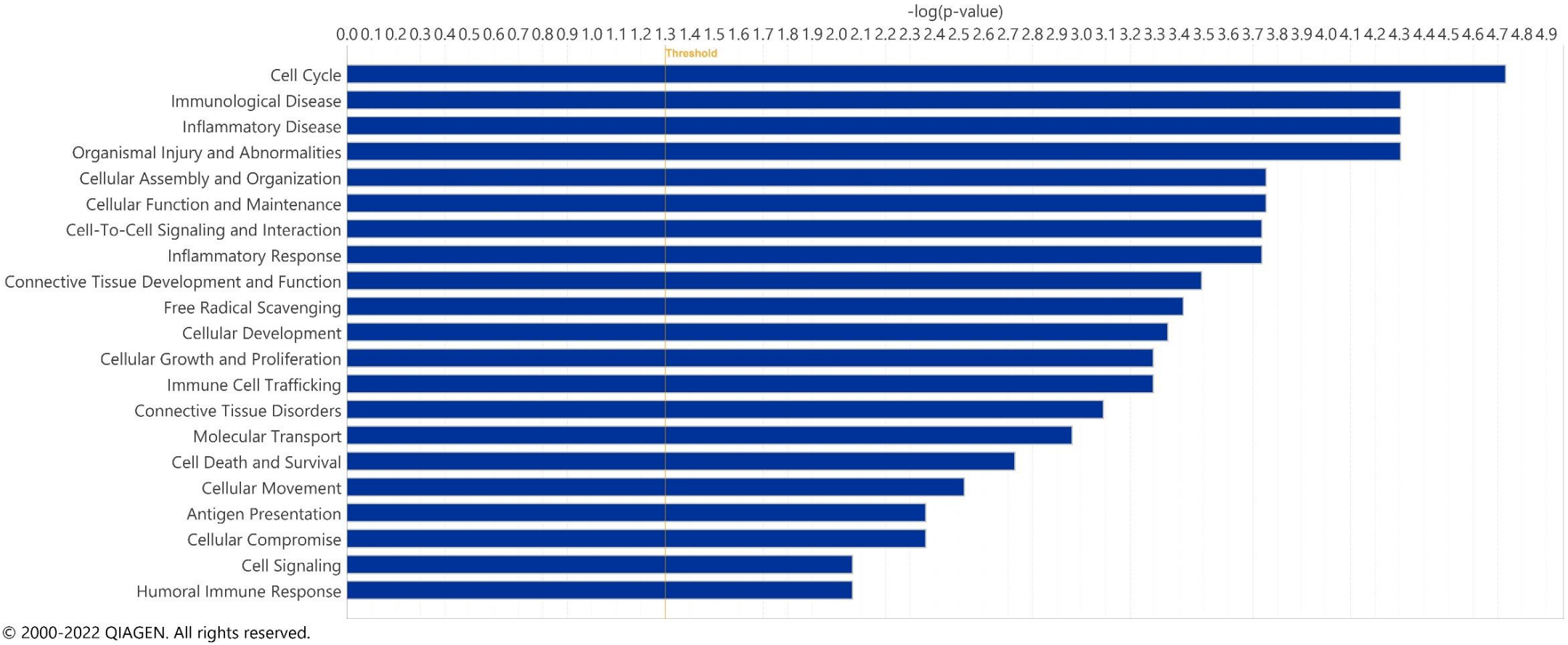


Figure 8.15 – Bar chart (focused on cellular function and immune response) displaying significant biological functions and diseases (P=<0.01), predicted to be affected by the differentially expressed genes within the Graves' disease dataset.

The identified biological impact predicted to be influenced by the DEG include inflammatory and immunological functions that are likely to be relevant in the context of an autoimmune disease, such as GD. With particular focus on B cells, relevant processes include 'Antigen presentation' (P= 0.004) and the 'Humoral immune response' (P= 0.009).

An example of how antigen presentation is affected by some of the DEG upregulated in the GD relapse patients is presented in Figure 8.16.

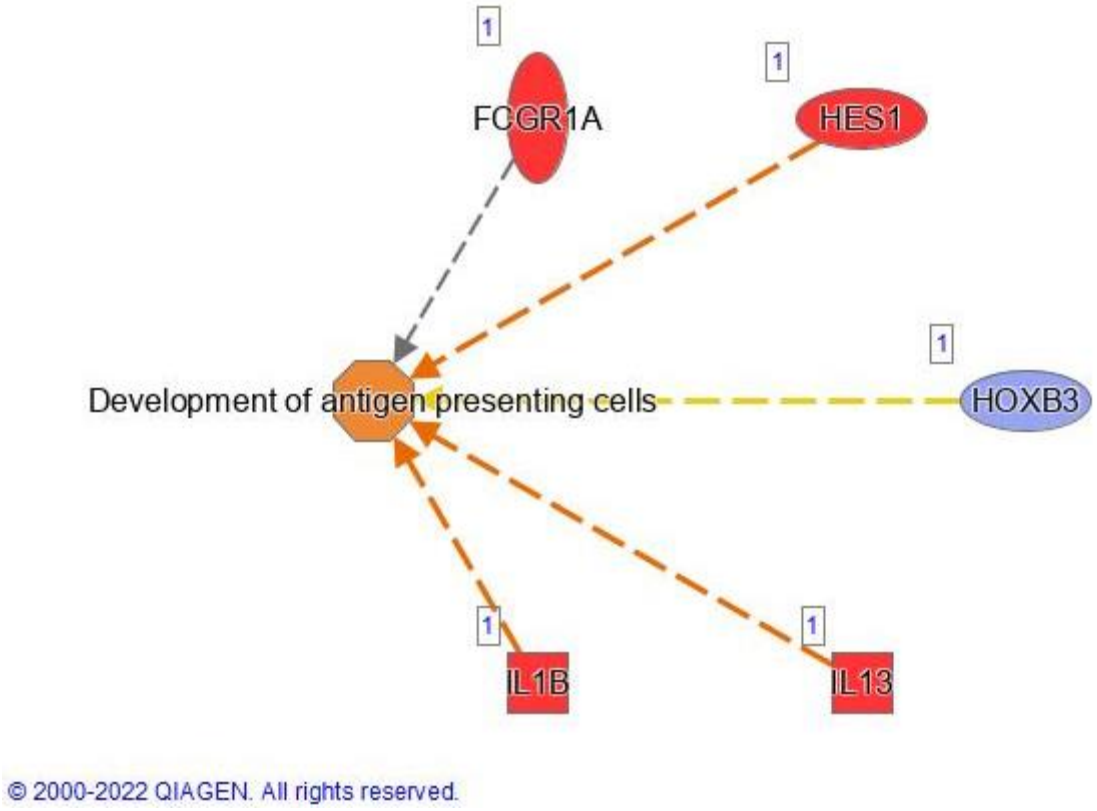


Figure 8.16 – The influence of differentially expressed genes in the dataset on the development of antigen presenting cells.

The most significant disease type associated with the DEG observed in the GD dataset was 'Immunological Disease' (Figure 8.15). The top three diseases associated with the observed pattern of DEG were SLE ($P = 4.97 \times 10^{-5}$), an active stage immunological disorder ($P = 5.7 \times 10^{-4}$), and refractory chronic primary immune thrombocytopenia ($P = 8.14 \times 10^{-4}$). The 19 immunological diseases and their associated genes are presented in the Appendix D. A graphical representation of how the DEG in the GD dataset are related to SLE is presented in Figure 8.17.

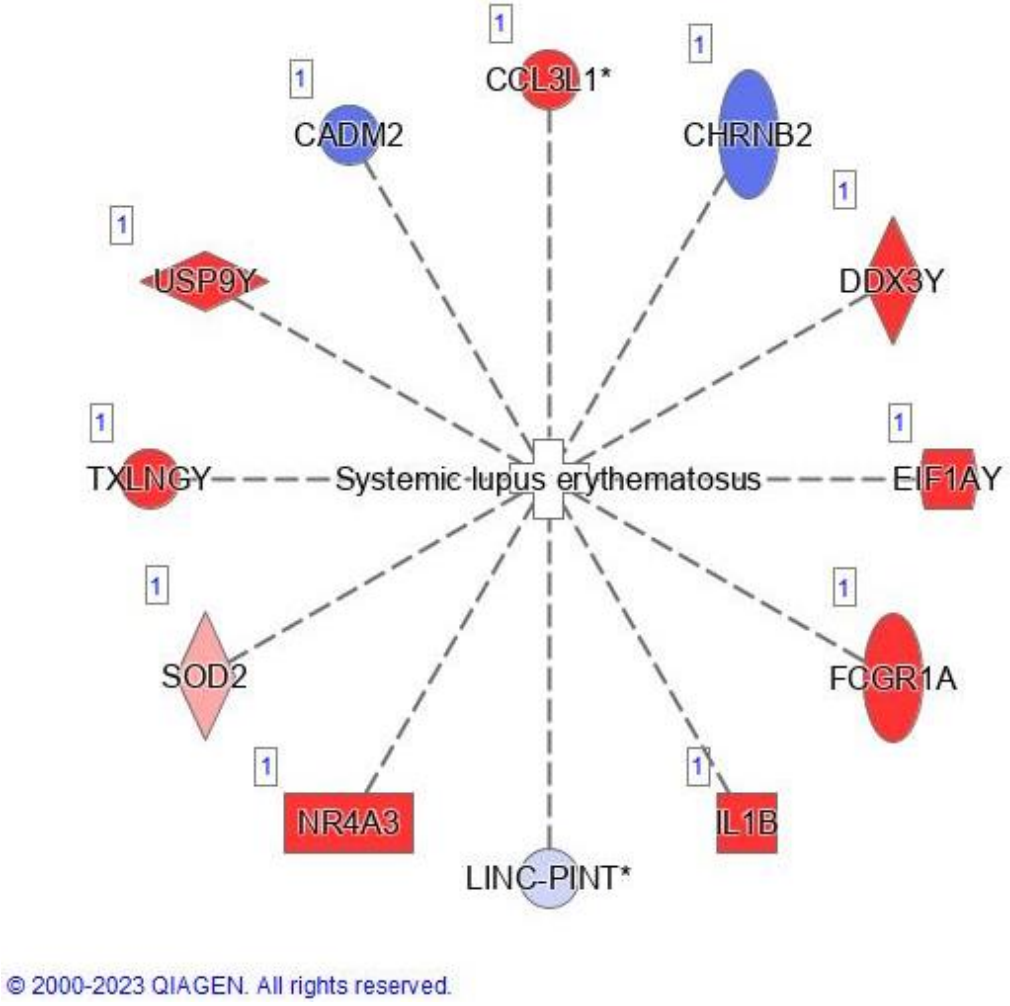


Figure 8.17 – Association of the differentially expressed genes observed in the Graves' disease dataset with systemic lupus erythematosus. The significant overlap of DEG in the GD dataset with another B cell mediated autoimmune disease, SLE, may help distinguish which specific genes are involved in the pathological autoimmune humoral immune response.

8.3.7 Molecular network analysis

Using the Ingenuity Pathway Analysis (IPA™) software, molecular network analysis was performed to explore the interactions between the DEG in the dataset and gain insight into the biologically relevant molecular networks and pathways leading to GD relapse. The top identified network was defined by the functions "Cell-To-Cell Signaling and Interaction, Cellular Function and Maintenance, Haematological System Development and Function, and Inflammatory Response" ($P=2.43 \times 10^{-7}$). A graphical hierarchical representation of this network is presented in Figure 8.18.

The findings from this network analysis indicate that IL-β is a key upstream regulator for many of the other genes differentially expressed in the GD dataset, and therefore it may be an essential component of molecular pathways leading to GD relapse.

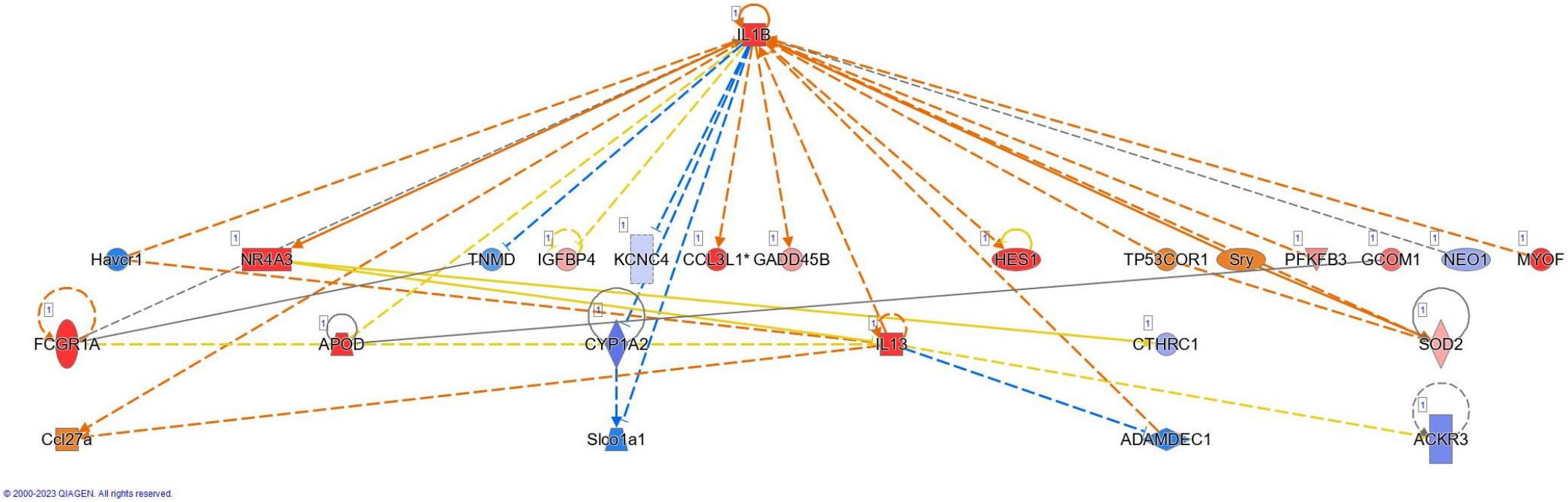


Figure 8.18 – IPA molecular network of differentially expressed genes within the ‘Cell-To-Cell Signalling and Interaction, Cellular Function and Maintenance, Haematological System Development and Function, and Inflammatory Response’ network between patients who relapsed their Graves’ disease and those that went into remission.

8.4 Survival analysis

The aim of survival analysis was to explore the association between DEG at the timepoint of stopping ATD and the time to relapse of GD.

After performing a univariate cox regression analysis, 11 genes were significantly associated with time to GD relapse, using an unadjusted p-value threshold of <0.001 (Table 8.4). However, none of the genes were significant after Benjamini-Hochberg multiple test correction. The proportional hazard assumption was supported by a non-significant relationship between residuals and time for each gene.

Out of these 11 genes associated with time-to-relapse, 2 genes (*SOD2* and *POC1B-AS1*) and a further 7 genes (*CCL3L3, DDX11L17, TMEM86A, GADD45B, NRADDP, PFKFB3, FCGR1A*) were also significantly (unadjusted P=<0.001, unadjusted P=<0.01, respectively) differentially expressed between the relapse and remission patients (Table 8.1).

The 11 genes that were significantly associated with time to GD relapse on univariate analysis were entered into a multivariate Cox regression analysis. Following the multivariate analysis, only one gene, *DDX11L17*, remained significant (unadjusted P=0.01).

Ensembl gene ID	HGNC symbol	Gene description	$HR_{relapse}$ (CI 95%)	B	Unadjusted P value	Adjusted P value
ENSG00000276085	*CCL3L3*	C-C motif chemokine ligand 3 like 1	1.4 (1.2-1.6)	0.32	4×10^{-5}	0.79
ENSG00000279928	*DDX11L17*	DEAD/H-box helicase 11 like 17 (pseudogene)	1.5 (1.2-1.8)	0.41	3.5×10^{-11}	0.79
ENSG00000151117	*TMEM86A*	transmembrane protein 86A	1.1 (1.02-1.1)	0.05	2.2×10^{-4}	0.98
ENSG00000099860	*GADD45B*	growth arrest and DNA damage inducible beta	1 (1.001-1.002)	0.001	2.8×10^{-4}	0.98
ENSG00000170345	*FOS*	Fos proto-oncogene, AP-1 transcription factor subunit	1.01 (1-1.01)	0.0001	5.3×10^{-4}	0.98
ENSG00000236409	*NRADDP*	neurotrophin receptor associated death domain, pseudogene	1.1 (1.05-1.2)	0.12	5.7×10^{-4}	0.98
ENSG00000170525	*PFKFB3*	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	1 (1.002-1.006)	0.004	5.8×10^{-4}	0.99
ENSG00000270344	*POC1B-AS1*	POC1B antisense RNA 1	0.99 (0.98-0.99)	-0.014	6.963×10^{-4}	0.99
ENSG00000112096	*SOD2*	superoxide dismutase 2	1 (1.001-1.002)	0.0013	7.088×10^{-4}	0.99
ENSG00000150337	*FCGR1A*	Fc gamma receptor Ia	1 (1.01-1.05)	0.03	7.699×10^{-4}	0.99
ENSG00000197272	*IL27*	interleukin 27	1.1 (1.05-1.2)	0.11	9.874×10^{-4}	0.99

Table 8.4 – Association between gene expression at time of stopping ATD and time-to-relapse of Graves' disease by univariate Cox regression analysis, using an unadjusted significance threshold of P<0.001.

8.4.1 Composite gene scores to predict outcome in Graves' disease

The performance of selected genes to predict GD relapse was assessed using a composite biomarker score. The genes selected to further explore their potential role in predicting GD relapse included all the 33 DEG genes (unadjusted P=<0.001) and additionally *DDX11L17* which was the only gene observed to be significant on multivariate cox regression (unadjusted P=0.01).

Firstly, lasso regression was performed (as described in Chapter 3: Methods) to select the genes with the greatest predictive value to simplify the model and enhance prediction accuracy. Following lasso regression, 28 genes were removed from the analysis leaving 6 genes remaining (*ANKRD20A3*: Ankyrin Repeat Domain 20 Family Member A3, *DDX11L17:* DEAD/H-Box Helicase 11 Like 17, *POC1B-AS1*; POC1B Antisense RNA 1, *SCO1*: Synthesis Of Cytochrome C Oxidase 1, *SOD2*: Superoxide dismutase 2, *TEC*: Tec Protein Tyrosine Kinase). The coefficients generated from the lasso regression provided a balanced weight for each gene, which was then multiplied by the respective gene value and added to the other genes to create a composite score. An example of the formula used to generate the regression-weighted composite scores is presented in Chapter 3: Methods.

8.4.2 ROC curve analysis using gene composite scores

The composite scores were used to analyse different combinations of the 6 genes and compare their predictive utility for GD relapse with each individual gene alone (63 potential combinations). ROC curve analysis was undertaken which generated an AUC score to determine the best gene(s) to predict the outcome of relapse vs. remission. The R programme 'cutpointr' generated the optimal composite score cut-off which produced the highest sensitivity and specificity of a specific gene or combination of genes for the prediction of relapse or remission (Chapter 3: Methods). The different combinations of variables and their respective ROC_{AUC} are presented in Table 8.5. The complete list of variable combinations and their respective ROC_{AUC} is presented in Appendix F.

POC1B-AS1	*SOD2*	*TEC*	*ANKRD20A3*	*SCO1*	DDX11L17	ROC_{AUC}
✓	✓	✓	✓	✓	✓	**0.95**
✓	✓	✓	✓	✓	×	**0.95**
✓	×	✓	✓	✓	✓	**0.95**
×	✓	✓	✓	✓	✓	**0.95**
✓	✓	✓	✓	×	×	**0.95**
✓	×	✓	✓	✓	×	**0.94**
×	×	✓	✓	✓	✓	**0.94**
✓	✓	✓	✓	×	✓	**0.94**
✓	✓	✓	×	✓	✓	**0.94**
✓	✓	×	✓	✓	✓	**0.94**
✓	×	✓	✓	×	✓	**0.94**
✓	✓	×	✓	✓	×	**0.93**
✓	×	✓	✓	×	×	**0.93**
✓	×	×	✓	✓	✓	**0.93**
✓	×	×	✓	✓	×	**0.93**
✓	✓	×	✓	×	✓	**0.92**
×	×	✓	✓	✓	×	**0.92**
×	✓	×	✓	✓	✓	**0.92**
×	✓	✓	✓	✓	×	**0.92**
✓	✓	✓	×	×	✓	**0.91**

Table 8.5 – The top 20 different combinations of biomarker composite scores ranked by their ROC_{AUC}.

There were 7 different combinations of genes that had the highest score for predicting relapse in GD (ROC_{AUC} = 0.95), including the composite score that incorporated all 6 genes and three scores that included only 4 of the genes. Therefore, excluding the combinations of *DDX11L17 +/- SCO1* or excluding *POC1B-*

AS1 + SOD2 didn't affect the ROC_{AUC}. The ROC curves of the top performing biomarker composite scores that incorporate differing numbers of genes (2 – 6 gene combinations) for comparison are presented in Figure 8.19.

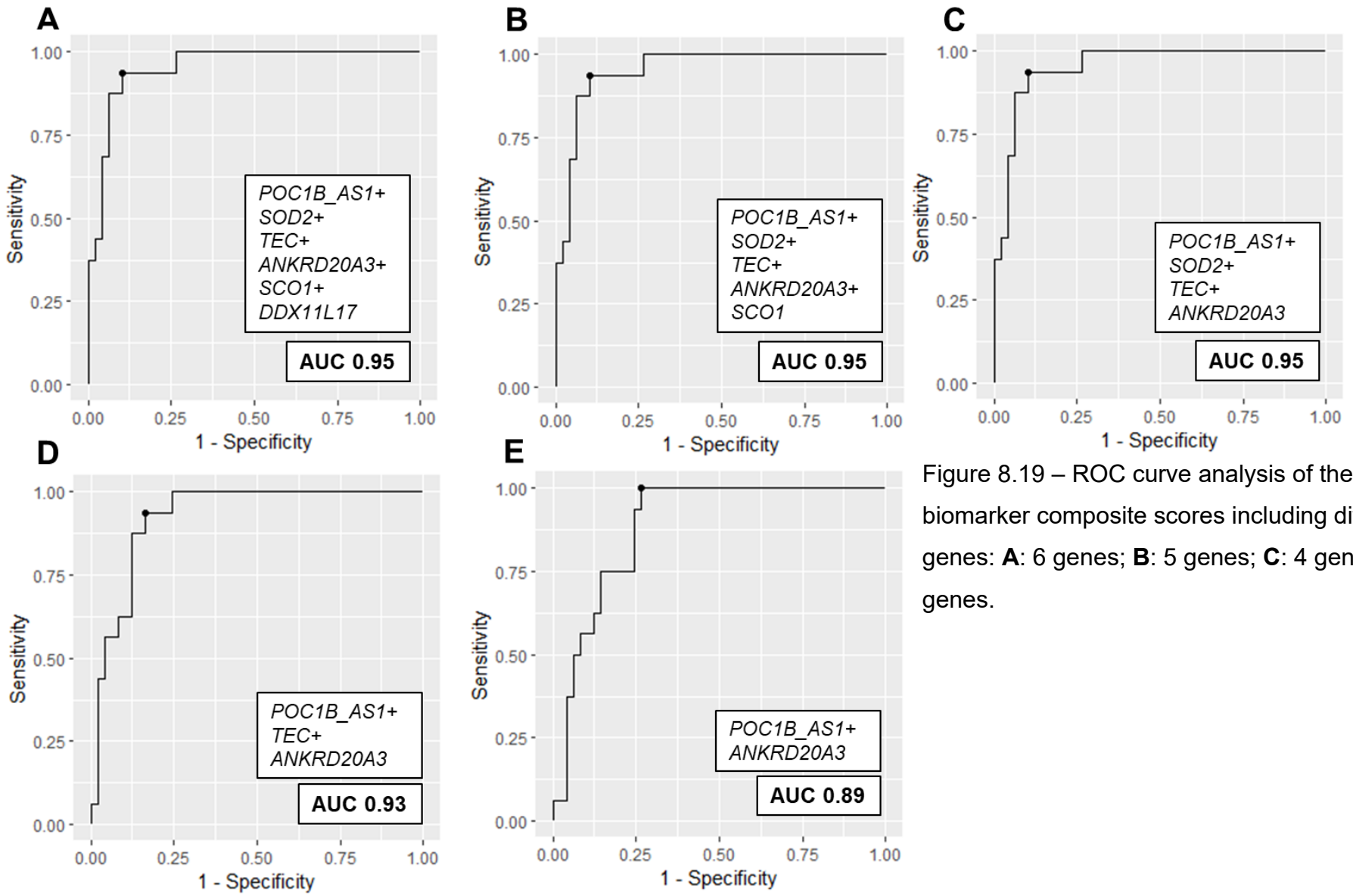


Figure 8.19 – ROC curve analysis of the top performing biomarker composite scores including different numbers of genes: **A**: 6 genes; **B**: 5 genes; **C**: 4 genes; **D**: 3 genes; **E**: 2 genes.

The difference observed in ROC_{AUC} between the 6 gene score (0.95) and 2 gene (0.89) is minimal despite the lower number of genes included. The single genes all had similar ROC_{AUC} scores between 0.71-0.79, with *SCO1* reported to be the best performing single gene. However, there was a significantly improved ROC_{AUC} (0.95) if these genes were combined with additional genes. Furthermore, the positive predictive value (PPV) fell with each gene removed from the composite score (0.83 (5-gene) vs. 0.55 (2-gene)). The predictive performance of the gene combinations incorporating different numbers of genes including sensitivity, specificity, and positive and negative predictive values is presented in Table 8.6.

No. of genes included	Genes	Threshold value	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	ROC_{AUC} (95% CI)
6	POC1B-AS1 +SOD2+TEC+ ANKRD20A3 +SCO1+ DDX11L17	-0.2105	0.94 (0.85 – 1)	0.90 (0.79 – 0.98)	0.75 (0.57 – 1)	0.98 (0.94 – 1)	**0.95** (0.91 – 0.99)
5	POC1B-AS1 +SOD2+TEC+ ANKRD20A3 +SCO1	-0.2163	0.94 (0.83 – 1)	**0.94** (0.86 – 1)	**0.83** (0.65 – 1)	0.98 (0.94 – 1)	**0.95** (0.90 – 0.99)
4	POC1B-AS1 +SOD2+TEC+ ANKRD20A3	-0.0766	**1.0** (1 – 1)	0.88 (0.81 – 0.96)	0.73 (0.59 – 0.90)	**1.0** (1 – 1)	**0.95** (0.90 – 0.98)
3	POC1B-AS1 +TEC+ ANKRD20A3	-0.1607	0.94 (0.88 – 1)	0.84 (0.72 – 0.94)	0.65 (0.36 – 0.86)	0.98 (0.95 – 1)	0.93 (0.87 – 0.97)
2	POC1B-AS1+ ANKRD20A3	-0.1319	**1.0** (0.88 – 1)	0.73 (0.65 – 0.89)	0.55 (0.42 – 0.75)	**1.0** (0.96 – 1)	0.89 (0.82 – 0.95)

Table 8.6 – Predictive value of the top (highest ROC_{AUC}) performing six, five, four, three and two gene combination composite scores, including the sensitivity, specificity, positive predictive (PPV) and negative predictive value (NPV).

Overall, the highest performing 4-gene (*POC1B-AS1+SOD2+TEC+ANKRD20A3*) combination with a ROC_{AUC} 0.95 had the highest sensitivity (1.0) and negative predictive value (1.0) while retaining a satisfactory PPV (0.73) and specificity (0.88) compared to other scores including those with a greater or fewer number of genes.

8.4.3 Survival analysis using gene composite scores

The survival distributions of the top (highest ROC_{AUC}) performing gene combinations with differing numbers of genes (2-6) as presented in Table 8.6 were analysed by the log-rank test. The Kaplan-Meier plots are presented in Figure 8.20. All the survival distributions for these gene combinations were significant (P=<0.0001).

Those assigned to the ‘positive’ cohort indicated that the patient’s gene score was above the optimal composite score cut-off threshold as determined by ‘cutpointr’ e.g., -0.2105 (for the 6-gene score) with any under this threshold considered ‘negative’. The thresholds used are presented in Table 8.6.

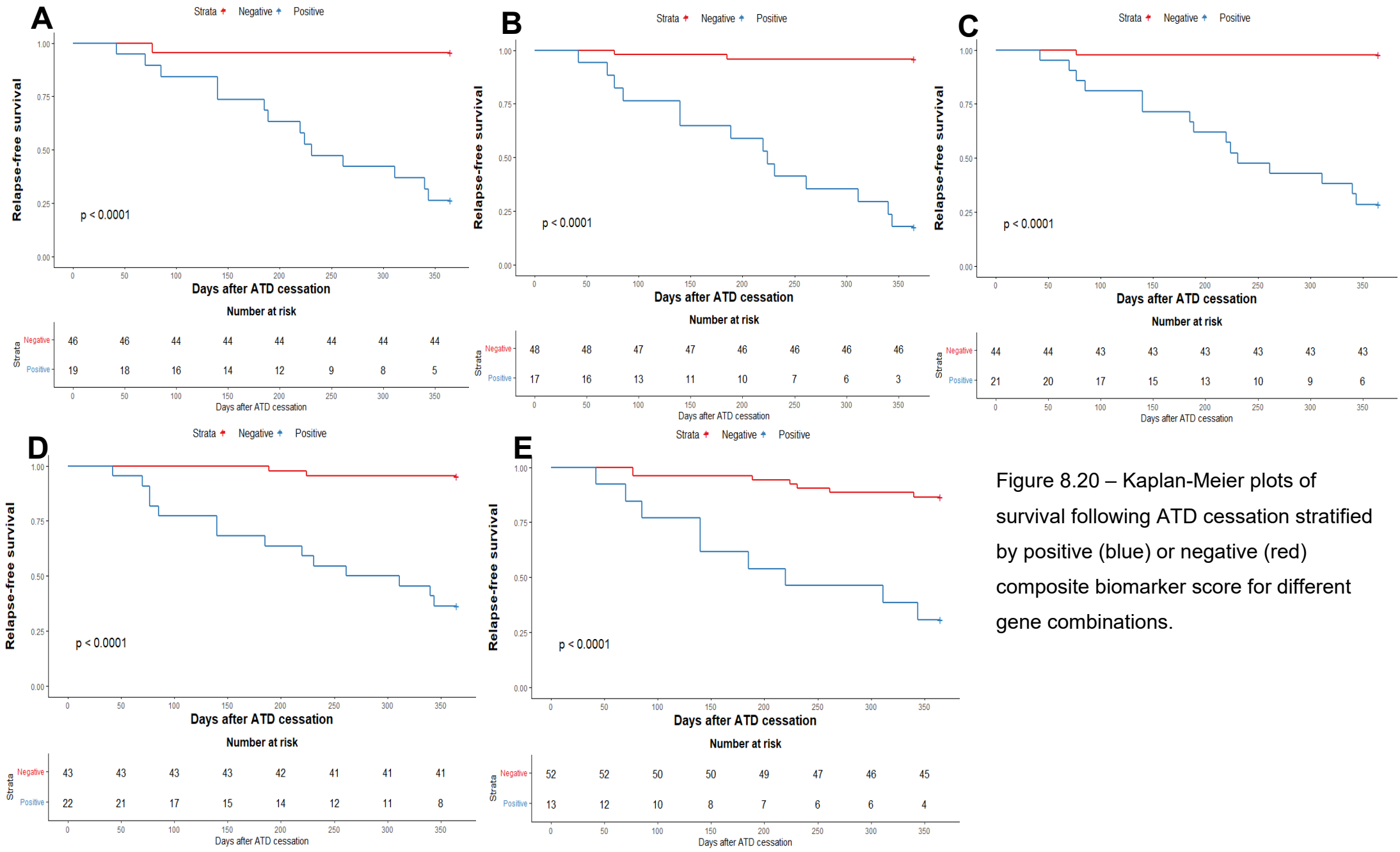


Figure 8.20 – Kaplan-Meier plots of survival following ATD cessation stratified by positive (blue) or negative (red) composite biomarker score for different gene combinations.

8.5 Discussion

The primary aim of the gene expression analysis was to determine whether there was differential gene expression between the relapse and remission patients at the time of stopping ATD. However, only two of the DEG, both Y-linked genes, were robust to multiple test correction. The lack of transcriptomic variation observed in the B cells of the GD study patients may reflect the fact that peripheral B cells are largely quiescent in patients taking ATD, and thus ATD are effective in suppressing the B cell activation that is presumably present at the onset of hyperthyroid GD.

Secondary analysis was performed to accept an unadjusted P value of <0.001 which identified 33 DEG. Of these 33 genes, 28 had a fold change ≥ 1.6 as previously specified in the power calculation. In addition, pathway analysis was performed on 258 DEG with an unadjusted P value of <0.01.

A review of the literature for the 33 DEGs with P<0.001 was undertaken to investigate their functional and biological characteristics with relevant findings discussed below. Further investigation of contributing genes involved in the pathway and network analysis is also discussed.

8.5.1. Differential gene expression of relapse and remission patients at ATD withdrawal

Y-linked genes

Two Y-linked genes remained significantly differentially expressed after multiple test correction, namely DEAD-box helicase 3 Y-linked *(DDX3Y)* and Ubiquitin specific peptidase 9 Y-linked (*USP9Y*). Although GD is commoner in females, men are more likely to relapse (*Suzuki N et al. 2021*). Therefore, although these genes may represent sex-biased gene expression rather than true differential gene expression, given that sexually dimorphic immune responses are well-described (*Gal-Oz ST et al. 2019*) and gender is associated with outcome in GD, further investigation of Y-linked genes is warranted.

USP9Y is a member of the peptidase C19 family and prevents the degradation of specific proteins through the removal of conjugated ubiquitin, with a role in the TGF (transforming growth factor)-beta(β)/BMP(bone morphogenetic protein) signalling cascade (*UniProt Consortium (2023)).* TGF-β is a key cytokine with a crucial

immunoregulatory role causing both suppression and activation, including specific effects on humoral immunity that involve the inhibition of B cell activation and reduction of antibody synthesis and production (*Kardalas E et al. 2021*). The deficiency of TGF-β has been implicated in promoting GD by causing an increase in IFN (interferon) γ-induced expression of MHC (*Kardalas E et al. 2021*). Interestingly, USP9Y has been identified as a potential diagnostic biomarker for multiple sclerosis in males and found to be positively correlated to memory B cells (*Zhang S et al. 2024*). Therefore, the upregulation of *USP9Y* observed in relapse patients may result in disruption of the TGF-β signalling pathway resulting in dysregulated TGF-β and an increase in B cell autoreactivity and a breakdown of immune tolerance.

DDX3Y is a member of the DEAD-box RNA helicase family and has also been reported to have an immunoregulatory role by enhancing IFN- β1 expression via IRF (Interferon regulatory factor) 3/IRF7 pathway (*UniProt Consortium (2023)).* Due to the aberrant production of type-I interferon, IRF7 is implicated in the pathogenesis of a number of autoimmune diseases including SLE and type I diabetes (*Fu Q et al. 2011, Heinig M 2010, Honda K et al. 2005*). Indeed, in the disease and functional downstream analysis using my dataset, all four Y-linked differentially expressed genes were identified to be related to and upregulated in SLE. These Y-linked genes may have important links to immunoregulatory signalling pathways and therefore their upregulation in relapsed patients could have a potential role in GD relapse and provide mechanistic insights into the increased relapse risk in males.

SOD2

SOD2 (Superoxide dismutase 2) is an isoform of the SOD antioxidant enzymes and is expressed in human mitochondria (*UniProt Consortium (2023)).* In my dataset it was found to be significantly upregulated in the relapse group with a Log_2FC of 0.5 (FC 1.4). Reactive oxidative species (ROS) produced because of oxidative stress are neutralised by these enzymes through catalysing the conversion of superoxide anion radicals to hydrogen peroxide and oxygen (*Palma FR et al. 2020*). Therefore, they have a protective role in preventing the accumulation of ROS and consequent cellular and tissue damage.

ROS are closely linked to thyroid hormones and can both affect and be affected by thyroid hormone status. ROS are essential in the initial stage of thyroid hormone synthesis during iodide oxidation (*Massart C et al. 2011*), and the hypermetabolic

state of thyrotoxicosis consumes antioxidants and promotes the accumulation of ROS in peripheral tissues and the thyroid. This may potentially cause damage to thyroid cells and drive thyroid autoimmunity through exposure of thyroid autoantigens (*Saban et al. 2022*). Indeed, GD and GO have been associated with increased markers of oxidative stress (*Saban et al. 2022, Diana T et al. 2018, Tsai C et al. 2009*), and the increased synthesis of ROS has been observed in GD compared to healthy controls with serum TRAbs demonstrated to augment the generation of ROS (*Diana T et al. 2018*). Furthermore, increased synthesis of ROS has been observed in Graves' disease relapse compared to euthyroid GD (*Dudina et al. 2022*).

Consistent with the finding of increased oxidative stress in GD, several studies have described significantly greater SOD activity in patients with hyperthyroidism compared with control groups, which gradually decreased with ATD therapy (*Komosinska-Vassev K et al.* 2000*, Zarković M. 2012, Guerra LN et al. 2001).* A polymorphism in *SOD2* (rs2758332) has been associated with increased susceptibility to SLE in children (*Abd El Azeem R et al. 2021*).

Therefore, SOD2 is a biomarker of oxidative stress with elevated levels suggesting the increased presence of ROS. The upregulation of *SOD2* in the relapse group in my dataset may directly reflect a hypermetabolic state caused by higher thyroid hormones concentrations in these patients. However, as discussed previously, thyroid inflammation may be a contributing factor to relapse in GD and high oxidative stress may induce and augment inflammation in the thyroid, leading to thyroid cell damage, the release of pro-inflammatory cytokines and the breakdown of self-tolerance, ultimately driving relapse (*Marcocci C et al. 2012, Ranneh Y et al. 2017*).

IL-1β

IL-1β (interleukin-1 beta) is a key proinflammatory cytokine produced by various immune cells including B cells and has been shown to induce the production of another inflammatory cytokine implicated in GD, IL-6 (*Tosato G et al. 1990).* In my dataset IL-1β was found to be significantly upregulated in the relapse group with a Log_2FC of 1.3 (FC 2.5) and it was also a key representative gene in the top significantly enriched canonical pathways in my dataset.

IL-1β beta has been described to disrupt immune tolerance through expansion of CD25+ effector T cells (*O'Sullivan BJ et al. 2006*), and elevated serum levels of IL-1β have been found in several autoimmune diseases, including GD (*Buldygina YV et al.*

2021, Pedro AB et al. 2011). IL-1β levels have been demonstrated to normalise following ATD treatment (*Pedro AB et al. 2011),* suggesting a potential role for IL-1β as a biomarker of GD activity. Therapeutics that involve the inhibition of IL-1β have been used successfully to reduce disease severity in autoinflammatory diseases e.g., the IL-1 receptor antagonist, Anakinra (*Kaneko N et al. 2019*). Furthermore, *IL-1*β polymorphisms (-511 C/T and rs16944, rs1143634) have been associated with susceptibility to GD (*Chen RH et al. 2005, Liu N et al. 2010, Shehjar F et al. 2018*).

The increased expression of *IL-1β* in the relapse patients in my dataset may reflect the presence of an inflammatory environment which is driving the breakdown of immune tolerance. Indeed, a previous RNA sequencing study of PBMCs in GD demonstrated an increased expression of *IL-1β* in patients with persistent TRAbs (>5 years) (*Chen Z et al. 2021*). Furthermore, IL-1β has been shown to stimulate intracellular SOD activity in Graves' retroorbital fibroblast cultures in a dose-dependent manner (*Zarković M. 2012*), suggesting a potential role for IL-1β in upregulating SOD2 in the context of inflammatory-induced oxidative stress perpetuating autoimmunity as discussed above. Interestingly, both *IL-1β* and *SOD2* are part of a distinct cluster of 6 genes in the heatmap analysis that contain all the relapse patients again suggesting a potential contributing role for these genes in the relapse of GD.

POC1B-AS1

POC1B-AS1 (POC1B Antisense RNA 1) is a long non-coding RNA gene (lncRNA), that was significantly downregulated in my dataset with a Log_2FC of -0.5 (FC -1.4). lncRNA is an important regulator of gene expression in immune cells and several lncRNAs have been identified to have immunoregulatory roles including receptor signalling, immune cell proliferation and differentiation and cytokine expression (*Atianand MK et al. 2014, Aune TM et al. 2017, Wang J et al. 2016*). Indeed, a microarray profile of the B cells in GD has found lncRNA to be associated with modulating B cell proliferation and survival (*Jiang X et al. 2020*). Furthermore, a lncRNA (*HCP5*) has been implicated in susceptibility to GD and associated with age of GD onset (*Lane LC et al. 2020*).

Interestingly, *POC1B-AS1* has been demonstrated to be positively correlated to Treg cells (*Liu J et al. 2022*). In GD a lower percentage of circulating Treg cells has been

associated with a lower remission rate due to persisting TRAbs (*Teniente-Serra A et al. 2019, Chen Z et al. 2021*). Therefore, reduced expression of *POC1B-AS1* in my dataset may suggest a lower Treg population in the relapsing patients which highlights their potential role in the immune dysregulation driving relapse.

VSTM4

VSTM4 (V-set and transmembrane domain-containing 4) is part of an immunoglobulin superfamily of 8 transmembrane proteins. It was found to be significantly downregulated in my dataset, with a Log_2FC of -2.2 (FC -4.5). It has recently been demonstrated to act as a negative regulator of T cell activation and significantly reduce the T cell production of pro-inflammatory cytokines IFN-γ, IL-2, and IL-17 (*Wang J et al. 2019*). Therefore, the reduced expression of *VSTM4* observed in the relapse cohort in my dataset may reflect a lack of immunosuppressive activity in these patients contributing to a resurgence of autoimmunity.

LEAP2

LEAP2 (liver-expressed antimicrobial peptide 2) is produced in the liver and small intestine and is described to have antibacterial activity and act as an endogenous antagonist of the ghrelin receptor (*Ge X et al. 2018*). It was found to be significantly downregulated in the relapse patients my dataset, with a Log_2FC of -0.9 (FC -1.9). Contrary to this, elevated serum levels of LEAP2 have been found in RA and shown to correlate with the inflammatory markers, CRP and IL-6, suggesting that LEAP2 is associated with inflammatory status in RA patients and could be a potential biomarker of disease activity (*Francisco V et al. 2020*). *LEAP2* has not previously been studied in GD but did demonstrate reduced expression in the relapse patients in my dataset, which unlike my findings above is not necessarily consistent with an inflammatory environment.

HMGB1

HMGB1 is a member of the High-mobility group (HMG) family of proteins expressed in most human cells (*Ge Y et al. 2021*). It was found to be significantly downregulated in my dataset, with a Log_2FC of -2.3 (FC -4.8). HMGB1 functions as a damage-associated molecular pattern (DAMP) molecule which activates Toll-like receptors (TLRs) expressed on immune cells resulting in secretion of pro-inflammatory

cytokines via nuclear factor (NF)-κB signaling (*Ge Y et al. 2021*). HMGB1 has also been reported to form synergistic complexes with inflammatory cytokines, including IL-1β, to enhance proinflammatory activity (*Sha Y et al. 2008*). Both HMGB1 and TLRs have been suggested as potential therapeutic targets for autoimmune disease (*Ren W et al. 2023*). Elevated levels of circulating HMGB1 have been reported in several autoimmune diseases, including autoimmune thyroid disease (*Li, C et al. 2017, Schierbeck H et al. 2013*). Increased mRNA expression of *HMGB1* has been demonstrated in a murine model of autoimmune thyroiditis with reduced thyroid inflammation and lymphocytic infiltrate following the addition of a direct inhibitor of HMGB1 (*Li, C et al. 2017*).

HMGB1 was found to be downregulated in the relapse patients in my dataset. In the context of a potentially inflammatory environment in these relapsing patients, it may be expected that *HMGB1* expression would be increased in this group compared to the remission patients. The functional activity of HMGB1 has been shown to be closely related to oxidative stress and ROS have been demonstrated to directly promote the release of HMGB1 and perpetuate inflammation (*Zhang B et al. 2017*). To counteract this, antioxidants are released which have been demonstrated to inhibit oxidative stress-induced HMGB1 release through a feedback loop (*Zhang B et al. 2017, Tang D et al. 2011, Janko C et al. 2014*). Therefore, the increased expression of the antioxidant *SOD2* observed in the relapse patients in my dataset may alter the redox state and prevent the accumulation of ROS, thereby reducing *HMGB1* expression particularly if inflammation is chronic.

8.5.2 Canonical Pathway Analysis: enriched genes

Cytokines are a fundamental part of the humoral immune response in GD. Consistent with this, the canonical pathway analysis revealed that the 'Cytokine Signalling' category formed the most significant pathways modulated by the DEG in my dataset. The genes represented in these pathways include *IL-1*β and *IL-13* which are both found to be present in most of the top enriched canonical pathways. *IL-1*β was also found to be differentially expressed in the relapse patients and is discussed above.

The increased expression of both *IL-1*β and *IL-13* in persisting GD (the presence of TRAb > 5 years), has been demonstrated on RNA sequencing of PBMCs suggesting an inflammatory role in driving autoimmunity (*Chen Z et al. 2021*). IL-13 has been implicated as a B cell stimulating factor, with a role in upregulating MHC II expression

and directing naïve B cells to switch to IgG4 (*Defrance T et al. 1994, Punnonen J et al. 1993*). Elevated levels of serum IL-13 has been associated with GD recurrence (*Komiya I et al. 2001*). Furthermore, a Japanese study found an *IL-13* polymorphism (rs1800925) was associated with susceptibility to GD (*Hiromatsu Y et al. 2005*), with the T allele of this polymorphism found to be more frequent in controls and GD remission patients (*Inoue N et al. 2011*). Conversely, the T allele has been observed to be associated with enhanced IL-13 promoter activity in T helper type 2 lymphocytes (*Cameron L et al. 2006*). IL-13 has been described as an 'anti-inflammatory' cytokine because of its capacity to inhibit proinflammatory cytokines (*de Vries JE 1998*), which may explain why there could be increased IL-13 activity in some GD remission patients.

Another significant cytokine signalling canonical pathway identified involves the IL-17 cytokines, IL-17A, IL17D and IL-17F. Both IL-17A and IL-17F are produced by T helper (Th)17 cells which differentiate from naïve $CD4^+$ T cells following stimulation from TGF-β, IL-6, IL-23, and IL-1β (*Veldhoen M et al. 2006*). The pathogenic role of Th17 cells in autoimmune disease has been described as maintaining and amplifying inflammatory reactions (*Wang Y et al. 2023*), and they have been shown to directly provide B cell help by driving B cell proliferation and promoting the formation of germinal centres and antibody production (*Mitsdoerffer M et al. 2010*). Several studies have found increased proportions of Th17 cells in GD compared to controls, with GD patients demonstrating elevated serum IL-17A with positive associations to both circulating TRAb concentration and intractable GD (*Nanba T et al. 2009, Peng et al. 2013, Li JR et al. 2013*). IL-1β has been described as the dominant cytokine driving the differentiation of naïve T cells into Th17 cells (*Acosta-Rodriguez EV et al. 2007, Lasigliè D et al. 2011*). The reported association of Th17 cells with GD and the upregulation of IL-1β observed in the relapse patients in my dataset may suggest a pathogenic role for the IL-1/IL-17 axis and the associated B cell activation and germinal centre development in GD relapse.

The *APOD* (Apolipoprotein D) gene which was also represented in the top canonical pathway encodes a small glycoprotein (ApoD) that is described to function as an antioxidant, modulating oxidative stress and inflammation, by binding to and blocking arachidonic acid signalling which reduces the production of pro-inflammatory molecules (*Fyfe-Desmarais G et al. 2023*). Transcriptomic and proteomic studies have found that *APOD* demonstrates Th17-specific expression and is upregulated in

Th17 cells suggesting ApoD as a potential biomarker of Th17-mediated inflammation (*Sałkowska A et al. 2020, Tripathi SK et al. 2019*). Given the potential role of Th17 cells in GD, ApoD could be relevant in the context of monitoring disease activity and suggests a potential mechanism for oxidative stress and inflammation in GD relapse.

Overall, the enriched cytokine signalling pathways and representative genes are suggestive of an active immune inflammatory environment with associated oxidative stress, providing potential mechanistic insight into the pathways driving GD relapse.

8.5.3 Causal analysis: upstream transcriptional regulators

In causal analysis of my dataset, IL-6 was the top predicted upstream transcriptional regulator. As previously discussed, IL-6 is a key component of the proinflammatory cytokine pattern observed in GD and has been associated with susceptibility and severity (Chapter 6: Results 3 – Cytokine and Chemokine data). Although the *IL-6* gene itself was not detected as differentially expressed in the RNAseq dataset, this may reflect the sensitivity of the causal analysis (*Krämer A et al. 2014)*.

IL-6 activates both the pro-inflammatory cytokine IL-1β and the antioxidant enzyme SOD2, suggesting a pattern of perpetuating inflammation and a compensatory response of increasing SOD2 expression to counteract the accompanying inflammatory-mediated oxidative stress. IL-6 has also been shown to inhibit downstream ADAMTS13 (a disintegrin and metalloprotease with a thrombospondin type 1 motif, member 13). Autoantibodies against ADAMTS13 lead to the development of the autoimmune disorder, thrombotic thrombocytopenic purpura (TTP) and there are rare case reports of an association between TTP and GD occurring concurrently (*Lhotta K et al. 2018*). ADAMTS13 has been proposed to have an anti-inflammatory role as decreased levels are associated with systemic inflammation and MS (*Ziliotto N et al. 2018, Kashiwakura Y et al. 2006*), and exogenous ADAMTS13 treatment in a murine model has been demonstrated to ameliorate the inflammatory response in experimental autoimmune encephalomyelitis (*Lu K et al. 2020*). Therefore, the inhibition of ADAMTS13 by IL-6 would reduce the regulation of inflammation.

GADD45β (Growth Arrest And DNA Damage Inducible Beta) which is activated downstream from IL-6 has been implicated in regulating autoimmunity and has a role in perpetuating inflammatory signals in T cells and promoting the release of pro-inflammatory cytokines (*Li RN et al. 2019, Lu B et al. 2004, Liu L et al. 2005*).

GADD45β is also expressed on germinal centre B cells and has been demonstrated to promote CD40-induced prosurvival signalling resulting in reduced apoptosis of activated B cells (*Zazzeroni F et al 2003*). In GD this has the potential to promote the survival of activated autoantibody-secreting B cells.

This upstream analysis highlights potential pathways of immune dysregulation within the adaptive immune response caused by IL-6 and its interaction with other downstream inflammatory mediators that result in a cycle of perpetual inflammation.

8.5.4 Molecular network analysis

Network analysis determined the top network to involve “Cell-To-Cell Signalling and Interaction, Cellular Function and Maintenance, Haematological System Development and Function, and Inflammatory Response”. The connected pathways identified in this network provide biological insights into the relationships between various DEG in the RNAseq dataset. The network highlights IL-1β as a key upstream regulator, and includes upregulated genes discussed previously such as *IL-13* and *SOD2*. Furthermore, both *IL-1β* and *IL-13* are also identified in the downstream functional analysis as key genes involved in the biological process of ‘Antigen presentation’ and the development of antigen-presenting cells, highlighting their potential pathogenic role in the humoral immune response and relapse of GD. Indeed, a single cell transcriptomic study of thyroid follicular cells in GD has highlighted the enrichment of genes and pathways involved with antigen processing and presentation (*Álvarez-Sierra D et al. 2023*).

CCL3L1 (chemokine ligand 3-like1) is also identified to be upregulated in the network. This chemokine binds to pro-inflammatory cytokine receptors to enhance inflammatory responses and increased *CCL3L1* expression has been demonstrated in RA and SLE (*McKinney C et al. 2008, Mamtani M et al. 2008*), although a study of copy number variations in *CCL3L1* was not associated with the development of GD (*Song RH et al. 2017*).

8.5.5 Gene composite scores survival analysis

The gene composite score analysis demonstrated that generally the gene combinations had high ROC_{AUC} (0.71-0.95). When translating a predictive score into clinical practice, as well as the effectiveness of the score, it is important to consider the cost effectiveness. By reducing the gene score to include only 4 out of the 6

genes, the ROC_{AUC} remained high (0.95) and retained good sensitivity, specificity and PPV/NPV. Survival analysis of the gene composite scores also demonstrated significant survival distributions.

Some of the genes included in the composite score are described to have potential immunoregulatory roles as discussed above (*SOD2* and *POC1B-AS1*) and therefore could be biomarkers of disease activity and outcome. *ANKRD20A3* and *DDX11L17* are both pseudogenes, however the role of pseudogenes exerting biological effects by regulating expression of coding genes is well described (*Pink RC et al. 2011*). There is no current literature on the role of these pseudogenes or that of the *TEC* gene (ENSG00000280216) in immune disease states, however other pseudogenes have been found associated with autoimmune thyroid disease so they may have undiscovered functional capabilities (*Li CW et al. 2021*). Although *SCO1* is also not described to have any direct effects on the immune system, it is reported to have a role regulating copper homeostasis which may indirectly affect the immune response (*Festa RA et al. 2012*).

Further analysis aiming to improve the composite score by integrating other significant variables from my study is presented in Chapter 9: Results 6 – Integrative Analysis.

8.6 Limitations

A key limitation of this RNA sequencing study is that only two of the DEG, both Y-linked genes, were robust to multiple test correction. This is likely related to the small sample size and magnitude of differential gene expression between the relapse and remission groups. In addition, although the model was corrected for gender, there is likely to have remained some sex-biased gene expression and distinguishing between whether differential gene expression is related to GD relapse, for which gender is described as a risk factor, and not just due to sexual dimorphism is challenging. Despite this, on further analysis many immunoregulatory biological molecules and pathways were identified that are relevant in the context of autoimmunity.

The sample analysed by RNA sequencing was taken whilst all the patients remained on ATD. It is unknown whether there is any specific effect of ATD on the $CD19^+$ B cell transcriptome, however ATD is described to potentially have an

immunomodulatory effect and therefore could potentially alter immune-related gene expression. However, it was important to sequence the B cells prior to stopping ATD treatment as this is the time a clinical decision would be made regarding management and therefore would allow the results to have greater translational potential.

Bulk RNA sequencing of the $CD19^+$ B cells provides useful information on the activity of the total population of $CD19^+$ B cells. However, this doesn't allow insight into the individual expression and functional activity of the specific B cell subpopulations. The advancement in technology now enables the use of single-cell and spatial transcriptomic approaches which may provide more accurate data and identify precise novel pathways and therapeutic targets.

8.7 Summary

This study of the $CD19^+$ B cell transcriptome highlights components of the humoral immune system that may be relevant in advancing our understanding of the mechanisms driving GD relapse. Although few genes were robust to multiple test correction, pathway, and network analysis of the DEG has identified potentially pertinent immune mediators and pathways involved in cytokine cell signalling, inflammation, and oxidative stress. The B cell signature highlights the balance of various humoral immune components that are functioning in synergy or antagonistically in reaction to inflammatory environmental stimuli to maintain immune homeostasis. Many cytokines and inflammatory mediators have the capacity to induce each other in a perpetuating cycle of inflammation that may serve to enhance autoimmunity.

Analysis of the gene composite scores in predicting GD relapse highlights the potential for the inclusion of genes in a predictive model. Further analysis using these genes in conjunction with other markers of humoral immunity studied is presented in the next chapter.

Chapter 9: Results 6 – Integrative Analysis

In the previous Results Chapters I have presented the results from the specific clinical, biochemical, immunological, and transcriptomic analyses. The aim of this chapter is to integrate the results from the previous Results Chapters to investigate how the individual components of the humoral immune system interact and use this information to form a composite biomarker score that encompasses the significant factors associated with outcome in GD, with the potential to translate to clinical practice.

9.1 Cytokines and monocytes

Monocytes are key components of the innate immune system, but they also have an important role in the adaptive immune system activating B and T cells through their ability to present antigen. They are also responsible for producing proinflammatory cytokines studied in this thesis, such as BAFF and IL-6.

In the previous Results chapters of this thesis, an elevated BAFF concentration and MLR were both independently observed to be associated with time-to-relapse in GD using the multivariate Cox regression model. Therefore, further analysis was undertaken to determine whether there was any relationship between BAFF levels and the absolute monocyte count or MLR at baseline and 6-10 weeks. However, the BAFF concentration was not observed to be significantly correlated to the absolute monocyte count or MLR at baseline or 6-10 weeks.

On the contrary, IL-6 levels were found to be positively associated with the absolute monocyte count both at baseline ($P=6.2 \times 10^{-4}$, rs 0.41) and at 6-10 weeks ($P=3.1 \times 10^{-3}$, rs 0.38) (Figure 9.1) and there was a positive correlation between IL-6 and MLR at baseline (P=0.026, rs 0.28) (Figure 9.2).

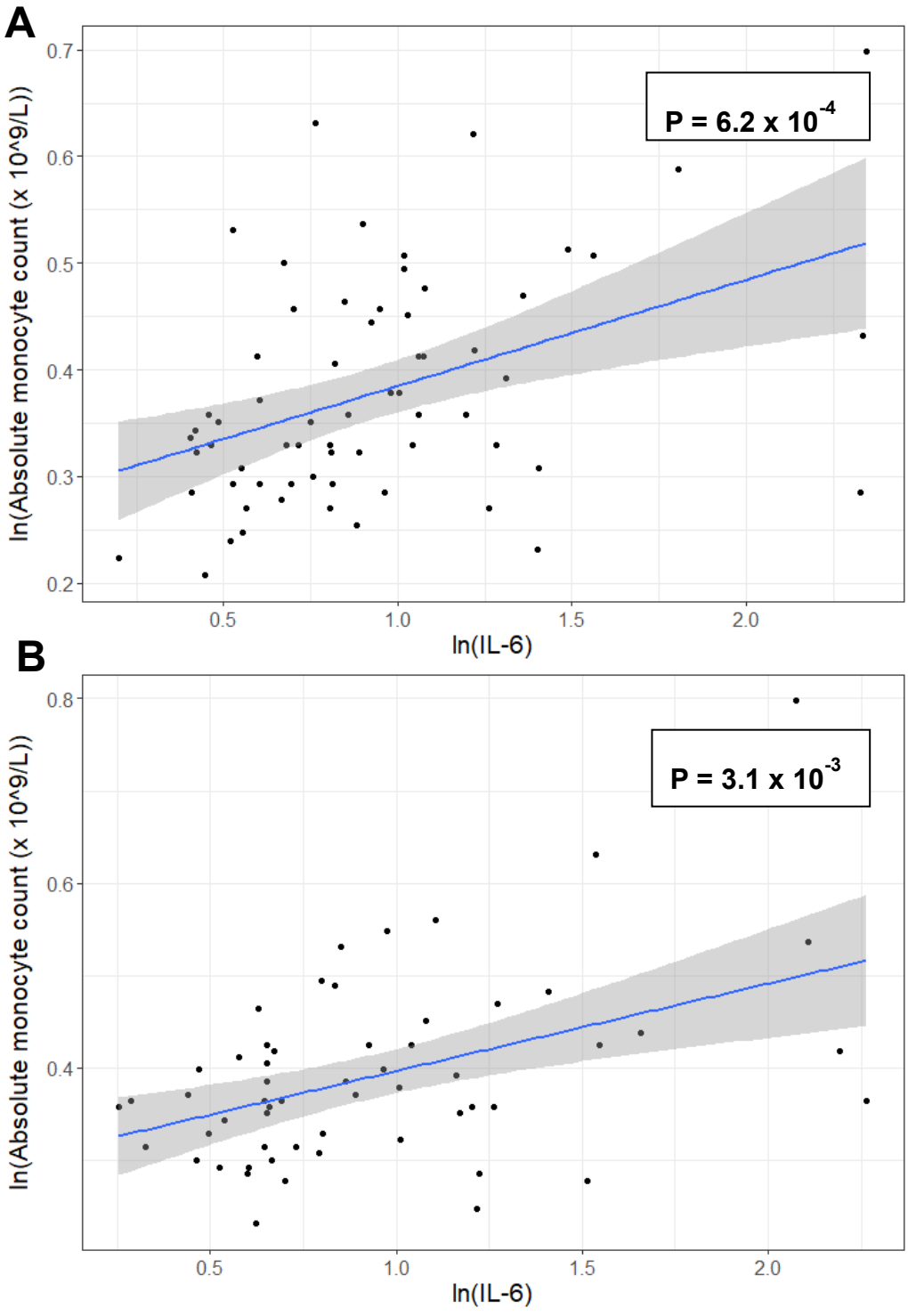


Figure 9.1 – Association between log transformed (ln) serum IL-6 levels and absolute monocyte count at ATD withdrawal (**A**) and 6-10 weeks later (**B**).

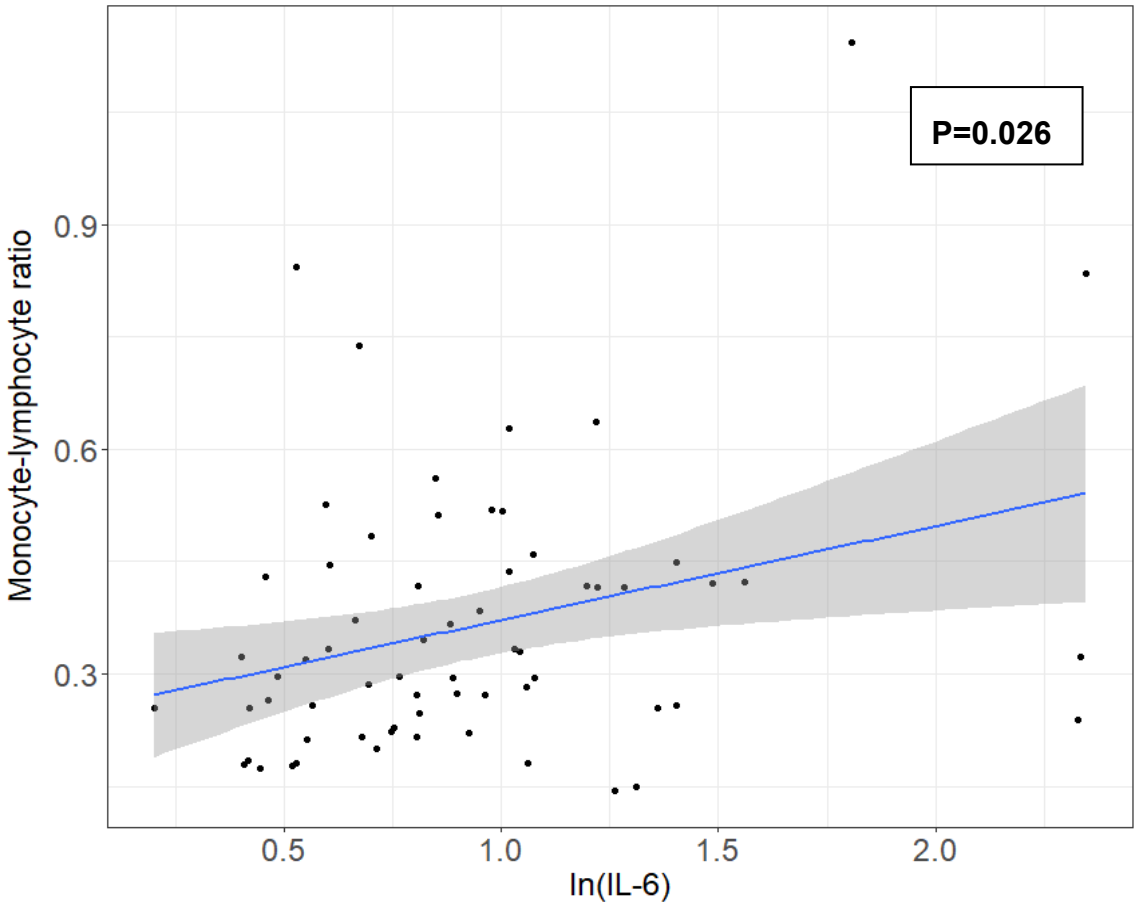


Figure 9.2 – Association between log transformed (ln) serum IL-6 levels and monocyte-lymphocyte ratio at ATD withdrawal.

Given the positive association between IL-6 and monocytes, alongside the previous observation of a positive association between IL-6 and CRP (Chapter 6: Results 3 – Cytokine and Chemokine data), the relationship between monocytes and CRP was also investigated. There was a positive association observed between CRP and absolute monocyte count ($P=6.04 \times 10^{-6}$, rs 0.40) and the MLR (P=0.018, rs 0.21) (Figure 9.3).

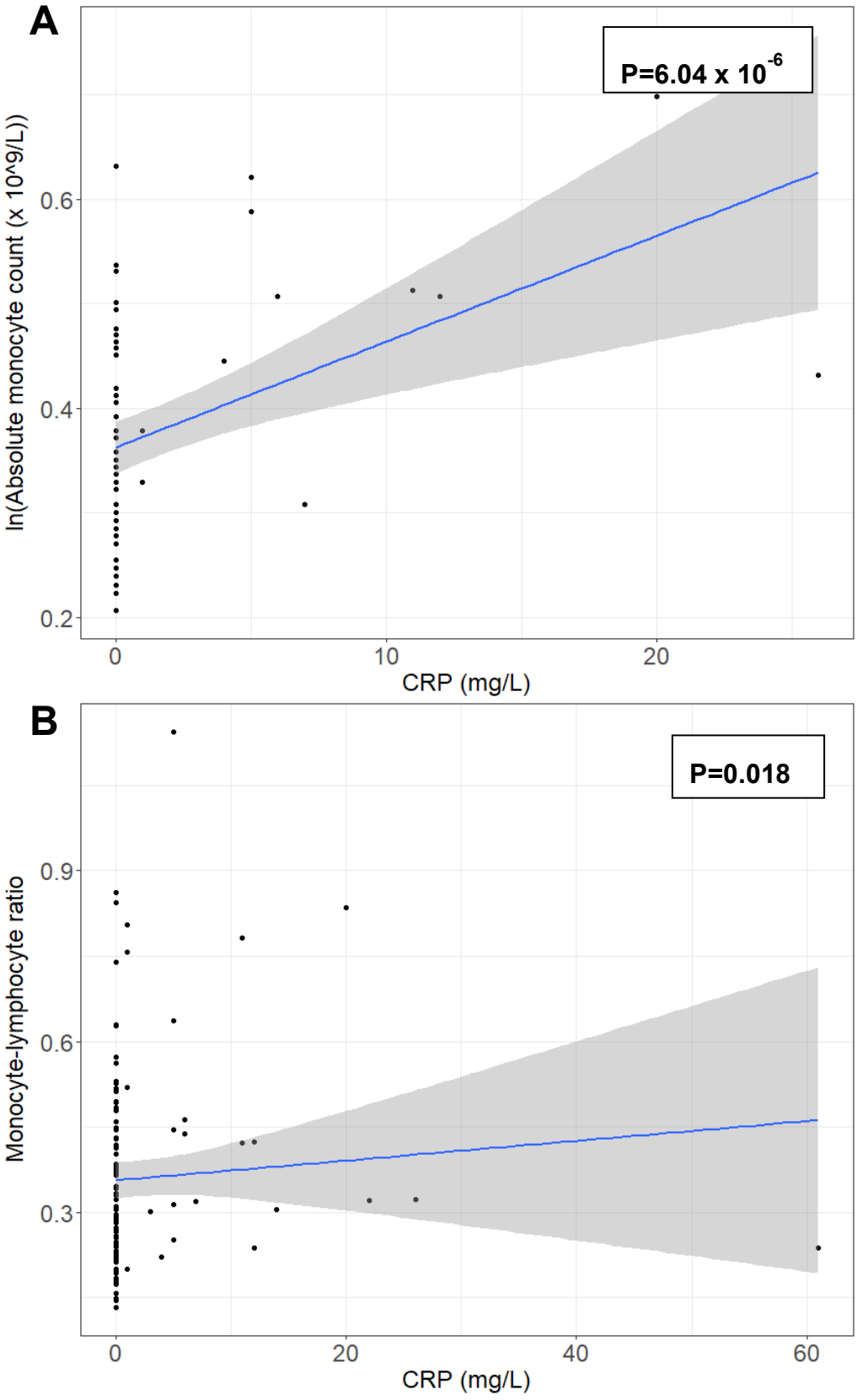


Figure 9.3 – Association between serum CRP levels and absolute monocyte count (**A**) and the monocyte-lymphocyte ratio (**B**). **9.2 Cytokines and B cell subpopulations**

Analysis was undertaken to determine the relationships between the peripherally circulating B cell subpopulations and cytokine concentrations at baseline prior to ATD withdrawal. Serum BAFF was positively correlated to the transitional cell population and the B regulatory cell 2 phenotype (P=0.0002, rs 0.44; and P=0.017, rs; 0.29, respectively) but was negatively correlated to the B regulatory cell 1 phenotype (P=0.012, rs;-0.31) (Figure 9.4). There was a weak positive and negative correlation between IL-6 and naïve B cells (P=0.04, rs 0.25) and switched memory cells (P=0.03, rs-0.27), respectively. There was a positive correlation observed between sTACI and the total $CD19^{+}$ B cell count (P=0.04, rs 0.25) (Figure 9.4). There was no observed association between CXCL13, APRIL, sBCMA or sTACI and any B cell subpopulation.

Following multiple test correction (Benjamini-Hochberg), the correlation between BAFF and transitional cells remained significant (P=0.01; rs 0.44).

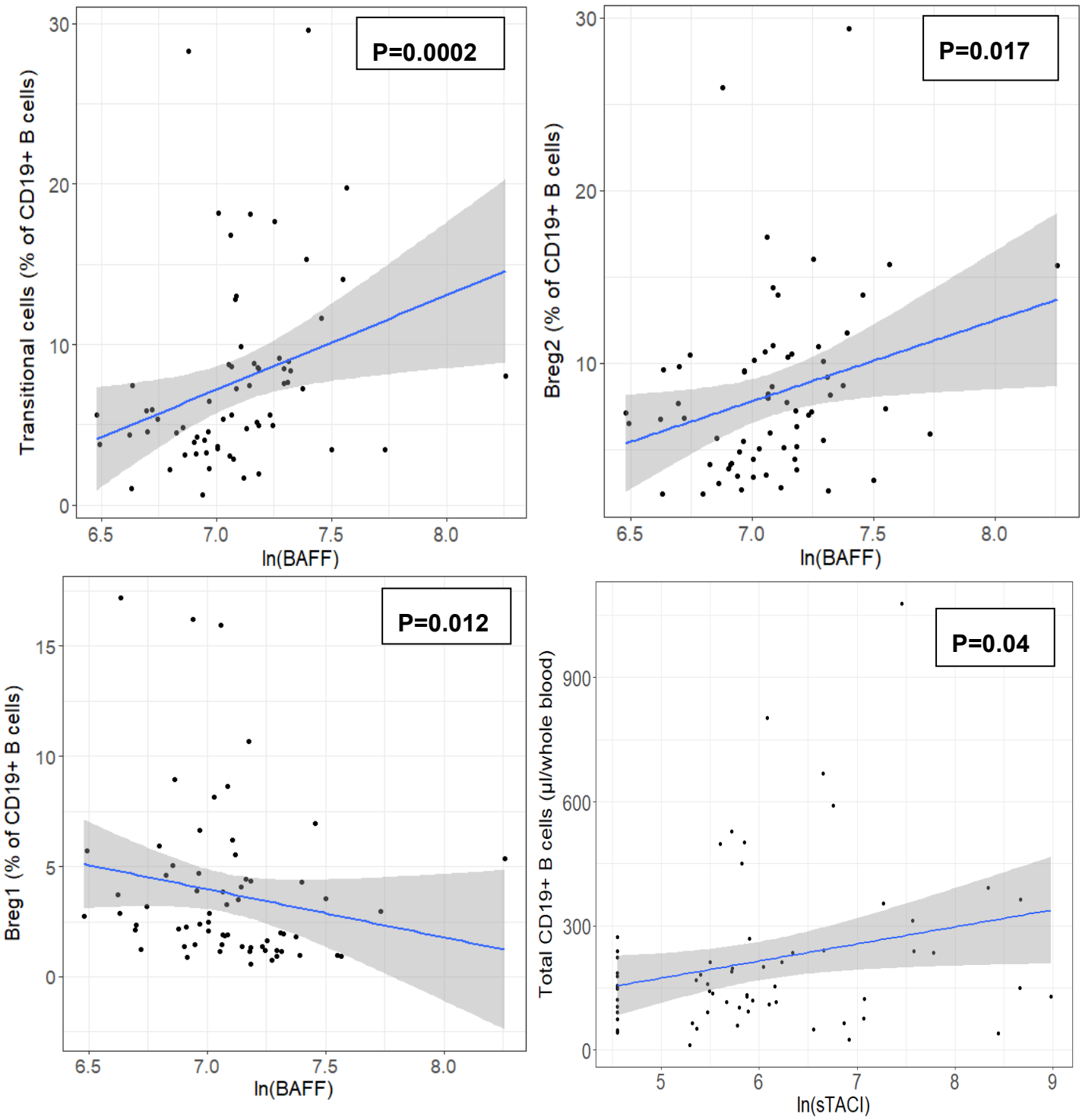


Figure 9.4 – Correlations between the different B cell subpopulations and log transformed (ln) cytokines.

9.3 Cytokine heatmap integrated analysis

The cytokine heatmap presented in Chapter 6: Results 3 – Cytokine and Chemokine data, highlighted two clustered groups of relapse patients (Group 1 and 2). The aim of examining separate groups of relapsing patients was to determine if there were any cytokine biomarkers that could discriminate relapse subtypes which could translate into a greater understanding of clinical prognosis. The main difference noted visually between the groups was the increased expression of sTACI and sBCMA in Group 1 patients. Indeed, further analysis showed elevated levels of both sTACI (P=0.006) and sBCMA (P=0.02) in Group 1 compared to Group 2 (Figure 9.5).

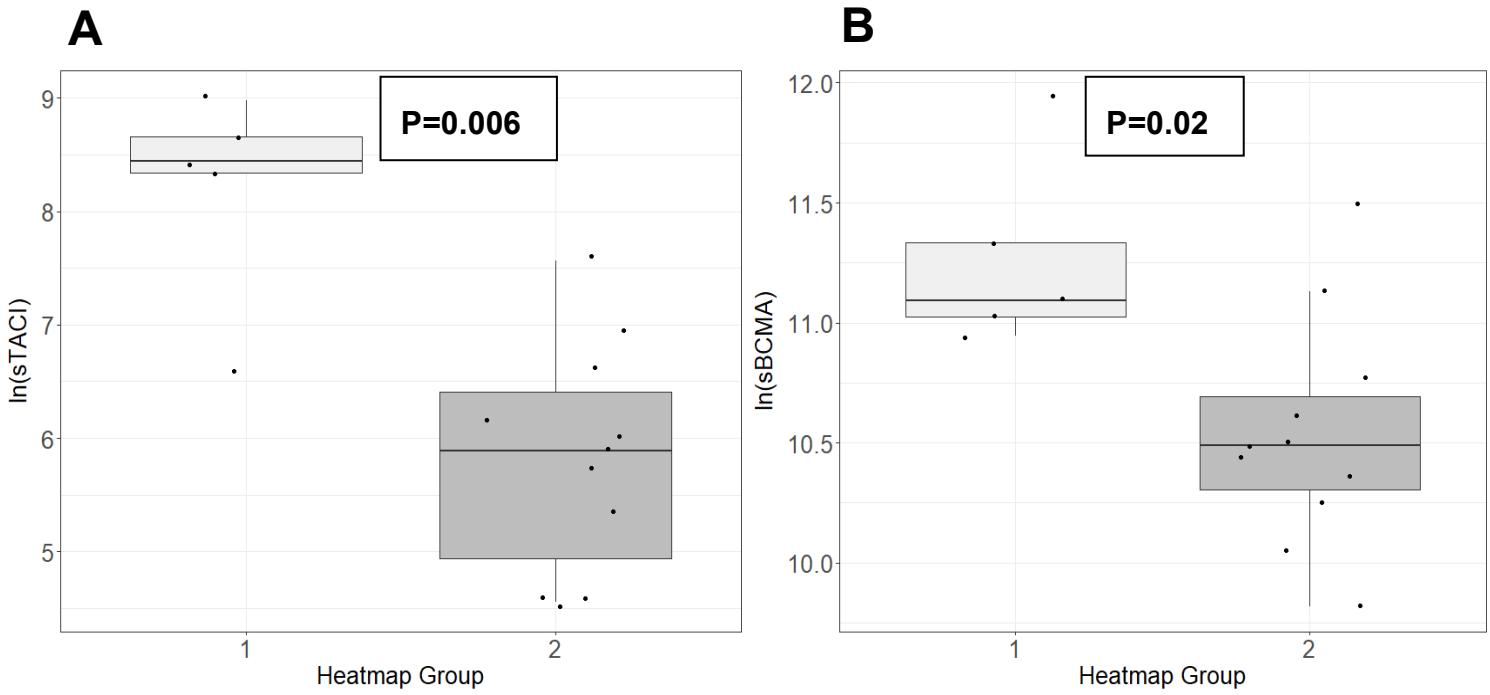


Figure 9.5 – Association between two groups of relapse patients (as clustered on cytokine heatmap analysis) with log transformed (ln) sTACI (**A**) sBCMA (**B**)

To further characterise these two separate identified clusters of relapse patients, clinical and biochemical parameters (patients demographics, thyroid function, B cell subpopulations and KRECs) were compared between the two groups.

There was no observed difference in most demographic and clinical characteristics including age, gender, smoking status, goitre size, CRP, or duration of ATD treatment. However, patients in Group 1 were observed to have lower levels of FT3 at the time of ATD withdrawal (P=0.01) (Figure 9.6), with no significant difference in

FT4 (P=0.46) or TRAb (P=0.10) concentrations. Group 1 was also observed to have a lower switched B memory cell (P=0.023) and plasmablast (P=0.025) population, and a higher naïve B cell population (P=0.04) (Figure 9.6). Time-to-relapse did not differ between the two groups (Group 1 = 189 days (median), Group 2 = 185 days (median), P= 1). As there was an expanded naïve cell population in Group 1, KRECs were assessed for any variation between the groups, however there was no difference in KRECs observed (P=0.47).

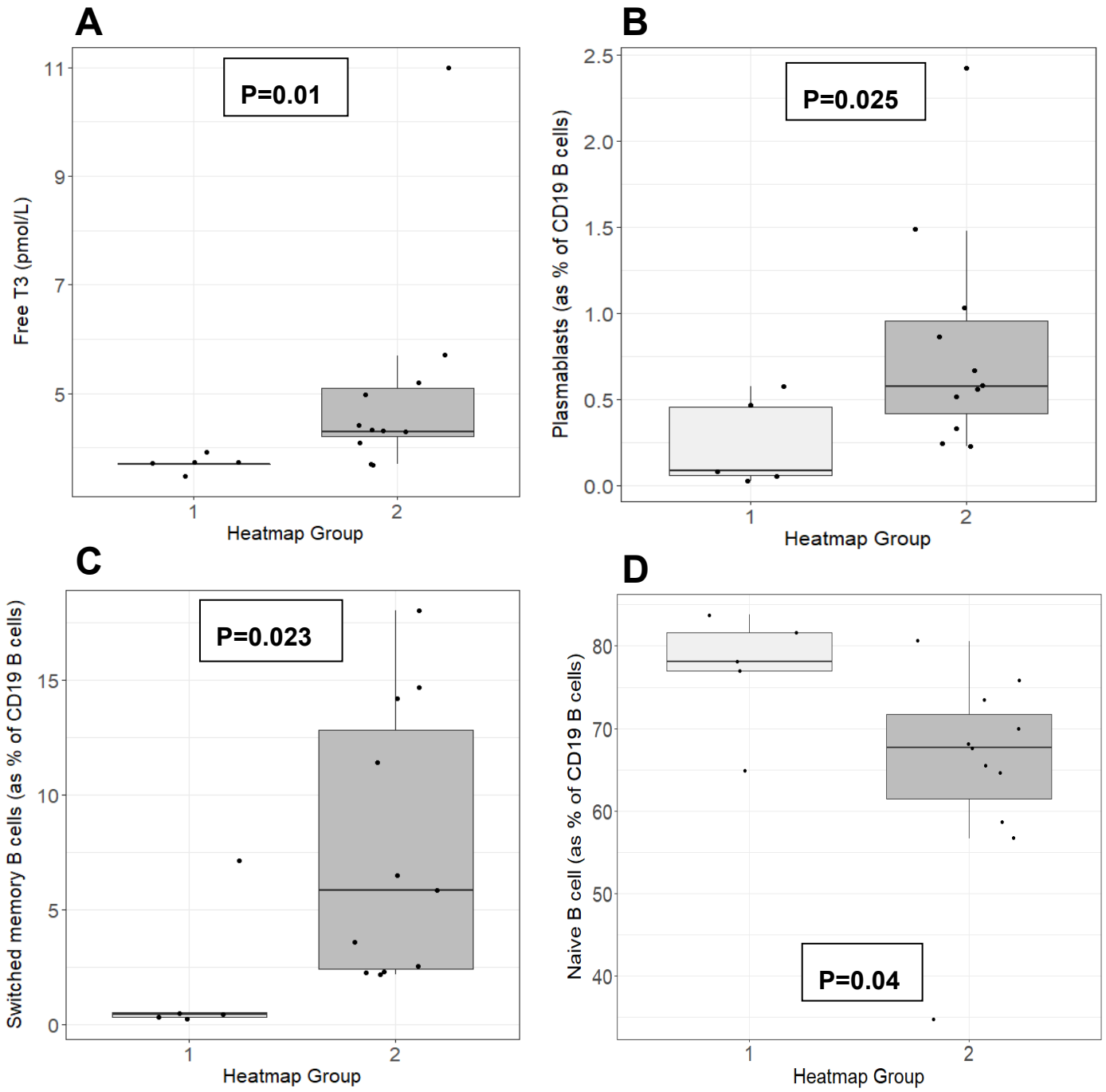


Figure 9.6 – Association between two subgroups of relapse patients (as clustered on heatmap analysis) with Free T3 (**A**) plasmablasts (**B**) switched memory B cells (**C**) and naïve B cells (**D**)

The gene expression levels of the 33 most DEG observed between relapse and remission groups were investigated in the two groups of clustered relapsed patients

in the heatmap for variations that may distinguish relapse subtypes, however no significant differences were observed.

9.4 Gene expression heatmap integrated analysis

The gene expression heatmap presented in Chapter 8: Results 5 – CD19+ B cell RNAseq highlighted two clustered groups of patients (Group 1 and 2), which appeared to have variations in their gene expression levels in two distinct clusters of genes. All the relapse patients were found in Group 2 and therefore comparisons were made between the remission patients in the two groups to determine if there is a specific gene expression profile associated with other components of the humoral immune system in patients that remain immune tolerant at one year after ATD withdrawal. This may provide mechanistic insights that could translate into a greater understanding of clinical prognosis.

To further characterise these two separate identified clusters of remission patients, clinical and biochemical parameters (patients demographics, thyroid function, B cell subpopulations and KRECs) were compared between the two groups.

The duration of taking ATD treatment (months) was significantly shorter in the remission patients in ‘Group 1’ (median 15.7) compared to ‘Group 2’ (median 18.2) (P=0.0008) (Figure 9.7). Furthermore, the patients in ‘Group 1’ had more circulating $CD19^{+}$ B cells (/µl whole blood) (median 103.6 vs. 211.5: P=0.0015) and sTACI (median 224.6 vs. 354 :P=0.03) (Figure 9.8). There was no observed difference in clinical or demographic characteristics including age, gender, smoking status, goitre size, CRP, thyroid status, KRECs, or B cell subpopulations.

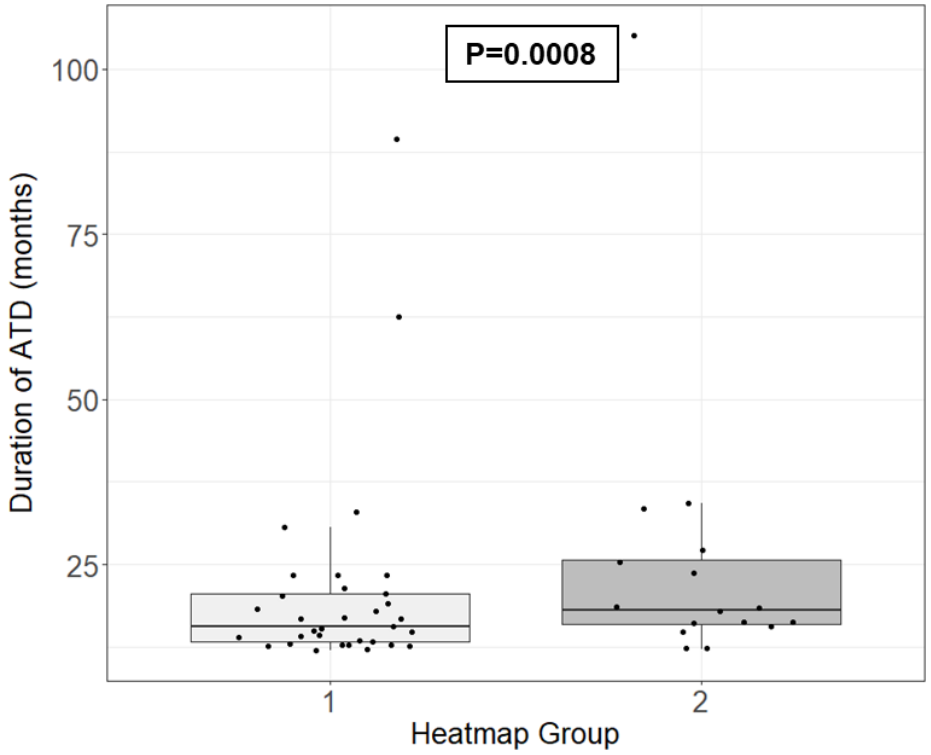


Figure 9.7 – Association between two groups of patients (as clustered on gene expression heatmap analysis) with duration of antithyroid drug (ATD) treatment (months)

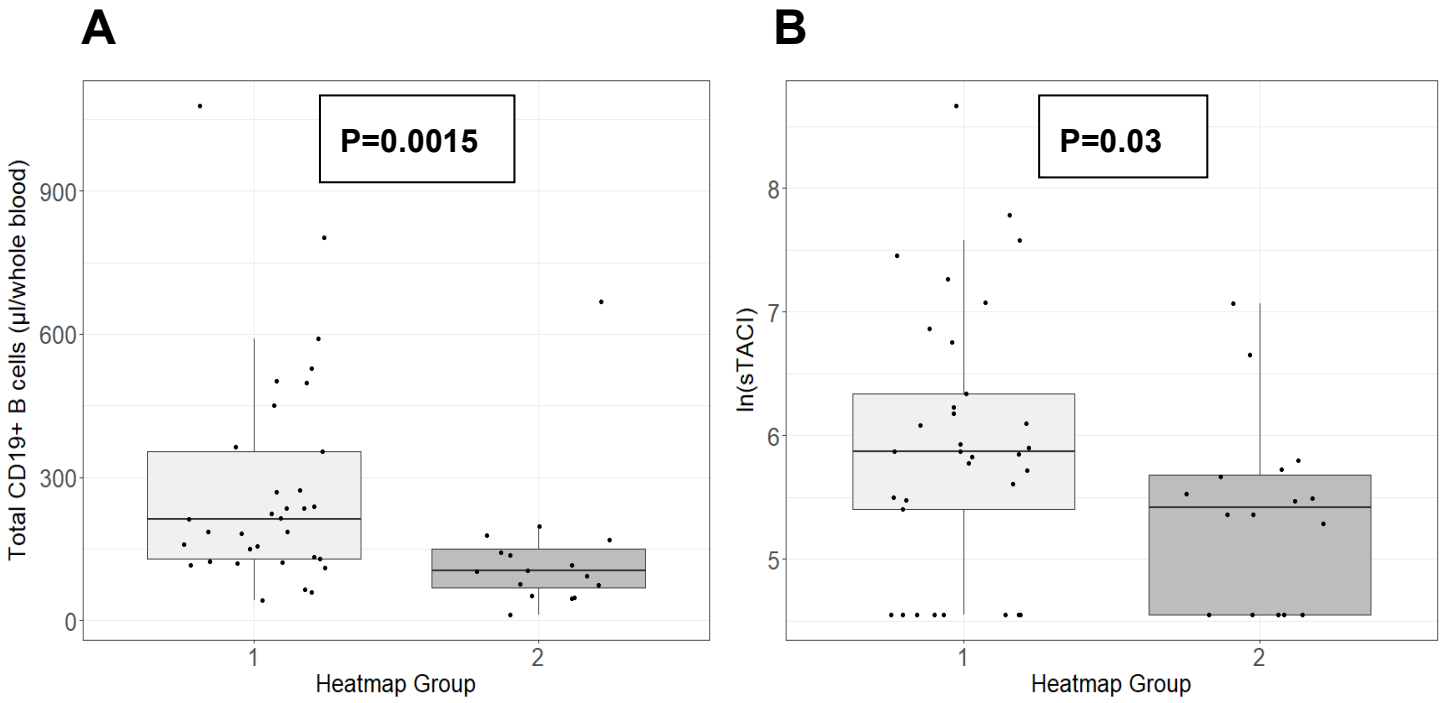


Figure 9.8 – Association between two groups of remission patients (as clustered on gene expression heatmap analysis) with total CD19⁺ B cell count (**A**) and log transformed (ln) sTACI (**B**)

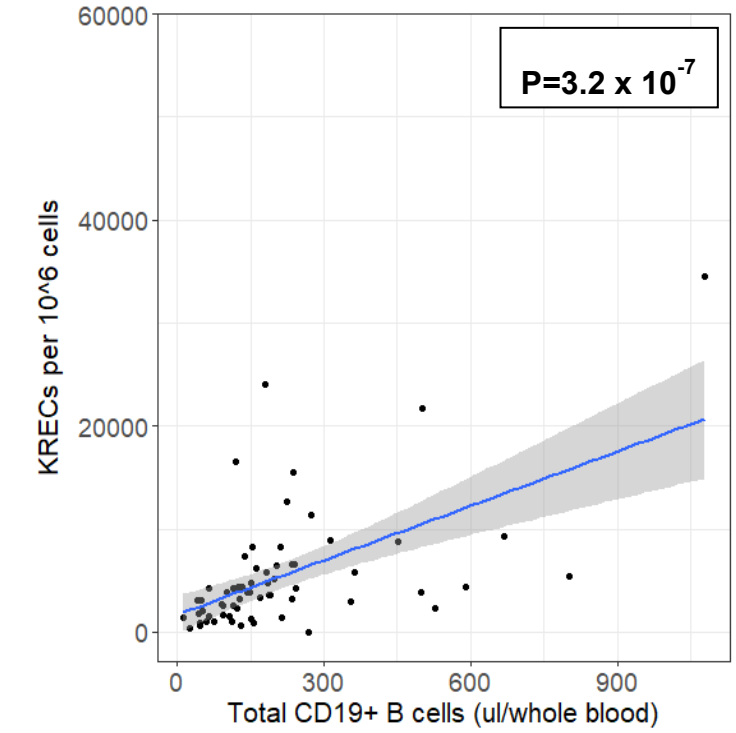
9.5 KRECs and B cell subpopulations

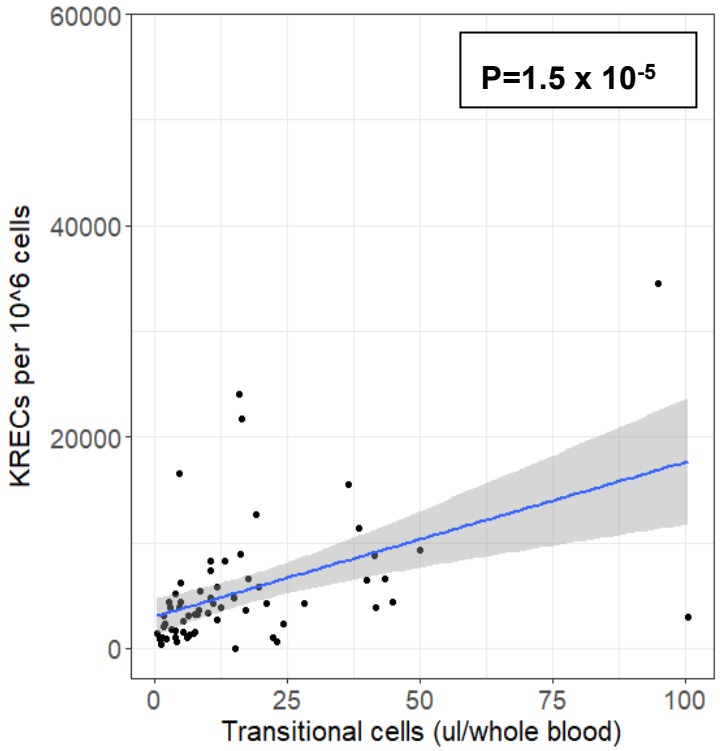
The KRECs analysis in Chapter 5: Results 2 – Kappa-deleting recombination excision circles (KRECs) highlighted the importance of B cells in GD and demonstrated a robust association between KRECs and GD activity. Further analysis was undertaken to determine the relationships between the quantity of KRECs and the peripherally circulating B cell subpopulations. KRECs were positively correlated to the total $CD19^+$ B cell count ($P=3.2 \times 10^{-7}$) and with most of the B cell subpopulations, excluding the plasmablast population (Table 9.1 and Figure 9.9).

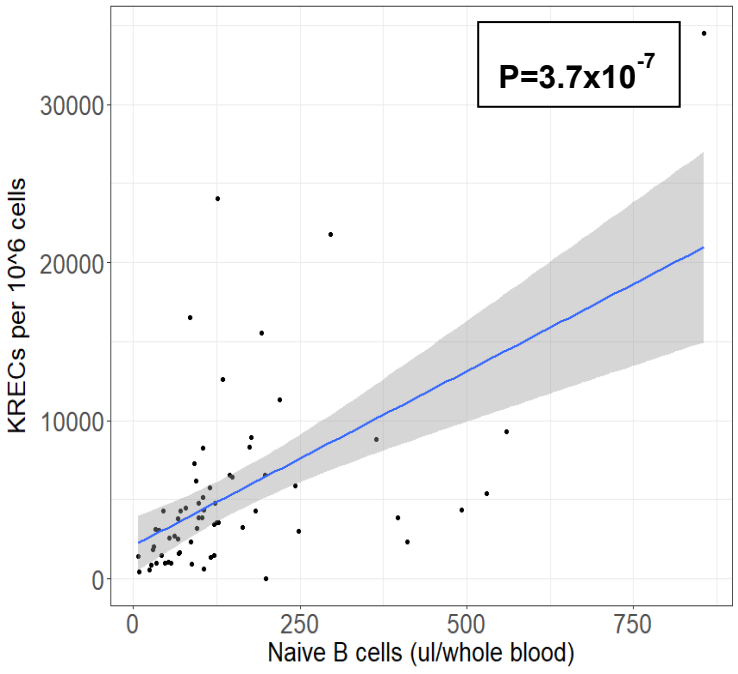
B cell subpopulation (per µl whole blood)		**KRECs per 10^6 cells**
Transitional	rs	0.52
	P	**1.5×10^{-5}***
Naïve	rs	0.60
	P	**3.7×10^{-7}***
Unswitched memory	rs	0.53
	P	**9.9×10^{-6}***
Switched memory	rs	0.28
	P	**0.03***
Plasmablast	rs	0.17
	P	0.17
Double negative	rs	0.34
	P	**0.006***
B reg cells 1	rs	0.27
	P	**0.03***
B reg cells 2	rs	0.52
	P	**1.4×10^{-5}***

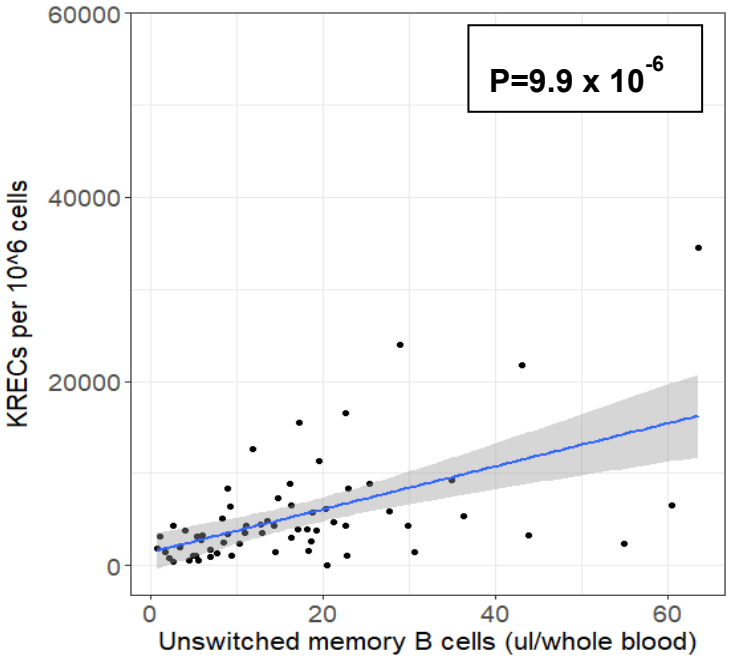
Table 9.1 – Spearman's correlation of each B cell subpopulation and KRECs quantity per 10^6 cells.

The strongest correlations were observed between KRECs and the early-stage B cell populations (transitional and naïve B cell subsets). Although there was no direct correlation observed between KRECs and plasmablasts, both the switched and unswitched memory B cell subsets were positively associated with plasmablasts (unswitched: P = 1.3 × $10^{-}5$, rs 0.52, switched: P = 1.7 × $10^{-}8$, rs 0.65). Additionally, the double negative (DN) B cells were positively associated with both KRECs (P = .006, rs 0.34) and plasmablasts (P = .018, rho rs 0.29).









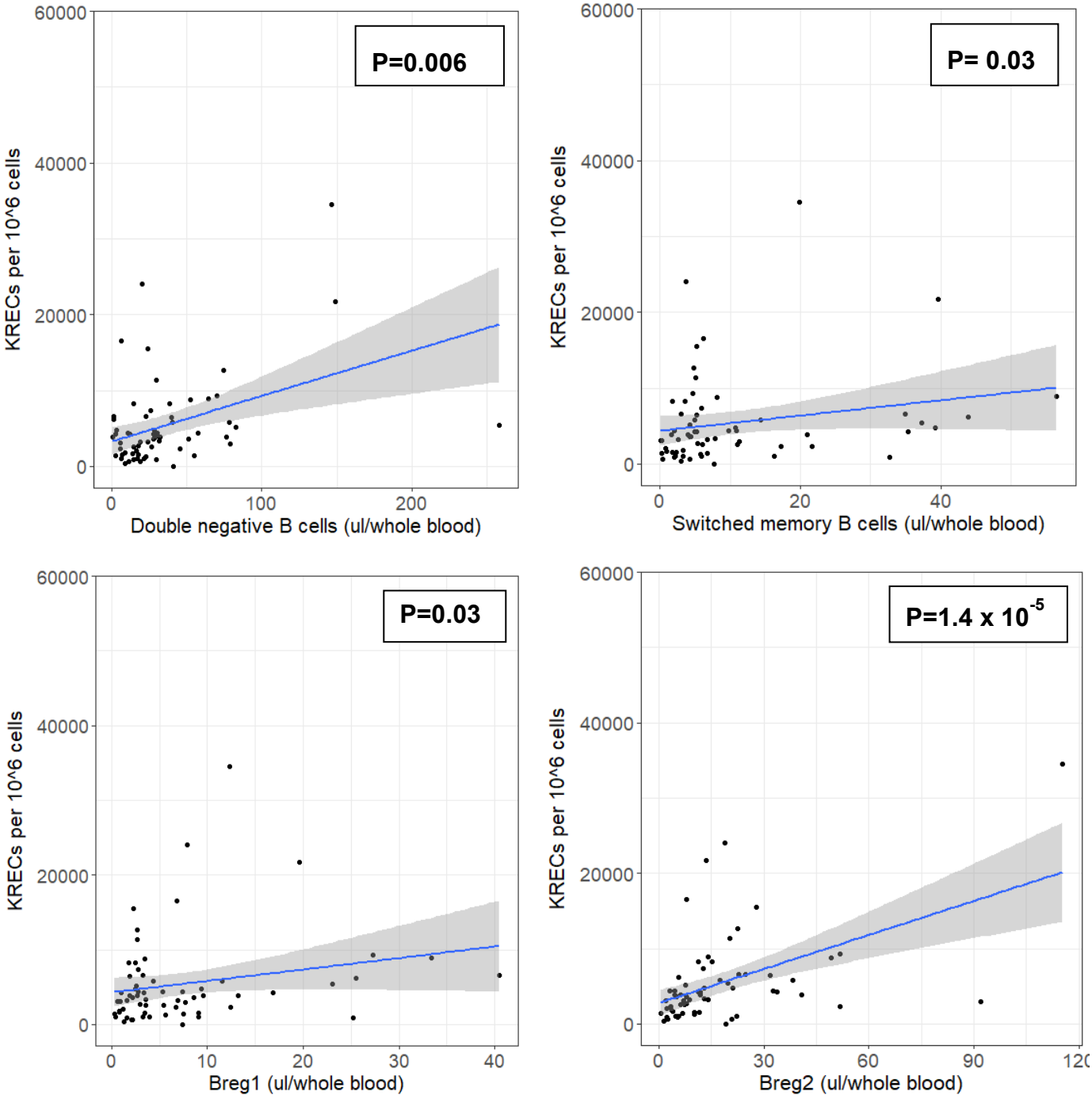


Figure 9.9 – Association between the total $CD19^+$ B cell count and the various B cell subpopulations with KRECs quantity per 10^6 cells.

There was no association between the KREC quantity and any of the cytokines measured in this study.

9.6 Composite biomarker scores to predict outcome in Graves' disease (excluding gene expression data)

After ATD has been started, the ideal timepoint to predict whether a patient is going to relapse their GD is after completing a 12-18 month course of ATD when the clinical decision is made about whether to stop treatment or continue. If a test, or combination of tests, existed that could reliably predict the likelihood of a patient relapsing after stopping ATD it would help to inform patient management and ultimately may prevent the patient relapsing.

A composite biomarker score, excluding the gene expression data, was used to analyse whether a combination of the significant variables demonstrated in this thesis could have predictive value for GD relapse. The advantage of excluding the $CD19^{+}$ B cell gene expression data in this initial analysis, which involves taking a larger volume of blood and requires costly processing, is that it makes the score more translational to clinical practice.

The variables selected to compose the biomarker score were based on their statistical significance (P=<0.05) in the previous multivariate logistic and/or cox regression analyses (Chapter 4: Results 1 – Clinical data and Chapter 6: Results 3 – Cytokine and Chemokine data). In total, six variables at baseline reached the level of significance to be selected for integrative analysis (Table 9.2). No variables were included from the flow cytometry or KRECs data as none of these were significantly associated on regression analyses.

Variable type	Baseline variable	Regression Test	Multivariate P value
Clinical	MLR	Logistic and Cox	**0.04***
	TRAb	Logistic and Cox	**0.027***
Cytokine	sTACI	Logistic and Cox	**0.007***
	BAFF	Cox	**0.025***
	sBCMA	Logistic	**0.029***
	IL-6	Cox	**0.04***

Table 9.2 – The six variables selected for integrative analysis based on their significance (P=<0.05) in the previously reported multivariate logistic and/or cox regression analysis.

9.6.1 Generating a composite biomarker score

The performance of these combined biomarkers to predict GD relapse was assessed using a composite biomarker score. Firstly, all cytokines were tested for significance by using lasso regression (as described in Chapter 3: Methods) to select the biomarkers with the greatest predictive value to simplify the model and enhance prediction accuracy. Following lasso regression, IL-6 was removed from the analysis. The coefficients generated from the lasso regression provided a balanced weight for each variable, which was then multiplied by the respective biomarker value and added to the other variables to create a composite score. An example of the formula used to generate the regression-weighted composite scores is presented in Chapter 3: Methods.

9.6.2 ROC curve analysis using biomarker composite scores

The composite scores were used to analyse different combinations of five variables (31 potential combinations) and analyse their ability to predict GD relapse. ROC curve analysis was undertaken which generated an AUC score to determine the best biomarker(s) to predict the outcome of relapse vs. remission. The R programme ‘cutpointr’ generated the optimal composite score cut-off which produced the highest sensitivity and specificity of a specific biomarker for the prediction of relapse or remission (Chapter 3: Methods).

The top 10 different combinations of variables and their respective ROC_{AUC} are presented in Table 9.3. The complete list of variable combinations and their respective ROC_{AUC} is presented in Appendix G.

ln(MLR)	ln(TRAb)	ln(sTACI)	ln(BAFF)	ln(sBCMA)	ROC_{AUC}
✓	✓	✓	✓	✓	**0.86**
×	✓	✓	✓	✓	**0.86**
✓	✓	✓	✓	×	**0.86**
×	✓	✓	✓	×	**0.85**
✓	✓	✓	×	✓	**0.84**
✓	✓	✓	×	×	**0.84**
×	✓	✓	×	×	**0.84**
×	✓	✓	×	✓	**0.82**
✓	×	✓	✓	×	**0.82**
✓	×	✓	✓	✓	**0.81**

Table 9.3 – The top 10 different combinations of biomarker composite scores ranked by their ROC_{AUC}.

The highest scoring composite biomarker scores (ROC_{AUC} = 0.86) for predicting relapse in GD included the score that incorporated all five variables and two scores that included four of the variables. Excluding MLR or sBCMA did not affect the ROC_{AUC}. The ROC curves of the highest scoring composite scores are presented in Figure 9.10, alongside the top (highest ROC_{AUC}) 3-variable and 2-variable scores for comparison. The difference observed between the ROC_{AUC} between these composite scores is minimal despite the lower number of variables included.

Notably, TRAb was the best performing single biomarker (ROC_{AUC} 0.72) (currently the gold standard for predicting relapse in GD), however there was a significantly improved ROC_{AUC} (0.86) when combined with the additional biomarkers. Furthermore, TRAb alone as a single predictive biomarker had lower sensitivity (0.63 vs. 0.88), and positive (0.5 vs. 0.57) and negative (0.87 vs. 0.95) predictive values

compared to the best performing composite score (Table 9.4). TRAb as a single biomarker did have a slightly higher specificity (0.80 vs. 0.79).

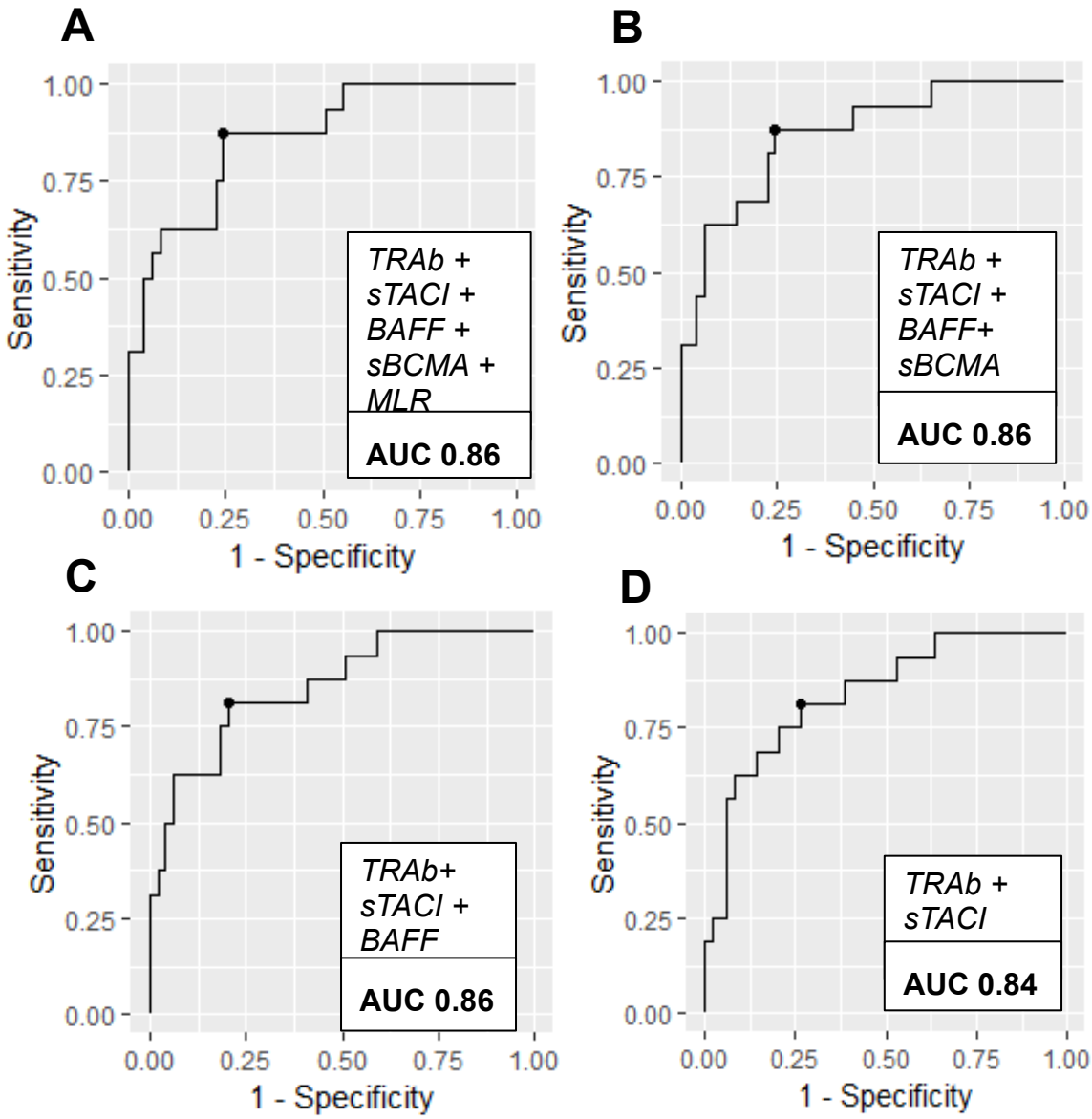


Figure 9.10 – ROC curve analysis of the top two performing biomarker composite scores

The predictive performance of the biomarker composite scores, including the sensitivity, specificity and positive and negative predictive values is detailed in Table 9.4.

Composite score	Variables included	Threshold value	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	ROC_{AUC} (95% CI)

5-variable	TRAb + sTACI + BAFF + sBCMA+ MLR	0.4912	**0.88** (0.64– 1.0)	0.75 (0.64– 1.00)	0.54 (0.40 – 0.87)	**0.95** (0.87 – 1.0)	**0.86** (0.77 – 0.94)
4-variable	TRAb + sTACI + BAFF+ sBCMA	0.4117	**0.88** (0.64– 1.0)	0.75 (0.67 – 0.97)	0.54 (0.41 – 0.88)	**0.95** (0.88 – 1.0)	**0.86** (0.77 – 0.94)
3-variable	TRAb + sTACI + BAFF	0.3925	0.81 (0.60 – 1.0)	**0.79** (0.61 – 0.98)	**0.57** (0.40 – 0.92)	0.93 (0.85 – 1.0)	**0.86** (0.75 – 0.93)
2-variable	TRAb + sTACI +	0.1374	0.81 (0.58 – 1.00)	0.73 (0.60 – 0.96)	0.5 (0.40 – 1.0)	0.92 (0.38 – 0.95)	0.84 (0.75 – 0.9)

Table 9.4 – Predictive performance of the top (highest ROC_{AUC}) performing 2-5 variable combination composite scores including the sensitivity, specificity, positive predictive (PPV) and negative predictive value (NPV).

Although the ROC_{AUC} is similar for all the variable composite scores in Table 9.4, the five- and four-variables scores had the highest sensitivity and negative predictive values, with slightly lower specificity and positive predictive values than the 3-variable score. Given the very similar predictive performance of the five- and four-variables scores there appears to be minimal benefit in including the MLR +/- sBCMA in the composite score. It is also worth noting that all the composite scores have low PPV (maximum of 0.57 in the 3-variable score).

9.6.3 Survival analysis using biomarker composite scores

The survival distributions of the high performing composite biomarker scores presented in Table 9.4 were analysed by the log-rank test. The Kaplan-Meier plots of these scores are presented in Figure 9.11. The survival distributions and log-rank P values were all significant (P = <0.0001). Being assigned to the 'positive' cohort indicated that the patient's biomarker score was above the optimal composite score cut-off threshold as determined by 'cutpointr' e.g., 0.4912 (for the 5-variable score) with any under this threshold considered 'negative'. The thresholds used are presented in Table 9.4.

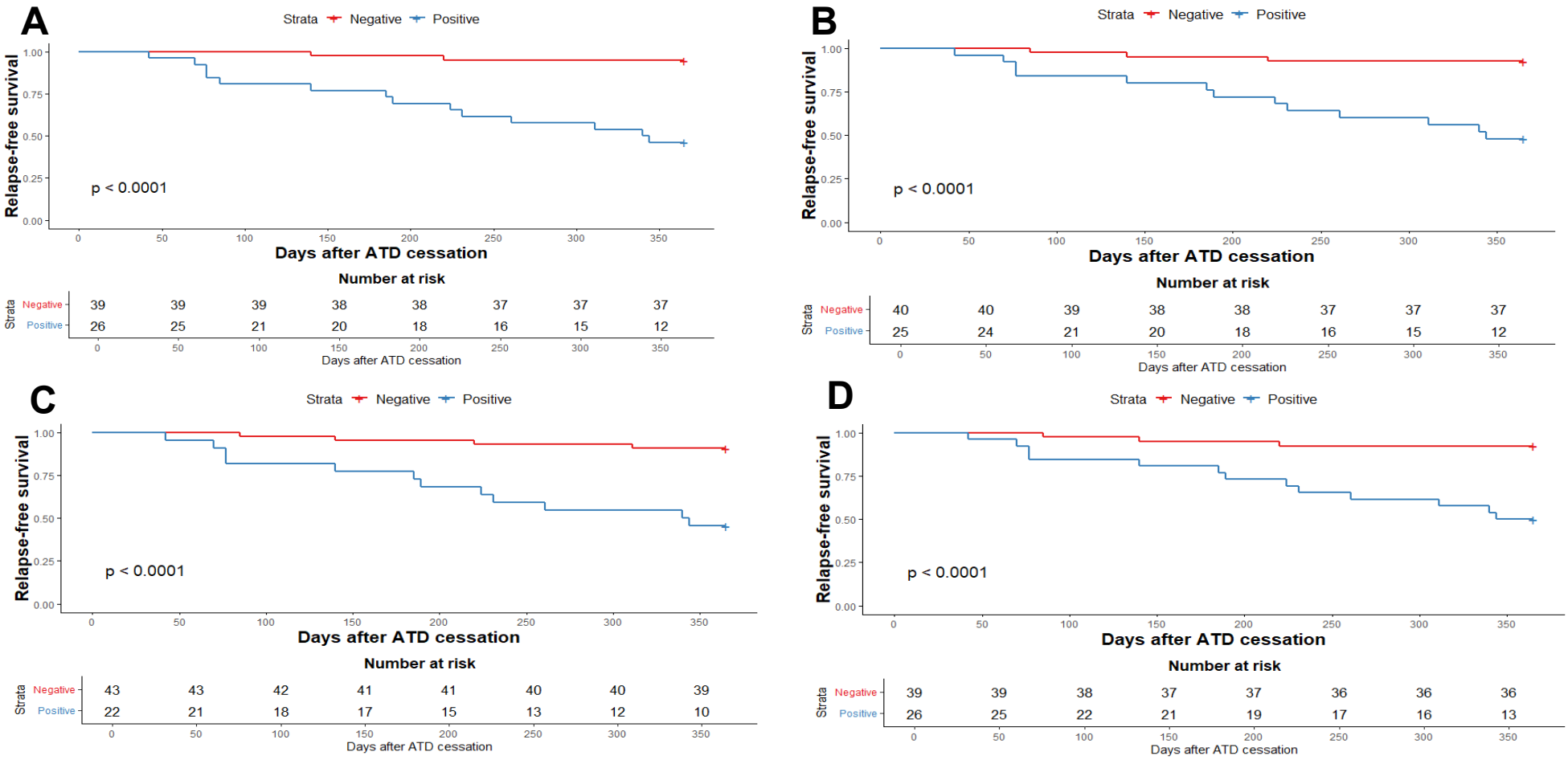


Figure 9.11 – Kaplan-Meier plot of survival following ATD cessation stratified by positive (blue) or negative (red) composite biomarker score.

9.7 Composite biomarker score to predict outcome in Graves' disease (including gene expression data)

A composite biomarker score, including the gene expression data and biomarkers, was used to analyse whether a combination of the significant variables demonstrated to be significant in this thesis could have predictive value for GD relapse. The aim of using the genetic data as well as the biomarkers was to examine whether there was any additional predictive value of including genes in the composite score.

The genes included in the initial analysis were all the 33 DEG and the additional gene that was positive on multivariate cox regression analysis (*DDX11L17).* The biomarkers included alongside the genes were those found significantly associated with outcome as discussed in the previous section (MLR, TRAb, BAFF, sTACI, sBCMA, IL-6) (Section 9.6). Therefore, in total 40 variables at baseline were selected for the initial analysis to generate a predictive model.

9.7.1 Generating a composite biomarker score

The performance of these combined variables to predict GD relapse was assessed using a composite biomarker score. Firstly, all the biomarkers and genes were tested for significance by using lasso regression (as described in Chapter 3: Methods). Following lasso regression, 32 variables were removed from the analysis and 8 remained in the model. A relaxed lasso regression was then performed, and the coefficients generated from this provided a balanced weight for each variable, which was then multiplied by the respective biomarker/gene value and added to the other variables to create a composite score. An example of the formula used to generate the regression-weighted composite scores is presented in Chapter 3: Methods.

9.7.2 ROC curve analysis using biomarker composite scores

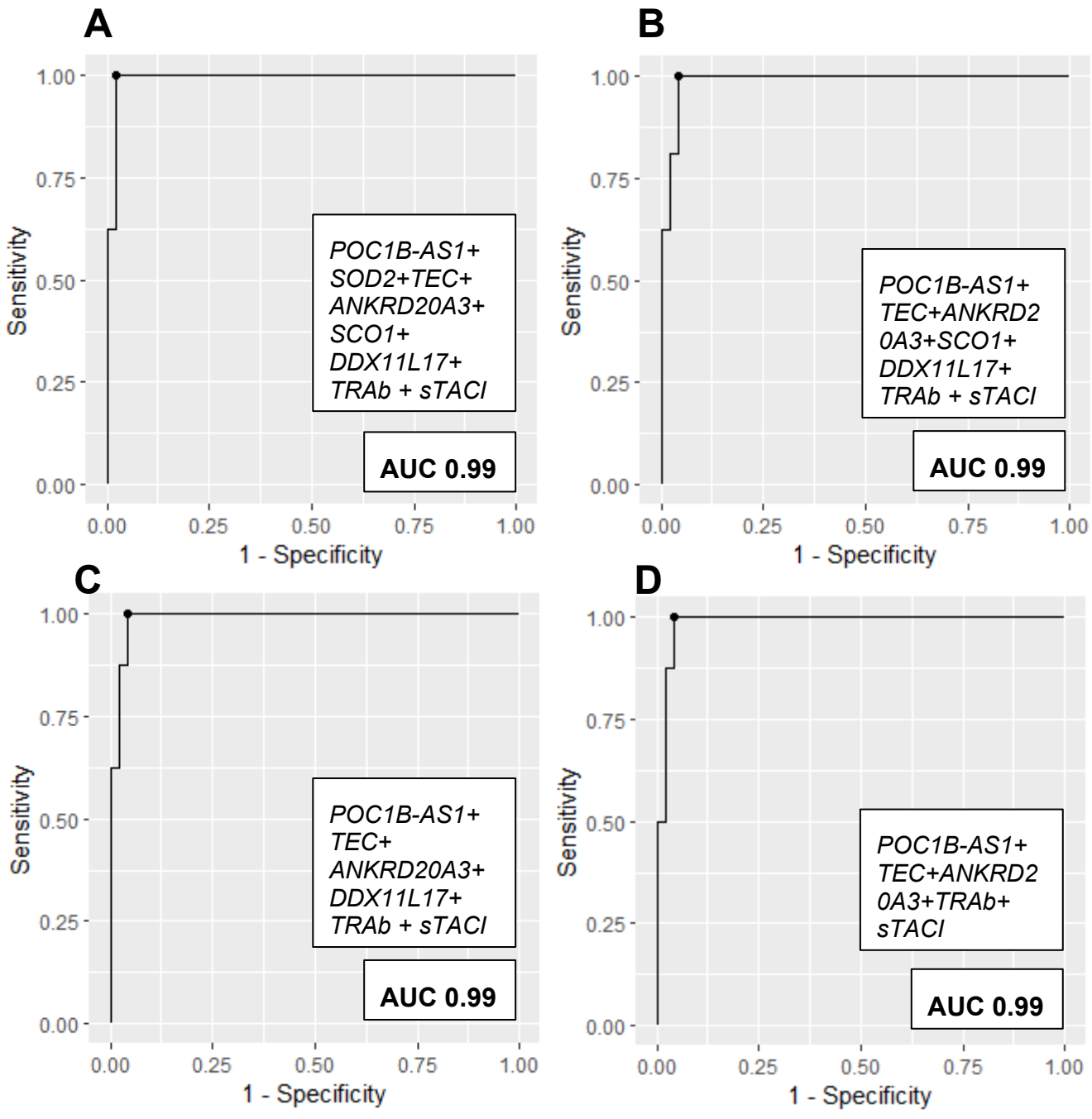
The composite scores were used to analyse different combinations of the eight remaining variables in the predictive model (255 potential combinations) and analyse their ability to predict GD relapse. ROC curve analysis was undertaken which generated an AUC score to determine the best combination of biomarker(s) and genes to predict the outcome of relapse vs. remission. The R programme 'cutpointr' was used as discussed previously (Chapter 3: Methods).

The composite scores that included combinations of 1-8 variables ranked by their respective ROC_{AUC} are presented in Table 9.5. The complete list of variable combinations and their respective ROC_{AUC} is presented in Appendix H.

POC1B-AS1	*SOD2*	*TEC*	*ANKRD20A3*	*SCO1*	*DDX11L17*	TRAb	sTACI	ROC_{AUC}
✓	✓	✓	✓	✓	✓	✓	✓	**0.99**
✓	×	✓	✓	✓	✓	✓	✓	**0.99**
✓	×	✓	✓	×	✓	✓	✓	**0.99**
✓	×	✓	✓	×	×	✓	✓	**0.99**
×	×	✓	✓	×	×	✓	✓	**0.96**
✓	×	✓	✓	×	×	×	×	**0.92**
✓	×	×	✓	×	×	×	×	**0.88**
✓	×	×	×	×	×	×	×	**0.78**

Table 9.5 – The top combinations of gene/biomarker composite scores including 1-8 variables ranked by their ROC_{AUC}.

There were 7 different combinations of gene/biomarkers that scored with a ROC_{AUC} = 0.99 for predicting relapse in GD including the combination that incorporated all eight variables. However, a ROC_{AUC} = 0.99 was also achieved when only 5 variables were included (*SOD2*, *SCO1*, *DDX11L17* removed). Therefore, excluding these variables did not affect the ROC_{AUC}. The ROC curves of these eight highest scoring combinations of genes and biomarkers that included different numbers of variables are presented in Figure 9.12. There is minimal difference in the ROC_{AUC} between the eight (0.99) and 4 (0.96) variable combinations.



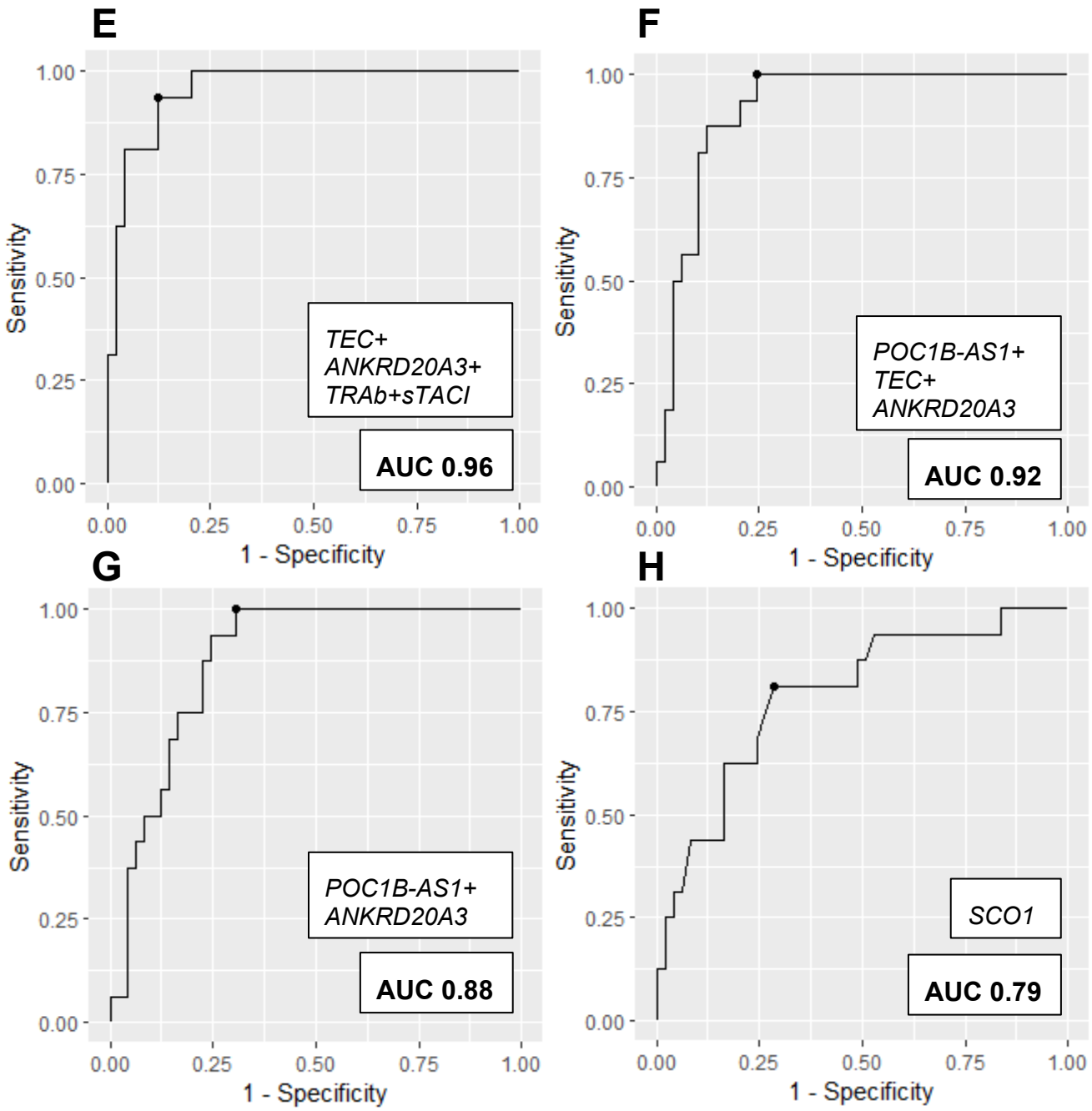


Figure 9.12 – ROC curve analysis of the top performing 1-8 variable gene/biomarker composite scores

The predictive performance of the gene/biomarker composite scores, including the sensitivity, specificity and positive and negative predictive values is detailed in Table 9.6.

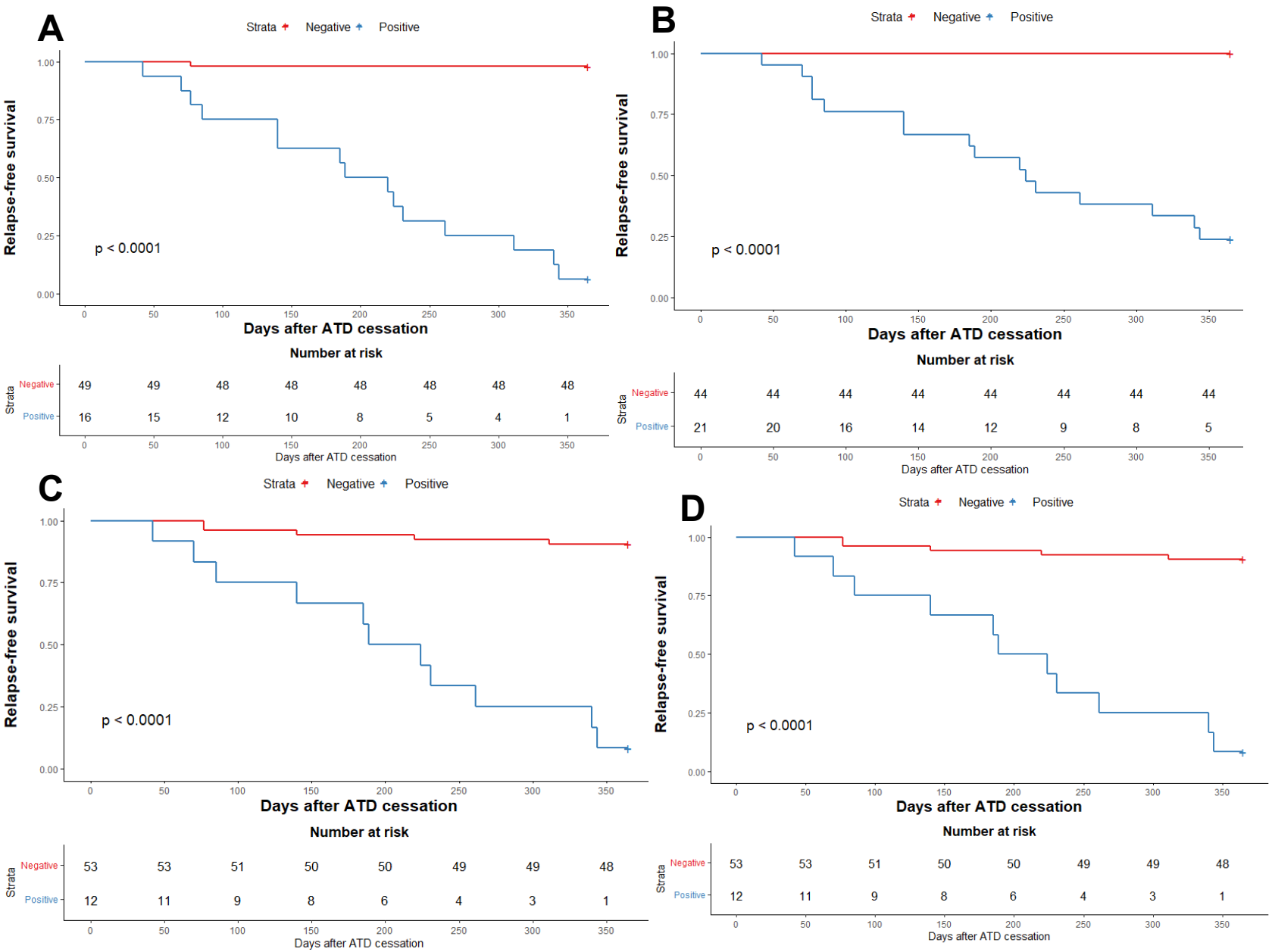
Composite score	Variables included	Threshold value	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	ROC_{AUC} (95% CI)
8-variable	*POC1B-AS1+ SOD2+TEC+ ANKRD20A3+ SCO1+ DDX11L17+ TRAb + sTACI*	-0.4649	**1.0** (0.60– 1.0)	**0.98** (0.60– 1.0)	**0.94** (0.75 – 1.0)	**1.0** (0.88 – 1.0)	**0.99** (0.98 – 1.0)
7-variable	*POC1B-AS1+ TEC+ ANKRD20A3+ SCO1+ DDX11L17+ TRAb + sTACI*	-0.5141	**1.0** (0.57– 1.0)	0.96 (0.89 – 1.0)	0.89 (0.67 – 1.0)	**1.0** (0.87 – 1.0)	**0.99** (0.96 – 0.94)
6-variable	*POC1B-AS1+ TEC+ ANKRD20A3+ DDX11L17+ TRAb + sTACI*	-0.3493	**1.0** (0.60 – 1.0)	0.96 (0.89 – 1.0)	**0.89** (0.67 – 1.0)	**1.0** (0.87 – 1.0)	**0.99** (0.96 – 1.0)
5-variable	*POC1B-AS1+ TEC+ ANKRD20A3+ TRAb + sTACI*	-0.3493	**1.0** (0.60 – 1.0)	0.96 (0.89 – 1.0)	0.89 (0.67 – 1.0)	**1.0** (0.87 – 1.0)	**0.99** (0.96 – 1.0)
4-variable	*TEC+ ANKRD20A3+ TRAb + sTACI*	-0.0881	0.94 (0.50 – 1.0)	0.88 (0.70 – 1.0)	0.71 (0.42 – 1.0)	0.98 (0.83 – 1.0)	0.96 (0.90 – 1.0)
3-variable	*POC1B-AS1+ TEC+ ANKRD20A3+*	-0.575	**1.0** (0.50 – 1.0)	0.76 (0.64 – 0.95)	0.57 (0.33 – 0.86)	**1.0** (0.82 – 1.0)	0.92 (0.83 – 0.99)
2-variable	*POC1B-AS1+ ANKRD20A3*	-0.3961	**1.0** (0.43 – 1.0)	0.69 (0.56 – 0.89)	0.52 (0.27 – 0.73)	**1.0** (0.80 – 1.0)	0.88 (0.78 – 0.97)
1-variable	*SCO1*	-0.17	0.81 (0.25 – 1.0)	0.71 (0.44 – 0.92)	0.48 (0.20 – 0.70)	0.92 (0.74 – 1.0)	0.79 (0.63 – 0.94)

Table 9.6 – Predictive performance of the top (highest ROC_{AUC}) performing one-eight variable combination composite scores, including the sensitivity, specificity, positive predictive (PPV) and negative predictive value (NPV).

The combination that included all 8 variables had the highest ROC_{AUC} and the highest level of sensitivity, specificity, PPV and NPV. However, the five variables score that had the same high ROC_{AUC} demonstrated similar predictive performance with only slightly reduced specificity (0.96 vs. 0.98) and PPV (0.89 vs. 0.94).

9.7.3 Survival analysis using biomarker composite scores

The survival distributions of the eight different gene/biomarker combinations presented in Table 9.6 were analysed by the log-rank test. The Kaplan-Meier plots of these scores are presented in Figure 9.13. The survival distributions and log-rank P values were all significant (P = <0.0001). Being assigned to the ‘positive’ cohort indicated that the patient’s biomarker score was above the optimal composite score cut-off threshold as determined by ‘cutpointr’ e.g., -0.3493 (for the 5-variable score) with any under this threshold considered ‘negative’. The thresholds used are presented in Table 9.6.


A
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 49 49 48 48 48 48 48 48
Positive 16 15 12 10 8 5 4 1
0 50 100 150 200 250 300 350
Days after ATD cessation
B
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 44 44 44 44 44 44 44 44
Positive 21 20 16 14 12 9 8 5
0 50 100 150 200 250 300 350
Days after ATD cessation
C
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 53 53 51 50 50 49 49 48
Positive 12 11 9 8 6 4 3 1
0 50 100 150 200 250 300 350
Days after ATD cessation
D
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 53 53 51 50 50 49 49 48
Positive 12 11 9 8 6 4 3 1
0 50 100 150 200 250 300 350
Days after ATD cessation

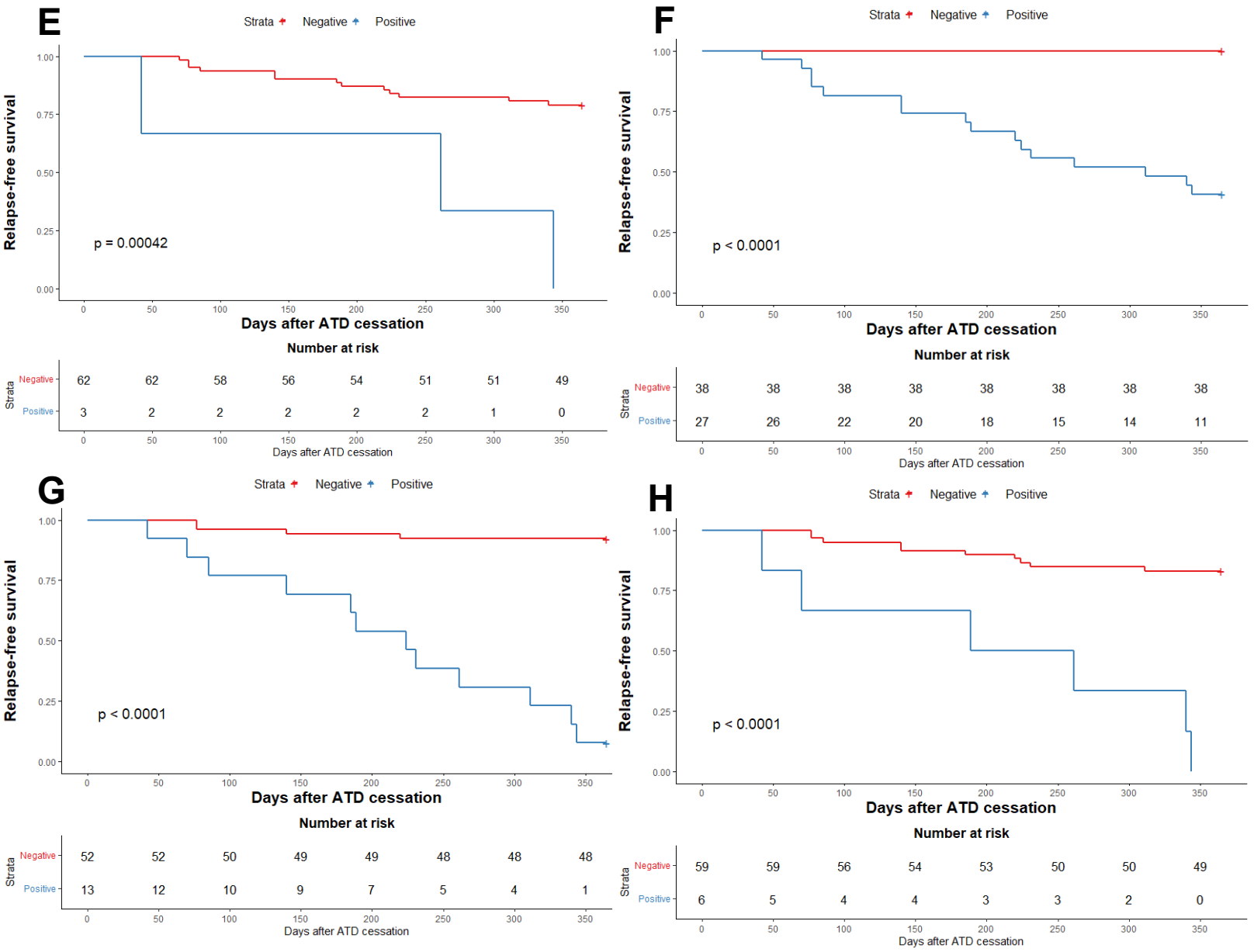

E
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p = 0.00042
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 62 62 58 56 54 51 51 49
Positive 3 2 2 2 2 2 1 0
0 50 100 150 200 250 300 350
Days after ATD cessation
F
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 38 38 38 38 38 38 38 38
Positive 27 26 22 20 18 15 14 11
0 50 100 150 200 250 300 350
Days after ATD cessation
G
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 52 52 50 49 49 48 48 48
Positive 13 12 10 9 7 5 4 1
0 50 100 150 200 250 300 350
Days after ATD cessation
H
Strata Negative Positive
Relapse-free survival
1.00
0.75
0.50
0.25
0.00
p < 0.0001
0 50 100 150 200 250 300 350
Days after ATD cessation
Number at risk
Strata
Negative 59 59 56 54 53 50 50 49
Positive 6 5 4 4 3 3 2 0
0 50 100 150 200 250 300 350
Days after ATD cessation

Figure 9.13 – Kaplan-Meier plot of survival following ATD cessation stratified by positive (blue) or negative (red) composite biomarker score.

9.8 Discussion

9.8.1 Association of IL-6 and CRP with monocytes

As previously discussed in Chapter 4: Results 1 – Clinical data, circulating monocytes have been observed to be elevated in GD, associated with disease activity, and positively correlated to TRAb titres (*Turan 2019, Chen et al. 2021*). Monocytes have a role in synthesising and inducing IL-6 in response to inflammatory mediators, including IL-1β (*Tosato G et al. 1990*). The positive association demonstrated in my study between the monocyte count and IL-6 concentration suggests that circulating monocytes could potentially be contributing to IL-6 production. Indeed, previous studies in RA have found a similar correlation and have also demonstrated that CD14+ monocytes in RA are metabolically re-programmed to promote a pro-inflammatory response (*McGarry T et al. 2021, Tsukamoto M et al. 2017*).

CRP has been demonstrated to induce pro-inflammatory cytokines, including IL-6 and IL-1β, from monocytes in a dose-dependent manner (*Ballou SP et al. 1992*). This is consistent with the finding in my dataset of a positive association between CRP, IL-6, and monocytes, suggesting that CRP may be inducing IL-6 production by monocytes in the GD patients. Neither CRP nor IL-6 was independently associated with outcome of GD, however IL-6 was associated with time-to-relapse on multivariate analysis suggesting a potential role in the relapse process. Only a few patients demonstrated CRP positivity and therefore further studies may need to include high sensitivity CRP to detect more subtle changes for analysis.

Both CRP and IL-6 were also positively associated with the MLR which was independently associated with relapse and time-to-relapse on multivariate analysis (Chapter 4: Results 1 – Clinical data). MLR is described as a marker of systemic inflammation, and therefore the concurrent finding of elevated inflammatory cytokines would be expected and suggests the presence of an inflammatory immune environment in relapsing patients. These findings could provide mechanistic insight into the interplay of cytokines and monocytes in the relapse process.

9.8.2 Cytokines and B cell subpopulations

There is a complex interplay between cytokines and B cell subpopulations which aims to maintain or restore immune homeostasis. In my dataset, BAFF was positively correlated with the transitional B cell subpopulation, consistent with previous studies in SLE and MS (*Landolt-Marticorena C et al. 2011, Smets I et al. 2021*). Transitional cells comprise a higher proportion of autoreactive B cells, with negative selection providing a peripheral checkpoint promoting the apoptosis of autoreactive transitional B cells (*Carsetti R et al. 1995*). Both transitional cells and BAFF have been found to be elevated in autoimmune disease, including GD (*Van der Weerd K et al. 2013, Wang X et al. 2021*). The dominant BAFF receptor, BAFF-R, is expressed at high levels on transitional cells, and BAFF has been demonstrated to mediate the survival of peripheral transitional B cells, many of which are autoreactive (*Rowland SL et al. 2010, Meyer-Bahlburg A et al. 2008, Batten M et al. 2000*). Specifically in GD there is also reported to be an increased expression of BAFF-R on peripheral B cells suggesting that enhanced BAFF-R receptor signalling activity could be promoting BAFF-induced B cell survival, including that of autoreactive transitional cells where BAFF-R is the predominant BAFF receptor (*Wang X et al. 2021*). Indeed, anergy of autoreactive B cells was restored in SLE patients treated with the anti-BAFF antibody, Belimumab, suggesting that BAFF could be promoting autoreactive B cell survival through blocking anergy induction of autoreactive B cells (*Malkiel et al. 2016*).

In my dataset, BAFF was associated with timing of GD relapse following ATD withdrawal using a multivariate Cox regression model and there was a larger proportion of transitional cells present in the relapse group (8.8% vs. 5.4%). Increased BAFF-R expression on transitional cells promoting the BAFF-induced survival of autoreactive B cells may explain the positive correlation observed between BAFF and transitional cells and suggest a potential pathway for the re-emergence of TRAbs in relapsing GD.

The transitional cell population is enriched with B regulatory cells, capable of producing immunosuppressive cytokines such as IL-10. Changes in the immune microenvironment are proposed to induce B regulatory cells through activation of B cells via receptors such as BAFF-R (*Mauri C et al. 2012*). In a murine model, BAFF has been demonstrated to induce IL-10-producing B cells specifically with a CD1d(hi)CD5(+) phenotype (*Yang M et al. 2010*). This is consistent with the finding

in my dataset of a positive correlation between BAFF and Breg2 ($CD19^+$ $CD5^+CD1d^+$) cells. However, conversely BAFF was found to be negatively correlated to the Breg1 phenotype ($CD19^+$ $CD27^+$ $CD24^{++}$). The role of B regulatory cells in GD remains controversial with studies observing both elevated and reduced circulating B regulatory cells with/without functional impairment, using a variety of phenotypes (*Grubczak K et al. 2021, Rosser EC et al. 2015, Kristensen B et al. 2015, Zha B et al 2012*). Functionally defective Bregs have also been suggested as responsible for the pathogenesis of GO (*Chen G et al. 2019*). This therefore highlights the need for further work to accurately immunophenotype B regulatory cells to investigate their role and interaction with cytokines in GD.

sTACI is cleaved and shed from the membrane of activated B cells and therefore functions as a soluble biomarker of B cell activation (*Meinl E et al. 2021).* Although sTACI wasn't associated with any of the B cell subpopulations in my dataset, it was weakly positively correlated with the total CD19+ B cell population and was independently associated with relapse and time-to-relapse of GD. sTACI has been proposed as an immunoregulator, functioning as a decoy to block BAFF-APRIL-mediated B cell survival to regulate the expansion of activated B cells (*Hoffman et al. 2015*) and TACI receptor deficient mice have elevated B cells (*Yan M et al. 2001*). This suggests that the B cell population should fall with increasing sTACI attenuating B cell survival via BAFF and APRIL. However, switched memory B cells are relatively independent of BAFF or APRIL for survival (*Smulski CR et al. 2018, Benson MJ et al. 2008*). This is demonstrated by the administration of soluble TACI-Ig (Atacicept), a fusion protein which blocks BAFF and APRIL, that has been demonstrated to target mature B cells and antibody-producing plasma cells but spare memory cells in RA (*Bracewell C et al. 2008*). Therefore, in an activated B cell state sTACI may affect selected B cell populations but the switched memory B cells may continue to survive and differentiate into plasma cells, increasing the B cell population.

9.8.3 Cytokine heatmap integrated analysis

The heatmap of cytokine expression at the time of ATD withdrawal demonstrated distinct clusters of relapsing patients, as presented in Chapter 6: Results 3 – Cytokine and Chemokine data, (Group 1 and Group 2), the former of which was noted to visually have greater expression of sTACI and sBCMA. These two groups

were explored in detail with other humoral immune markers and clinical/demographic features to further characterise and explore potential relapsing subtypes.

The Group 1 relapsing cohort were confirmed to have significantly higher levels of sTACI and sBCMA compared to Group 2. Group 1 was also observed to have a significantly increased proportion of naïve B cells, with a lower switched memory B cell and plasmablast population. Neither sTACI nor sBCMA were found to be directly correlated with any of these B cell subpopulations. As well as indirectly affecting the survival of B cell populations through their interaction with BAFF, elevated sTACI and sBCMA reflect a state of B cell activation, and therefore there may be fewer circulating memory cells and plasmablasts as they are differentiating into unmeasured antibody-secreting plasma cells.

Three patients on the heatmap were noted to have relatively higher levels of IL-6, however only one of these patients relapsed at one year. Out of these 3 patients, 2 also had an elevated CRP (20,26 mg/L) which is consistent with the positive correlation demonstrated earlier in this Chapter between IL-6 and CRP and may indicate an acute phase inflammatory response in these patients.

9.9 Gene expression heatmap integrated analysis

The heatmap of gene expression at the time of ATD withdrawal presented in Chapter 8: Results 5 – CD19+ B cell RNAseq demonstrated one cluster containing remission patients only (Group 1) and a second cluster containing all the relapse patients and 16 of the remission patients (Group 2). These two groups were explored in detail in conjunction with other measured clinical and humoral immune markers as previously described to examine whether there were differences in the 16 remission patients clustered with the relapse patients.

The remission patients in ‘Group 2’ had lower circulating $CD19^{+}$ B cells and sTACI which was unexpected considering they were grouped with the rest of the relapse patients for whom sTACI was an independent biomarker of relapse. However, this group of remission patients had also received a longer duration of ATD compared to the remission patients in ‘Group 1’ which suggests they had experienced a longer disease course and had a greater exposure to ATD which may have affected the

quantity of circulating B cells and their level of activation considering the changes in immunophenotypes demonstrated after treatment with ATD (*Torimoto K et al. 2022*).

9.9.1 KRECs and B cell subpopulations

Data presented in Chapter 5: Results 2 – Kappa-deleting recombination excision circles demonstrated that KRECs have a potential role as a marker of activity and outcome in GD. Further analysis demonstrated a strong association between KRECs and the early-stage B cell populations that are present shortly after emigration from the bone marrow (transitional and naïve B cell subsets).

Although there was no direct correlation observed between KRECs and plasmablasts, both the switched and unswitched memory B cell subsets were positively associated with plasmablasts. Additionally, the DN B cells were positively associated with both KRECs and plasmablasts. This indicates that although the plasmablasts are not directly correlated to KRECs, the positive association of memory and DN B cells with KRECs, may be consistent with their roles as precursors of antibody-secreting plasma cells.

Previous evidence has suggested that patients with GD, particularly in the untreated hyperthyroid phase, have an expanded immature B cell population, with increased numbers of $CD5^+$ transitional cells and naïve B cells (*Iwatani Y et al. 1989, Torimoto K et al. 2022*). Within this immature B cell population, even after passing the central tolerance checkpoint in the bone marrow, around 50% of transitional cells continue to demonstrate autoreactivity which only reduces once they differentiate into mature B cells (*Wardemann H et al. 2003, Giltiay NV et al. 2012*). This circulating pool of autoreactive B cells was studied in SLE, where analysis of antinuclear antibodies (ANA) demonstrated that they were mostly of the naïve B cell subset with the potential to become activated due to a failure of anergy induction (*Malkiel S et al. 2016*). This was observed to have a direct effect on disease activity in SLE, as increased frequencies of anergic ANA^+ B cells was associated with decreased disease activity (*Malkiel S et al. 2016*). Additional studies have demonstrated the association of autoimmune conditions, such as JDM and systemic sclerosis, with an expanded immature transitional B cell population (*Piper CJM et al. 2018, Liu M et al. 2019, Taher TE et al. 2018*). The increased KRECs and their strongly positive association with early-stage B cells observed in this study of GD patients, may reflect

this expanded autoreactive immature B cell population that ultimately gives rise to TRAb-producing B lymphocytes.

The differentiation of B lymphocytes into IgG^+ plasma and memory cells in autoimmunity is classically described to occur through the germinal centre pathway (*Suurmond J et al. 2019*). However, the activation and differentiation of naïve B lymphocytes into plasma cells has also been demonstrated to occur directly through the extrafollicular pathway in SLE (*Suurmond J et al. 2019*). The germinal centre response results in memory cell formation and is also reported to produce higher affinity plasma cells (*Jenks SA et al. 2019*), however it is slower to develop and therefore may be less important during an acute relapse of GD. Indeed, a large proportion of the circulating plasma cells during an active flare of SLE have been observed to be clonally related to naïve B cells rather than IgG^+ memory B cells, suggesting they originate from an extrafollicular, rather than germinal centre, response (*Jenks SA et al. 2019, Tipton CM et al. 2015*). Therefore, the extrafollicular pathway may activate the expanded immature naïve B lymphocytes in GD into TRAb-producing plasma cells and play a critical role in the humoral immune response during active and acutely relapsing GD. Thus, the association of KRECs with early-stage B lymphocytes suggests that they could represent a biomarker for disease activity in GD, with increased levels associated with active or relapsing GD.

DN B cells, characterised by the absence of both naïve (IgD) and conventional memory markers (CD27), have been observed to be an expanded cell population in SLE, MS and RA (*You X et al. 2020, Moura RA et al. 2017, Fraussen J et al 2019*). Furthermore, DN B cells have been demonstrated to generate antibody-secreting plasmablasts in response to antigenic stimulation (*Ruschil C et al. 2020*). In particular, the 'DN2' subtype (phenotype IgD^-CD27^- $CD38^-$ $CD24^-$) is described as having a specific role in the extrafollicular pathway, as an extrafollicular antibody-secreting cell precursor (*Stewart A et al. 2021, Sanz I et al. 2019*). B cells expressing the surface marker CD11c, which is found highly expressed on DN2 cell subsets, have been demonstrated to be of increased frequency and correlate with serum TRAb levels in GD (*Cao Y et al. 2022*). The positive association of DN B cells with KRECs and plasmablasts observed in this study could be consistent with the differentiation of the naïve B cell subset into DN B cells via the extrafollicular pathway, which further differentiate into TRAb-secreting plasma cells.

Breg cells are immunosuppressive cells that help maintain immunological tolerance (*Rosser EC et al. 2015*). In this study, Breg cells as defined by $CD24^{++}CD27^{+}$ expression (otherwise known as 'B10' cells (*Kalampokis I et al. 2013*), were positively associated with KRECs. There is reported to be a higher frequency and increased suppressive ability of these 'B10' Breg cells in autoimmune disease in response to an inflammatory environment (*Rosser EC et al. 2014, Iwata Y et al. 2011*). Furthermore, FT3 levels have been demonstrated to be positively correlated with B cells that express interleukin 10 (IL-10) in GD (*Kristensen B et al. 2015*). However, various phenotypes of Breg cells are described, highlighting their unclear lineage of origin, and suggesting that it is environmental stimuli that drives the differentiation of B cells into a regulatory phenotype. Indeed, stimulation with pro-inflammatory cytokines such as IL-6 and IL-1β, is demonstrated to directly promote the differentiation of IL-10 producing B cells (*Rosser EC et al. 2014*). Therefore, the positive association of Bregs and KRECs observed in this study could reflect the presence of an inflammatory autoimmune environment, inducing the differentiation of newly produced B cells into a regulatory phenotype.

These findings demonstrate the relationship between KRECs and B cell subsets and provide mechanistic insights into the possible immunopathology driving GD relapse.

9.9.2 Predictive composite scores

There is currently no robust method of accurately predicting outcome for most people with GD. The purpose of developing a composite biomarker +/- gene score was to translate the results generated by this thesis into a method of predicting outcome in GD at the timepoint of stopping ATD.

Using a score that incorporates a low number of biomarkers and excludes the genes would be advantageous in terms of cost and laboratory processing time. The highest performing combination biomarker score excluding the genes was the 4-variable score that included TRAb, sTACI, BAFF and sBCMA. This combination of biomarkers had a good sensitivity (0.88) and high ROC_{AUC} (0.86) for predicting relapse but like the other top scoring biomarker combinations had a poor PPV (0.57). Therefore, using this test in isolation would result in only around half of those testing 'positive' proceeding to relapse their GD at one year. It is also notable that the 2-variable score that includes only TRAb and sTACI had a very similar predictive value (ROC_{AUC} =0.84) with minimal compromise of sensitivity, specificity, PPV and NPV compared to

the 4-variable score. The involvement of these two specific biomarkers in a predictive model of relapse is consistent with the earlier findings in my thesis that an elevated TRAb and sTACI at the time of stopping ATD were the most significant biomarkers for predicting relapse and time-to-relapse occurring.

An elevated TRAb level at the end of ATD treatment is consistently demonstrated to be a sensitive biomarker of GD outcome and is currently the 'gold standard' for predicting relapse (*Kahaly GJ et al. 2018*). However, a negative TRAb on ATD withdrawal does not necessarily accurately predict long-term outcome (*Carella C et al. 2006*). My findings demonstrate that although TRAb was the best performing single biomarker (ROC_{AUC} 0.72), when incorporated with sTACI the predictive performance of the score was enhanced (ROC_{AUC} 0.84), with improved sensitivity, specificity, PPV and NPV. sTACI can be measured using a small volume of serum and is therefore a good potential candidate biomarker to be used in addition to TRAb.

Further analysis of my data to produce a composite score that included genes as well as immunological biomarkers was performed with the aim of improving the predictive value of the model and reducing the risk of false positives. The only validated predictive scoring tool available for GD that incorporates genetic data is the GREAT+ score that incorporates *PTPN22* C/T polymorphism and *HLA* subtypes DQB1*02, DQA1*05 and DRB1*03 (*Vos et al. 2016*). However, these genotypes were not found to be useful for predicting outcome in those with a low or high risk of disease reoccurrence (*Vos et al. 2016*). Furthermore, when the GREAT score was applied to my dataset (without the genetic data) this did not accurately predict outcome in this cohort (Chapter 4: Results 1 – Clinical data).

The initial analysis performed to construct a combined biomarker and genetic predictive model included all the 33 DEG, the one gene found positive on multivariate cox regression and all the 6 significant biomarkers discussed earlier in this Chapter to ensure that all relevant variables were included. The 8 variables which remained after independent variable selection by lasso regression included 6 genes (*POC1B-AS1, SOD2, TEC, ANKRD20A3, SCO1, DDX11L17*) alongside TRAb and sTACI. The highest performing combined gene/biomarker score was the 8-variable score (ROC_{AUC} 0.99), however if three of the variables were removed from the model it continued to demonstrate the same high ROC_{AUC} with similar predictive performance. In addition to TRAb and sTACI, the genes in this 5-variable score were *POC1B-AS1*,

TEC and *ANKRD20A3*. These genes were all found to be downregulated in the GD relapse cohort.

Variants in these genes are not known to have any direct association with GD, although the positive association of downregulated *POC1B-AS1* with circulating Tregs may indicate a reduced circulating population of Tregs in those that relapse their GD. These genes require further investigation to determine possible functional or mechanistic association with GD relapse. Furthermore, this predictive score would need to be tested in another larger population of GD patients to validate the findings.

The composite score in its current form including the genetic data may be challenging to implement in NHS laboratories and from a clinical perspective, due to the requirement for a large volume of blood and the high cost of RNA sequencing. However, the cost of RNA sequencing continues to fall, and this may become more feasible in future practice. The composite score excluding the $CD19^{+}$ B cell gene expression data would be more translational to clinical practice at the current time, as it would only require a small volume of blood and procedures that are already routinely undertaken in NHS laboratories.

9.10 Summary

This Chapter highlights the complex interplay between various components of the humoral immune system. The association of MLR, an independent predictor of outcome, with acute phase inflammatory markers suggests the presence of an inflammatory environment preceding relapse. The development of a predictive model for GD relapse including three genes and two biomarkers demonstrated good predictive value and if validated, has translational potential to help facilitate clinical decision making on ATD withdrawal.

Chapter 10: General Discussion and Future perspectives

This Chapter summaries the key findings from my thesis and discusses the potential for future work aimed at gaining greater insight into the determinants of GD outcome.

The key findings are summarised in the points as below:

1. TRAb is well-described as the most effective biomarker currently available to predict outcome in GD at the timepoint of stopping ATD. This is consistent with the findings from my thesis which also demonstrates TRAb (at the end of ATD treatment) as an important independent predictive biomarker for outcome, evidenced by its significance in all predictive models and inclusion in the overall final predictive model including genetic factors. This thesis also demonstrates the significance of TRAb as a biomarker in predicting timing of GD relapse, with higher TRAb levels resulting in a quicker onset of relapse. However, as previously reported and observed in this thesis, the presence of a negative TRAb at the timepoint of stopping ATD does not accurately predict relapse in all patients (nearly 40% with a negative TRAb had relapsed by one year in this thesis). A positive TRAb likely reflects ongoing GD activity and therefore if the ATD is withdrawn while the disease remains active relapse ensues. Those testing negative may have circulating TRAb not detectable on the currently available assays and therefore still have the potential to relapse, albeit at a later timepoint as demonstrated in this thesis. The limitations of TRAb in isolation as a predictor of GD outcome highlight the need for additional predictive biomarkers.
2. This thesis revealed sTACI to be a good predictor of GD outcome, and like TRAb, an elevated sTACI at the timepoint of stopping ATD was consistently associated with relapse. sTACI also featured in the final predictive model alongside the genetic factors. sTACI has not previously been studied in GD but is an important biomarker reflecting B cell activation. The association of sTACI with relapse at the timepoint of stopping ATD highlights that even at a time of seemingly ‘quiet’ immunity, there is an undercurrent of previously unmeasured B cell activity present in those that will relapse by one year. sTACI can be measured with a small sample of serum and a simple laboratory test similar to the routine measurement of TRAb prior to stopping ATD, giving it great translatable potential to the clinical setting as a biomarker of outcome.

3. The importance of humoral activity in GD was also highlighted in this thesis by KRECs which provided insight into the effect of thyroid status on B cell activity, with hyperthyroidism associated with increased central B cell output and reduced peripheral B cell proliferation. The strong association between KRECs and early-stage B cell populations suggests an expanded autoreactive immature B cell population in patients with elevated KRECs that may ultimately give rise to the TRAb-producing B cells. Indeed, elevated KRECs 6-10 weeks after stopping ATD was associated with relapse indicating an emergence of B cell activity after withdrawal of treatment that could distinguish which patients will relapse by one year. This observation of increased B cell output centrally from the bone marrow after ATD withdrawal provides possible mechanistic insight into the relapse process, however the inability of KRECs to detect relapse prior to stopping ATD means that, unlike TRAb and sTACI, it is not an ideal biomarker of predicting relapse that would help facilitate clinical decision making.
4. The B cell RNA sequencing revealed minimal B cell transcriptomic variation at the timepoint of stopping ATD between outcome groups. This may be because these genes were measured at a time of ‘quiet’ immunity prior to the ATD being withdrawn and therefore I may have observed greater DEG between groups if captured whilst the relapse process was occurring. However, the purpose of RNA sequencing in this thesis was to generate a B cell transcriptomic signature that could predict relapse at the time when the clinical decision about ATD treatment is made, which is prior to the treatment stopping. Using unadjusted P values there were 33 DEG which provided some potential molecular mechanistic insights into signalling and pathways involved in GD relapse. Some of the top DEG were sex-specific genes which highlighted the possibility of the model failing to adequately correct for gender. However, considering the sexual dimorphism of the immune system and the known gender-related risk of relapse, the expression of sex-specific genes is important to consider as potential pathways involved in the relapse process.
5. Analysis of over 250 of the top DEG revealed that the most enriched canonical pathway was associated with cytokine signalling, with significant humoral and cellular immune response pathways also implicated, indicating the potential immune mechanisms associated with relapse. Investigation of the functional and biological characteristics of the DEG revealed a possible pathway of

oxidative stress (upregulated *SOD2*) potentially resulting from an inflammatory environment (upregulated *IL-B*) or vice versa, that results in a perpetual cycle of inflammatory-induced oxidative stress. This may involve thyroid inflammation with associated thyroid cellular damage and the exposure of autoantigen driving the breakdown of self-tolerance (Figure 10.1).

6. The final 5-variable predictive model that provided good predictive value (ROC_{AUC} = 0.99), sensitivity, specificity, PPV and NPV, included TRAb and sTACI, and 3 downregulated genes (*POC1B-AS1, TEC, ANKRD20A3*). It is unclear how *TEC* or *ANKRD20A3* may contribute to GD relapse as there is no functional data to support them having an immune-related role or otherwise. However, the previously observed positive association of *POC1B-AS1* with circulating Tregs suggests down-regulated *POC1B-AS1* in relapse patients may indicate a reduced circulating population of Tregs which has previously been associated with a greater risk of GD relapse. Further work to characterise the biological functions of these genes and explore their association with GD relapse is warranted.

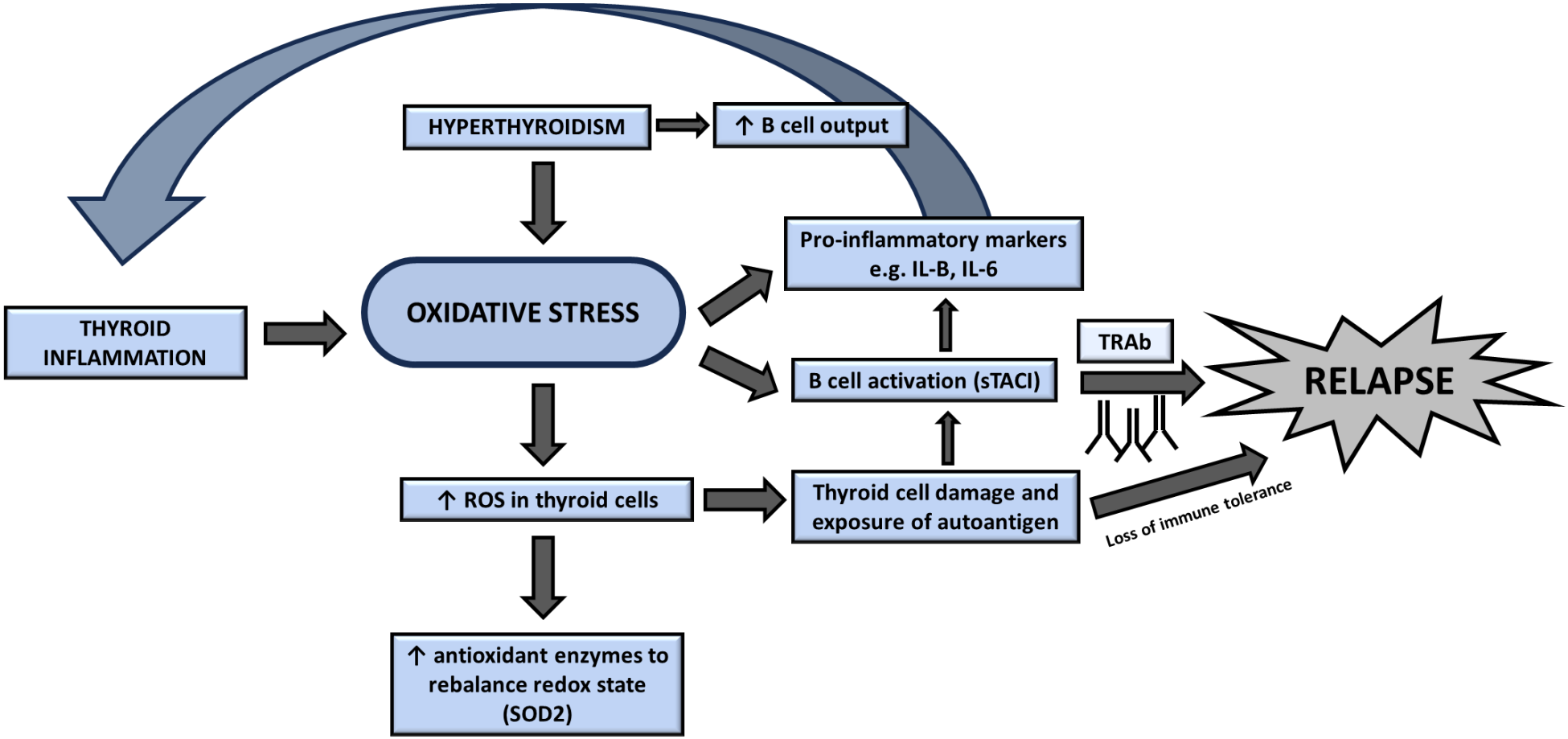


Figure 10.1 – Summary of proposed pathways to Graves' disease relapse.

Future work

1. Applying the predictive model in another larger cohort of GD patients is essential to validate the findings from this thesis and ensure good predictive value is maintained.
2. The DEG generated by the sequencing analysis could be validated by qPCR, however RNA-sequencing methods and data analysis are usually robust enough to not require validation, unlike previously used microarray methods (*Coenye T. 2021*)
3. Genomic association studies examining B cell eQTLs in the top identified DEG could be further investigated to explain variation in expression levels. The B cell eQTLs identified in the top DEG are presented in Appendix I.
4. The RNA sequencing canonical pathway analysis revealed other potentially relevant cytokines that could be studied in relation to GD outcome such as IL-1B, IL-A, IL-13, and IL-17.
5. Functional work of the DEG identified in this thesis may enable greater mechanistic insight into the pathogenesis of thyroid autoimmunity and specifically assist in understanding the immune pathways involved in GD relapse.
6. The application of single-cell RNA sequencing would allow greater insight into the contribution of individual B cell phenotypes in GD relapse and therefore provide a more in-depth understanding of the complex B cell dynamics and additionally may help to identify specific novel therapeutic targets.

In summary, there is clinical imperative for an effective predictive tool that can be used at the end of ATD treatment to support the clinical decision on timing of ATD withdrawal and enable a ‘precision-medicine’ approach to the management of GD. The predictive score developed in this thesis, if validated, could provide a tool that can translate to the clinical setting and assist in personalised decisions about the timing of ATD withdrawal.

Appendix A: Laboratory reagents and equipment

The following list details the laboratory reagents and equipment used in this study, including the manufacturer and catalogue number, where applicable.

Serum separation		
BD Vacutainer Serum Tube, 10ml, Red	Becton Dickinson	MD367895
Megafuge™ 16R Centrifuge	ThermoFisher Scientific	75004270
DNA extraction and quantification		
BD Vacutainer K2EDTA Tube, 10ml, Lavender	Becton Dickinson	MD367525
Gibco™ Trypsin-EDTA (0.05%), phenol red	ThermoFisher Scientific	25300054
NanoDrop™ 8000 Spectrophotometer	ThermoFisher Scientific	ND-8000-GL
Kappa light chain excision circles (KRECs)		
KREC CJ/SJ primers and probes	Integrated DNA technologies	Custom
TCRAC primers and probe	ThermoFisher Scientific	Custom
TaqMan™ Universal PCR Master Mix	ThermoFisher Scientific	4304437
QuantStudio™ 7 Flex Real-Time PCR System	ThermoFisher Scientific	4485701
PBMC isolation		
Hanks' Balanced Salt Solution (HBSS) with phenol red without Calcium/Magnesium 500 ml	Lonza	LZBE10-543F
Lymphoprep™	STEMCELL Technologies	07851
SepMate™ tubes	STEMCELL Technologies	85450
EASYstrainer™ sterile cell strainer 70 µM	Greiner Bio-One	542070
$CD19^+$ B cell isolation		

CD19 magnetic Microbeads®	Miltenyi Biotec	130-050-301
MACS® LS columns	Miltenyi Biotec	130-042-401
QuadroMACS™ Separator	Miltenyi Biotec	130-090-976
Acrodisc ® 0.2µm Syringe Filter	ThermoFisher Scientific	NC0784274
$CD19^+$ B cell lysis		
RNase*Zap*™	ThermoFisher Scientific	AM9780
TRIzol™	ThermoFisher Scientific	15596026
$CD19^+$ B cell RNA extraction		
RNeasy® Mini Kits	Qiagen	74106
Qubit™ 4 Fluorometer	ThermoFisher Scientific	Q33238
Flow cytometry		
Anti-Mouse Ig, κ/Negative Control Compensation Particles Set	BD Biosciences	552843
Fluorescein isothiocyanate (FITC) Mouse anti-human CD19 Antibody (clone HIB19)	BioLegend	302206
Pacific Blue (PE) Mouse anti-human CD1d Antibody (clone 51.1)	BioLegend	350306
Phycoerythrin (PE)/ Cyanine7 (Cy7) Mouse anti-human CD5 (clone L17F12)	BioLegend	364008
Allophycocyanin (APC) Mouse anti-human CD24 Antibody (clone ML5)	BioLegend	311118
APC/Cyanine7 Mouse anti-human CD20 Antibody (clone 2H7)	BioLegend	302314
Brilliant Violet 650™ Mouse anti-human CD27 Antibody (clone O323)	BioLegend	302828
Peridinin Chlorophyll Protein Complex (PerCP)/Cyanine5.5 Mouse	BioLegend	348208

anti-human IgD Antibody (clone IA6-2)		
PE Mouse anti-human CD14 Antibody (clone M5E2)	BioLegend	301806
APC Mouse anti-human CD4 Antibody (clone OKT4)	BioLegend	317416
Brilliant Violet 421™ Mouse anti-human CD3 Antibody (clone OKT3)	BioLegend	317344
Zombie Aqua™ Fixable Viability Kit	BioLegend	423101
LSRFortessa™ Cell Analyzer	BD Biosciences	23-11947-01
Cytokine ELISAs		
Human BAFF/BLyS/TNFSF13B Quantikine ELISA Kit	R&D Systems	DBLYS0B
Human CXCL13/BLC/BCA-1 Quantikine ELISA Kit	R&D Systems	DCX130
APRIL Human ELISA Kit	ThermoFisher Scientific	BMS2008
Human TACI/TNFRSF13B DuoSet ELISA	R&D Systems	DY174
Human BCMA/TNFRSF17 DuoSet ELISA	R&D Systems	DY193
Human IL-6 HS ELISA - Quantikine	R&D Systems	HS600C
Varioskan™ LUX multimode microplate reader	ThermoFisher Scientific	VL0000D0
mRNA-sequencing		
Tapestation™ 4200	Agilent	G2991BA
NEBNext® Low Input RNA Library Prep Kit for Illumina®	New England Biolabs	E6420L
NovaSeq 6000 S2 flow cell	Ilumina	20012850

Appendix B: Probes and primers used in the KRECs experiment

Target	Forward primer (5'-3')	Reverse primer (5'-3')	Probe (FAM/VIC-TAMRA) (5'-3')
KREC-SJ	TCAGCGCCCATTACGTTTCT	GTGAGGGACACGCAGCC	CCAGCTCTTACCCTAGAGTTTCTGCACGG
KREC-CJ	CCCGATTAATGCTGCCGTAG	CCTAGGGAGCAGGGAGGCTT	AGCTGCATTTTTGCCATATCCACTATTTGGAGT
TCRAC	TGGCCTAACCCTGATCCTCTT	GGATTTAGAGTCTCTCAGCTGGTACAC	TCCCACAGATATCCAGAACCCTGACCC

Appendix C: Canonical Pathway Analysis: relapse vs. remission

Ingenuity Canonical Pathway	Ingenuity Canonical Pathway Category (s)	-log (P-value)	Representative gene
Granulocyte Adhesion and Diapedesis	Cellular immune response	2.04	*ACKR3, CCL3L1, CLDN4, IL1B*
Hepatic Cholestasis	Ingenuity toxicity pathway	2.02	*ABCG8, IL13, IL17D, IL1B*
Agranulocyte Adhesion and Diapedesis	Cellular immune response	1.88	*ACKR3, CCL3L1, CLDN4, IL1B*
LXR/RXR Activaton	Nuclear receptor signalling	1.79	*ABCG8, APOD, IL1B*
FXR/RXR Activation	Nuclear receptor signalling	1.76	*ABCG8, APOD, IL1B*
Atherosclerosis Signaling	Cardiovascular signalling	1.7	*APOD, IL1B, PLAAT1*
GADD45 Signaling	Cell cycle regulation	1.56	*GADD45B, IL1B*
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	Cellular immune response	1.52	*IL13, IL17D, IL1B*
Superoxide Radicals Degradation	Detoxification	1.47	*SOD2*
HMGB1 Signaling	Humoral immune response	1.44	*IL13, IL17D, IL1B*

Erythropoietin Signaling Pathway	Growth factor signalling	1.38	*IL13, IL17D, IL1B*
Embryonic Stem Cell Differentiation into Cardiac Lineages	Organismal growth and development	1.37	*MESP1*
Tumor Microenvironment Pathway	Cancer	1.37	*FGF12, IL13, IL1B*
Role of MAPK Signaling in Inhibiting the Pathogenesis of Influenza	Pathogen influenced signalling	1.34	*IL1B, PLAAT1*
IL-17 Signaling	Cellular immune response	1.32	*IL13, IL17D, IL1B*
Macrophage Classical Activation Signaling Pathway	Cellular immune response	1.31	*IL13, IL17D, IL1B*
NOD1/2 Signaling Pathway	Cellular stress and injury	1.31	*IL13, IL17D, IL1B*

Table – Canonical Pathways and categories as predicted by Ingenuity Pathway Analysis. The table shows the significantly enriched canonical pathways (P <0.05) across the differentially expressed genes identified between the relapse and remission groups with an unadjusted P value of <0.01. Those representing the 'Cytokine Signaling' and 'Disease-specific' categories are found in Chapter 8: Table 8.2.

Appendix D: Disease and Function Analysis (downstream analysis)

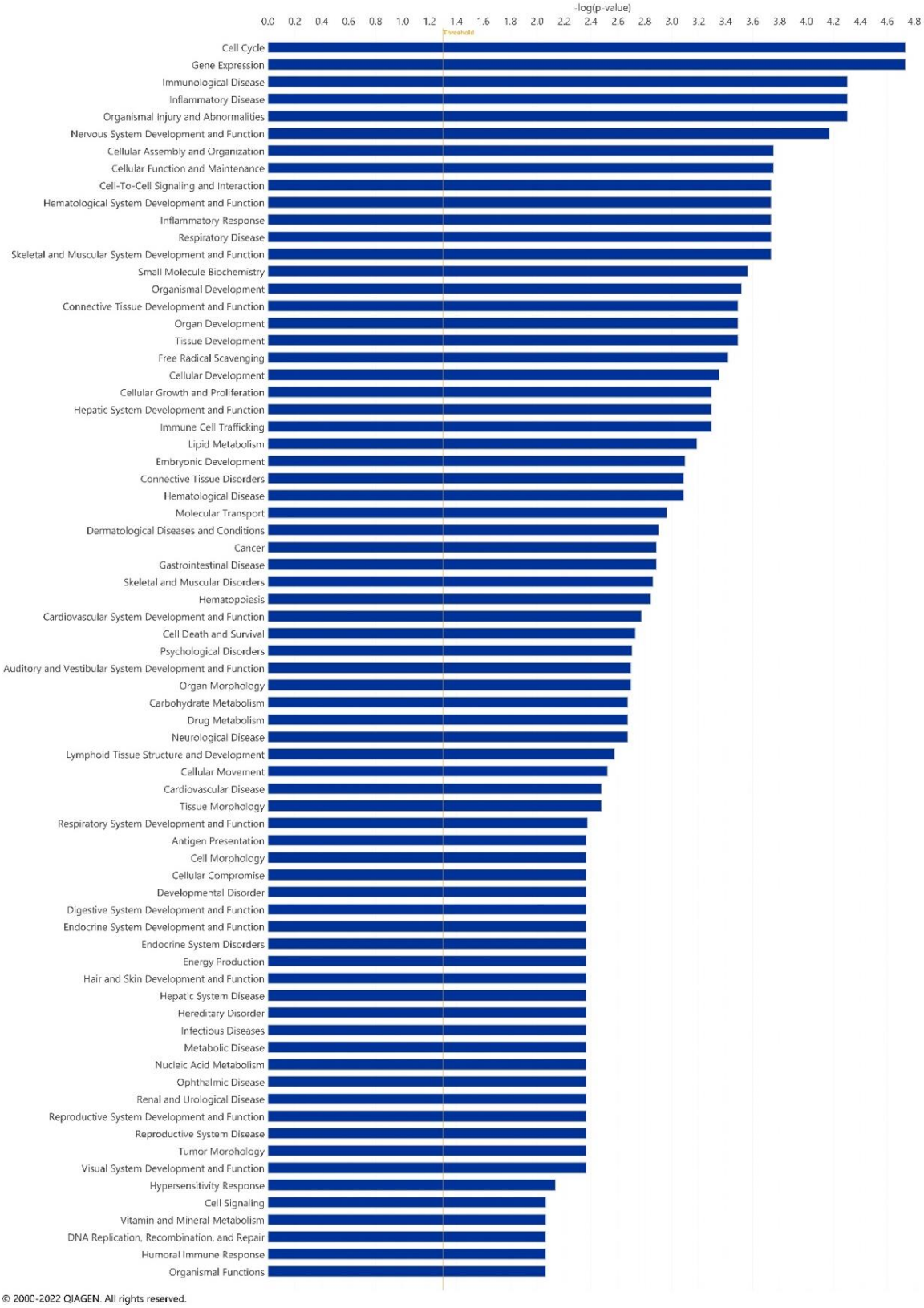


Figure – Bar chart displaying the most significant biological functions predicted to be affected by the differentially expressed genes within the Graves' disease dataset. The orange horizontal threshold line denotes the cut-off for significance (P-value 0.05). Figure generated using QIAGEN Ingenuity Pathway Analysis™ software.

Table – Biological functions and diseases (focused on cellular function and immune response) predicted to be significantly (P=<0.01) associated with the pattern of differentially expressed genes in the Graves' disease dataset

Biological function/ Disease	P-value	Representative genes
Cell Cycle	1.85E-05	GADD45B,HES1,HOXB3,IL13,IL1B,MPL,NR4A3,PFKFB3,SMC1B,SOD2
Immunological Disease	4.97E-05	B3GALT2,CADM2,CCL3L1,CHRNB2,CYP1A2,DDX3Y,EIF1AY,FAM153A/FAM153B,FCGR1A,GADD45B,GPR18, IGFBP4,IL13,IL1B,LINC-PINT,MPL,NEO1,NR4A3,SOD2,TXLNGY,USP9Y
Inflammatory Disease	4.97E-05	ACKR3,CADM2,CCL3L1,CHRNA9,CHRNB2,CTHRC1,CYP1A2,DDX3Y,EIF1AY,FAM153A/FAM153B,FCGR1A, GADD45B,IGFBP4,IL13,IL1B,LINC-PINT,MPL,NR4A3,SOD2,TXLNGY,USP9Y
Organismal Injury and Abnormalities	4.97E-05	ABCG8,ABRA,ACKR3,ACOT1,ADAMTS13,ALG10B,APOD,ARPP21,B3GALT2,C18orf54,C1orf43,CADM2, CCDC148,CCDC22,CCL3L1,CFAP251,CHRNA9,CHRNB2,CLDN4,CTHRC1,CYP1A2,DCLRE1B,DDX3Y, DENND2A,DENND2B,DMAC2,DNAH3,EIF1AY,EPB41L4AAS1,FAM153A/FAM153B,FAXC,FCGR1A,FGF12, FOXC1,FRMD3,GADD45B,GCOM1,GPC5,GPR137C,GPR18,HELB,HES1,HOXB3,IGFBP4,IL13,IL17D,IL1B ,KCNC4,KCTD16,KCTD2,LEAP2,LHFPL4,LINC-PINT,LRRC75B,MESP1,MPL,MRPL11,MYOF,NBEAL1,NEDD4L, NEO1,NOG,NR4A3,OMD,PFKFB3,RBPMS,RFTN2,RUFY4,SCO1,SDE2,SERTAD1,SETSIP,SLC10A3,SLC41A3, SMC1B,SMG1,SOD2,SPX,TANC1,TDRD1,TMEM40,TMEM86A,TRABD2A,TSEN34,TSPYL6,TTBK2,TXLNGY, USP9Y,VSTM4,ZBTB8A,ZIM2,ZNF474,ZNF578,ZNF609,ZNF630,ZNF783,ZNF846
Cellular Assembly and Organization	1.76E-04	CHRNA9,CHRNB2,CLDN4,CTHRC1,FGF12,HES1,IL13,IL1B,MYOF,NEDD4L,NEO1,TANC1
Cellular Function and Maintenance	1.76E-04	CCDC22,CCL3L1,CHRNA9,CHRNB2,CLDN4,CTHRC1,FCGR1A,FGF12,GPR18,HES1,IL13,IL17D,IL1B,MPL, NEDD4L,NEO1,PFKFB3,SCO1,SOD2

Cell-To-Cell Signaling and Interaction	1.83E-04	ADAMTS13,CCL3L1,CHRNB2,CLDN4,CTHRC1,CYP1A2,FCGR1A,GPR18,HES1,IL13,IL1B,KCNC4,MYOF, NEDD4L,NEO1,NOG,NR4A3,PFKFB3,SOD2,TANC1
Inflammatory Response	1.83E-04	ACKR3,ACOT1,ADAMTS13,CCL3L1,CHRNA9,CHRNB2,CYP1A2,FCGR1A,GADD45B,GPR18,IL13,IL17D,IL1B, LINC-PINT,MPL,NEDD4L,NEO1,NR4A3,PFKFB3,SOD2
Connective Tissue Development and Function	3.23E-04	ACKR3,ADAMTS13,CTHRC1,CYP1A2,FCGR1A,FOXC1,GADD45B,HES1,HOXB3,IGFBP4,IL13,IL1B,MYOF, NOG,NR4A3,PFKFB3, PLAAT1,SERTAD1,SOD2
Free Radical Scavenging	3.83E-04	CYP1A2,IL13,IL1B,SOD2
Cellular Development	4.44E-04	ACKR3,APOD,ARPP21,CLDN4,CTHRC1,CYP1A2,FCGR1A,FGF12,FOXC1,GADD45B,HES1,HOXB3, IGFBP4,IL13,IL1B,MESP1,MPL,MYOF,NEO1,NOG,NR4A3,PFKFB3,PLAAT1,SERTAD1,SETSIP,SOD2, ZNF609
Cellular Growth and Proliferation	5.09E-04	ACKR3,APOD,CHRNB2,CLDN4,CTHRC1,CYP1A2,FCGR1A,FGF12,FOXC1,GADD45B,HES1,HOXB3, IGFBP4,IL13,IL17D,IL1B,MESP1,MPL,MYOF,NEDD4L,NOG,NR4A3,PFKFB3,PLAAT1,SERTAD1,SOD2, ZNF609
Immune Cell Trafficking	5.09E-04	ACKR3,ADAMTS13,CCL3L1,CYP1A2,IL13,IL1B,NEDD4L,NR4A3,SOD2
Connective Tissue Disorders	8.14E-04	ABCG8,ADAMTS13,CCL3L1,FCGR1A,GADD45B,HOXB3,IGFBP4,IL1B,MPL,NOG,NR4A3,SOD2
Molecular Transport	1.09E-03	ABCG8,APOD,CCDC22,CHRNB2,CLDN4,CYP1A2,FCGR1A,FGF12,IL13,IL1B,KCNC4,NEDD4L, NEO1,PFKFB3,PLAAT1,SCO1,SOD2
Cell Death and Survival	1.87E-03	ABCG8,ACKR3,ADAMTS13,CLDN4,FCGR1A,GADD45B,IL13,IL1B,MPL,NR4A3,PFKFB3,SOD2

Lymphoid Tissue Structure and Development	2.67E-03	ACKR3,CYP1A2,FCGR1A,GPR18,HES1,HOXB3,IL13,IL1B,MPL
Cellular Movement	3.01E-03	ACKR3,ADAMTS13,CCL3L1,CYP1A2,HES1,IL13,IL1B,NEDD4L,NOG,NR4A3,SOD2
Antigen Presentation	4.32E-03	FCGR1A
Cell Morphology	4.32E-03	CHRNA9,CYP1A2,HES1,IL13,IL1B,MESP1
Cellular Compromise	4.32E-03	FCGR1A,IL13,IL1B,NR4A3,PFKFB3,SOD2
Infectious Diseases	4.32E-03	B3GALT2,CCL3L1,FAM153A/FAM153B,IL13,NEO1
Cell Signaling	8.61E-03	CLDN4,FCGR1A,IL1B
Humoral Immune Response	8.62E-03	ACKR3

Table – Specific immunological diseases predicted to be significantly associated with the differentially expressed genes in the Graves' disease dataset

Immunological Disease	P-value	Representative genes
Systemic lupus erythematosus	0.0000497	CADM2,CCL3L1,CHRNB2,DDX3Y,EIF1AY,FCGR1A,IL1B,LINC-PINT,NR4A3,SOD2,TXLNGY,USP9Y
Active stage immunological disorder	0.000571	DDX3Y,EIF1AY,IL1B,TXLNGY,USP9Y
Refractory chronic primary immune thrombocytopenia	0.000814	FCGR1A,MPL
Active stage systemic lupus erythematosus	0.000903	DDX3Y,EIF1AY,TXLNGY,USP9Y
Polyarticular juvenile rheumatoid arthritis	0.00138	CCL3L1,GADD45B,IL1B,NR4A3
Immune thrombocytopenic purpura	0.00334	FCGR1A,MPL
Secondary immune thrombocytopenia	0.00432	MPL
Familial CD64 phagocytic deficiency	0.00432	FCGR1A
Susceptibility to HIV/AIDS	0.00432	CCL3L1
Antisynthetase syndrome	0.00432	IL1B

Primary immune thrombocytopenic purpura	0.00432	MPL
Systemic autoimmune syndrome	0.0045	CADM2,CCL3L1,CHRNB2,CYP1A2,DDX3Y,EIF1AY,FCGR1A,GADD45B,GPR18,IGFBP4,IL13,IL1B,LINC-PINT,MPL,NR4A3,SOD2,TXLNGY,USP9Y
Bullous pemphigoid	0.0052	FCGR1A,IL13
Severe splenomegaly	0.00862	MPL
Mild splenomegaly	0.00862	MPL
Immune mediated inflammatory disease	0.00901	CADM2,CCL3L1,CHRNB2,CYP1A2,DDX3Y,EIF1AY,FAM153A/FAM153B,FCGR1A,GADD45B,IGFBP4,IL13,IL1B,LINC-PINT,MPL,NR4A3,SOD2,TXLNGY,USP9Y
Severe hypersensitive reaction	0.01	FCGR1A,IL13
X-linked immunodeficiency with magnesium defect, Epstein-Barr virus infection and neoplasia	0.0107	B3GALT2,FAM153A/FAM153B,NEO1
Continuously active disease stage systemic juvenile idiopathic arthritis	0.0129	IL1B

Appendix E: Full list of molecular shapes used in QIAGEN Ingenuity Pathway Analysis™ canonical pathways

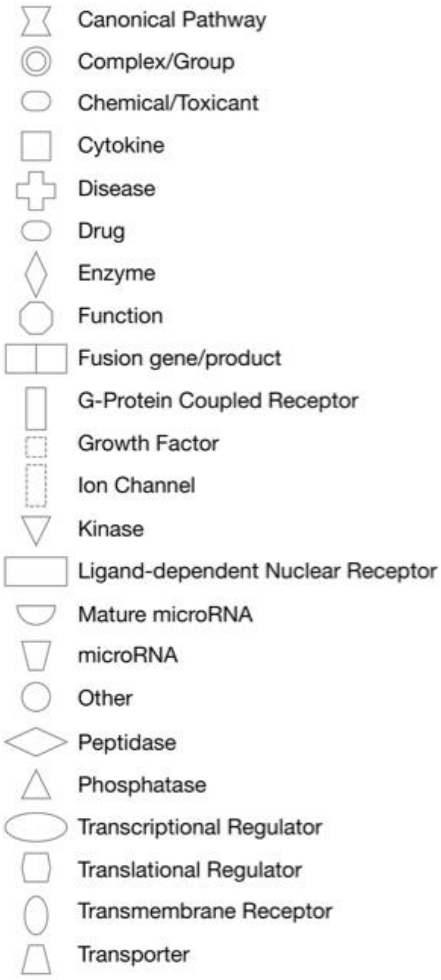


Figure – Legend describing the node shapes used to represent different molecules in the canonical pathways generated by the QIAGEN Ingenuity Pathway Analysis™ software. IPA legend downloaded from QIAGEN, https://qiagen.my.salesforce-sites.com/KnowledgeBase/articles/Knowledge/Legend. Accessed 19.08.2024.

Appendix F: Full list of composite gene scores ranked by ROC_{AUC}

POC1B-AS1	*SOD2*	*TEC*	*ANKRD20A3*	*SCO1*	DDX11L17	ROC_{AUC}
✓	✓	✓	✓	✓	✓	0.95
✓	✓	✓	✓	✓	×	0.95
✓	×	✓	✓	✓	✓	0.95
×	✓	✓	✓	✓	✓	0.95
✓	✓	✓	✓	×	×	0.95
✓	×	✓	✓	✓	×	0.94
×	×	✓	✓	✓	✓	0.94
✓	✓	✓	✓	×	✓	0.94
✓	✓	✓	×	✓	✓	0.94
✓	✓	×	✓	✓	✓	0.94
✓	×	✓	✓	×	✓	0.94
✓	✓	×	✓	✓	×	0.93
✓	×	✓	✓	×	×	0.93
✓	×	×	✓	✓	✓	0.93
✓	×	×	✓	✓	×	0.93
✓	✓	×	✓	×	✓	0.92
×	×	✓	✓	✓	×	0.92
×	✓	×	✓	✓	✓	0.92
×	✓	✓	✓	✓	×	0.92
✓	✓	✓	×	×	✓	0.91
✓	×	✓	✓	×	✓	0.91
✓	✓	×	✓	×	×	0.91
×	✓	✓	✓	×	✓	0.91
×	✓	✓	×	✓	✓	0.90
✓	✓	✓	×	✓	×	0.90
✓	×	✓	×	✓	×	0.90
✓	×	✓	×	×	✓	0.90
×	×	✓	×	✓	✓	0.90
×	×	×	✓	✓	✓	0.90
×	✓	✓	✓	×	×	0.89
✓	✓	×	×	✓	✓	0.89
×	×	✓	✓	×	✓	0.89
✓	×	×	✓	×	✓	0.89

✓	×	×	✓	×	×	0.89
✓	✓	✓	×	×	×	0.88
×	✓	×	✓	✓	×	0.88
×	✓	✓	×	✓	×	0.88
✓	×	×	×	✓	✓	0.88
✓	✓	×	×	×	✓	0.87
×	×	×	✓	✓	×	0.87
×	×	✓	×	✓	×	0.87
×	✓	×	✓	×	✓	0.87
×	✓	✓	×	×	✓	0.87
✓	✓	×	×	✓	×	0.86
×	✓	×	×	✓	✓	0.86
×	✓	×	×	✓	✓	0.86
✓	×	✓	×	×	×	0.86
✓	×	×	×	✓	×	0.86
×	×	✓	✓	×	×	0.85
×	✓	✓	×	×	×	0.83
×	×	✓	×	×	✓	0.83
×	×	×	×	✓	✓	0.83
×	✓	×	✓	×	×	0.83
✓	✓	×	×	×	×	0.82
✓	×	×	×	×	✓	0.82
×	✓	×	×	✓	×	0.81
×	×	×	✓	×	✓	0.81
×	✓	×	×	×	✓	0.79
×	×	×	×	✓	×	0.79
✓	×	×	×	×	×	0.78
×	×	✓	×	×	×	0.77
×	×	×	✓	×	×	0.77
×	✓	×	×	×	×	0.72
×	×	×	×	×	✓	0.71

Appendix G: Full list of composite biomarker scores ranked by ROC_{AUC}

ln(MLR)	ln(TRAb)	ln(sTACI)	ln(BAFF)	ln(sBCMA)	ROC_{AUC}
✓	✓	✓	✓	✓	0.86
×	✓	✓	✓	✓	0.86
✓	✓	✓	✓	×	0.86
×	✓	✓	✓	×	0.85
✓	✓	✓	×	✓	0.84
✓	✓	✓	×	×	0.84
×	✓	✓	×	×	0.84
×	✓	✓	×	✓	0.82
✓	×	✓	✓	×	0.82
✓	×	✓	✓	✓	0.81
×	✓	×	✓	✓	0.79
✓	✓	×	×	✓	0.79
✓	✓	×	✓	✓	0.78
×	✓	×	×	✓	0.78
✓	✓	×	✓	×	0.77
×	×	✓	✓	✓	0.77
×	×	✓	✓	×	0.76
✓	×	✓	×	×	0.76
×	✓	×	✓	×	0.75
✓	×	×	✓	✓	0.74
✓	×	✓	×	✓	0.73
✓	✓	×	×	×	0.73
×	✓	×	×	×	0.72
✓	×	×	✓	×	0.70
×	×	×	✓	✓	0.70
✓	×	×	×	✓	0.70
×	×	✓	×	×	0.66
✓	×	×	×	×	0.66
×	×	✓	×	✓	0.65
×	×	×	✓	×	0.64
×	×	×	×	✓	0.64

Appendix H: Full list of composite gene/biomarker scores ranked by ROC_{AUC}

POC1B-AS1	*SOD2*	*TEC*	*ANKRD20A3*	*SCO1*	*DDX11L17*	TRAb	sTACI	ROC_{AUC}
✓	✓	✓	✓	✓	✓	✓	✓	0.99
✓	✓	✓	✓	×	✓	✓	✓	0.99
✓	×	✓	✓	✓	✓	✓	✓	0.99
✓	✓	✓	✓	×	×	✓	✓	0.99
✓	×	✓	✓	✓	×	✓	✓	0.99
✓	×	✓	✓	×	✓	✓	✓	0.99
✓	×	✓	✓	×	×	✓	✓	0.99
✓	✓	✓	✓	✓	✓	✓	×	0.97
✓	✓	✓	✓	✓	✓	×	✓	0.97
✓	✓	✓	✓	✓	×	✓	×	0.97
×	✓	✓	✓	✓	✓	✓	✓	0.97
✓	×	✓	✓	✓	✓	✓	×	0.97
×	✓	✓	✓	✓	×	✓	✓	0.97
×	✓	✓	✓	×	✓	✓	✓	0.97
×	×	✓	✓	✓	✓	✓	✓	0.97
✓	✓	×	✓	×	×	✓	✓	0.97
✓	×	✓	✓	✓	×	✓	×	0.97

×	✓	✓	✓	×	×	✓	✓	0.97
×	×	✓	✓	✓	×	✓	✓	0.97
×	×	✓	✓	×	×	✓	✓	0.96
✓	✓	×	✓	✓	✓	✓	✓	0.96
✓	✓	✓	✓	✓	×	×	✓	0.96
✓	✓	✓	✓	×	✓	✓	×	0.96
✓	✓	✓	✓	×	✓	×	✓	0.96
✓	✓	×	✓	✓	×	✓	✓	0.96
✓	✓	✓	×	✓	×	✓	✓	0.96
✓	✓	×	✓	×	✓	✓	✓	0.96
✓	×	×	✓	✓	✓	✓	✓	0.96
✓	×	✓	✓	✓	✓	×	✓	0.96
✓	×	✓	✓	×	✓	✓	×	0.96
✓	×	✓	×	✓	×	✓	✓	0.96
✓	×	×	✓	✓	×	✓	✓	0.96
✓	×	×	✓	×	✓	✓	✓	0.96
×	×	✓	✓	×	✓	✓	✓	0.96
✓	×	✓	✓	×	×	✓	×	0.96
✓	×	×	✓	×	×	✓	✓	0.96
✓	✓	✓	×	✓	✓	✓	✓	0.95
✓	✓	✓	✓	✓	✓	×	×	0.95

✓	✓	✓	×	×	✓	✓	✓	0.95
✓	×	✓	×	✓	✓	✓	✓	0.95
×	✓	✓	✓	✓	✓	✓	×	0.95
×	✓	✓	✓	✓	✓	×	✓	0.95
✓	✓	✓	✓	×	×	✓	×	0.95
✓	✓	✓	✓	×	×	×	✓	0.95
✓	✓	✓	✓	✓	×	×	×	0.95
✓	✓	✓	×	×	×	✓	✓	0.95
✓	✓	×	✓	✓	✓	×	×	0.95
✓	×	✓	✓	✓	✓	×	×	0.95
✓	×	✓	✓	×	✓	×	✓	0.95
✓	×	✓	✓	✓	×	×	✓	0.95
✓	×	✓	×	×	✓	✓	✓	0.95
✓	×	✓	×	×	×	✓	✓	0.95
✓	✓	×	✓	✓	✓	✓	×	0.94
✓	✓	×	✓	✓	✓	×	✓	0.94
✓	✓	✓	✓	×	✓	×	×	0.94
×	✓	✓	✓	✓	✓	×	×	0.94
×	✓	✓	✓	✓	×	✓	×	0.94
×	×	✓	✓	✓	✓	✓	×	0.94
×	×	✓	✓	✓	✓	×	✓	0.94

✓	✓	✓	✓	×	×	×	×	0.94
✓	×	✓	✓	✓	×	×	×	0.94
✓	×	✓	✓	×	×	×	✓	0.94
×	×	✓	✓	✓	×	✓	×	0.94
✓	✓	✓	×	✓	✓	×	✓	0.93
×	✓	✓	×	✓	✓	✓	✓	0.93
×	✓	×	✓	✓	✓	✓	✓	0.93
✓	✓	✓	×	✓	×	×	✓	0.93
✓	✓	×	✓	✓	×	✓	×	0.93
✓	✓	×	✓	✓	×	×	✓	0.93
✓	×	×	✓	✓	✓	✓	×	0.93
✓	×	×	✓	✓	✓	×	✓	0.93
✓	✓	×	✓	×	✓	✓	×	0.93
✓	✓	×	✓	×	✓	×	✓	0.93
✓	✓	✓	×	×	✓	×	✓	0.93
✓	×	✓	×	✓	✓	×	✓	0.93
×	✓	✓	×	✓	×	✓	✓	0.93
×	✓	×	✓	✓	×	✓	✓	0.93
✓	✓	✓	×	×	×	×	✓	0.93
✓	×	✓	✓	×	✓	×	×	0.93
✓	×	✓	×	✓	×	×	✓	0.93

✓	×	×	✓	✓	✓	×	×	0.93
✓	×	×	✓	✓	×	✓	×	0.93
×	✓	✓	✓	×	×	✓	×	0.93
×	✓	✓	×	×	×	✓	✓	0.93
×	×	✓	×	✓	×	✓	✓	0.93
×	×	✓	✓	✓	×	×	✓	0.93
×	×	✓	✓	×	✓	✓	×	0.93
✓	✓	×	×	✓	✓	✓	✓	0.92
✓	✓	✓	×	✓	✓	×	×	0.92
✓	✓	×	×	✓	×	✓	✓	0.92
✓	✓	×	×	×	✓	✓	✓	0.92
✓	×	✓	×	✓	✓	✓	×	0.92
✓	×	×	×	✓	✓	✓	✓	0.92
×	✓	✓	✓	×	✓	×	✓	0.92
×	✓	✓	✓	×	✓	✓	×	0.92
×	✓	✓	×	✓	✓	×	✓	0.92
×	✓	✓	×	×	✓	✓	✓	0.92
×	✓	×	✓	×	✓	✓	✓	0.92
×	×	✓	×	✓	✓	✓	✓	0.92
×	×	×	✓	✓	✓	✓	✓	0.92
✓	✓	×	✓	×	×	×	✓	0.92

✓	✓	×	×	×	×	✓	✓	0.92
✓	×	✓	×	✓	✓	×	×	0.92
✓	×	×	✓	×	✓	✓	×	0.92
×	✓	×	✓	×	×	✓	✓	0.92
×	×	✓	✓	✓	✓	×	×	0.92
×	×	✓	×	×	✓	✓	✓	0.92
×	×	×	✓	✓	×	✓	✓	0.92
✓	×	✓	✓	×	×	×	×	0.92
✓	×	×	✓	✓	×	×	×	0.92
✓	×	×	✓	×	×	✓	×	0.92
×	×	✓	✓	×	×	✓	×	0.92
×	×	✓	×	×	×	✓	✓	0.92
✓	✓	✓	×	✓	✓	✓	×	0.91
✓	✓	✓	×	✓	×	✓	×	0.91
✓	✓	✓	×	×	✓	✓	×	0.91
✓	✓	×	✓	×	×	✓	×	0.91
✓	×	✓	×	✓	×	✓	×	0.91
✓	×	✓	×	×	✓	✓	×	0.91
✓	×	✓	×	×	✓	×	✓	0.91
✓	×	×	✓	✓	×	×	✓	0.91
✓	×	×	✓	×	✓	×	✓	0.91

✓	×	×	×	✓	×	✓	✓	0.91
✓	×	×	×	×	✓	✓	✓	0.91
×	✓	✓	✓	✓	×	×	×	0.91
×	✓	✓	×	✓	×	×	✓	0.91
×	×	✓	✓	×	✓	×	✓	0.91
×	×	×	✓	×	✓	✓	✓	0.91
✓	×	✓	×	×	×	×	✓	0.91
✓	×	×	×	×	×	✓	✓	0.91
✓	✓	×	×	✓	✓	✓	×	0.90
✓	✓	✓	×	✓	×	×	×	0.90
×	✓	×	✓	✓	✓	✓	×	0.90
✓	✓	✓	×	×	×	✓	×	0.90
×	✓	✓	✓	✓	×	×	×	0.90
×	✓	✓	✓	×	×	×	✓	0.90
×	✓	✓	×	✓	✓	×	×	0.90
×	✓	✓	×	×	✓	×	✓	0.90
×	✓	×	✓	✓	✓	×	×	0.90
×	×	✓	×	✓	✓	×	✓	0.90
✓	✓	×	✓	×	×	×	×	0.90
✓	×	✓	×	×	×	✓	×	0.90
✓	×	×	✓	×	✓	×	×	0.90

×	×	✓	✓	✓	×	×	×	0.90
×	×	✓	✓	×	×	×	✓	0.90
×	×	✓	×	✓	×	×	✓	0.90
×	×	×	✓	×	×	✓	✓	0.90
✓	✓	×	×	✓	✓	×	✓	0.89
×	✓	✓	×	✓	✓	✓	×	0.89
×	✓	×	✓	✓	✓	×	✓	0.89
✓	✓	×	×	✓	×	✓	×	0.89
✓	×	×	×	✓	✓	✓	×	0.89
×	×	×	✓	✓	✓	✓	×	0.89
✓	×	✓	×	✓	×	×	×	0.89
✓	×	✓	×	×	✓	×	×	0.89
✓	×	×	✓	×	×	×	✓	0.89
×	×	✓	✓	×	✓	×	×	0.89
×	✓	×	×	✓	✓	✓	✓	0.88
✓	✓	×	×	✓	×	×	✓	0.88
✓	✓	×	×	×	✓	✓	×	0.88
×	✓	✓	×	✓	×	✓	×	0.88
×	✓	×	✓	✓	×	✓	×	0.88
×	✓	×	✓	✓	×	×	✓	0.88
×	✓	×	✓	×	✓	✓	×	0.88

×	✓	×	✓	×	✓	×	✓	**0.88**
×	×	×	✓	✓	✓	×	✓	**0.88**
✓	✓	✓	×	×	×	×	×	**0.88**
✓	✓	×	×	✓	×	×	×	**0.88**
×	×	✓	×	✓	✓	✓	×	**0.88**
✓	×	×	×	✓	×	✓	×	**0.88**
✓	×	×	×	×	✓	✓	×	**0.88**
×	✓	✓	×	×	×	×	✓	**0.88**
×	×	×	✓	✓	✓	×	×	**0.88**
×	×	×	✓	✓	×	✓	×	**0.88**
×	×	✓	×	✓	✓	×	×	**0.88**
✓	×	×	✓	×	×	×	×	**0.88**
✓	✓	×	×	×	✓	×	✓	**0.87**
✓	×	×	×	✓	✓	×	✓	**0.87**
×	✓	×	×	✓	×	✓	✓	**0.87**
×	✓	×	×	×	✓	✓	✓	**0.87**
×	×	×	×	✓	✓	✓	✓	**0.87**
✓	✓	×	×	×	×	✓	×	**0.87**
×	✓	✓	✓	×	×	×	×	**0.87**
×	✓	✓	×	✓	×	×	×	**0.87**
×	✓	×	✓	×	×	✓	×	**0.87**

×	×	✓	×	×	✓	×	✓	0.87
×	✓	✓	×	×	✓	✓	×	0.86
✓	✓	×	×	×	×	×	✓	0.86
✓	×	×	×	✓	✓	×	×	0.86
✓	×	×	×	✓	×	×	✓	0.86
×	✓	✓	×	×	✓	×	×	0.86
×	✓	×	✓	×	✓	×	×	0.86
×	✓	×	×	×	×	✓	✓	0.86
×	×	✓	×	✓	×	✓	×	0.86
×	×	✓	×	×	✓	✓	×	0.86
×	×	×	✓	×	✓	✓	×	0.86
×	×	×	✓	×	✓	×	✓	0.86
✓	×	✓	×	×	×	×	×	0.86
✓	×	×	×	×	×	✓	×	0.86
×	×	✓	×	×	×	×	✓	0.86
×	✓	×	×	✓	✓	×	✓	0.85
✓	✓	×	×	×	✓	×	×	0.85
✓	×	×	×	×	✓	×	✓	0.85
×	✓	×	✓	✓	×	×	×	0.85
×	✓	×	×	✓	✓	×	×	0.85
×	×	×	×	✓	×	✓	✓	0.85

×	×	×	×	×	✓	✓	✓	0.85
×	×	×	✓	✓	×	×	✓	0.85
×	×	✓	✓	×	×	×	×	0.85
×	×	×	✓	×	×	✓	×	0.85
×	✓	✓	×	×	×	✓	×	0.84
×	✓	×	×	✓	×	✓	×	0.84
✓	×	×	×	✓	×	×	×	0.84
×	×	×	×	×	×	✓	✓	0.84
×	✓	×	×	✓	✓	✓	×	0.83
×	✓	×	✓	×	×	×	✓	0.83
×	×	×	×	✓	✓	✓	×	0.83
✓	×	×	×	×	✓	×	×	0.83
✓	×	×	×	×	×	×	✓	0.83
×	✓	✓	×	×	×	×	×	0.83
×	×	✓	×	✓	×	×	×	0.83
×	×	✓	×	×	✓	×	×	0.83
×	×	✓	×	×	×	✓	×	0.83
×	×	×	✓	✓	×	×	×	0.83
×	×	×	✓	×	✓	×	×	0.83
×	×	×	×	✓	✓	×	×	0.83
✓	✓	✓	×	×	✓	×	×	0.82
×	✓	×	✓	×	×	×	×	0.82

×	×	×	×	✓	✓	×	✓	0.82
✓	✓	×	×	×	×	×	×	0.82
×	×	×	×	✓	×	✓	×	0.82
✓	✓	×	×	✓	✓	×	×	0.81
×	✓	×	×	×	✓	✓	×	0.81
×	✓	×	×	✓	×	×	×	0.81
×	×	×	✓	×	×	×	✓	0.81
✓	✓	×	✓	×	✓	×	×	0.80
×	✓	×	×	×	×	✓	×	0.80
✓	✓	×	✓	✓	×	×	×	0.79
×	✓	×	×	✓	×	×	✓	0.79
×	✓	×	×	×	✓	×	×	0.79
×	×	×	×	✓	×	×	×	0.79
✓	×	×	×	×	×	×	×	0.78
×	✓	×	×	×	✓	×	✓	0.78
×	×	×	×	✓	×	×	✓	0.77
×	×	✓	×	×	×	×	×	0.77
×	×	×	✓	×	×	×	×	0.77
×	✓	×	×	×	×	×	✓	0.72
×	×	×	×	×	✓	✓	×	0.72
×	×	×	×	×	✓	×	✓	0.72
×	✓	×	×	×	×	×	×	0.72

×	×	×	×	×	×	✓	×	**0.72**
×	×	×	×	×	✓	×	×	**0.71**
×	×	×	×	×	×	×	✓	**0.66**

Appendix I: B cell eQTLs

This table details B-cell specific eQTLs associated with the top DEG (unadjusted P= <0.001) in my thesis that could be explored to explain the variation in gene expression (*Fairfax BP et al. 2012*)

Gene	P value DEG (unadjusted)	eQTL	MAF	Alleles	eQTL P value
ZNF783	<0.001	rs4727046	0.5	T/C	4.46E-06
		rs13236978	0.37	A/G	0.0001225
SOD2	<0.001	rs1789990	0.47	G/T	4.17E-05
FRMD3	<0.001	rs11139944	0.2	A/G	0.0009875
SCO1	<0.001	rs2322788	0.41	C/T	0.0003765
MPL	<0.001	rs11210339	0.28	A/G	0.0006442

eQTL; expression quantitative trait loci, MAF; minor allele frequency

Of note, there are also single nucleotide polymorphisms in *SOD2* (rs2758332) and *IL-1B (*rs1143634, rs1143627, rs16944, rs2853550, rs1143643*)* (both *SOD2* and *IL-1B* are DEG, unadjusted P= <0.001) that have been observed to be associated with risk of autoimmunity (SLE, RA and GD) (*Abd El Azeem RA et al. 2021, Heidari Z et al. 2021, Rong, H et al. 2020*), including an *IL-1B* polymorphism (rs1143627) that is associated specifically with GD prognosis (*Hayashi F et al. 2009).*

Appendix J: Relevant publications

Grixti, L, Lane, LC, & Pearce, SH (2024). The genetics of Graves' disease. *Reviews in endocrine & metabolic disorders*, 25(1), 203–214.

Lane, LC, Cheetham, TD, Razvi, S, Allinson, K, & Pearce, SHS (2023). Expansion of the immature B lymphocyte compartment in Graves' disease. *European journal of endocrinology*, 189(2), 208–216.

Lane, LC, Wood, CL, & Cheetham, T (2023). Graves' disease: moving forwards. *Archives of disease in childhood*, 108(4), 276–281.

Lane, LC, Pearce, SH, Mitchell AL. Autoimmune Thyroid disease (Chapter 70). *Clinical Immunology Principles and Practice 6th Edition.* Published August 2022.

Lane, LC, Cheetham, TD, Perros, P, & Pearce, SHS (2020). New Therapeutic Horizons for Graves' Hyperthyroidism. *Endocrine reviews*, 41(6), 873–884.

Lane, LC, Rankin, J, & Cheetham, T (2021). A survey of the young person's experience of Graves' disease and its management. *Clinical endocrinology*, 94(2), 330–340.

Lane, LC, Kuś, A, Bednarczuk, T, Bossowski, A, Daroszewski, J, Jurecka-Lubieniecka, B, Cordell, HJ, Pearce, SHS, Cheetham, T, & Mitchell, AL (2020). An Intronic HCP5 Variant Is Associated With Age of Onset and Susceptibility to Graves' Disease in UK and Polish Cohorts. *The Journal of clinical endocrinology and metabolism*, 105(9), e3277–e3284.

Lane, LC, & Cheetham, T. (2020). Graves' disease: developments in first-line antithyroid drugs in the young. *Expert review of endocrinology & metabolism*, 15(2), 59–69.

Lane, LC, Allinson, KR, Campbell, K, Bhatnagar, I, Ingoe, L, Razvi, S, Cheetham, T, Cordell, HJ, Pearce, SH, & Mitchell, AL (2019). Analysis of BAFF gene polymorphisms in UK Graves' disease patients. *Clinical endocrinology*, 90(1), 170–174.

Oral presentations

‘Can systemic cytokines predict relapse of Graves' Disease?’ Presented at the British Endocrine Society meeting, November 2022, Harrogate, UK.

‘Can lymphocyte subsets and B cell cytokines predict clinical response to Rituximab in paediatric graves' disease?’. Presented at the British Society for Paediatric Endocrinology and Diabetes meeting, November 2022, Belfast, UK. (Awarded prize for Best Basic Abstract)

‘Chronic inflammation and relapse in Graves’ disease’. Presented at the British Society of Immunology Northeast symposium, June 2022, Newcastle-upon-Tyne, UK. (Awarded prize for Best short oral presentation)

‘Does chronic inflammation drive relapse in Graves’ disease?’. Presented at the British Thyroid Association meeting, May 2022, London, UK.

References

Abd El Azeem, R. A., Zedan, M. M., Saad, E. A., Mutawi, T. M., & Attia, Z. R. (2021). Single-nucleotide polymorphisms (SNPs) of antioxidant enzymes SOD2 and GSTP1 genes and SLE risk and severity in an Egyptian pediatric population. *Clinical biochemistry*, *88*, 37–42.

Abraham-Nordling, M., Törring, O., Hamberger, B., Lundell, G., Tallstedt, L., Calissendorff, J., & Wallin, G. (2005). Graves' disease: a long-term quality-of-life follow up of patients randomized to treatment with antithyroid drugs, radioiodine, or surgery. *Thyroid*, 15(11), 1279–1286.

Abraham, P., Avenell, A., McGeoch, S. C., Clark, L. F., & Bevan, J. S. (2010). Antithyroid drug regimen for treating Graves' hyperthyroidism. *The Cochrane database of systematic reviews*, 2010(1), CD003420.

Acosta-Rodriguez, E. V., Napolitani, G., Lanzavecchia, A., & Sallusto, F. (2007). Interleukins 1beta and 6 but not transforming growth factor-beta are essential for the differentiation of interleukin 17-producing human T helper cells. *Nature immunology*, *8*(9), 942–949.

Adlowitz, D. G., Barnard, J., Biear, J. N., Cistrone, C., Owen, T., Wang, W., Palanichamy, A., Ezealah, E., Campbell, D., Wei, C., Looney, R. J., Sanz, I., & Anolik, J. H. (2015). Expansion of Activated Peripheral Blood Memory B Cells in Rheumatoid Arthritis, Impact of B Cell Depletion Therapy, and Biomarkers of Response. *PloS one*, 10(6), e0128269.

Aggarwal, N., Tee, S. A., Saqib, W., Fretwell, T., Summerfield, G. P., & Razvi, S. (2016). Treatment of hyperthyroidism with antithyroid drugs corrects mild neutropenia in Graves' disease. *Clinical endocrinology*, 85(6), 949–953.

Alfaro, R., Llorente, S., Martinez, P., Jimenez-Coll, V., Martínez-Banaclocha, H., Galián, J. A., Botella, C., Moya-Quiles, M. R., de la Peña-Moral, J., Minguela, A., Legaz, I., & Muro, M. (2022). Monitoring of Soluble Forms of BAFF System (BAFF, APRIL, sR-BAFF, sTACI and sBCMA) in Kidney Transplantation. *Archivum immunologiae et therapiae experimentalis*, 70(1), 21.

Allahabadia, A., Daykin, J., Holder, R. L., Sheppard, M. C., Gough, S. C., & Franklyn, J. A. (2000). Age and gender predict the outcome of treatment for Graves'

hyperthyroidism. *The Journal of clinical endocrinology and metabolism*, 85(3), 1038–1042.

Álvarez-Sierra, D., Rodríguez-Grande, J., Gómez-Brey, A., Bello, I., Caubet, E., González, Ó., Zafón, C., Iglesias, C., Moreno, P., Ruiz, N., Marín-Sánchez, A., Colobran, R., & Pujol-Borrell, R. (2023). Single cell transcriptomic analysis of Graves' disease thyroid glands reveals the broad immunoregulatory potential of thyroid follicular and stromal cells and implies a major re-interpretation of the role of aberrant HLA class II expression in autoimmunity. *Journal of autoimmunity*, 139, 103072.

Anagnostis, P., Adamidou, F., Polyzos, S. A., Katergari, S., Karathanasi, E., Zouli, C., Panagiotou, A., & Kita, M. (2013). Predictors of long-term remission in patients with Graves' disease: a single center experience. *Endocrine*, 44(2), 448–453.

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, 31(2), 166–169.

Andersen, S. L., Olsen, J., Wu, C. S., & Laurberg, P. (2013). Birth defects after early pregnancy use of antithyroid drugs: a Danish nationwide study. *The Journal of clinical endocrinology and metabolism*, *98*(11), 4373–4381.

Andersen, M. N., Al-Karradi, S. N., Kragstrup, T. W., & Hokland, M. (2016). Elimination of erroneous results in flow cytometry caused by antibody binding to Fc receptors on human monocytes and macrophages. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, 89(11), 1001–1009.

Andrews, S. (2018). FastQC: A quality control tool for high throughput sequence data. Version 0.11.5. Available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (Accessed: 1st June 2024).

Arkatkar, T., Du, S. W., Jacobs, H. M., Dam, E. M., Hou, B., Buckner, J. H., Rawlings, D. J., & Jackson, S. W. (2017). B cell-derived IL-6 initiates spontaneous germinal center formation during systemic autoimmunity. *The Journal of experimental medicine*, 214(11), 3207–3217.

Atianand, M. K., & Fitzgerald, K. A. (2014). Long non-coding RNAs and control of gene expression in the immune system. *Trends in molecular medicine*, *20*(11), 623–631.

Aune, T. M., Crooke, P. S., 3rd, Patrick, A. E., Tossberg, J. T., Olsen, N. J., & Spurlock, C. F., 3rd (2017). Expression of long non-coding RNAs in autoimmunity and linkage to enhancer function and autoimmune disease risk genetic variants. *Journal of autoimmunity*, 81, 99–109.

Aust, G., Sittig, D., Becherer, L., Anderegg, U., Schütz, A., Lamesch, P., & Schmücking, E. (2004). The role of CXCR5 and its ligand CXCL13 in the compartmentalization of lymphocytes in thyroids affected by autoimmune thyroid diseases. *European journal of endocrinology*, 150(2), 225–234.

Autilio, C., Morelli, R., Locantore, P., Pontecorvi, A., Zuppi, C., & Carrozza, C. (2018). Stimulating TSH receptor autoantibodies immunoassay: analytical evaluation and clinical performance in Graves' disease. *Annals of clinical biochemistry*, 55(1), 172–177.

Autilio, C., Morelli, R., Locantore, P., Pontecorvi, A., Zuppi, C., & Carrozza, C. (2018). Stimulating TSH receptor autoantibodies immunoassay: analytical evaluation and clinical performance in Graves' disease. *Annals of clinical biochemistry*, 55(1), 172–177.

AXELROD, A. R., & BERMAN, L. (1951). The bone marrow in hyperthyroidism and hypothyroidism. *Blood*, 6(5), 436–453.

Badenhoop, K., Donner, H., Braun, J., Siegmund, T., Rau, H., & Usadel, K. H. (1996). Genetic markers in diagnosis and prediction of relapse in Graves' disease. *Experimental and clinical endocrinology & diabetes*, 104 Suppl 4, 98–100.

Badr, G., Borhis, G., Lefevre, E. A., Chaoul, N., Deshayes, F., Dessirier, V., Lapree, G., Tsapis, A., & Richard, Y. (2008). BAFF enhances chemotaxis of primary human B cells: a particular synergy between BAFF and CXCL13 on memory B cells. *Blood*, 111(5), 2744–2754.

Bahn, R. S., Dutton, C. M., Natt, N., Joba, W., Spitzweg, C., & Heufelder, A. E. (1998). Thyrotropin receptor expression in Graves' orbital adipose/connective tissues: potential autoantigen in Graves' ophthalmopathy. *The Journal of clinical endocrinology and metabolism*, 83(3), 998–1002.

Ballou, S. P., & Lozanski, G. (1992). Induction of inflammatory cytokine release from cultured human monocytes by C-reactive protein. *Cytokine*, *4*(5), 361–368.

Bankó, Z., Pozsgay, J., Szili, D., Tóth, M., Gáti, T., Nagy, G., Rojkovich, B., & Sármay, G. (2017). Induction and Differentiation of IL-10-Producing Regulatory B Cells from Healthy Blood Donors and Rheumatoid Arthritis Patients. *Journal of immunology (Baltimore, Md. : 1950)*, 198(4), 1512–1520.

Bano, A., Gan, E., Addison, C., Narayanan, K., Weaver, J. U., Tsatlidis, V., & Razvi, S. (2019). Age May Influence the Impact of TRAbs on Thyroid Function and Relapse-Risk in Patients With Graves Disease. *The Journal of clinical endocrinology and metabolism*, 104(5), 1378–1385.

Barbesino, G., & Tomer, Y. (2013). Clinical review: Clinical utility of TSH receptor antibodies. *The Journal of clinical endocrinology and metabolism*, 98(6), 2247–2255.

Bartalena, L., & Fatourechi, V. (2014). Extrathyroidal manifestations of Graves' disease: a 2014 update. *Journal of endocrinological investigation*, 37(8), 691–700.

Bartalena, L., Marcocci, C., Bogazzi, F., Manetti, L., Tanda, M. L., Dell'Unto, E., Bruno-Bossio, G., Nardi, M., Bartolomei, M. P., Lepri, A., Rossi, G., Martino, E., & Pinchera, A. (1998). Relation between therapy for hyperthyroidism and the course of Graves' ophthalmopathy. *The New England journal of medicine*, 338(2), 73–78.

Batten, M., Groom, J., Cachero, T. G., Qian, F., Schneider, P., Tschopp, J., Browning, J. L., & Mackay, F. (2000). BAFF mediates survival of peripheral immature B lymphocytes. *The Journal of experimental medicine*, *192*(10), 1453–1466.

Bauer A. J. (2011). Approach to the pediatric patient with Graves' disease: when is definitive therapy warranted?. *The Journal of clinical endocrinology and metabolism*, 96(3), 580–588.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.

Benson, M. J., Dillon, S. R., Castigli, E., Geha, R. S., Xu, S., Lam, K. P., & Noelle, R. J. (2008). Cutting edge: the dependence of plasma cells and independence of memory B cells on BAFF and APRIL. *Journal of immunology (Baltimore, Md. : 1950)*, *180*(6), 3655–3659.

Bossowski, A., Urban, M., & Stasiak-Barmuta, A. (2003). Analysis of changes in the percentage of B (CD19) and T (CD3) lymphocytes, subsets CD4, CD8 and their

memory (CD45RO), and naive (CD45RA) T cells in children with immune and non-immune thyroid diseases. *Journal of pediatric endocrinology & metabolism*, 16(1), 63–70.

Bracewell, C., Isaacs, J. D., Emery, P., & Ng, W. F. (2009). Atacicept, a novel B cell-targeting biological therapy for the treatment of rheumatoid arthritis. *Expert opinion on biological therapy*, *9*(7), 909–919.

Brandt, F., Almind, D., Christensen, K., Green, A., Brix, T. H., & Hegedüs, L. (2012). Excess mortality in hyperthyroidism: the influence of preexisting comorbidity and genetic confounding: a danish nationwide register-based cohort study of twins and singletons. *The Journal of clinical endocrinology and metabolism*, 97(11), 4123–4129.

Brandt, F., Thvilum, M., Hegedüs, L., & Brix, T. H. (2015). Hyperthyroidism is associated with work disability and loss of labour market income. A Danish register-based study in singletons and disease-discordant twin pairs. *European journal of endocrinology*, 173(5), 595–602.

Brix, T. H., Kyvik, K. O., Christensen, K., & Hegedüs, L. (2001). Evidence for a major role of heredity in Graves' disease: a population-based study of two Danish twin cohorts. *The Journal of clinical endocrinology and metabolism*, 86(2), 930–934.

Brokken, L. J., Scheenhart, J. W., Wiersinga, W. M., & Prummel, M. F. (2001). Suppression of serum TSH by Graves' Ig: evidence for a functional pituitary TSH receptor. *The Journal of clinical endocrinology and metabolism*, 86(10), 4814–4817.

Bugatti, S., Manzo, A., Vitolo, B., Benaglio, F., Binda, E., Scarabelli, M., Humby, F., Caporali, R., Pitzalis, C., & Montecucco, C. (2014). High expression levels of the B cell chemoattractant CXCL13 in rheumatoid synovium are a marker of severe disease. *Rheumatology (Oxford, England)*, 53(10), 1886–1895.

Buldygina YV, Zamotaieva, HA, Terekhova, GM, Stepura NM, Klochkova VM, Fed'ko TV. Role of cytokines, IL1-β and IL-10,in the development of endocrine ophthalmopathy in Graves' disease. (2021). *Journal of Ophthalmology (Ukraine).* Number 4 (501).

Cai, T. T., Muhali, F. S., Song, R. H., Qin, Q., Wang, X., Shi, L. F., Jiang, W. J., Xiao, L., Li, D. F., & Zhang, J. A. (2015). Genome-wide DNA methylation analysis in Graves' disease. *Genomics*, 105(4), 204–210.

Cameron, L., Webster, R. B., Strempel, J. M., Kiesler, P., Kabesch, M., Ramachandran, H., Yu, L., Stern, D. A., Graves, P. E., Lohman, I. C., Wright, A. L., Halonen, M., Klimecki, W. T., & Vercelli, D. (2006). Th2 cell-selective enhancement of human IL13 transcription by IL13-1112C>T, a polymorphism associated with allergic inflammation. *Journal of immunology (Baltimore, Md. : 1950)*, 177(12), 8633–8642.

Campi, I., Tosi, D., Rossi, S., Vannucchi, G., Covelli, D., Colombo, F., Trombetta, E., Porretti, L., Vicentini, L., Cantoni, G., Currò, N., Beck-Peccoz, P., Bulfamante, G., & Salvi, M. (2015). B Cell Activating Factor (BAFF) and BAFF Receptor Expression in Autoimmune and Nonautoimmune Thyroid Diseases. *Thyroid*, 25(9), 1043–1049.

Cao, Y., Zhao, X., You, R., Zhang, Y., Qu, C., Huang, Y., Yu, Y., Gong, Y., Cong, T., Zhao, E., Zhang, L., Gao, Y., & Zhang, J. (2022). CD11c+ B Cells Participate in the Pathogenesis of Graves' Disease by Secreting Thyroid Autoantibodies and Cytokines. *Frontiers in immunology*, 13, 836347.

Cappelli, C., Gandossi, E., Castellano, M., Pizzocaro, C., Agosti, B., Delbarba, A., Pirola, I., De Martino, E., & Rosei, E. A. (2007). Prognostic value of thyrotropin receptor antibodies (TRAb) in Graves' disease: a 120 months prospective study. *Endocrine journal*, 54(5), 713–720.

Carella, C., Mazziotti, G., Sorvillo, F., Piscopo, M., Cioffi, M., Pilla, P., Nersita, R., Iorio, S., Amato, G., Braverman, L. E., & Roti, E. (2006). Serum thyrotropin receptor antibodies concentrations in patients with Graves' disease before, at the end of methimazole treatment, and after drug withdrawal: evidence that the activity of thyrotropin receptor antibody and/or thyroid response modify during the observation period. *Thyroid*, *16*(3), 295–302.

Carsetti, R., Köhler, G., & Lamers, M. C. (1995). Transitional B cells are the target of negative selection in the B cell compartment. *The Journal of experimental medicine*, *181*(6), 2129–2140.

Carvalho, D. P., & Dupuy, C. (2017). Thyroid hormone biosynthesis and release. *Molecular and cellular endocrinology*, 458, 6–15.

Cassese, G., Arce, S., Hauser, A. E., Lehnert, K., Moewes, B., Mostarac, M., Muehlinghaus, G., Szyska, M., Radbruch, A., & Manz, R. A. (2003). Plasma cell survival is mediated by synergistic effects of cytokines and adhesion-dependent signals. *Journal of immunology (Baltimore, Md. : 1950)*, 171(4), 1684–1690.

Chaker, L., Cooper, D. S., Walsh, J. P., & Peeters, R. P. (2024). Hyperthyroidism. *Lancet (London, England),* 403(10428), 768–780.

Chazenbalk, G. D., Pichurin, P., Chen, C. R., Latrofa, F., Johnstone, A. P., McLachlan, S. M., & Rapoport, B. (2002). Thyroid-stimulating autoantibodies in Graves disease preferentially recognize the free A subunit, not the thyrotropin holoreceptor. *The Journal of clinical investigation*, 110(2), 209–217.

Cheetham, T. D., Cole, M., Abinun, M., Allahabadia, A., Barratt, T., Davies, J. H., Dimitri, P., Drake, A., Mohamed, Z., Murray, R. D., Steele, C. A., Zammitt, N., Carnell, S., Prichard, J., Watson, G., Hambleton, S., Matthews, J. N. S., & Pearce, S. H. S. (2022). Adjuvant Rituximab-Exploratory Trial in Young People With Graves Disease. *The Journal of clinical endocrinology and metabolism*, 107(3), 743–754.

Chen, G., Ding, Y., Li, Q., Li, Y., Wen, X., Ji, X., Bi, S., Chen, J., Xu, J., Chen, R., Ye, H., Wei, L., & Yang, H. (2019). Defective Regulatory B Cells Are Associated With Thyroid-Associated Ophthalmopathy. *The Journal of clinical endocrinology and metabolism*, *104*(9), 4067–4077.

Chen, R. H., Chen, W. C., Chang, C. T., Tsai, C. H., & Tsai, F. J. (2005). Interleukin-1-beta gene, but not the interleukin-1 receptor antagonist gene, is associated with Graves' disease. *Journal of clinical laboratory analysis*, *19*(4), 133–138.

Chen, X., Hu, Z., Liu, M., Li, H., Liang, C., Li, W., Bao, L., Chen, M., & Wu, G. (2018). Correlation between CTLA-4 and CD40 gene polymorphisms and their interaction in graves' disease in a Chinese Han population. *BMC medical genetics*, 19(1), 171.

Chen, Z., Liu, Y., Hu, S., Zhang, M., Shi, B., & Wang, Y. (2021). Decreased Treg Cell and TCR Expansion Are Involved in Long-Lasting Graves' Disease. *Frontiers in endocrinology*, *12*, 632492.

Cheng, C. W., Fang, W. F., Tang, K. T., & Lin, J. D. (2021). Possible interplay between estrogen and the BAFF may modify thyroid activity in Graves' disease. *Scientific reports*, 11(1), 21350.

Choo, Y. K., Yoo, W. S., Kim, D. W., & Chung, H. K. (2010). Hypothyroidism during antithyroid drug treatment with methimazole is a favorable prognostic indicator in patients with Graves' disease. *Thyroid*, 20(9), 949–954.

Chu, X., Yang, M., Song, Z. J., Dong, Y., Li, C., Shen, M., Zhu, Y. Q., Song, H. D., Chen, S. J., Chen, Z., & Huang, W. (2018). Fine mapping MHC associations in Graves' disease and its clinical subtypes in Han Chinese. *Journal of medical genetics*, 55(10), 685–692.

Codaccioni, J. L., Orgiazzi, J., Blanc, P., Pugeat, M., Roulier, R., & Carayon, P. (1988). Lasting remissions in patients treated for Graves' hyperthyroidism with propranolol alone: a pattern of spontaneous evolution of the disease. *The Journal of clinical endocrinology and metabolism*, 67(4), 656–662.

Coenye T. (2021). Do results obtained with RNA-sequencing require independent verification?. *Biofilm*, *3*, 100043.

Conrad, N., Misra, S., Verbakel, J. Y., Verbeke, G., Molenberghs, G., Taylor, P. N., Mason, J., Sattar, N., McMurray, J. J. V., McInnes, I. B., Khunti, K., & Cambridge, G. (2023). Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. *Lancet (London, England)*, 401(10391), 1878–1890.

Corrales, J. J., Orfao, A., López, A., Ciudad, J., & Mories, M. T. (1996). Serial analysis of the effects of methimazole therapy on circulating B cell subsets in Graves' disease. *The Journal of endocrinology*, 151(2), 231–240.

Creative Biobanks Antibody. Sandwich ELISA technique. Available at https://www.antibody-creativebiolabs.com/sandwich-elisa-with-streptavidin-biotin-detection.htm. (Accessed:12 December 2023).

Dales, L. G., & Ury, H. K. (1978). An improper use of statistical significance testing in studying covariables. *International journal of epidemiology*, *7*(4), 373–375.

Dasouki, M., Jabr, A., AlDakheel, G., Elbadaoui, F., Alazami, A. M., Al-Saud, B., Arnaout, R., Aldhekri, H., Alotaibi, I., Al-Mousa, H., & Hawwari, A. (2020). TREC and KREC profiling as a representative of thymus and bone marrow output in patients with various inborn errors of immunity. *Clinical and experimental immunology*, 202(1), 60–71.

de Vries J. E. (1998). The role of IL-13 and its receptor in allergy and inflammatory responses. *The Journal of allergy and clinical immunology*, *102*(2), 165–169.

Defrance, T., Carayon, P., Billian, G., Guillemot, J. C., Minty, A., Caput, D., & Ferrara, P. (1994). Interleukin 13 is a B cell stimulating factor. *The Journal of experimental medicine*, *179*(1), 135–143.

de Gruijter, N. M., Jebson, B., & Rosser, E. C. (2022). Cytokine production by human B cells: role in health and autoimmune disease. *Clinical and experimental immunology*, 210(3), 253–262.

Department of Health and Social Care. North-East smoking prevalence (2023). https://fingertips.phe.org.uk/static-reports/health-trends-in-england/North_East/smoking.html (Accessed 12.03.2025).

Diana, T., Daiber, A., Oelze, M., Neumann, S., Olivo, P. D., Kanitz, M., Stamm, P., & Kahaly, G. J. (2018). Stimulatory TSH-Receptor Antibodies and Oxidative Stress in Graves Disease. *The Journal of clinical endocrinology and metabolism*, *103*(10), 3668–3677.

Donovan, P. J., McLeod, D. S., Little, R., & Gordon, L. (2016). Cost-utility analysis comparing radioactive iodine, anti-thyroid drugs and total thyroidectomy for primary treatment of Graves' disease. *European journal of endocrinology*, 175(6), 595–603.

Dudina M, Savchenko A, Dogadin SA, Gvozdev I. (2022). The level of neutrophil reactive oxygen species in euthyroid and relapse patients with Graves’ hyperthyroidism. *Endocrine Abstracts* 81 EP954.

Eckstein, A. K., Plicht, M., Lax, H., Neuhäuser, M., Mann, K., Lederbogen, S., Heckmann, C., Esser, J., & Morgenthaler, N. G. (2006). Thyrotropin receptor autoantibodies are independent risk factors for Graves' ophthalmopathy and help to predict severity and outcome of the disease. *The Journal of clinical endocrinology and metabolism*, 91(9), 3464–3470.

El Fassi, D., Nielsen, C. H., Bonnema, S. J., Hasselbalch, H. C., & Hegedüs, L. (2007). B lymphocyte depletion with the monoclonal antibody rituximab in Graves' disease: a controlled pilot study. *The Journal of clinical endocrinology and metabolism*, 92(5), 1769–1772.

Fairfax, B. P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., Langford, C., Vannberg, F. O., & Knight, J. C. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nature genetics*, *44*(5), 502–510.

Fallahi, P., Ferrari, S. M., Elia, G., Ragusa, F., Paparo, S. R., Patrizio, A., Camastra, S., Miccoli, M., Cavallini, G., Benvenga, S., & Antonelli, A. (2021). Cytokines as Targets of Novel Therapies for Graves' Ophthalmopathy. *Frontiers in endocrinology*, 12, 654473.

Favero V, Currò N, Campi I, Lazzaroni E, Covelli D, Vannucchi G, Dazzi D, Muller I, Minorini V, Guastella C, Arosio M & Salvi M (2021) Efficacy of the Anti-BAFF monoclonal antibody belimumab vs methylprednisolone in active moderate-severe graves' orbitopathy: Preliminary analysis of a randomized controlled trial. *Endocrine Abstracts*, 73 YI2

Feng, R., Zhao, J., Sun, F., Miao, M., Sun, X., He, J., & Li, Z. (2022). Comparison of the deep immune profiling of B cell subsets between healthy adults and Sjögren's syndrome. *Annals of medicine*, 54(1), 472–483.

Festa, R. A., & Thiele, D. J. (2012). Copper at the front line of the host-pathogen battle. *PLoS pathogens*, *8*(9), e1002887.

Foster, M. P., Montecino-Rodriguez, E., & Dorshkind, K. (1999). Proliferation of bone marrow pro-B cells is dependent on stimulation by the pituitary/thyroid axis. *Journal of immunology (Baltimore, Md. : 1950)*, 163(11), 5883–5890.

Francisco, V., Tovar, S., Conde, J., Pino, J., Mera, A., Lago, F., González-Gay, M. A., Dieguez, C., & Gualillo, O. (2020). Levels of the Novel Endogenous Antagonist of Ghrelin Receptor, Liver-Enriched Antimicrobial Peptide-2, in Patients with Rheumatoid Arthritis. *Nutrients*, *12*(4), 1006.

Franklyn, J. A., & Boelaert, K. (2012). Thyrotoxicosis. *Lancet (London, England)*, 379(9821), 1155–1166.

Frasca, D., Diaz, A., Romero, M., & Blomberg, B. B. (2017). Human peripheral late/exhausted memory B cells express a senescent-associated secretory phenotype and preferentially utilize metabolic signaling pathways. *Experimental gerontology,* 87(Pt A), 113–120.

Fraussen, J., Marquez, S., Takata, K., Beckers, L., Montes Diaz, G., Zografou, C., Van Wijmeersch, B., Villar, L. M., O'Connor, K. C., Kleinstein, S. H., & Somers, V. (2019). Phenotypic and Ig Repertoire Analyses Indicate a Common Origin of IgD^{-} CD27^{-}Double Negative B Cells in Healthy Individuals and Multiple Sclerosis Patients. *Journal of immunology (Baltimore, Md. : 1950)*, *203*(6), 1650–1664.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software*, 33(1), 1–22.

Fu, Q., Zhao, J., Qian, X., Wong, J. L., Kaufman, K. M., Yu, C. Y., Hwee Siew Howe, Tan Tock Seng Hospital Lupus Study Group, Mok, M. Y., Harley, J. B., Guthridge, J. M., Song, Y. W., Cho, S. K., Bae, S. C., Grossman, J. M., Hahn, B. H., Arnett, F. C., Shen, N., & Tsao, B. P. (2011). Association of a functional IRF7 variant with systemic lupus erythematosus. *Arthritis and rheumatism*, 63(3), 749–754.

Furmaniak, J., Sanders, J., Sanders, P., Li, Y., & Rees Smith, B. (2022). TSH receptor specific monoclonal autoantibody K1-70TM targeting of the TSH receptor in subjects with Graves' disease and Graves' orbitopathy-Results from a phase I clinical trial. *Clinical endocrinology*, 96(6), 878–887.

Fyfe-Desmarais, G., Desmarais, F., Rassart, É., & Mounier, C. (2023). Apolipoprotein D in Oxidative Stress and Inflammation. *Antioxidants (Basel, Switzerland)*, *12*(5), 1027.

Gal-Oz, S. T., Maier, B., Yoshida, H., Seddu, K., Elbaz, N., Czysz, C., Zuk, O., Stranger, B. E., Ner-Gaon, H., & Shay, T. (2019). ImmGen report: sexual dimorphism in the immune system transcriptome. *Nature communications*, 10(1), 4295.

Gallego Romero, I., Pai, A. A., Tung, J., & Gilad, Y. (2014). RNA-seq: impact of RNA degradation on transcript quantification. *BMC biology*, 12, 42.

Ganser, G. H., & Hewett, P. (2010). An accurate substitution method for analyzing censored data. *Journal of occupational and environmental hygiene*, 7(4), 233–244.

Ge, X., Yang, H., Bednarek, M. A., Galon-Tilleman, H., Chen, P., Chen, M., Lichtman, J. S., Wang, Y., Dalmas, O., Yin, Y., Tian, H., Jermutus, L., Grimsby, J., Rondinone, C. M., Konkar, A., & Kaplan, D. D. (2018). LEAP2 Is an Endogenous Antagonist of the Ghrelin Receptor. *Cell metabolism*, *27*(2), 461–469.e6.

Ge, Y., Huang, M., & Yao, Y. M. (2021). The Effect and Regulatory Mechanism of High Mobility Group Box-1 Protein on Immune Cells in Inflammatory Diseases. *Cells*, *10*(5), 1044.

Gianoukakis, A. G., Khadavi, N., & Smith, T. J. (2008). Cytokines, Graves' disease, and thyroid-associated ophthalmopathy. *Thyroid*, 18(9), 953–958.

Gilbert, J. A., Kalled, S. L., Moorhead, J., Hess, D. M., Rennert, P., Li, Z., Khan, M. Z., & Banga, J. P. (2006). Treatment of autoimmune hyperthyroidism in a murine model of Graves' disease with tumor necrosis factor-family ligand inhibitors suggests a key role for B cell activating factor in disease pathology. *Endocrinology*, 147(10), 4561–4568.

Giltiay, N. V., Chappell, C. P., & Clark, E. A. (2012). B-cell selection and the development of autoantibodies. *Arthritis research & therapy*, *14 Suppl 4*(Suppl 4), S1.

Glaser, N. S., Styne, D. M., & Organization of Pediatric Endocrinologists of Northern California Collaborative Graves' Disease Study Group (2008). Predicting the likelihood of remission in children with Graves' disease: a prospective, multicenter study. *Pediatrics*, 121(3), e481–e488.

Glowacka, D., Loesch, C., Johnson, K. T., Mann, K., Esser, J., Morgenthaler, N. G., Siffert, W., Schmid, K. W., & Eckstein, A. K. (2009). The T393C polymorphism of the Galphas gene (GNAS1) is associated with the course of Graves' disease. *Hormone and metabolic research*, 41(6), 430–435.

Grebe, S. K., Feek, C. M., Ford, H. C., Fagerström, J. N., Cordwell, D. P., Delahunt, J. W., & Toomath, R. J. (1998). A randomized trial of short-term treatment of Graves' disease with high-dose carbimazole plus thyroxine versus low-dose carbimazole. *Clinical endocrinology*, 48(5), 585–592.

Grixti, L., Lane, L. C., & Pearce, S. H. (2024). The genetics of Graves' disease. *Reviews in endocrine & metabolic disorders*, 25(1), 203–214.

Grubeck-Loebenstein, B., Buchan, G., Chantry, D., Kassal, H., Londei, M., Pirich, K., Barrett, K., Turner, M., Waldhausl, W., & Feldmann, M. (1989). Analysis of intrathyroidal cytokine production in thyroid autoimmune disease: thyroid follicular

cells produce interleukin-1 alpha and interleukin-6. *Clinical and experimental immunology*, 77(3), 324–330.

Grubczak, K., Starosz, A., Stożek, K., Bossowski, F., Moniuszko, M., & Bossowski, A. (2021). Regulatory B Cells Involvement in Autoimmune Phenomena Occurring in Pediatric Graves' Disease Patients. *International journal of molecular sciences*, 22(20), 10926.

Guerra, L. N., Moiguer, S., Karner, M., de Molina, M. C., Sreider, C. M., & Burdman, J. A. (2001). Antioxidants in the treatment of Graves disease. *IUBMB life*, *51*(2), 105–109.

Gümüş, P., Nizam, N., Lappin, D. F., & Buduneli, N. (2014). Saliva and serum levels of B-cell activating factors and tumor necrosis factor-α in patients with periodontitis. *Journal of periodontology*, 85(2), 270–280.

Hadlow, N. C., Rothacker, K. M., Wardrop, R., Brown, S. J., Lim, E. M., & Walsh, J. P. (2013). The relationship between TSH and free T_4 in a large population is complex and nonlinear and differs by age and sex. *The Journal of clinical endocrinology and metabolism*, 98(7), 2936–2943.

Hamidi, S., Chen, C. R., McLachlan, S. M., & Rapoport, B. (2011). Insight into thyroid-stimulating autoantibody interaction with the thyrotropin receptor N-terminus based on mutagenesis and re-evaluation of ambiguity in this region of the receptor crystal structure. *Thyroid*, 21(9), 1013–1020.

Hart, S. N., Therneau, T. M., Zhang, Y., Poland, G. A., & Kocher, J. P. (2013). Calculating sample size estimates for RNA sequencing data. *Journal of computational biology: a journal of computational molecular cell biology*, 20(12), 970–978.

Hastie, T., Tibshirani, T., & Tibshirani, R. (2020). Best Subset, Forward Stepwise or Lasso? Analysis and Recommendations Based on Extensive Comparisons. *Statistical Science*. 35(4), 579–592.

Hayashi, F., Watanabe, M., Nanba, T., Inoue, N., Akamizu, T., & Iwatani, Y. (2009). Association of the -31C/T functional polymorphism in the interleukin-1beta gene with the intractability of Graves' disease and the proportion of T helper type 17 cells. *Clinical and experimental immunology*, *158*(3), 281–286.

Hayashi, M., Kouki, T., Takasu, N., Sunagawa, S., & Komiya, I. (2008). Association of an A/C single nucleotide polymorphism in programmed cell death-ligand 1 gene with Graves' disease in Japanese patients. *European journal of endocrinology*, 158(6), 817–822.

Hegele, R. A., & Volpé, R. (1985). Relapse of Graves' disease 23 years after treatment with radioactive iodine (131I). *Journal of clinical & laboratory immunology*, 18(2), 103–105.

Heidari, Z., Salimi, S., Rokni, M., Rezaei, M., Khalafi, N., Shahroudi, M. J., Dehghan, A., & Saravani, M. (2021). Association of IL-1*β*, NLRP3, and COX-2 Gene Polymorphisms with Autoimmune Thyroid Disease Risk and Clinical Features in the Iranian Population. *BioMed research international*, *2021*, 7729238.

Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., Li, Y., Sarwar, R., Langley, S. R., Bauerfeind, A., Hummel, O., Lee, Y. A., Paskas, S., Rintisch, C., Saar, K., Cooper, J., Buchan, R., Gray, E. E., Cyster, J. G., Cardiogenics Consortium, … Cook, S. A. (2010). A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature*, 467(7314), 460–464.

Hirano T. (2021). IL-6 in inflammation, autoimmunity and cancer. *International immunology*, 33(3), 127–148.

Hirai, N., Watanabe, M., Inoue, N., Kinoshita, R., Ohtani, H., Hidaka, Y., & Iwatani, Y. (2019). Association of IL6 gene methylation in peripheral blood cells with the development and prognosis of autoimmune thyroid diseases. *Autoimmunity*, 52(7-8), 251–255.

Hiromatsu, Y., Fukutani, T., Ichimura, M., Mukai, T., Kaku, H., Nakayama, H., Miyake, I., Shoji, S., Koda, Y., & Bednarczuk, T. (2005). Interleukin-13 gene polymorphisms confer the susceptibility of Japanese populations to Graves' disease. *The Journal of clinical endocrinology and metabolism*, *90*(1), 296–301.

Hoffmann, F. S., Kuhn, P. H., Laurent, S. A., Hauck, S. M., Berer, K., Wendlinger, S. A., Krumbholz, M., Khademi, M., Olsson, T., Dreyling, M., Pfister, H. W., Alexander, T., Hiepe, F., Kümpfel, T., Crawford, H. C., Wekerle, H., Hohlfeld, R., Lichtenthaler, S. F., & Meinl, E. (2015). The immunoregulator soluble TACI is released by ADAM10 and reflects B cell activation in autoimmunity. *Journal of immunology (Baltimore, Md. : 1950)*, 194(2), 542–552.

Honda, K., Yanai, H., Negishi, H., Asagiri, M., Sato, M., Mizutani, T., Shimada, N., Ohba, Y., Takaoka, A., Yoshida, N., & Taniguchi, T. (2005). IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature*, *434*(7034), 772–777.

Hornung, R. W., & Reed, L. D. (1990). Estimation of Average Concentration in the Presence of Nondetectable Values. *Applied Occupational and Environmental Hygiene*, 5(1), 46–51.

Huang, M. J., & Liaw, Y. F. (1995). Clinical associations between thyroid and liver diseases. *Journal of gastroenterology and hepatology*, 10(3), 344–350.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oleś, A. K., Pagès, H., … Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods*, 12(2), 115–121.

Hussain, Y. S., Hookham, J. C., Allahabadia, A., & Balasubramanian, S. P. (2017). Epidemiology, management and outcomes of Graves' disease-real life data. *Endocrine*, 56(3), 568–578.

Inoue, N., Watanabe, M., Morita, M., Tatusmi, K., Hidaka, Y., Akamizu, T., & Iwatani, Y. (2011). Association of functional polymorphisms in promoter regions of IL5, IL6 and IL13 genes with development and prognosis of autoimmune thyroid diseases. *Clinical and experimental immunology*, *163*(3), 318–323.

Iwata, Y., Matsushita, T., Horikawa, M., Dilillo, D. J., Yanaba, K., Venturi, G. M., Szabolcs, P. M., Bernstein, S. H., Magro, C. M., Williams, A. D., Hall, R. P., St Clair, E. W., & Tedder, T. F. (2011). Characterization of a rare IL-10-competent B-cell subset in humans that parallels mouse regulatory B10 cells. *Blood*, 117(2), 530–541.

Iwatani, Y., Amino, N., Kaneda, T., Ichihara, K., Tamaki, H., Tachi, J., Matsuzuka, F., Fukata, S., Kuma, K., & Miyai, K. (1989). Marked increase of CD5 + B cells in hyperthyroid Graves' disease. *Clinical and experimental immunology*, 78(2), 196–200.

Jacobs, M., Verschraegen, S., Salhi, B., Anckaert, J., Mestdagh, P., Brusselle, G. G., & Bracke, K. R. (2022). IL-10 producing regulatory B cells are decreased in blood from smokers and COPD patients. *Respiratory research*, 23(1), 287.

Jaeger, M., Sloot, Y. J. E., Horst, R. T., Chu, X., Koenen, H. J. P. M., Koeken, V. A. C. M., Moorlag, S. J. C. F. M., de Bree, C. J., Mourits, V. P., Lemmers, H., Dijkstra, H., Medici, M., van Herwaarden, A. E., Joosten, I., Joosten, L. A. B., Li, Y., Smit, J. W. A., Netea, M. G., & Netea-Maier, R. T. (2021). Thyrotrophin and thyroxine support immune homeostasis in humans. *Immunology,* 163(2), 155–168.

Janko, C., Filipović, M., Munoz, L. E., Schorn, C., Schett, G., Ivanović-Burmazović, I., & Herrmann, M. (2014). Redox modulation of HMGB1-related signaling. *Antioxidants & redox signaling*, *20*(7), 1075–1085.

Jansen, H. I., Heuveling van Beek, C., Bisschop, P. H., Heijboer, A. C., Bruinstroop, E., & Boelen, A. (2024). The need for the GREAT+ score to predict relapse in Graves' disease: a questionnaire among patients and internal medicine specialists. *Journal of endocrinological investigation*, 10.1007/s40618-024-02358-7.

Jansen, K., Cevhertas, L., Ma, S., Satitsuksanoa, P., Akdis, M., & van de Veen, W. (2021). Regulatory B cells, *A to Z. Allergy*, 76(9), 2699–2715.

Jenks, S. A., Cashman, K. S., Zumaquero, E., Marigorta, U. M., Patel, A. V., Wang, X., Tomar, D., Woodruff, M. C., Simon, Z., Bugrovsky, R., Blalock, E. L., Scharer, C. D., Tipton, C. M., Wei, C., Lim, S. S., Petri, M., Niewold, T. B., Anolik, J. H., Gibson, G., Lee, F. E., … Sanz, I. (2018). Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity*, 49(4), 725–739.e6.

Jenks, S. A., Cashman, K. S., Woodruff, M. C., Lee, F. E., & Sanz, I. (2019). Extrafollicular responses in humans and SLE. *Immunological reviews*, *288*(1), 136–148.

Jiang, X., Wang, Y., Li, X., He, L., Yang, Q., Wang, W., Liu, J., & Zha, B. (2020). Microarray profile of B cells from Graves' disease patients reveals biomarkers of proliferation. *Endocrine connections*, *9*(5), 405–417.

Kaguelidou, F., Alberti, C., Castanet, M., Guitteny, M. A., Czernichow, P., Léger, J., & French Childhood Graves' Disease Study Group (2008). Predictors of autoimmune

hyperthyroidism relapse in children after discontinuation of antithyroid drug treatment. *The Journal of clinical endocrinology and metabolism*, 93(10), 3817–3826.

Kahaly, G. J., Bartalena, L., Hegedüs, L., Leenhardt, L., Poppe, K., & Pearce, S. H. (2018). 2018 European Thyroid Association Guideline for the Management of Graves' Hyperthyroidism. *European thyroid journal*, 7(4), 167–186.

Kalampokis, I., Yoshizaki, A., & Tedder, T. F. (2013). IL-10-producing regulatory B cells (B10 cells) in autoimmune disease. *Arthritis research & therapy*, *15 Suppl 1*(Suppl 1), S1.

Kaneko, N., Kurata, M., Yamamoto, T., Morikawa, S., & Masumoto, J. (2019). The role of interleukin-1 in general pathology. *Inflammation and regeneration*, *39*, 12.

Kany, S., Vollrath, J. T., & Relja, B. (2019). Cytokines in Inflammatory Disease. *International journal of molecular sciences*, 20(23), 6008.

Kardalas E, Maraka S, Papagianni M, Paltoglou G, Siristatidis C, Mastorakos G. TGF-β Physiology as a Novel Therapeutic Target Regarding Autoimmune Thyroid Diseases: Where Do We Stand and What to Expect. *Medicina (Kaunas*). 2021 Jun 14;57(6):621.

Karim, M. R., & Wang, Y. F. (2019). Phenotypic identification of CD19+CD5+CD1d+ regulatory B cells that produce interleukin 10 and transforming growth factor β1 in human peripheral blood. *Archives of medical science: AMS*, 15(5), 1176–1183.

Kashiwai, T., Hidaka, Y., Takano, T., Tatsumi, K. I., Izumi, Y., Shimaoka, Y., Tada, H., Takeoka, K., & Amino, N. (2003). Practical treatment with minimum maintenance dose of anti-thyroid drugs for prediction of remission in Graves' disease. *Endocrine journal*, 50(1), 45–49.

Kashiwakura, Y., Ishiwata, A., Takano, K., Ohmori, T., & Sakata, Y. (2006). Severe secondary deficiency of von Willebrand factor-cleaving protease (ADAMTS13) in patients with sepsis-induced disseminated intravascular coagulation: its correlation with development of renal failure. *Blood*, *107*(2), 528–534.

Khadzhieva, M. B., Kalinina, E. V., Larin, S. S., Sviridova, D. A., Gracheva, A. S., Chursinova, J. V., Stepanov, V. A., Redkin, I. V., Avdeikina, L. S., Rumyantsev, A. G., Kuzovlev, A. N., & Salnikova, L. E. (2021). TREC/KREC Levels in Young COVID-19 Patients. *Diagnostics (Basel, Switzerland)*, 11(8), 1486.

Khamisi, S., Lundqvist, M., Rasmusson, A. J., Engström, B. E., Karlsson, F. A., & Ljunggren, Ö. (2023). Vitamin D and bone metabolism in Graves' disease: a prospective study. *Journal of endocrinological investigation*, 46(2), 425–433.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods*, 12(4), 357–360.

Kim, K. W., Park, Y. J., Kim, T. Y., Park, D. J., Park, K. S., & Cho, B. Y. (2007). Susceptible alleles of the CD40 and CTLA-4 genes are not associated with the relapse after antithyroid withdrawal in Graves' disease. *Thyroid*, 17(12), 1229–1234.

Kimball, L. E., Kulinskaya, E., Brown, B., Johnston, C., & Farid, N. R. (2002). Does smoking increase relapse rates in Graves' disease?. *Journal of endocrinological investigation*, 25(2), 152–157.

Klein, S. L., & Flanagan, K. L. (2016). Sex differences in immune responses. Nature reviews. *Immunology*, 16(10), 626–638.

Kokuina, E., Breff-Fonseca, M. C., Villegas-Valverde, C. A., & Mora-Díaz, I. (2019). Normal Values of T, B and NK Lymphocyte Subpopulations in Peripheral Blood of Healthy Cuban Adults. *MEDICC review*, 21(2-3), 16–21.

Komiya, I., Yamada, T., Sato, A., Kouki, T., Nishimori, T., & Takasu, N. (2001). Remission and recurrence of hyperthyroid Graves' disease during and after methimazole treatment when assessed by IgE and interleukin 13. *The Journal of clinical endocrinology and metabolism*, *86*(8), 3540–3544.

Komosinska-Vassev, K., Olczyk, K., Kucharz, E. J., Marcisz, C., Winsz-Szczotka, K., & Kotulska, A. (2000). Free radical activity and antioxidant defense mechanisms in patients with hyperthyroidism due to Graves' disease during therapy. *Clinica chimica acta*, 300(1-2), 107–117.

Krämer, A., Green, J., Pollard, J., Jr, & Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics (Oxford, England)*, 30(4), 523–530.

Kristensen, B., Hegedüs, L., Lundy, S. K., Brimnes, M. K., Smith, T. J., & Nielsen, C. H. (2015). Characterization of Regulatory B Cells in Graves' Disease and Hashimoto's Thyroiditis. *PloS one*, *10*(5), e0127949.

Kuiper, G. G., Kester, M. H., Peeters, R. P., & Visser, T. J. (2005). Biochemical mechanisms of thyroid hormone deiodination. *Thyroid*, 15(8), 787–798.

Kuś, A., Radziszewski, M., Glina, A., Szymański, K., Jurecka-Lubieniecka, B., Pawlak-Adamska, E., Kula, D., Wawrusiewicz-Kurylonek, N., Kuś, J., Miśkiewicz, P., Płoski, R., Bolanowski, M., Daroszewski, J., Jarząb, B., Bossowski, A., & Bednarczuk, T. (2019). Paediatric-onset and adult-onset Graves' disease share multiple genetic risk factors. *Clinical endocrinology*, 90(2), 320–327.

Kverneland, A. H., Streitz, M., Geissler, E., Hutchinson, J., Vogt, K., Boës, D., Niemann, N., Pedersen, A. E., Schlickeiser, S., & Sawitzki, B. (2016). Age and gender leucocytes variances and references values generated using the standardized ONE-Study protocol. Cytometry. *Part A : the journal of the International Society for Analytical Cytology*, 89(6), 543–564.

Landolt-Marticorena, C., Wither, R., Reich, H., Herzenberg, A., Scholey, J., Gladman, D. D., Urowitz, M. B., Fortin, P. R., & Wither, J. (2011). Increased expression of B cell activation factor supports the abnormal expansion of transitional B cells in systemic lupus erythematosus. *The Journal of rheumatology*, *38*(4), 642–651.

Lane, L. C., Rankin, J., & Cheetham, T. (2021). A survey of the young person's experience of Graves' disease and its management. *Clinical endocrinology*, 94(2), 330–340.

Lane, L. C., Allinson, K. R., Campbell, K., Bhatnagar, I., Ingoe, L., Razvi, S., Cheetham, T., Cordell, H. J., Pearce, S. H., & Mitchell, A. L. (2019). Analysis of BAFF gene polymorphisms in UK Graves' disease patients. *Clinical endocrinology*, 90(1), 170–174.

Lane, L. C., Kuś, A., Bednarczuk, T., Bossowski, A., Daroszewski, J., Jurecka-Lubieniecka, B., Cordell, H. J., Pearce, S. H. S., Cheetham, T., & Mitchell, A. L. (2020). An Intronic HCP5 Variant Is Associated With Age of Onset and Susceptibility to Graves Disease in UK and Polish Cohorts. *The Journal of clinical endocrinology and metabolism*, 105(9), e3277–e3284.

Lane, L. C., Cheetham, T. D., Perros, P., & Pearce, S. H. S. (2020). New Therapeutic Horizons for Graves' Hyperthyroidism. *Endocrine reviews*, 41(6), 873–884.

Lane, L. C., Cheetham, T. D., Razvi, S., Allinson, K., & Pearce, S. H. S. (2023). Expansion of the immature B lymphocyte compartment in Graves' disease. *European journal of endocrinology*, 189(2), 208–216.

Laurent, S. A., Hoffmann, F. S., Kuhn, P. H., Cheng, Q., Chu, Y., Schmidt-Supprian, M., Hauck, S. M., Schuh, E., Krumbholz, M., Rübsamen, H., Wanngren, J., Khademi, M., Olsson, T., Alexander, T., Hiepe, F., Pfister, H. W., Weber, F., Jenne, D., Wekerle, H., Hohlfeld, R., … Meinl, E. (2015). γ-Secretase directly sheds the survival receptor BCMA from plasma cells. *Nature communications*, 6, 7333.

Lasigliè, D., Traggiai, E., Federici, S., Alessio, M., Buoncompagni, A., Accogli, A., Chiesa, S., Penco, F., Martini, A., & Gattorno, M. (2011). Role of IL-1 beta in the development of human T(H)17 cells: lesson from NLPR3 mutated patients. *PloS one*, *6*(5), e20014.

Lasko, T. A., Bhagwat, J. G., Zou, K. H., & Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of biomedical informatics*, 38(5), 404–415.

Lee, S., Ko, Y., & Kim, T. J. (2020). Homeostasis and regulation of autoreactive B cells. *Cellular & molecular immunology*, 17(6), 561–569.

Lee, H. J., Lombardi, A., Stefan, M., Li, C. W., Inabnet, W. B., 3rd, Owen, R. P., Concepcion, E., & Tomer, Y. (2017). CD40 Signaling in Graves Disease Is Mediated Through Canonical and Noncanonical Thyroidal Nuclear Factor κB Activation. *Endocrinology*, 158(2), 410–418.

Léger, J., & Carel, J. C. (2017). MANAGEMENT OF ENDOCRINE DISEASE: Arguments for the prolonged use of antithyroid drugs in children with Graves' disease. *European journal of endocrinology*, 177(2), R59–R67.

Léger, J., Gelwane, G., Kaguelidou, F., Benmerad, M., Alberti, C., & French Childhood Graves' Disease Study Group (2012). Positive impact of long-term antithyroid drug treatment on the outcome of children with Graves' disease: national long-term cohort study. *The Journal of clinical endocrinology and metabolism*, 97(1), 110–119.

Lepez, T., Vandewoestyne, M., & Deforce, D. (2012). Fetal microchimeric cells in blood and thyroid glands of women with an autoimmune thyroid disease. *Chimerism*, 3(1), 21–23.

Levy-Mendelovich, S., Lev, A., Aviner, S., Rosenberg, N., Kaplinsky, C., Sharon, N., Miskin, H., Dvir, A., Kenet, G., Schushan, I. E., & Somech, R. (2017). Quantification of specific T and B cells immunological markers in children with chronic and transient ITP. *Pediatric blood & cancer*, 64(12), 10.1002/pbc.26646.

Lhotta, K., Zitt, E., Sprenger-Mähr, H., Loacker, L., & Becherer, A. (2018). Treatment of Concurrent Thrombotic Thrombocytopenic Purpura and Graves' Disease: A Report on Two Cases. *Case reports in endocrinology*, *2018*, 5747969.

Li, C., Peng, S., Liu, X., Han, C., Wang, X., Jin, T., Liu, S., Wang, W., Xie, X., He, X., Zhang, H., Shan, L., Fan, C., Shan, Z., & Teng, W. (2017). Glycyrrhizin, a Direct HMGB1 Antagonist, Ameliorates Inflammatory Infiltration in a Model of Autoimmune Thyroiditis via Inhibition of TLR2-HMGB1 Signaling. *Thyroid*, *27*(5), 722–731.

Li, C. W., Sachidanandam, R., Jayaprakash, A., Yi, Z., Zhang, W., Stefan-Lifshitz, M., Concepcion, E., & Tomer, Y. (2021). Identification of New Rare Variants Associated With Familial Autoimmune Thyroid Diseases by Deep Sequencing of Linked Loci. *The Journal of clinical endocrinology and metabolism*, *106*(11), e4680–e4687.

Li, L., Ding, X., Wang, X., Yao, Q., Shao, X., An, X., Yan, N., Jiang, Y., Wang, W., Shi, L., Qin, Q., Song, R., Zhang, J. A., & Sun, P. (2018). Polymorphisms of IKZF3 Gene and Autoimmune Thyroid Diseases: Associated with Graves' Disease but Not with Hashimoto's Thyroiditis. *Cellular physiology and biochemistry: international journal of experimental cellular physiology, biochemistry, and pharmacology*, 45(5), 1787–1796.

Li, J., Liu, Z., Zhang, P., Lin, W., Lu, H., Peng, Y., Peng, L., Zhou, J., Wang, M., Chen, H., Zhao, L., Wang, L., Qin, C., Hu, C., Zeng, X., Zhao, Y., Fei, Y., & Zhang, W. (2021). Peripheral B-Cell Immunophenotyping Identifies Heterogeneity in IgG4-Related Disease. *Frontiers in immunology*, 12, 747076.

Li, J. R., Hong, F. Y., Zeng, J. Y., & Huang, G. L. (2013). Functional interleukin-17 receptor A are present in the thyroid gland in intractable Graves disease. *Cellular immunology*, *281*(1), 85–90.

Li, R. N., Lin, Y. Z., Pan, Y. C., Lin, C. H., Tseng, C. C., Sung, W. Y., Wu, C. C., Ou, T. T., Tsai, W. C., & Yen, J. H. (2019). GADD45a and GADD45b Genes in Rheumatoid Arthritis and Systemic Lupus Erythematosus Patients. *Journal of clinical medicine*, *8*(6), 801.

Lillevang-Johansen, M., Abrahamsen, B., Jørgensen, H. L., Brix, T. H., & Hegedüs, L. (2019). Duration of Hyperthyroidism and Lack of Sufficient Treatment Are Associated with Increased Cardiovascular Risk. *Thyroid*, 29(3), 332–340.

Limbach, M., Saare, M., Tserel, L., Kisand, K., Eglit, T., Sauer, S., Axelsson, T., Syvänen, A. C., Metspalu, A., Milani, L., & Peterson, P. (2016). Epigenetic profiling in CD4+ and CD8+ T cells from Graves' disease patients reveals changes in genes associated with T cell receptor signaling. *Journal of autoimmunity*, 67, 46–56.

Lin, J. D., Wang, Y. H., Fang, W. F., Hsiao, C. J., Chagnaadorj, A., Lin, Y. F., Tang, K. T., & Cheng, C. W. (2016). Serum BAFF and thyroid autoantibodies in autoimmune thyroid disease. *Clinica chimica acta; international journal of clinical chemistry*, 462, 96–102.

Liu, J., Geng, R., Ni, S., Cai, L., Yang, S., Shao, F., & Bai, J. (2022). Pyroptosis-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with UCEC. *Molecular therapy. Nucleic acids*, *27*, 1036–1055.

Liu, L., Lu, H., Liu, Y., Liu, C., & Xun, C. (2016). Predicting relapse of Graves' disease following treatment with antithyroid drugs. *Experimental and therapeutic medicine*, 11(4), 1453–1458.

Liu, L., Tran, E., Zhao, Y., Huang, Y., Flavell, R., & Lu, B. (2005). Gadd45 beta and Gadd45 gamma are critical for regulating autoimmunity. *The Journal of experimental medicine*, *202*(10), 1341–1347.

Liu, M., Guo, Q., Wu, C., Sterlin, D., Goswami, S., Zhang, Y., Li, T., Bao, C., Shen, N., Fu, Q., & Zhang, X. (2019). Type I interferons promote the survival and proinflammatory properties of transitional B cells in systemic lupus erythematosus patients. *Cellular & molecular immunology*, *16*(4), 367–379.

Liu, N., Li, X., Liu, C., Zhao, Y., Cui, B., & Ning, G. (2010). The association of interleukin-1alpha and interleukin-1beta polymorphisms with the risk of Graves'

disease in a case-control study and meta-analysis. *Human immunology*, *71*(4), 397–401.

Liu, S., Miao, J. J., Zhou, X., Sun, Q., & Mao, X. M. (2022). High levels of thyroid hormones promote recurrence of Graves' disease via overexpression of B-cell-activating factor. *Journal of clinical laboratory analysis*, 36(10), e24701.

Liu, W., Wang, H. N., Gu, Z. H., Yang, S. Y., Ye, X. P., Pan, C. M., Zhao, S. X., Xue, L. Q., Xie, H. J., Yu, S. S., Guo, C. C., Du, W. H., Liang, J., Zhang, X. M., Yuan, G. Y., Li, C. G., Su, Q., Gao, G. Q., Song, H. D., & China Consortium for the Genetics of Autoimmune Thyroid Disease (2014). Identification of BACH2 as a susceptibility gene for Graves' disease in the Chinese Han population based on a three-stage genome-wide association study. *Human genetics*, 133(5), 661–671.

Liu, X., Qiang, W., Liu, X., Liu, L., Liu, S., Gao, A., Gao, S., & Shi, B. (2015). A second course of antithyroid drug therapy for recurrent Graves' disease: an experience in endocrine practice. *European journal of endocrinology*, 172(3), 321–326.

Liu, Y., Feng, Y., Tang, S., Zhang, L., Huang, Z., Shi, X., Fang, Y., Yang, J., Deng, X., Wang, L., Liu, X., & Yuan, H. (2022). Aberrant expression of inhibitory receptors on B cells in patients with Graves' disease. *Human immunology*, 83(2), 144–152.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), 550.

Lu, B., Ferrandino, A. F., & Flavell, R. A. (2004). Gadd45beta is important for perpetuating cognate and inflammatory signals in T cells. *Nature immunology*, *5*(1), 38–44.

Lu, K., Liu, L., Xu, X., Zhao, F., Deng, J., Tang, X., Wang, X., Zhao, B. Q., Zhang, X., & Zhao, Y. (2020). ADAMTS13 ameliorates inflammatory responses in experimental autoimmune encephalomyelitis. *Journal of neuroinflammation*, *17*(1), 67.

Macauley, M., Shawgi, M., Ali, T., Curry, A., Howe, K., Howell, E., Jefferson, E., James, A., Perros, P., & Petrides, G. S. (2018). Assessment of normal reference values for thyroid uptake of technetium-99m pertechnetate in a single centre UK population. *Nuclear medicine communications*, 39(9), 834–838.

Magri, F., Zerbini, F., Gaiti, M., Capelli, V., Ragni, A., Rotondi, M., & Chiovato, L. (2016). GENDER INFLUENCES THE CLINICAL PRESENTATION AND LONG-TERM OUTCOME OF GRAVES DISEASE. *Endocrine practice*, 22(11), 1336–1342.

Malkiel, S., Jeganathan, V., Wolfson, S., Manjarrez Orduño, N., Marasco, E., Aranow, C., Mackay, M., Gregersen, P. K., & Diamond, B. (2016). Checkpoints for Autoreactive B Cells in the Peripheral Blood of Lupus Patients Assessed by Flow Cytometry. *Arthritis & rheumatology (Hoboken, N.J.)*, *68*(9), 2210–2220.

Mamtani, M., Rovin, B., Brey, R., Camargo, J. F., Kulkarni, H., Herrera, M., Correa, P., Holliday, S., Anaya, J. M., & Ahuja, S. K. (2008). CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythaematosus. *Annals of the rheumatic diseases*, *67*(8), 1076–1083.

Marcocci, C., Leo, M., & Altea, M. A. (2012). Oxidative stress in graves' disease. *European thyroid journal*, *1*(2), 80–87.

Masetti, G., & Ludgate, M. (2020). Microbiome and Graves' Orbitopathy. *European thyroid journal*, 9(Suppl 1), 78–85.

Masiello, E., Veronesi, G., Gallo, D., Premoli, P., Bianconi, E., Rosetti, S., Cusini, C., Sabatino, J., Ippolito, S., Piantanida, E., Tanda, M. L., Chiovato, L., Wiersinga, W. M., & Bartalena, L. (2018). Antithyroid drug treatment for Graves' disease: baseline predictive models of relapse after treatment for a patient-tailored management. *Journal of endocrinological investigation*, 41(12), 1425–1432.

Massart, C., Hoste, C., Virion, A., Ruf, J., Dumont, J. E., & Van Sande, J. (2011). Cell biology of H2O2 generation in the thyroid: investigation of the control of dual oxidases (DUOX) activity in intact ex vivo thyroid tissue and cell lines. *Molecular and cellular endocrinology*, 343(1-2), 32–44.

Matsubara, T., Soh, J., Morita, M., Uwabo, T., Tomida, S., Fujiwara, T., Kanazawa, S., Toyooka, S., & Hirasawa, A. (2020). DV200 Index for Assessing RNA Integrity in Next-Generation Sequencing. *BioMed research international*, 2020, 9349132.

Matthews, D. C., & Syed, A. A. (2011). The role of TSH receptor antibodies in the management of Graves' disease. *European journal of internal medicine*, 22(3), 213–216.

Mauri, C., & Bosma, A. (2012). Immune regulatory function of B cells. *Annual review of immunology*, *30*, 221–241.

Mazza, E., Carlini, M., Flecchia, D., Blatto, A., Zuccarini, O., Gamba, S., Beninati, S., & Messina, M. (2008). Long-term follow-up of patients with hyperthyroidism due to Graves' disease treated with methimazole. Comparison of usual treatment schedule with drug discontinuation vs continuous treatment with low methimazole doses: a retrospective study. *Journal of endocrinological investigation*, 31(10), 866–872.

McGarry, T., Hanlon, M. M., Marzaioli, V., Cunningham, C. C., Krishna, V., Murray, K., Hurson, C., Gallagher, P., Nagpal, S., Veale, D. J., & Fearon, U. (2021). Rheumatoid arthritis $CD14^{+}$ monocytes display metabolic and inflammatory dysfunction, a phenotype that precedes clinical manifestation of disease. *Clinical & translational immunology*, *10*(1), e1237.

McKee, A., & Peyerl, F. (2012). TSI assay utilization: impact on costs of Graves' hyperthyroidism diagnosis. *The American journal of managed care*, 18(1), e1–e14.

McKinney, C., Merriman, M. E., Chapman, P. T., Gow, P. J., Harrison, A. A., Highton, J., Jones, P. B., McLean, L., O'Donnell, J. L., Pokorny, V., Spellerberg, M., Stamp, L. K., Willis, J., Steer, S., & Merriman, T. R. (2008). Evidence for an influence of chemokine ligand 3-like 1 (CCL3L1) gene copy number on susceptibility to rheumatoid arthritis. *Annals of the rheumatic diseases*, *67*(3), 409–413.

McLachlan, S. M., & Rapoport, B. (2014). Breaking tolerance to thyroid antigens: changing concepts in thyroid autoimmunity. *Endocrine reviews*, 35(1), 59–105.

Meinl, E., & Krumbholz, M. (2021). Endogenous soluble receptors sBCMA and sTACI: biomarker, immunoregulator and hurdle for therapy in multiple myeloma. *Current opinion in immunology*, *71*, 117–123.

Meling Stokland, A. E., Austdal, M., Nedrebø, B. G., Carlsen, S., Hetland, H. B., Breivik, L., Ueland, H. O., Watt, T., Cramon, P. K., Løvås, K., Husebye, E. S., & Ueland, G. Å. (2024). Outcomes of Patients With Graves Disease 25 Years After Initiating Antithyroid Drug Therapy. *The Journal of clinical endocrinology and metabolism*, 109(3), 827–836.

Meyer-Bahlburg, A., Andrews, S. F., Yu, K. O., Porcelli, S. A., & Rawlings, D. J. (2008). Characterization of a late transitional B cell population highly sensitive to

BAFF-mediated homeostatic proliferation. *The Journal of experimental medicine*, *205*(1), 155–168.

Mickey, R. M., & Greenland, S. (1989). The impact of confounder selection criteria on effect estimation. *American journal of epidemiology*, 129(1), 125–137.

Miltenyi® Microbead technology (*Magnetic cell separation strategies with MACS® Technology)*. Available at https://www.miltenyibiotec.com/GB-en/products/macs-cell-separation/macs-cell-separation-strategies.html (Accessed: 12 May 2024).

Mitchell, A. L., Cordell, H. J., Soemedi, R., Owen, K., Skinningsrud, B., Wolff, A. B., Ericksen, M., Undlien, D., Husebye, E., & Pearce, S. H. (2009). Programmed death ligand 1 (PD-L1) gene variants contribute to autoimmune Addison's disease and Graves' disease susceptibility. *The Journal of clinical endocrinology and metabolism*, 94(12), 5139–5145.

Mitsdoerffer, M., Lee, Y., Jäger, A., Kim, H. J., Korn, T., Kolls, J. K., Cantor, H., Bettelli, E., & Kuchroo, V. K. (2010). Proinflammatory T helper type 17 cells are effective B-cell helpers. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14292–14297.

Mizutori, Y., Chen, C. R., Latrofa, F., McLachlan, S. M., & Rapoport, B. (2009). Evidence that shed thyrotropin receptor A subunits drive affinity maturation of autoantibodies causing Graves' disease. *The Journal of clinical endocrinology and metabolism*, 94(3), 927–935.

Mohlin, E., Filipsson Nyström, H., & Eliasson, M. (2014). Long-term prognosis after medical treatment of Graves' disease in a northern Swedish population 2000-2010. *European journal of endocrinology*, 170(3), 419–427.

Morbach, H., Eichhorn, E. M., Liese, J. G., & Girschick, H. J. (2010). Reference values for B cell subpopulations from infancy to adulthood. *Clinical and experimental immunology*, 162(2), 271–279.

Mouat, I. C., Goldberg, E., & Horwitz, M. S. (2022). Age-associated B cells in autoimmune diseases. *Cellular and molecular life sciences : CMLS*, 79(8), 402.

Moudgil, K. D., & Choubey, D. (2011). Cytokines in autoimmunity: role in induction, regulation, and treatment. *Journal of interferon & cytokine research*, 31(10), 695–703.

Moura, R. A., Quaresma, C., Vieira, A. R., Gonçalves, M. J., Polido-Pereira, J., Romão, V. C., Martins, N., Canhão, H., & Fonseca, J. E. (2017). B-cell phenotype and IgD-CD27- memory B cells are affected by TNF-inhibitors and tocilizumab treatment in rheumatoid arthritis. *PloS one*, *12*(9), e0182927.

Murtagh F and Legendre P (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?. *Journal of Classification*, 31(3), 274-295.

Nakamura, H., Miyauchi, A., Miyawaki, N., & Imagawa, J. (2013). Analysis of 754 cases of antithyroid drug-induced agranulocytosis over 30 years in Japan. *The Journal of clinical endocrinology and metabolism*, 98(12), 4776–4783.

Nanba, T., Watanabe, M., Inoue, N., & Iwatani, Y. (2009). Increases of the Th1/Th2 cell ratio in severe Hashimoto's disease and in the proportion of Th17 cells in intractable Graves' disease. *Thyroid*, *19*(5), 495–501.

New England Biolabs. NEBNext® Low Input RNA Library Preparation (*NEBNext® Single Cell/Low Input RNA Library Prep Kit Protocol for Illumina).* Available at https://international.neb.com/-/media/nebus/files/manuals/manuale6420.pdf?rev=7731c29186c3494380bfbcec5e538827&hash=9255FF80686B984DE9367866C933CFA8 (Accessed: 1st June 2024).

Ngo, S. T., Steyn, F. J., & McCombe, P. A. (2014). Gender differences in autoimmune disease. *Frontiers in neuroendocrinology*, 35(3), 347–369.

NICE (2019) *Resource impact report: Thyroid disease: assessment and management (NG145)*. Available at: https://www.nice.org.uk/guidance/ng145/resources/resource-impact-report-pdf-6967385389 (Accessed: 11 May 2024).

Nordyke, R. A., Gilbert, F. I., Jr, & Harada, A. S. (1988). Graves' disease. Influence of age on clinical findings. *Archives of internal medicine*, 148(3), 626–631.

Nygaard, B., Hegedüs, L., Gervil, M., Hjalgrim, H., Hansen, B. M., Søe-Jensen, P., & Hansen, J. M. (1995). Influence of compensated radioiodine therapy on thyroid volume and incidence of hypothyroidism in Graves' disease. *Journal of internal medicine*, 238(6), 491–497.

O'Sullivan, B. J., Thomas, H. E., Pai, S., Santamaria, P., Iwakura, Y., Steptoe, R. J., Kay, T. W., & Thomas, R. (2006). IL-1 beta breaks tolerance through expansion of CD25+ effector T cells. *Journal of immunology (Baltimore, Md. : 1950)*, 176(12), 7278–7287.

Ochfeld, E., Hans, V., Marin, W., Ahsan, N., Morgan, G., Pachman, L. M., & Khojah, A. (2022). Coding joint: kappa-deleting recombination excision circle ratio and B cell activating factor level: predicting juvenile dermatomyositis rituximab response, a proof-of-concept study. *BMC rheumatology*, 6(1), 36.

Office for National statistics. Regional Ethnic Diversity (2022). https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest/#areas-of-england-and-wales-by-ethnicity (Accessed 13.03.2025)

O'Neill, S. K., Shlomchik, M. J., Glant, T. T., Cao, Y., Doodes, P. D., & Finnegan, A. (2005). Antigen-specific B cells are required as APCs and autoantibody-producing cells for induction of severe autoimmune arthritis. *Journal of immunology (Baltimore, Md. : 1950)*, 174(6), 3781–3788.

Owen, C. J., Kelly, H., Eden, J. A., Merriman, M. E., Pearce, S. H., & Merriman, T. R. (2007). Analysis of the Fc receptor-like-3 (FCRL3) locus in Caucasians with autoimmune disorders suggests a complex pattern of disease association. *The Journal of clinical endocrinology and metabolism*, 92(3), 1106–1111.

Palma FR, He C, Danes JM, Paviani V, Coelho DR, Gantner BN, Bonini MG. Mitochondrial Superoxide Dismutase: What the Established, the Intriguing, and the Novel Reveal About a Key Cellular Redox Switch. *Antioxid Redox Signal*. 2020 Apr 1;32(10):701-714.

Pan, Z., Zhu, T., Liu, Y., & Zhang, N. (2022). Corrigendum: Role of the CXCL13/CXCR5 axis in autoimmune diseases. *Frontiers in immunology*, 13, 1061939.

Pearce, S. H. S., Dayan, C., Wraith, D. C., Barrell, K., Olive, N., Jansson, L., Walker-Smith, T., Carnegie, C., Martin, K. F., Boelaert, K., Gilbert, J., Higham, C. E., Muller, I., Murray, R. D., Perros, P., Razvi, S., Vaidya, B., Wernig, F., & Kahaly, G. J. (2019). Antigen-Specific Immunotherapy with Thyrotropin Receptor Peptides in Graves' Hyperthyroidism: A Phase I Study. *Thyroid*, 29(7), 1003–1011.

Pedro, A. B., Romaldini, J. H., & Takei, K. (2011). Changes of serum cytokines in hyperthyroid Graves' disease patients at diagnosis and during methimazole treatment. *Neuroimmunomodulation*, 18(1), 45–51.

Peng, D., Xu, B., Wang, Y., Guo, H., & Jiang, Y. (2013). A high frequency of circulating th22 and th17 cells in patients with new onset graves' disease. *PloS one*, *8*(7), e68446.

Petersone, L., Edner, N. M., Ovcinnikovs, V., Heuts, F., Ross, E. M., Ntavli, E., Wang, C. J., & Walker, L. S. K. (2018). T Cell/B Cell Collaboration and Autoimmunity: An Intimate Relationship. *Frontiers in immunology*, 9, 1941.

Pichurin, P., Aliesky, H., Chen, C. R., Nagayama, Y., Rapoport, B., & McLachlan, S. M. (2003). Thyrotrophin receptor-specific memory T cell responses require normal B cells in a murine model of Graves' disease. *Clinical and experimental immunology*, 134(3), 396–402.

Pink, R. C., Wicks, K., Caley, D. P., Punch, E. K., Jacobs, L., & Carter, D. R. (2011). Pseudogenes: pseudo-functional or key regulators in health and disease?. *RNA (New York, N.Y.)*, *17*(5), 792–798.

Pinto, T. N. C., da Silva, C. C. B. M., Pinto, R. M. C., da Silva Duarte, A. J., Benard, G., & Fernandes, J. R. (2024). Tobacco exposure, but not aging, shifts the frequency of peripheral blood B cell subpopulations. *GeroScience*, 46(2), 2729–2738.

Piper, C. J. M., Wilkinson, M. G. L., Deakin, C. T., Otto, G. W., Dowle, S., Duurland, C. L., Adams, S., Marasco, E., Rosser, E. C., Radziszewska, A., Carsetti, R., Ioannou, Y., Beales, P. L., Kelberman, D., Isenberg, D. A., Mauri, C., Nistala, K., & Wedderburn, L. R. (2018). $CD19^{+}CD24^{hi}CD38^{hi}$ B Cells Are Expanded in Juvenile Dermatomyositis and Exhibit a Pro-Inflammatory Phenotype After Activation Through Toll-Like Receptor 7 and Interferon-α. *Frontiers in immunology*, *9*, 1372.

Płoski, R., Brand, O. J., Jurecka-Lubieniecka, B., Franaszczyk, M., Kula, D., Krajewski, P., Karamat, M. A., Simmonds, M. J., Franklyn, J. A., Gough, S. C., Jarząb, B., & Bednarczuk, T. (2010). Thyroid stimulating hormone receptor (TSHR) intron 1 variants are major risk factors for Graves' disease in three European Caucasian cohorts. *PloS one*, 5(11), e15512.

Punnonen, J., Aversa, G., Cocks, B. G., McKenzie, A. N., Menon, S., Zurawski, G., de Waal Malefyt, R., & de Vries, J. E. (1993). Interleukin 13 induces interleukin 4-independent IgG4 and IgE synthesis and CD23 expression by human B cells. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(8), 3730–3734.

Qiagen, RNeasy® process to extract RNA (*RNeasy Mini Handbook, Qiagen)*. Available at https://www.qiagen.com/us/resources/resourcedetail ?id=14e7cf6e-521a-4cf7-8cbc-bf9f6fa33e24&lang=en (Accessed:12 December 2022)).

Qiu, F., Liang, C. L., Liu, H., Zeng, Y. Q., Hou, S., Huang, S., Lai, X., & Dai, Z. (2017). Impacts of cigarette smoking on immune responsiveness: Up and down or upside down?. *Oncotarget*, 8(1), 268–284.

Quadbeck, B., Hoermann, R., Hahn, S., Roggenbuck, U., Mann, K., & Janssen, O. E. (2005). Binding, stimulating and blocking TSH receptor antibodies to the thyrotropin receptor as predictors of relapse of Graves' disease after withdrawal of antithyroid treatment. *Hormone and metabolic research*, 37(12), 745–750.

R Core Team (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org. Accessed: 1st June 2024.

Rabon, S., Burton, A. M., & White, P. C. (2016). Graves' disease in children: long-term outcomes of medical therapy. *Clinical endocrinology*, 85(4), 632–635.

Rapoport, B., & McLachlan, S. M. (2016). TSH Receptor Cleavage Into Subunits and Shedding of the A-Subunit; A Molecular and Clinical Perspective. *Endocrine reviews*, 37(2), 114–134.

Ranneh, Y., Ali, F., Akim, A.M., Hamid, H.A., Khazaai, H., & Fadel, A. (2017). Crosstalk between reactive oxygen species and pro-inflammatory markers in developing various chronic diseases: a review. *Appl Biol Chem* 60, 327–338.

Rea, I. M., Gibson, D. S., McGilligan, V., McNerlan, S. E., Alexander, H. D., & Ross, O. A. (2018). Age and Age-Related Diseases: Role of Inflammation Triggers and Cytokines. *Frontiers in immunology*, 9, 586.

Reed, G. F., Lynn, F., & Meade, B. D. (2002). Use of coefficient of variation in assessing variability of quantitative assays. *Clinical and diagnostic laboratory immunology*, 9(6), 1235–1239.

Reinwein, D., Benker, G., Lazarus, J. H., & Alexander, W. D. (1993). A prospective randomized trial of antithyroid drug dose in Graves' disease therapy. European Multicenter Study Group on Antithyroid Drug Treatment. *The Journal of clinical endocrinology and metabolism*, 76(6), 1516–1521.

Ren, W., Zhao, L., Sun, Y., Wang, X., & Shi, X. (2023). HMGB1 and Toll-like receptors: potential therapeutic targets in autoimmune diseases. *Molecular medicine (Cambridge, Mass.)*, *29*(1), 117.

Riguetto, C. M., Neto, A. M., Tambascia, M. A., & Zantut-Wittmann, D. E. (2019). The relationship between quality of life, cognition, and thyroid status in Graves' disease. *Endocrine*, 63(1), 87–93.

Rivkees, S. A., & Mattison, D. R. (2009). Propylthiouracil (PTU) Hepatoxicity in Children and Recommendations for Discontinuation of Use. *International journal of pediatric endocrinology*, 2009, 132041.

Rodríguez-Carrio, J., Alperi-López, M., López, P., Ballina-García, F. J., & Suárez, A. (2018). Profiling of B-Cell Factors and Their Decoy Receptors in Rheumatoid Arthritis: Association With Clinical Features and Treatment Outcomes. *Frontiers in immunology*, 9, 2351.

Rojas, M., Restrepo-Jiménez, P., Monsalve, D. M., Pacheco, Y., Acosta-Ampudia, Y., Ramírez-Santana, C., Leung, P. S. C., Ansari, A. A., Gershwin, M. E., & Anaya, J. M. (2018). Molecular mimicry and autoimmunity. *Journal of autoimmunity*, 95, 100–123.

Rong, H., He, X., Wang, L., Bai, M., Jin, T., Wang, Y., Yang, W., He, Y., & Yuan, D. (2020). Association between IL1B polymorphisms and the risk of rheumatoid arthritis. *International immunopharmacology*, *83*, 106401.

Rosser, E. C., & Mauri, C. (2021). The emerging field of regulatory B cell immunometabolism. *Cell metabolism*, 33(6), 1088–1097.

Rosser, E. C., & Mauri, C. (2015). Regulatory B cells: origin, phenotype, and function. *Immunity*, 42(4), 607–612.

Rosser, E. C., Oleinika, K., Tonon, S., Doyle, R., Bosma, A., Carter, N. A., Harris, K. A., Jones, S. A., Klein, N., & Mauri, C. (2014). Regulatory B cells are induced by gut microbiota-driven interleukin-1β and interleukin-6 production. *Nature medicine*, *20*(11), 1334–1339.

Rousseeuw, P.J. (1987) Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.

Rowland, S. L., Leahy, K. F., Halverson, R., Torres, R. M., & Pelanda, R. (2010). BAFF receptor signaling aids the differentiation of immature B cells into transitional B cells following tonic BCR signaling. *Journal of immunology (Baltimore, Md. : 1950)*, *185*(8), 4570–4581.

Ruschil, C., Gabernet, G., Lepennetier, G., Heumos, S., Kaminski, M., Hracsko, Z., Irmler, M., Beckers, J., Ziemann, U., Nahnsen, S., Owens, G. P., Bennett, J. L., Hemmer, B., & Kowarik, M. C. (2020). Specific Induction of Double Negative B Cells During Protective and Pathogenic Immune Responses. *Frontiers in immunology*, *11*, 606338.

Saeed, M. H., Kurosh, K., Zahra, A., Hossein, D. M., Davood, R., & Ataollahi, M. R. (2021). Decreased serum levels of IL-27and IL-35 in patients with Graves disease. *Archives of endocrinology and metabolism*, 64(5), 521–527.

Salazar-Camarena, D. C., Palafox-Sánchez, C. A., Cruz, A., Marín-Rosales, M., & Muñoz-Valle, J. F. (2020). Analysis of the receptor BCMA as a biomarker in systemic lupus erythematosus patients. *Scientific reports*, *10*(1), 6236.

Sałkowska, A., Karaś, K., Karwaciak, I., Walczak-Drzewiecka, A., Krawczyk, M., Sobalska-Kwapis, M., Dastych, J., & Ratajewski, M. (2020). Identification of Novel Molecular Markers of Human Th17 Cells. *Cells*, *9*(7), 1611.

Salvi, M., Colucci, G., Masetti, G., Covelli, D., Muller, I., Koehling, H., & Biscarini, F. (2019). The randomised probiotic trial of indigo study (investigation of novel biomarkers and definition of role of the microbiome in Graves. *Endocrine Abstracts* (Vol. 63). Bioscientifica.

Salvi, M., Girasole, G., Pedrazzoni, M., Passeri, M., Giuliani, N., Minelli, R., Braverman, L. E., & Roti, E. (1996). Increased serum concentrations of interleukin-6

(IL-6) and soluble IL-6 receptor in patients with Graves' disease. *The Journal of clinical endocrinology and metabolism*, 81(8), 2976–2979.

Saint-André, V., Charbit, B., Biton, A., Rouilly, V., Possémé, C., Bertrand, A., Rotival, M., Bergstedt, J., Patin, E., Albert, M. L., Quintana-Murci, L., Duffy, D., & Milieu Intérieur Consortium (2024). Smoking changes adaptive immunity with persistent effects. *Nature*, 626(8000), 827–835.

Sanz, I., Wei, C., Jenks, S. A., Cashman, K. S., Tipton, C., Woodruff, M. C., Hom, J., & Lee, F. E. (2019). Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Frontiers in immunology*, *10*, 2458.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, OP., Tiwari, A., Er, MJ., Ding, W., Lin, CT. (2017). A review of clustering techniques and developments. *Neurocomputing*. 267, 664-681.

Schierbeck, H., Pullerits, R., Pruunsild, C., Fischer, M., Holzinger, D., Laestadius, Å., Sundberg, E., & Harris, H. E. (2013). HMGB1 levels are increased in patients with juvenile idiopathic arthritis, correlate with early onset of disease, and are independent of disease duration. *The Journal of rheumatology*, *40*(9), 1604–1613.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., & Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology*, 7, 3.

Schleusener, H., Schwander, J., Fischer, C., Holle, R., Holl, G., Badenhoop, K., Hensen, J., Finke, R., Bogner, U., & Mayr, W. R. (1989). Prospective multicentre study on the prediction of relapse after antithyroid drug treatment in patients with Graves' disease. *Acta endocrinologica*, 120(6), 689–701.

Segundo, C., Rodríguez, C., García-Poley, A., Aguilar, M., Gavilán, I., Bellas, C., & Brieva, J. A. (2001). Thyroid-infiltrating B lymphocytes in Graves' disease are related to marginal zone and memory B cell compartments. *Thyroid*, 11(6), 525–530.

Sha, Y., Zmijewski, J., Xu, Z., & Abraham, E. (2008). HMGB1 develops enhanced proinflammatory activity by binding to cytokines. *Journal of immunology (Baltimore, Md. : 1950)*, *180*(4), 2531–2537.

Shehjar, F., Afroze, D., Misgar, R. A., Malik, S. A., & Laway, B. A. (2018). Association of polymorphic variants of IL-1β and IL-1RN genes in the development of Graves' disease in Kashmiri population (North India). *Human immunology*, *79*(4), 228–232.

Simmonds, M. J., Kavvoura, F. K., Brand, O. J., Newby, P. R., Jackson, L. E., Hargreaves, C. E., Franklyn, J. A., & Gough, S. C. (2014). Skewed X chromosome inactivation and female preponderance in autoimmune thyroid disease: an association study and meta-analysis. *The Journal of clinical endocrinology and metabolism*, 99(1), E127–E131.

Sjölin, G., Holmberg, M., Törring, O., Byström, K., Khamisi, S., de Laval, D., Abraham-Nordling, M., Calissendorff, J., Lantz, M., Hallengren, B., Filipsson Nyström, H., & Wallin, G. (2019). The Long-Term Outcome of Treatment for Graves' Hyperthyroidism. *Thyroid*, 29(11), 1545–1557.

Skov, J., Calissendorff, J., Eriksson, D., Magnusson, P., Kämpe, O., Bensing, S., & Kuja-Halkola, R. (2021). Limited Genetic Overlap Between Overt Hashimoto's Thyroiditis and Graves' Disease in Twins: A Population-based Study. *The Journal of clinical endocrinology and metabolism*, 106(4), 1101–1110.

Smets, I., Prezzemolo, T., Imbrechts, M., Mallants, K., Mitera, T., Humblet-Baron, S., Dubois, B., Matthys, P., Liston, A., & Goris, A. (2021). Treatment-Induced BAFF Expression and B Cell Biology in Multiple Sclerosis. *Frontiers in immunology*, *12*, 676619.

Smulski, C. R., & Eibel, H. (2018). BAFF and BAFF-Receptor in B Cell Selection and Survival. *Frontiers in immunology*, *9*, 2285.

Song, H., Fang, F., Tomasson, G., Arnberg, F. K., Mataix-Cols, D., Fernández de la Cruz, L., Almqvist, C., Fall, K., & Valdimarsdóttir, U. A. (2018). Association of Stress-Related Disorders With Subsequent Autoimmune Disease. *JAMA*, 319(23), 2388–2400.

Song, R. H., Shao, X. Q., Li, L., Wang, W., & Zhang, J. A. (2017). Copy number variations exploration of multiple genes in Graves' disease. *Medicine*, *96*(4), e5866.

Sottini, A., Serana, F., Bertoli, D., Chiarini, M., Valotti, M., Vaglio Tessitore, M., & Imberti, L. (2014). Simultaneous quantification of T-cell receptor excision circles

(TRECs) and K-deleting recombination excision circles (KRECs) by real-time PCR. *Journal of visualized experiments : JoVE*, (94), 52184.

Stefan, M., Jacobson, E. M., Huber, A. K., Greenberg, D. A., Li, C. W., Skrabanek, L., Conception, E., Fadlalla, M., Ho, K., & Tomer, Y. (2011). Novel variant of thyroglobulin promoter triggers thyroid autoimmunity through an epigenetic interferon alpha-modulated mechanism. *The Journal of biological chemistry*, 286(36), 31168–31179.

Steri, M., Orrù, V., Idda, M. L., Pitzalis, M., Pala, M., Zara, I., Sidore, C., Faà, V., Floris, M., Deiana, M., Asunis, I., Porcu, E., Mulas, A., Piras, M. G., Lobina, M., Lai, S., Marongiu, M., Serra, V., Marongiu, M., Sole, G., … Cucca, F. (2017). Overexpression of the Cytokine BAFF and Autoimmunity Risk. *The New England journal of medicine*, 376(17), 1615–1626.

Stewart, A., Ng, J. C., Wallis, G., Tsioligka, V., Fraternali, F., & Dunn-Walters, D. K. (2021). Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways. *Frontiers in immunology*, *12*, 602539.

Stożek, K., Grubczak, K., Marolda, V., Eljaszewicz, A., Moniuszko, M., & Bossowski, A. (2020). Lower proportion of CD19+IL-10+ and CD19+CD24+CD27+ but not CD1d+CD5+CD19+CD24+CD27+ IL-10+ B cells in children with autoimmune thyroid diseases. *Autoimmunity*, 53(1), 46–55.

Strawa, K., Markowska, A., Miśkiewicz, P., Kuś, A., Ambroziak, U., Szymański, K., Zbiec, R., Spólnicka, M., Krajewski, P., Bednarczuk, T., & Płoski, R. (2014). Increased concentration of T-cell receptor rearrangement excision circles (TREC) in peripheral blood in Graves' disease. *Clinical endocrinology*, 81(5), 769–774.

Struja, T., Kaeslin, M., Boesiger, F., Jutzi, R., Imahorn, N., Kutz, A., Bernasconi, L., Mundwiler, E., Mueller, B., Christ-Crain, M., Meienberg, F., Ebrahimi, F., Henzen, C., Fischli, S., Kraenzlin, M., Meier, C., & Schuetz, P. (2017). External validation of the GREAT score to predict relapse risk in Graves' disease: results from a multicenter, retrospective study with 741 patients. *European journal of endocrinology*, 176(4), 413–419.

Suurmond, J., Atisha-Fregoso, Y., Barlev, A. N., Calderon, S. A., Mackay, M. C., Aranow, C., & Diamond, B. (2019). Patterns of ANA+ B cells for SLE patient stratification. *JCI insight*, *4*(9), e127885.

Suzuki, N., Noh, J. Y., Yoshimura, R., Mikura, K., Kinoshita, A., Suzuki, A., Mitsumatsu, T., Hoshiyama, A., Fukushita, M., Matsumoto, M., Yoshihara, A., Watanabe, N., Sugino, K., & Ito, K. (2021). Does Age or Sex Relate to Severity or Treatment Prognosis in Graves' Disease? *Thyroid*, 31(9), 1409–1415.

Taher, T. E., Ong, V. H., Bystrom, J., Hillion, S., Simon, Q., Denton, C. P., Pers, J. O., Abraham, D. J., & Mageed, R. A. (2018). Association of Defective Regulation of Autoreactive Interleukin-6-Producing Transitional B Lymphocytes With Disease in Patients With Systemic Sclerosis. *Arthritis & rheumatology (Hoboken, N.J.)*, *70*(3), 450–461.

Takemura, S., Klimiuk, P. A., Braun, A., Goronzy, J. J., & Weyand, C. M. (2001). T cell activation in rheumatoid synovium is B cell dependent. *Journal of immunology (Baltimore, Md.: 1950)*, 167(8), 4710–4718.

Takeoka, K., Watanabe, M., Matsuzuka, F., Miyauchi, A., & Iwatani, Y. (2004). Increase of serum interleukin-10 in intractable Graves' disease. *Thyroid*, 14(3), 201–205.

Tang, D., Kang, R., Livesey, K. M., Zeh, H. J., 3rd, & Lotze, M. T. (2011). High mobility group box 1 (HMGB1) activates an autophagic response to oxidative stress. *Antioxidants & redox signaling*, *15*(8), 2185–2195.

Tangye, S. G., Avery, D. T., Deenick, E. K., & Hodgkin, P. D. (2003). Intrinsic differences in the proliferation of naive and memory human B cells as a mechanism for enhanced secondary immune responses. *Journal of immunology (Baltimore, Md.: 1950)*, 170(2), 686–694.

Taurog A. (1976). The mechanism of action of the thioureylene antithyroid drugs. *Endocrinology*, 98(4), 1031–1046.

Teniente-Serra, A., Soldevila, B., Quirant-Sánchez, B., Fernández, M. A., Ester Condins, A., Puig-Domingo, M., Pujol-Borrell, R., & Martínez-Cáceres, E. M. (2019). Distinct pattern of peripheral lymphocyte subsets in Graves' disease with persistency of anti-TSHR autoantibodies. *Autoimmunity*, *52*(5-6), 220–227.

Ter Horst, R., Jaeger, M., Smeekens, S. P., Oosting, M., Swertz, M. A., Li, Y., Kumar, V., Diavatopoulos, D. A., Jansen, A. F. M., Lemmers, H., Toenhake-Dijkstra, H., van Herwaarden, A. E., Janssen, M., van der Molen, R. G., Joosten, I., Sweep, F.

C. G. J., Smit, J. W., Netea-Maier, R. T., Koenders, M. M. J. F., Xavier, R. J., … Netea, M. G. (2016). Host and Environmental Factors Influencing Individual Human Cytokine Responses. *Cell*, 167(4), 1111–1124.e13.

Thaler, F. S., Laurent, S. A., Huber, M., Mulazzani, M., Dreyling, M., Ködel, U., Kümpfel, T., Straube, A., Meinl, E., & von Baumgarten, L. (2017). Soluble TACI and soluble BCMA as biomarkers in primary central nervous system lymphoma. *Neuro-oncology*, 19(12), 1618–1627.

Tighe, P. J., Ryder, R. R., Todd, I., & Fairclough, L. C. (2015). ELISA in the multiplex era: potentials and pitfalls. Proteomics. Clinical applications, 9(3-4), 406–422.

Tipton, C. M., Fucile, C. F., Darce, J., Chida, A., Ichikawa, T., Gregoretti, I., Schieferl, S., Hom, J., Jenks, S., Feldman, R. J., Mehr, R., Wei, C., Lee, F. E., Cheung, W. C., Rosenberg, A. F., & Sanz, I. (2015). Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nature immunology*, *16*(7), 755–765.

Torimoto, K., Okada, Y., Nakayamada, S., Kubo, S., Kurozumi, A., Narisawa, M., & Tanaka, Y. (2022). Comprehensive immunophenotypic analysis reveals the pathological involvement of Th17 cells in Graves' disease. *Scientific reports*, 12(1), 16880.

Tosato, G., & Jones, K. D. (1990). Interleukin-1 induces interleukin-6 production in peripheral blood monocytes. *Blood*, *75*(6), 1305–1310.

Tötterman, T. H., Karlsson, F. A., Bengtsson, M., & Mendel-Hartvig, I. (1987). Induction of circulating activated suppressor-like T cells by methimazole therapy for Graves' disease. *The New England journal of medicine*, 316(1), 15–22.

Toung, J. M., Morley, M., Li, M., & Cheung, V. G. (2011). RNA-sequence analysis of human B-cells. *Genome research*, 21(6), 991–998.

Tozzoli, R., Bagnasco, M., Giavarina, D., & Bizzaro, N. (2012). TSH receptor autoantibody immunoassay in patients with Graves' disease: improvement of diagnostic accuracy over different generations of methods. Systematic review and meta-analysis. *Autoimmunity reviews*, 12(2), 107–113.

Traianos, E. Y., Locke, J., Lendrem, D., Bowman, S., Hargreaves, B., Macrae, V., UK primary Sjögren’s syndrome registry, Tarn, J. R., & Ng, W. F. (2020). Serum CXCL13

levels are associated with lymphoma risk and lymphoma occurrence in primary Sjögren's syndrome. *Rheumatology international*, 40(4), 541–548.

Tripathi, S. K., Välikangas, T., Shetty, A., Khan, M. M., Moulder, R., Bhosale, S. D., Komsi, E., Salo, V., De Albuquerque, R. S., Rasool, O., Galande, S., Elo, L. L., & Lahesmaa, R. (2019). Quantitative Proteomics Reveals the Dynamic Protein Landscape during Initiation of Human Th17 Cell Polarization. *iScience*, 11, 334–355.

Tsai, C. C., Cheng, C. Y., Liu, C. Y., Kao, S. C., Kau, H. C., Hsu, W. M., & Wei, Y. H. (2009). Oxidative stress in patients with Graves' ophthalmopathy: relationship between oxidative DNA damage and clinical evolution. *Eye (London, England),* 23(8), 1725–1730.

Tsukamoto, M., Seta, N., Yoshimoto, K., Suzuki, K., Yamaoka, K., & Takeuchi, T. (2017). CD14^{bright}CD16+ intermediate monocytes are induced by interleukin-10 and positively correlate with disease activity in rheumatoid arthritis. *Arthritis research & therapy*, *19*(1), 28.

Tun, N. N., Beckett, G., Zammitt, N. N., Strachan, M. W., Seckl, J. R., & Gibb, F. W. (2016). Thyrotropin Receptor Antibody Levels at Diagnosis and After Thionamide Course Predict Graves' Disease Relapse. *Thyroid*, 26(8), 1004–1009.

Turan E. (2019). Evaluation of neutrophil-to-lymphocyte ratio and hematologic parameters in patients with Graves' disease. *Bratislavske lekarske listy*, *120*(6), 476–480.

UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* 51:D523–D531.

Vaidya, B., Imrie, H., Perros, P., Young, E. T., Kelly, W. F., Carr, D., Large, D. M., Toft, A. D., McCarthy, M. I., Kendall-Taylor, P., & Pearce, S. H. (1999). The cytotoxic T lymphocyte antigen-4 is a major Graves' disease locus. *Human molecular genetics*, *8*(7), 1195–1199.

van Balkum, M., Schreurs, M. W. J., Visser, W. E., Peeters, R. P., & Dik, W. A. (2023). Comparison of two different TSH-receptor antibody assays: A clinical practice study. *Heliyon*, 9(12), e22468.

Van der Weerd, K., Van Hagen, P. M., Schrijver, B., Kwekkeboom, D. J., De Herder, W. W., Ten Broek, M. R., Postema, P. T., Van Dongen, J. J., Staal, F. J., & Dik, W.

A. (2013). The peripheral blood compartment in patients with Graves' disease: activated T lymphocytes and increased transitional and pre-naive mature B lymphocytes. *Clinical and experimental immunology*, 174(2), 256–264.

van Lieshout, J. M., Mooij, C. F., van Trotsenburg, A. S. P., & Zwaveling-Soonawala, N. (2021). Methimazole-induced remission rates in pediatric Graves' disease: a systematic review. *European journal of endocrinology*, 185(2), 219–229.

van Zelm, M. C., Szczepanski, T., van der Burg, M., & van Dongen, J. J. (2007). Replication history of B lymphocytes reveals homeostatic proliferation and extensive antigen-induced B cell expansion. *The Journal of experimental medicine*, 204(3), 645–655.

Vannucchi, G., Covelli, D., Currò, N., Dazzi, D., Maffini, A., Campi, I., Bonara, P., Guastella, C., Pignataro, L., Ratiglia, R., Beck-Peccoz, P., & Salvi, M. (2012). Serum BAFF concentrations in patients with Graves' disease and orbitopathy before and after immunosuppressive therapy. *The Journal of clinical endocrinology and metabolism*, 97(5), E755–E759.

Vejrazkova, D., Vcelak, J., Vaclavikova, E., Vankova, M., Zajickova, K., Vrbikova, J., Duskova, M., Pacesova, P., Novak, Z., & Bendlova, B. (2021). Recurrence of Graves' Disease: What Genetics of HLA and PTPN22 Can Tell Us. *Frontiers in endocrinology*, 12, 761077.

Velaga, M. R., Wilson, V., Jennings, C. E., Owen, C. J., Herington, S., Donaldson, P. T., Ball, S. G., James, R. A., Quinton, R., Perros, P., & Pearce, S. H. (2004). The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease. *The Journal of clinical endocrinology and metabolism*, 89(11), 5862–5865.

Vigone, M. C., Peroni, E., Di Frenna, M., Mora, S., Barera, G., & Weber, G. (2020). "Block-and-replace" treatment in Graves' disease: experience in a cohort of pediatric patients. *Journal of endocrinological investigation*, 43(5), 595–600.

Villagelin, D., Cooper, D. S., & Burch, H. B. (2024). A 2023 International Survey of Clinical Practice Patterns in the Management of Graves' Disease: A Decade of Change. *The Journal of clinical endocrinology and metabolism*, dgae222.

Vincent, F. B., Kandane-Rathnayake, R., Koelmeyer, R., Hoi, A. Y., Harris, J., Mackay, F., & Morand, E. F. (2019). Analysis of serum B cell-activating factor from the tumor necrosis factor family (BAFF) and its soluble receptors in systemic lupus erythematosus. *Clinical & translational immunology*, 8(4), e01047.

Vita, R., Lapa, D., Trimarchi, F., & Benvenga, S. (2015). Stress triggers the onset and the recurrences of hyperthyroidism in patients with Graves' disease. *Endocrine*, 48(1), 254–263.

Vitti, P., Rago, T., Chiovato, L., Pallini, S., Santini, F., Fiore, E., Rocchi, R., Martino, E., & Pinchera, A. (1997). Clinical features of patients with Graves' disease undergoing remission after antithyroid drug treatment. *Thyroid*, 7(3), 369–375.

Veldhoen, M., Hocking, R. J., Atkins, C. J., Locksley, R. M., & Stockinger, B. (2006). TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity*, *24*(2), 179–189.

Vos, X. G., Endert, E., Zwinderman, A. H., Tijssen, J. G., & Wiersinga, W. M. (2016). Predicting the Risk of Recurrence Before the Start of Antithyroid Drug Therapy in Patients With Graves' Hyperthyroidism. *The Journal of clinical endocrinology and metabolism*, 101(4), 1381–1389.

Wahren-Herlenius, M., & Dörner, T. (2013). Immunopathogenic mechanisms of systemic autoimmune disease. *Lancet (London, England)*, 382(9894), 819–831.

Wang, M. T., Lee, W. J., Huang, T. Y., Chu, C. L., & Hsieh, C. H. (2014). Antithyroid drug-related hepatotoxicity in hyperthyroidism patients: a population-based cohort study. *British journal of clinical pharmacology*, 78(3), 619–629.

Wang, P. W., Chen, I. Y., Juo, S. H., Hsi, E., Liu, R. T., & Hsieh, C. J. (2013). Genotype and phenotype predictors of relapse of graves' disease after antithyroid drug withdrawal. *European thyroid journal*, 1(4), 251–258.

Wang, F., Yan, T., Chen, L., Chen, X., Liu, T., Shen, S., Li, T., Gao, L., Wang, T., Sun, J., Liu, C., Wu, H., Zhang, X., & Chen, L. (2012). Involvement of inducible costimulator ligand (ICOSL) expression in thyroid tissue in hyperthyroidism of Graves' disease patients. *Journal of clinical immunology*, 32(6), 1253–1261.

Wang, J., Peng, H., Tian, J., Ma, J., Tang, X., Rui, K., Tian, X., Wang, Y., Chen, J., Lu, L., Xu, H., & Wang, S. (2016). Upregulation of long noncoding RNA TMEVPG1

enhances T helper type 1 cell response in patients with Sjögren syndrome. *Immunologic research*, *64*(2), 489–496.

Wang, J., Manick, B., Renelt, M., Hansen, L., Person, A., Kalabokis, V., Suin, J., Wu, G. (2019). VSTM4 is a novel negative regulator of T cell activation. *J Immunol.* 202(1 supplement):124.4.

Wang, S., Wang, J., Kumar, V., Karnell, J. L., Naiman, B., Gross, P. S., Rahman, S., Zerrouki, K., Hanna, R., Morehouse, C., Holoweckyj, N., Liu, H., Autoimmunity Molecular Medicine Team, Manna, Z., Goldbach-Mansky, R., Hasni, S., Siegel, R., Sanjuan, M., Streicher, K., Cancro, M. P., … Ettinger, R. (2018). IL-21 drives expansion and plasma cell differentiation of autoreactive CD11chiT-bet+ B cells in SLE. *Nature communications*, 9(1), 1758.

Wang, X., Huang, J., Zhang, A., Fang, C., Ma, Q., & Jiang, P. (2021). Altered expression profile of BAFF receptors on peripheral blood B lymphocytes in Graves' disease. *BMC endocrine disorders*, *21*(1), 88.

Wang, Y., Fang, S., & Zhou, H. (2023). Pathogenic role of Th17 cells in autoimmune thyroid disease and their underlying mechanisms. *Best practice & research. Clinical endocrinology & metabolism*, *37*(2), 101743.

Wardemann, H., Yurasov, S., Schaefer, A., Young, J. W., Meffre, E., & Nussenzweig, M. C. (2003). Predominant autoantibody production by early human B cell precursors. *Science (New York, N.Y.)*, *301*(5638), 1374–1377.

Weetman, A. P., McGregor, A. M., & Hall, R. (1983). Methimazole inhibits thyroid autoantibody production by an action on accessory cells. *Clinical immunology and immunopathology*, 28(1), 39–45.

Weetman, A. P., Ratanachaiyavong, S., Middleton, G. W., Love, W., John, R., Owen, G. M., Darke, C., Lazarus, J. H., Hall, R., & McGregor, A. M. (1986). Prediction of outcome in Graves' disease after carbimazole treatment. *The Quarterly journal of medicine*, 59(228), 409–419.

Weetman, A. P., Ajjan, R. A., & Watson, P. F. (1997). Cytokines and Graves' disease. *Bailliere's clinical endocrinology and metabolism*, 11(3), 481–497.

Weetman, A. P., Tandon, N., & Morgan, B. P. (1992). Antithyroid drugs and release of inflammatory mediators by complement-attacked thyroid cells. *Lancet (London, England)*, 340(8820), 633–636.

Wenzek, C., Boelen, A., Westendorf, A. M., Engel, D. R., Moeller, L. C., & Führer, D. (2022). The interplay of thyroid hormones and the immune system - where we stand and why we need to know about it. *European journal of endocrinology*, 186(5), R65–R77.

Williamson, S., & Greene, S. A. (2010). Incidence of thyrotoxicosis in childhood: a national population based study in the UK and Ireland. *Clinical endocrinology*, 72(3), 358–363.

Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*, 7, 1338.

Wood, C. L., Cole, M., Donaldson, M., Dunger, D. B., Wood, R., Morrison, N., Matthews, J. N. S., Pearce, S. H. S., & Cheetham, T. D. (2020). Randomised trial of block and replace vs dose titration thionamide in young people with thyrotoxicosis. *European journal of endocrinology*, 183(6), 637–645.

Yang, M., Sun, L., Wang, S., Ko, K. H., Xu, H., Zheng, B. J., Cao, X., & Lu, L. (2010). Novel function of B cell-activating factor in the induction of IL-10-producing regulatory B cells. *Journal of immunology (Baltimore, Md. : 1950)*, 184(7), 3321–3325.

Yan, M., Wang, H., Chan, B., Roose-Girma, M., Erickson, S., Baker, T., Tumas, D., Grewal, I. S., & Dixit, V. M. (2001). Activation and accumulation of B cells in TACI-deficient mice. *Nature immunology*, *2*(7), 638–643.

Yin, X., Sachidanandam, R., Morshed, S., Latif, R., Shi, R., & Davies, T. F. (2014). mRNA-Seq reveals novel molecular mechanisms and a robust fingerprint in Graves' disease. *The Journal of clinical endocrinology and metabolism*, 99(10), E2076–E2083.

Yoshimoto, K., Tanaka, M., Kojima, M., Setoyama, Y., Kameda, H., Suzuki, K., Tsuzaka, K., Ogawa, Y., Tsubota, K., Abe, T., & Takeuchi, T. (2011). Regulatory mechanisms for the production of BAFF and IL-6 are impaired in monocytes of patients of primary Sjögren's syndrome. *Arthritis research & therapy*, 13(5), R170.

You, X., Zhang, R., Shao, M., He, J., Chen, J., Liu, J., Zhang, X., Liu, X., Jia, R., Sun, X., & Li, Z. (2020). Double Negative B Cell Is Associated With Renal Impairment in Systemic Lupus Erythematosus and Acts as a Marker for Nephritis Remission. *Frontiers in medicine*, *7*, 85.

Zeng, Y., Zhang, Y., Lin, Y., Wang, X., Chen, Q., Huang, Q., Wang, J., Jiang, L., & Xiao, Y. (2021). The CXCL13 chemokine serves as a potential biomarker to diagnose systemic lupus erythematosus with disease activity. *Clinical and experimental medicine*, 21(4), 611–619.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C. G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., To, J. K., Laird, M. R., … Flicek, P. (2018). Ensembl 2018. *Nucleic acids research*, 46(D1), D754–D761.

Zarković M. (2012). The role of oxidative stress on the pathogenesis of graves' disease. *Journal of thyroid research*, 2012, 302537.

Zazzeroni, F., Papa, S., Algeciras-Schimnich, A., Alvarez, K., Melis, T., Bubici, C., Majewski, N., Hay, N., De Smaele, E., Peter, M. E., & Franzoso, G. (2003). Gadd45 beta mediates the protective effects of CD40 costimulation against Fas-induced apoptosis. *Blood*, *102*(9), 3270–3279.

Zha, B., Wang, L., Liu, X., Liu, J., Chen, Z., Xu, J., Sheng, L., Li, Y., & Chu, Y. (2012). Decrease in proportion of CD19+ CD24(hi) CD27+ B cells and impairment of their suppressive function in Graves' disease. *PloS one*, *7*(11), e49835.

Zhang, B., Zhong, Q., Chen, X., Wu, X., Sha, R., Song, G., Zhang, C., & Chen, X. (2020). Neuroprotective Effects of Celastrol on Transient Global Cerebral Ischemia Rats via Regulating HMGB1/NF-κB Signaling Pathway. *Frontiers in neuroscience*, *14*, 847.

Zhang, S., Zhang, M., Zhang, L., Wang, Z., Tang, S., Yang, X., Li, Z., Feng, J., & Qin, X. (2024). Identification of Y–linked biomarkers and exploration of immune infiltration of normal-appearing gray matter in multiple sclerosis by bioinformatic analysis. *Heliyon*, *10*(6), e28085.

Ziliotto, N., Bernardi, F., Jakimovski, D., Baroni, M., Marchetti, G., Bergsland, N., Ramasamy, D. P., Weinstock-Guttman, B., Schweser, F., Zamboni, P., Ramanathan,

M., & Zivadinov, R. (2018). Hemostasis biomarkers in multiple sclerosis. *European journal of neurology*, *25*(9), 1169–1176.

Zuhur, S. S., Elbuken, G., Yildiz, I., Kadioglu, P., Erol, S., Sahin, S., Kilinc, F., Akbaba, G., Topcu, B., & Altuntas, Y. (2019). External Validation of the GREAT Score in Turkish Patients with Graves' Hyperthyroidism Treated with the Titration Regimen Method of Antithyroid Drugs: A Multicenter Study. *Hormone and metabolic research*, 51(10), 627–633.